MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3844-3846

EDGE INTELLIGENCE IN RESOURCE-CONSTRAINED IOT NETWORKS: A LATENCY–PRIVACY CO-DESIGN FRAMEWORK

Ms.Nandhakumari J
Assistant Professor, Department of Mechanical Engineering,
Kgisl institute of technology, Coimbatore, India
nandhakumari.j@kgkite.ac.in; nandhuswathi21@gmail.com
Dr.M.Indirani,
Assistant Professor, Department of Computer Science and Business Systems,
Vel Tech Multi Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Chennai, India mindirani2008@gmail.com
A. Prasanth,
Assistant Professor/Department of Management Science,
Hindusthan College of Engineering and Technology, Coimbatore, India prasantharuns@gmail.com
Orchid Id: 0000-0001-6394-9449
S. Shanmugapriya,
Assistant Professor, Department of Artificial intelligence and Data Science,
Nehru Institute of Engineering and Technology, Coimbatore, India shanmugapriyacse2008@gmail.com
Dr T Jayaprakash
Professor in Physics (S&H), Department of Science and Humanities,
Nehru Institute of Technology Coimbatore, India
nitjayaprakash@nehrucolleges.com
R. Kathiresh Ramalingam
Assistant Professor , Department of Computer Science and Engineering
VSB college of engineering technical campus, Coimbatore, India
kathiresh89@gmail.com

Abstract
Edge intelligence provides real-time processing of data in constrained resource IoT networks through minimising latency and maximising system performance. This paper will introduce a latency privacy co-design, whereby lightweight AI models are combined with privacy-preserving methods, to give ideal performance and guarantee secure data management in distributed settings.
Keywords: Edge Intelligence; IoT Networks; Privacy Preservation

1. Introduction

The fast development of the Internet of Things (IoT) has resulted in more than 15 billion devices worldwide, which is expected to reach over 29 billion by 2030, and is already becoming increasingly intense in terms of the real-time generation and processing of data (Statista, 2023). The majority of IoT devices are intrinsically resource-limited and have little computational power, memory, and energy capacity, which limits their capabilities in supporting complex analytics (Djigal *et al.,* 2022). Conventional cloud-based frameworks also compound these constraints with the addition of network latency, which can be measured in 100-200 ms on time-sensitive applications, and critical data privacy and security issues as a result of data transmission being centralised (Kong *et al.,* 2022). As an answer, edge intelligence, which involves the integration of artificial intelligence at the network edge, has become a promising remedy to this, enabling the reduction of latency by 40-60 per cent due to local processing (Tirupatamma *et al.,* 2024). However, edge deployment comes with an inherent trade-off between the efficiency of the performance and privacy, and such a trade-off requires a co-design framework that can optimally balance both.

