Artificial Intelligence Empowered Communication and Computation for Edge-Enabled IoT Systems


Huaming Wu


Chaogang Tang


Joel J. P. C. Rodrigues


Houbing Herbert Song


Huijun Tang


Lei Wang

The rapid advancement of Information and Communication Technologies (ICTs) has a profound impact on Internet of Things (IoT). However, the exponentially growing data traffic in the core or backbone network poses a significant challenge to achieving ultra-reliable low-latency communications among diverse communication entities. Moreover, various IoT devices equipped with multi-core processors possess the capability to not only acquire data but also perform various tasks. The increasingly complex tasks require more computing resources, which cannot be satisfied by the local computing capabilities of IoT devices themselves, so data offloading and task outsourcing have become prevalent in recent years.

Edge computing, a pivotal technology for enhancing IoT system performance, integrates communication and computing resources at the network edge to deliver ultra-reliable and low-latency responses for a wide range of IoT applications. Current research in edge-enabled IoT systems primarily focuses on task offloading from IoT devices and resource scheduling at the network edge. To simplify these processes, several assumptions are often made, limiting in-depth exploration of the topic. However, traditional cloud-based training and edge-based inference are increasingly inadequate for the dynamic nature of edge network environments and the complexity of user demands.

Artificial Intelligence (AI) technology has emerged as an effective solution to meet the stringent requirements of large-scale, real-time IoT systems, as AI-driven approaches offer flexible frameworks and models for communication and computation, free from many conventional limitations. Despite progress, challenges persist in integrating AI into edge-enabled IoT systems. Key issues include deploying lightweight AI algorithms for on-device learning, evaluating systems with realistic datasets, facilitating efficient collaboration in data processing and task execution within the tiered end-edge paradigm, and enabling distributed resource management in EC while ensuring data privacy. These problems remain largely unresolved. This special issue aims to explore research, solutions, and practical implementations related to AI-powered communication and computation for edge-enabled IoT systems.

This special issue of the IEEE Internet of Things Magazine (IEEE IoTM) seeks to showcase high-quality, innovative research, solutions and practical implementations related to AI-empowered communication and computation for edge-enabled IoT systems. It invites contributions on the following key topics:

- Architecture and modeling for AI-powered communication and computation in IoT systems
- Intelligent task offloading, resource management and service deployment in edge-enabled IoT system
- Design of lightweight AI algorithms for on-device learning in IoT systems
- Information model and protocols facilitating AI-enabled edge computing
- Cutting-edge AI techniques for edge computing, including federated learning and graph neural networks
- Prototypes and systems demonstrating AI-enabled edge computing
- Realistic dataset collection and synthetic dataset generation for AI-powered IoT systems
- Performance evaluation of edge network IoT applications using realistic or synthetic datasets
- AI-based resource allocation for ultra-reliable low-latency communication in edge computing
- Data privacy in distributed resource management for edge computing
- Performance analysis of AI-powered communication and computation strategies in realistic IoT systems
- Scalability and network management for large-scale, real-world IoT systems

This special issue received an unexpectedly high number of submissions. After a rigorous review process, seven research papers were accepted, covering topics such as modeling, resource allocation, and AI-enabled edge computing applications.

In [A1], the authors propose a Semi-Federated Learning (SemiFL) framework that integrates Centralized Learning (CL) and Federated Learning (FL) to achieve efficient model training across heterogeneous IoT devices. It enables all IoT devices to participate in global model training, regardless of their computational capabilities and data distributions. Experimental results demonstrate that Simultaneously Transmitting and Reflecting Reconfigurable Intelligent Surface (STAR-RIS) aided SemiFL outperforms over-the-air computation (AirComp)-based FL and fixed beamforming approaches in terms of both training loss and test accuracy.

The authors of [A2] introduce an AI-driven optimization framework that utilizes Unmanned Aerial Vehicles (UAVs) equipped with Intelligent Reflecting Surfaces (IRS) for Multi-access Edge Computing (MEC), aiming to optimize resource allocation in Next-Generation (Next-Gen) wireless networks. They design a hierarchical heuristic approach by decomposing the optimization problem into three sequential stages. Simulation results demonstrate that the proposed approach significantly improves data transmission efficiency, reduces latency, and optimizes energy consumption, while enhancing Non-Line-of-Sight (NLoS) communications and enabling adaptive and scalable management of MEC-enabled IoT networks for real-time applications.

In [A3], the authors introduce the new concept of *"No further delay"*, which defines the optimal tradeoff time instant between increased latency due to waiting for a more complete application-layer dataset vs. the latency reduction achieved with successful computation with an incomplete application-layer dataset. For a computer vision inference workload with progressively coded JPEG Images, the *No further delay* approach, combined with edge node offloading, achieves a three-fold speedup compared to local on-device execution. This paper defines a new class of AI mechanisms aimed at estimating the optimal *No further delay* time instant, e.g., based on historical success rates of application-layer tasks computed with partial datasets. These mechanisms can enable IoT nodes to leverage complex AI model inference more effectively.

