

Editorial

1 Introduction

The theme of this special section, “Low-Consumption Computing Technologies for Big Data Analytics (LCCT-BigDataAna)”, focuses on the critical challenge of balancing limited computing resources with the exponential growth of data. As big data analytics becomes increasingly essential, small and medium-sized enterprises, along with budget-constrained research institutions struggle to afford high-performance computing solutions^[1]. Consequently, the development of cost-effective and efficient methods for big data processing has become a pivotal research focus, attracting significant attention from both academia and industry.

LCCT-BigDataAna covers hardware and software optimizations for improving computational efficiency, including data preprocessing, storage management, and novel learning and modeling technologies. However, raw big data often contains noise, missing values, and outliers, which complicates large-scale big data analytics^[2]. Additionally, the widespread use of decentralized and cross-regional data storage^[3] necessitates the development of advanced analytical methods capable of extracting reliable insights from incomplete yet statistically representative data subsets. These challenges drive substantial improvements over traditional big data computation approaches.

Despite the increasing focus on big data analytics, there are still major limitations and challenges for low-consumption computing technologies. Some key areas of research are training memory-free learning models under the Spark computing framework, managing computing tasks for cross-regional stored big data, leveraging advanced statistical methods to obtain multi-sample computation results, and formalizing the relationship between big data analytics and digital economic activities^[4]. Consequently, there is a significant opportunity to explore the latest advancements in LCCT-BigDataAna^[5]. To disseminate the latest developments, current challenges, and

highlight potential applications of low-consumption computing technologies for big data analytics, we are delighted and honored to present this special section in *Tsinghua Science and Technology*.

2 Contribution to the Special Section

This special section features five outstanding works, carefully selected through a rigorous review process. These studies span various domains in big data analytics, including incremental learning, resource scheduling, flow prediction, cross-domain processing, and correlation filtering. Each work addresses critical aspects of low-consumption computing in big data environments, collectively contributing to the advancement of efficient computational methodologies for massive data processing.

Chen et al.^[6] proposed a resource-efficient solution for federated incremental learning, enabling the efficient removal of poisoned data while maintaining model performance. Leveraging membership inference attacks to detect poisoned data and applying targeted fine-tuning to erase the impact of such data, the approach aligns model outputs with clean data, achieving up to 21.05× faster forgetting time than traditional baselines. This innovation addresses low-consumption computing challenges in big data analytics, ensuring scalability, security, and efficiency in constrained computing environments. It advances sustainable federated learning, offering practical insights for resource-constrained applications in both academia and industry.

Cheng et al.^[7] introduced a Computing power network Resource Scheduling method with reward shaping (C2RS). It is an intelligent deep reinforcement learning (DRL)-based scheduler designed to optimize energy efficiency in computing power networks through big data analytics. The key innovation is a reward shaping mechanism that converts sparse rewards into dense feedback, enabling faster convergence and more efficient resource allocation. Experimental results confirm that C2RS outperforms

existing DRL-based scheduling methods in energy efficiency and system responsiveness, making it ideal for low-consumption computing. This work aligns with the growing demand for energy-aware scheduling in large-scale data processing, offering a practical and scalable solution for green big data infrastructures.

Niu et al.^[8] introduced a Dual-module Adaptive Transformer and Spatio-Temporal Attention Network (DAT-STAN), an efficient deep learning model for large-scale traffic prediction using low-consumption computing. It combines spatio-temporal attention and adaptive dynamic adjacency matrices to capture traffic patterns while minimizing computational costs. Unlike static models, it dynamically adjusts to real-time changes, improving both accuracy and efficiency. Experiments demonstrate that DAT-STAN outperforms existing methods in prediction accuracy and computational efficiency. The model balances accuracy with low computational overhead, making it ideal for smart city applications where energy efficiency and real-time performance are crucial. This work highlights the potential of integrating low-consumption technologies in big data analytics, offering green AI solutions for data-intensive urban computing.

Tan et al.^[9] proposed an energy-efficient hybrid P2P architecture for cloud-based big data analytics, addressing the high energy costs of traditional resource discovery methods. The solution combines three key innovations: (1) A service-oriented topology that reduces redundant data transfers through service-description matching; (2) an active replication protocol for efficient metadata synchronization among supernodes; and (3) a restricted flooding search that cuts message hops by 50% and network messages by 99%. Experimental results demonstrate scalable, low-power resource management while maintaining search efficiency, aligning with LCCT-BigDataAna objectives. This work provides small and medium-sized enterprises and research institutions with a practical energy-aware solution for distributed big data analytics, significantly reducing computational and network overheads in cross-regional deployments.