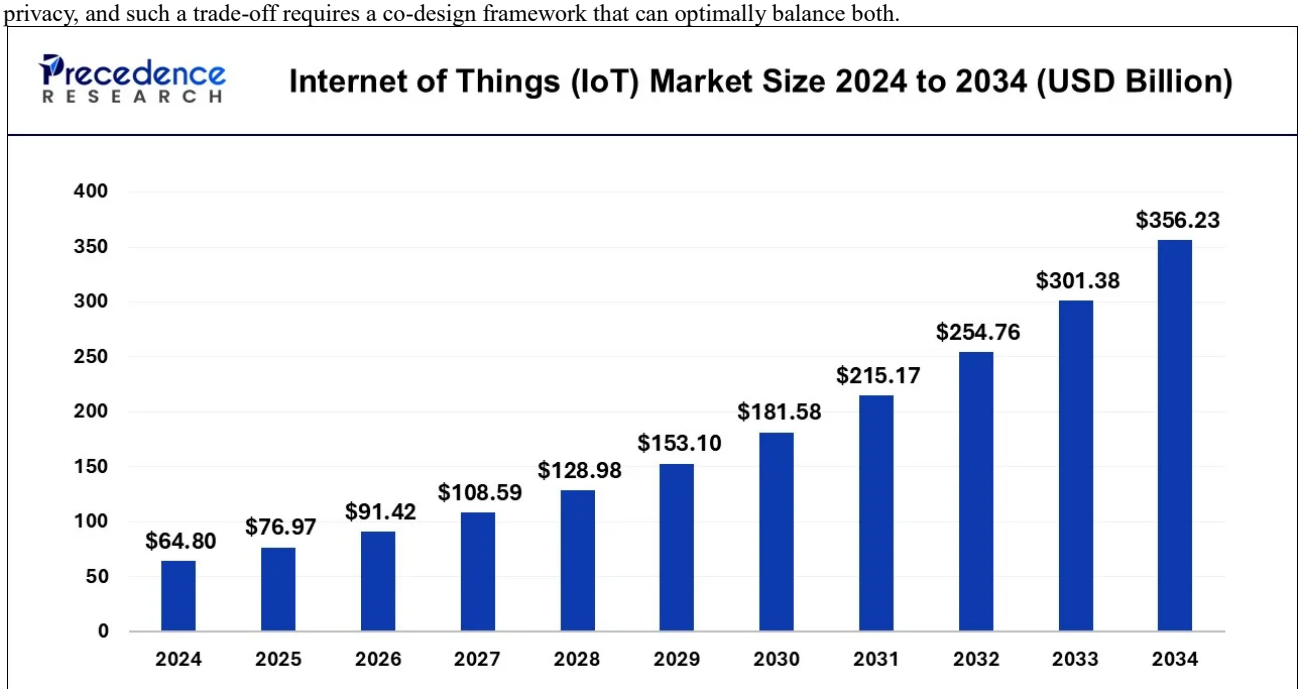

Figure one: "Internet of Things (IoT) Market Advancements Driving Smart Connectivity Solutions"
Source: Precedence research, 2025

2. Key Themes of the Study

2.1 Resource Limits and Edge Computing Problems

The IoT devices are being powered by extremely limited resources, such as limited processing power, memory, and battery capacity, and it is a major constraint on the implementation of complex AI models at the edge. Recent research points out that edge devices are not able to serve high computing loads in most cases, and thus, real-time analytics are challenging unless optimisation methods are applied (Arif *et al.,* 2025). Another severe issue is power consumption since data processing and communication may drain battery charges fast, particularly in remote or wearable

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3844-3846

IoT solutions. Limited bandwidth and poor connections also restrict effective data transfer, especially in large-scale IoT systems where enormous data streams are created (Fu *et al.,* 2021).

A practical example can be seen in the case of smart healthcare, where wearable sensors have to run the patient data on the device itself to provide a timely response and save on battery life. To solve these problems, recent studies focus on lightweight AI models, federated learning, and adaptive task offloading, which would be able to cut cloud reliance by 50 per cent and increase edge environments' energy efficiency (Hong *et al.,* 2024). The strategies play a vital role in making it possible to scale and have efficient edge intelligence in resource-constrained Internet of Things networks.

2.2 Latency Optimisation of Edge Intelligence

Low latency is a crucial aspect of modern IoT applications such as medical monitoring, self-driving cars, and smart cities, where even a slight delay may trigger the breakdown of the whole system or endanger lives. The latency is normally 120-200 ms in traditional cloud-based architectures because the data should be transmitted over long distances, and edge computing makes it about 10-30 ms because the data is processed nearer to the source (Irshad, 2024). Empirical research also shows that edge-based AI can reduce response times by more than 85 in healthcare monitoring and predictive maintenance, and other applications (Murthy *et al.,* 2025).

Local processing reduces communications delay as well as increasing responsiveness of real-time systems, and it makes the systems more reliable, especially in the context of latency-sensitive environments (Walani and Doorsamy, 2025). Efficient workload distribution is possible through task offloading strategies, in which time-sensitive processes are performed on the edge, and complicated calculations are transferred to the cloud. Edge devices in smart traffic management systems are used to receive and interpret real-time video near the large intersection to minimise congestion and accidents. To have optimal latency, though, there must be coordinated orchestration between edge devices, gateways and cloud systems. Hybrid edge cloud designs are becoming widely used to represent a compromise between responsiveness and computational scalability to provide efficient and reliable operation of IoTs (Irshad, 2024).

2.3 Distributed IoT Systems: Privacy Preservation

IoT systems produce enormous volumes of personal and work-related sensitive data, and privacy protection is one of the key issues in distributed settings. Data processing in a centralised manner increases the probability of a data breach, unauthorised access, and massive-scale attacks, especially in the healthcare sector or smart cities (Mahmud *et al.,* 2024). The risk of data exposure is particularly imminent at the time of transmission and aggregation, where attackers can use the vulnerability to reveal personal information based on shared information or model updates. Besides, the regulations, including the General Data Protection Regulation (GDPR), place tight restrictions on the data collection process, storage, and processing, making privacy-conscious system design a more challenging requirement (Yazdinejad *et al.,* 2024).