The authors of [A4] develop a practical framework that uses the Data Distribution Service (DDS) as the backbone for scalable, low-latency, and resilient Artificial Intelligence of Things (AIoT) deployments in smart eldercare environments. The framework supports real-time video and audio event detection at the edge, automated Quality of Service (QoS) tuning, plug-and-play device discovery, and containerized deployment for rapid prototyping and scaling. Experimental results indicate that DDS enhances runtime efficiency and reduces engineering complexity in managing heterogeneous IoT nodes within AI-powered home-care edge networks.

Integrating AI into edge computing is challenging, as cost and complexity constraints on edge devices and infrastructure may lead to trade-offs in computational capacity, limiting the ability to process data-intensive AI workloads and derive actionable insights. The authors of [A5] provide an overview of representative and emerging Edge AI use cases across various application domains, and outline the architectural and operational requirements and challenges that must be addressed for scalable, efficient, and reliable deployments of Edge AI.

In order to address the growing cybersecurity challenges in modern power distribution networks, the authors of [A6] explores AI-empowered communication techniques, including 5G virtual Local Area Networks (LANs), network digital twins, and radio frequency fingerprinting. These methods leverage the power of AI techniques to provide robust solutions for network protection and ensure the safe and efficient transmission of electrical energy. The integration of AI-empowered communication techniques strengthens the technical robustness of power distribution networks and boosts societal resilience by safeguarding against blackouts and ensuring a reliable power supply for critical infrastructure.

In [A7], the authors propose an AI-driven framework called Net0AICloud, which dynamically schedules workloads and allocates resources using Light Gradient Boosting Machine (LightGBM). Designed to optimize resource management and energy consumption, it also ensures QoS while minimizing carbon emissions to achieve net-zero emissions in cloud-edge IoT environments. Experimental validation through a real-world IoT healthcare application for disease prediction demonstrates its effectiveness in enhancing both sustainability and service quality.

We would like to express our sincere appreciation to the authors, reviewers, and the Editor-in-Chief (EiC) for their invaluable support in assisting the Guest Editors with the organization of this special issue. We hope that the work presented here will inspire and motivate researchers and professionals in both academia and industry to further explore and advance the application of AI-enhanced communication and computation in edge computing.

Huaming Wu, *Senior Member, IEEE*
Center for Applied Mathematics
Tianjin University
Tianjin 300072, China

Chaogang Tang, *Member, IEEE*
School of Computer Science and Technology
China University of Mining and Technology
Xuzhou 221116, China

Joel J. P. C. Rodrigues, *Fellow, IEEE*
Federal University of Piaui (UFPI), 64049-550 Teresina, Brazil
Artificial Intelligence Research Center (AIRC), Ajman University
Ajman, United Arab Emirates

Houbing Herbert Song, *Fellow, IEEE*
Department of Information Systems
University of Maryland, Baltimore County (UMBC)
Baltimore, MD 21250 USA

Huijun Tang, *Member, IEEE*
Department of Engineering
Durham University
DH1 3LE Durham, U.K.

Lei Wang, *Member, IEEE*
Tianjin Shenglian Intelligent Technology
Development Company Ltd., Tianjin 300457,
China

Appendix: Related Articles

[A1] W. Ni and H. Tian, "Joint communication and over-the-air computation for semi-federated learning towards scalable AI in computing-heterogeneous IoT systems," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 16–24, Mar. 2026.
[A2] A. Andreou, C. X. Mavromoustakis, H. Wang, E. Markakis, and G. Mastorakis, "Smart IRS resource allocation for multi-access edge computing in IoT networks," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 25–32, Mar. 2026.
[A3] J. Schulz, P. Seeling, M. Reisslein, and F. H. P. Fitzek, "No further delay: Making time an ally of edge computation of AI workloads," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 33–40, Mar. 2026.
[A4] C. Chen, C. Chang, C. Peng, and G. Lan, "Enabling AI-driven real-time communication and computation with DDS for IoT-assisted home care," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 41–48, Mar. 2026.
[A5] S. C. Shah et al., "AI at the edge: Use cases, requirements, and challenges," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 49–56, Mar. 2026.
[A6] J. Bai et al., "AI-empowered communication techniques for power distribution network protection business," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 57–63, Mar. 2026.
[A7] H. Wang and S. S. Gill, "Net0AICloud: AIoT-driven sustainable edge cloud computing towards net zero carbon emissions," *IEEE Internet Things Mag.*, vol. 9, no. 2, pp. 64–70, Mar. 2026.