Yu et al.^[10] presented an energy-efficient visual tracking method for big data analytics, enhancing traditional Discriminative Correlation Filters (DCF) with a correlation-based attention mechanism. To reduce computational redundancy in processing massive video streams, the method employs KNN

clustering to identify correlated feature channels and a p-Laplacian matrix to dynamically adjust channel weights. This approach minimizes redundant computations while improving adaptability across scenarios. Experiments show the method achieves higher tracking accuracy with lower energy consumption compared to conventional DCF, making it ideal for resource-constrained, large-scale video analytics. The work advances low-consumption computing in real-time object tracking, balancing performance and efficiency for sustainable big data applications.

Collectively, these works advance LCCT-BigDataAna by developing innovative solutions for energy-efficient computing across different domains. Together, they demonstrate significant progress in reducing computational overhead while maintaining performance, addressing critical challenges in resource-constrained big data analytics. All papers in this special section were rigorously peer-reviewed. We sincerely thank the authors for their valuable contributions and timely revisions, and the reviewers for their insightful comments. We hope this collection provides useful insights and inspiration for future research in low-consumption computing technologies for big data analytics.

Guest editors:

Yulin He, Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ), China.

Philippe Fournier-Viger, the Big Data Institute, College of Computer Science and Software Engineering, Shenzhen University, China.

References

- [1] P. Boncz, K. Salem, Y. Cao, and W. Fan, Data driven approximation with bounded resources, *Proc. VLDB Endow*, vol. 10, no. 9, pp. 973–984, 2017.
- [2] C. Yang, X. Xu, K. Ramamohanarao, and J. Chen, A scalable multi-data sources based recursive approximation approach for fast error recovery in big sensing data on cloud, *IEEE Trans. Knowl. Data Eng.*, vol. 32, no. 5, pp. 841–854, 2020.
- [3] A. Siddiqa, A. Karim, and A. Gani, Big data storage technologies: A survey, *Front. Inf. Technol. Electronic Eng.*, vol. 18, no. 8, pp. 1040–1070, 2017.
- [4] C. Li, W. He, and E. Cao, Impact of green data center pilots on the digital economy development: An empirical study based on dual machine learning methods, *Comput. Ind. Eng.*, vol. 201, p. 110914, 2025.

- [5] S. G. Paul, A. Saha, M. S. Arefin, T. Bhuiyan, A. A. Biswas, A. W. Reza, N. M. Alotaibi, S. A. Alyami, and M. A. Moni, A comprehensive review of green computing: Past, present, and future research, *IEEE Access*, vol. 11, pp. 87445–87494, 2023.
- [6] K. Chen, W. Li, J. Cao, B. Mi, and J. Shen, Optimizing federated incremental learning: Efficient malicious data removal for big data analytics, *Tsinghua Science and Technology*, vol. 32, no. 1, pp. 62–79, 2026.
- [7] Y. Cheng, Z. Hao, J. Z. Huang, and Z. Cai, Optimizing resource scheduling in computing power networks for low-consumption big data analytics, *Tsinghua Science and Technology*, vol. 32, no. 1, pp. 28–44, 2026.
- [8] K. Yin, X. Niu, J. Zhu, Y. Jiang, and F. Min, DAT-STAN: Dual-module adaptive transformer and spatio-temporal attention network for large-scale traffic flow prediction, *Tsinghua Science and Technology*, vol. 32, no. 1, pp. 45–61, 2026.
- [9] W. Wu, L. Tan, Z. Yu, P. Yang, D. Ye, and K. She, A hybrid P2P enabled big data resource discovery method in cloud environments, *Tsinghua Science and Technology*, vol. 32, no. 1, pp. 4–27, 2026.
- [10] Y. F. Yu, X. Tan, Y. Zhang, L. Chen, and W. Ding, K-nearest neighbor correlation filters learning with p -laplacian regularization for visual tracking, *Tsinghua Science and Technology*, vol. 32, no. 1, pp. 80–101, 2026.