Federated learning and other privacy-preserving mechanisms, like federated learning, allow local data processing but only model updates to be shared, which decreases the use of centralised data storage and removes privacy-related risks (Alqazzaz, 2025). Complementary methods, including either the use of differential privacy or secure multi-party computation, are aimed at protecting sensitive data by introducing noise to the data or encrypting the cooperation between the devices, respectively. Federated learning can be used in smart healthcare (IoMT) systems to enable hospitals to train AI models together without exchanging patient information, thereby improving privacy and compliance considerably (Sinthiya, 2025). The balancing between privacy, accuracy and computational efficiency in large-scale IoT applications is still a problem.

3. Discussion: Toward a Latency–Privacy Co-Design Framework

Latency and privacy are discussed separately in the traditional IoT architectures, resulting in poor performance of the system. Recent research points at the fact that co-design systems should be integrated to consolidate these objectives through the joint models (Telkar and Yogi, 2025). Latency and privacy can be considered simultaneously, allowing dynamically to trade-off systems depending on contextual needs, e.g. network overload or data sensitivity. Context-aware mechanisms also contribute to the level of adaptability, making it possible to make intelligent decisions at the edge and enhance the responsiveness and protection of data (Cheng *et al.,* 2018).

Hierarchical edge cloud architecture has been used extensively to enable latency privacy co-design. Data collection and initial filtering are done by the device layer to minimise unwanted data transmission. Real-time AI inference and local decision-making are done on the edge layer, which substantially minimises latency. Computationally-intensive procedures like model training and long-term analytics are supported by the cloud layer (Kairouz and McMahan, 2021). It is this multiple-layer solution that enables restricted sharing of data, such that any sensitive data does not leave the local area, and only necessary insights are sent, which preserves privacy and efficiency (Wang and Wang, 2021).

There are a number of techniques that facilitate co-design implementation. Pruning and quantisation are examples of model compression techniques that can be used to minimise the computational overhead and energy used, without impairing the accuracy (Alwarafy *et al.,* 2020). Federated learning allows training models using decentralisation, which means that raw data are stored on local machines, which increases privacy (Xu *et al.,* 2022). There are adaptive task offloading strategies which attempt to dynamically allocate workloads between edge and cloud based on the needs of latency and network conditions.

4. Conclusion

Edge intelligence is able to provide a viable answer to the constraints of IoT networks with limited resources by means of local processing of data, latency reduction, and enhanced real-time performance. There are a lot of privacy and security-related issues with a decentralised architecture. Latency privacy co-design framework emphasises achieving data protection and efficiency of systems by using lightweight AI models, federated learning, and adaptive architecture. To advance further in the future, researchers and practitioners are advised to pay more attention to scalable and standardised co-design frameworks that would be widely applicable to a range of IoT applications. More focus should also be laid on actual implementation, interoperability, and adherence to international data protection laws. Furthermore, it is also possible to enhance trust and reliability with the inclusion of emerging technologies such as explainable AI and secure hardware applications. The sustainability of IoT ecosystems will, ultimately, be determined by their capacity to provide services of low-latency, security, and intelligence in progressively more challenging and more constrained environments.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3844-3846

References

Alqazzaz, A., 2025. Federated learning with homomorphic encryption: A privacy-preserving solution for smart cities. *International Journal of Computational Intelligence Systems*, *18*(1), p.304.

Alwarafy, A., Al-Thelaya, K.A., Abdallah, M., Schneider, J. and Hamdi, M., 2020. A survey on security and privacy issues in edge-computing-assisted internet of things. *IEEE Internet of Things Journal*, *8*(6), pp.4004-4022.

Arif, A., Ali, M.Z.H., Haider, S.Z.Q. and Niaz, Q., 2025. AI-Driven Edge Computing for IoT: Revolutionizing Phishing Detection and Mitigation. *Journal of Computing & Biomedical Informatics*, *9*(01).

C. SINTHIYA, 2025. *Federated Learning Architectures for Privacy-Preserving Collaborative Intelligence in Distributed Networks*. https://doi.org/10.64137/31079458/IJCSEI-V1I1P104

Cheng, Y., Wang, D., Zhou, P. and Zhang, T., 2018. Model compression and acceleration for deep neural networks: The principles, progress, and challenges. *IEEE Signal Processing Magazine*, *35*(1), pp.126-136.