Biographies

Huaming Wu (Senior Member, IEEE) (whming@tju.edu.cn) received the B.E. and M.S. degrees from Harbin Institute of Technology, China, in 2009 and 2011, respectively, both in electrical engineering, and the Ph.D. degree with the highest honor in computer science from Freie Universität Berlin, Germany, in 2015. He is currently a Full Professor with the Center for Applied Mathematics, Tianjin University, China. He serves as the Associate Editors of IEEE Transactions on Dependable and Secure Computing, IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Network and Service Management, and IEEE Transactions on Consumer Electronics. He serves as a Conference Chair or a TPC Member in several international conferences. He has published 160+ papers in top-tier journals and conferences such as *Nature Computational Science*, *Nature Communications*, IEEE Transactions on Mobile Computing, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Services Computing, IEEE Transactions on Computers, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Information Forensics and Security, IEEE Transactions on Fuzzy Systems, IEEE Transactions on Industrial Informatics, IEEE Transactions on Intelligent Transportation Systems, IEEE Transactions on Cybernetics, IEEE Transactions on Neural Networks and Learning, AAAI, and INFOCOM. His research interests include mobile edge computing, Internet of Things, edge intelligence, and DNA-based data Storage.

CHAOGANG TANG (Member, IEEE) received the B.S. degree from Nanjing University of Aeronautics and Astronautics, Nanjing, China, and the Ph.D. degree from the School of Information Science and Technology, University of Science and Technology of China, Hefei, China, in 2012. He is currently with China University of Mining and Technology. His research interests include vehicular edge computing Internet of Things, and big data.

JOEL J. P. C. RODRIGUES (Fellow, IEEE) is with the Federal University of Piauí, Brazil, the AIRC, Ajman University, Ajman, United Arab Emirates, and the Leader of the Center for Intelligence at Fecomércio/CE, Brazil. He is an Highly Cited Researcher (Clarivate), N. 1 of the top scientists in computer science in Brazil (Research.com), a Distinguished PQ-1 Researcher of the National Council for Scientific and Technological Development (CNPq), a Leader of NetGNA Research Group (CNPq), the Chair of the IEEE Fellow Committee at the IEEE Region 9, Member of the IEEE COM Fellow Evaluation Committee, and the President of the scientific council at ParkUrbis Covilhã. He was a member of IEEE ComSoc Board of Governors, a Director for ComSoc Conference Development, an IEEE Distinguished Lecturer, a TAC Chair of IEEE ComSoc Latin America Region Board, a Past-Chair of the IEEE ComSoc TC on eHealth, and a TC on Communications Software. He is the EiC of the *International Journal of E-Health and Medical Communications* and editorial board member of high-reputed journals. He has authored more than 1350 contributions. He had been awarded several Outstanding Leadership and Outstanding Service Awards by IEEE Communications Society and several best papers awards.

HOUBING HERBERT SONG (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Virginia, Charlottesville, VA, in 2012. He is currently a Tenured Associate Professor, the Director of the NSF Center for Aviation Big Data Analytics (Planning), and the Associate Director for Leadership of the DoT Transportation Cybersecurity Center for Advanced Research and Education, University of Maryland, Baltimore County (UMBC), Baltimore, MD. He has served as one of the Coeditors-in-Chief for IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, and an Associate Editor for IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, IEEE INTERNET OF THINGS JOURNAL, and IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. He is an ACM Distinguished Member, an ACM Distinguished Speaker, and an IEEE Vehicular Technology Society Distinguished Lecturer. He was a recipient of 10+ Best Paper Awards from major international conferences, including IEEE CPSCom 2019, IEEE ICII 2019, IEEE/AIAA ICNS 2019, IEEE CBDCom 2020, WASA 2020, AIAA/IEEE DASC 2021, IEEE GLOBECOM 2021, and IEEE INFOCOM 2022.

HUIJUN TANG (Member, IEEE) received the B.Sc. degree from Jinan University, China, in 2016, and the M.S. and Ph.D. degrees from Tianjin University, China, in 2018 and 2022, respectively. She is currently a Postdoctoral Research Associate with Durham University, U.K. She has served as a TPC member for conferences, including INFOCOM FutureTelecoms, ICCCN, ICA3PP, TrustCom, ICCC, SmartData, ISPA, SustainCom, and ICPADS. She was a recipient of the 2025 IEEE HITC Award for Excellence (Early Career Researcher). Her research interests include Internet of Things, edge computing, and complex networks.

LEI WANG (Member, IEEE) received the bachelor's and master's degrees from Nankai University, Tianjin, in 2005 and 2014, respectively. He is currently working toward the Ph.D. degree with the College of Intelligence and Computing, Tianjin University and serves as the General Manager of Tianjin Shenglian Intelligent Technology Development Company, Ltd. His research fields are deep learning, traffic data analysis and application, complex spatiotemporal data mining, and multisource heterogeneous data analysis.