Djigal, H., Xu, J., Liu, L. and Zhang, Y., 2022. Machine and deep learning for resource allocation in multi-access edge computing: A survey. *IEEE Communications Surveys & Tutorials*, *24*(4), pp.2449-2494.

Fu, Y., Yang, X., Yang, P., Wong, A.K., Shi, Z., Wang, H. and Quek, T.Q., 2021. Energy-efficient offloading and resource allocation for mobile edge computing enabled mission-critical internet-of-things systems. *EURASIP Journal on Wireless Communications and Networking*, *2021*(1), p.26.

Hong, J., Hong, Y.G., de Foy, X., Kovatsch, M., Schooler, E. and Kutscher, D., 2024. Internet of Things (IoT) Edge Challenges and Functions. *Internet Research Task Force (IRTF)*, p.556.

Irshad, A., 2024. Latency Optimization in Edge vs. Cloud Computing: A Comparative Study for Real-Time Applications. *International Journal of Business & Computational Science*, *1*(1).

Irshad, A., 2024. Latency Optimization in Edge vs. Cloud Computing: A Comparative Study for Real-Time Applications. *International Journal of Business & Computational Science*, *1*(1).

Kairouz, P. and McMahan, H.B., 2021. Advances and open problems in federated learning. *Foundations and trends in machine learning*, *14*(1-2), pp.1-210.

Kong, X., Wu, Y., Wang, H. and Xia, F., 2022. Edge computing for internet of everything: A survey. *IEEE internet of things journal*, *9*(23), pp.23472-23485.

Mahmud, S.A., Islam, N., Islam, Z., Rahman, Z. and Mehedi, S.T., 2024. Privacy-preserving federated learning-based intrusion detection technique for cyber-physical systems. *Mathematics*, *12*(20), p.3194.

Murthy, V.S.N., Kumari, R., Goyal, M., Dubey, P., Manikandan, M. and Ramesh, S.P., 2025. Edge-AI in IoT: Leveraging Cloud Computing and Big Data for Intelligent Decision-Making. *Journal of Information Systems Engineering and Management*, *10*, pp.601-619.

Precedence research, 2025. *Internet of Things (IoT) Market Advancements Driving Smart Connectivity Solutions*. https://www.precedenceresearch.com/internet-of-things-market

Statista, 2023. *Number of IoT connected devices worldwide*. Available at: https://www.statista.com/statistics/1183457/iot-connected-devices-worldwide/ [Accessed on: 18th March 2026]

Telkar, S.S. and Yogi, M.K., 2025. A comprehensive review of differential privacy with federated meta-learning for privacy-preserving medical IoT. *ICCK Transactions on Wireless Networks*, *1*(1), pp.16-31.

Tirupatamma, N., Anjali, N., Keerthi, V.P.M., Atchi, H.S. and Sammy, F., 2024, July. An Exploration of Edge Computing: Emergence, Evolution, Challenges, Approaches. In *2024 8th International Conference on Inventive Systems and Control (ICISC)* (pp. 19-26). IEEE.

Walani, C.C. and Doorsamy, W., 2025. Edge vs. Cloud: Empirical Insights into Data-Driven Condition Monitoring. *Big Data and Cognitive Computing*, *9*(5), p.121.

Wang, J. and Wang, L., 2021. A Computing Resource Allocation Optimization Strategy for Massive Internet of Health Things Devices Considering Privacy Protection in Cloud Edge Computing Environment: A Computing Resource Allocation Optimization Strategy for Massive Internet of Health Things Devices Considering Privacy Protection in Cloud Edge Computing Environment. *Journal of Grid Computing*, *19*(2), p.17.

Xu, Y., Bhuiyan, M.Z.A., Wang, T., Zhou, X. and Singh, A.K., 2022. C-fdrl: Context-aware privacy-preserving offloading through federated deep reinforcement learning in cloud-enabled IoT. *IEEE Transactions on Industrial Informatics*, *19*(2), pp.1155-1164.

Yazdinejad, A., Dehghantanha, A., Srivastava, G., Karimipour, H. and Parizi, R.M., 2024. Hybrid privacy preserving federated learning against irregular users in next-generation Internet of Things. *Journal of Systems Architecture*, *148*, p.103088.