A Vision of an AI-enabled Cloud-Edge Continuum

Anna Triantafyllou, Ilias Siniosoglou, Athanasios Liatifis, Vasileios Argyriou*, Thomas Lagkas, Panagiotis Sarigiannidis
{atriantafyllou, isiniosoglou, aliatifis, tlagkas, psarigiannidis}@uowm.gr, University of Western Macedonia, Kozani, Greece
* vasileios.argyriou@kingston.ac.uk, Kingston University, United Kingdom

***Abstract*—This paper proposes a vision for the Next Generation Internet of Things (NG-IoT) and examines the current state-of-the-art technology enablers that support the transition of the current technological status quo to the era of Artificial Intelligence (AI)-enabled Cloud-Edge Continuum.**

I. Introduction and Motivation

Modern society shows an increasing need to establish a robust framework of digital operations in all application domains to facilitate all aspects of our everyday life, improve the quality of living and foster reduced human intervention on a worldwide scale. This is especially true for the Cloud-Edge domain where a vast number of sensors is expected to be utilised in all subsequent fields, where IoT is an integral part. Examples include precision agriculture and environmental monitoring for field and livestock observation, in healthcare for ensuring the treatment and quality of life for patients as well as for asset management, retail and manufacturing operations.

The NG-IoT is a technology that requires the use of a massive amount of data from edge devices and sensors. It should be able to efficiently and securely handle this information while offering advanced automation control and cognitive decision support. However, achieving reliability in smart and decentralised applications is a challenge due to preserving low-latency communication links, high throughput information processing and actuation.

Nowadays, with the introduction of cutting-edge AI, integrated into more and more fields in an unstoppable momentum, such as behavioural recognition, network automation, cognitive detection and decision support, AI-enabled Cloud-Edge infrastructures are all the more required to facilitate its cross-domain application. Though, as technological advancements are not always aligned, resource allocation and operational requirements are significant issues, with the supplementary need for energy consumption balancing for these demanding operations.

Current technology enablers of the NG-IoT display promising progress towards the realisation of the next in the IoT evolution: **a vision of a unified, diligent, tactile and virtual environment, which leverages the Cloud-Edge Continuum.** A vision engaging revolutionised 5G and NG-IoT infrastructures with the aim to develop and employ an intelligent Beyond 5G (B5G) infrastructure that is expected to emerge after 2025. This advanced facility will facilitate the deployment of the fifth industrial revolution (Industry 5.0) and other NG-IoT applications combined cutting edge technologies, towards reshaping everyday digital operations.

II. Current Status

The NG-IoT has revolutionised how organisations operate, offering efficient asset monitoring solutions and near real-time data collection mechanisms empowering them with new insights. Toward achieving these goals, Critical Infrastructures (CI) and Decision Support Systems (DSS), produce large data volumes that are both difficult to process and transfer.

Cloud computing is the de-facto computing paradigm for Big Data processing, though it involves moving large data volumes across many networks introducing delay and exposing data to third parties. To this end, AI-enabled smart offloading of data processing and localisation on both the edge and the cloud can prove a key factor to the smooth operation of the CI in all application domains. Edge computing on the other hand provides the ability to operate in a scalable environment, while reducing the amount of information sent to the cloud servers [1].

By utilising cutting-edge information and decentralised processing techniques, such as Evolutionary Federated Learning, Intelligent Edge-to-Cloud orchestration, Information decentralisation and localisation, CI can cross over to a more digitally sound and reliable stage [2].

In addition, the NG-IoT market needs an open ecosystem where stakeholders can freely and safely publish data without fear of exposing sensitive information. Blockchain, Smart contracts, and Distributed Ledger Technologies are promising solutions for ensuring data integrity and anonymity for both individual users and large infrastructures.

Furthermore, marketability of the produced research outcomes of the variety of stakeholders is a key motivator for the Cloud-Edge evolution. For example, EU funded projects, such as TERMINET (H2020) [3], pushes the limits of current E2C in a variety of usecase, like Smart Farming, with contributions like the AGROMINDS platform, enabling dentralised agro-operational management [4]. Such initiatives support the establishment of innovative solutions under the umbrella of new market entities (start ups, spin-offs) like MetaMind Innovations [5], responsible for the creation of the aforementioned dashboard, under the premise of TERMINET.

NG-IoT infrastructure promotes the deployment of autonomous systems without human control or programming. Digital twins, virtual representations of IoT physical objects, are the stepping stone of the Tactile Internet evolution. Digital twins can gather data from IoT sensors embedded into core services and raise research interest in advanced simulation and

emulation. AI algorithms are used to optimise Industry 5.0 and beyond.

III. Research Challenges

The typically dense and heterogeneous nature of NG-IoT deployments necessitates further research on existing technological gaps to cover a variety of challenges.

A. Architectural challenges

- Enabling a **cross-architecture Federated edge management** to enhance fairness and efficiency of B5G infrastructures, levering also QoS improvement [2].
- **Engaging open interfaces and whitebox hardware for the Next Generation SDN (NG-SDN)** [6]. SDN-enabled Middleware mechanisms open up a wider range of opportunities for enhancing the flexibility, scalability, and the capability of NG-IoT to support new use cases, while reducing the latency of the cloud-based application to the end users.
- As demonstrated in [1], there is a need to design **common standards for network integration** towards enabling **virtualised Multiaccess Edge Computing (vMEC)**, which will reduce the communication overhead, promote the design of new signaling schemes and focus on the synchronisation of MEC services. There are several open issues related to pricing policies, user mobility, and scalability.
- Producing **novel Federated datasets** [7], [8] in a variety of application domains to accommodate the heterogeneity of data produced by decentralised IoT devices. This will help establish Federated AI and security- and privacy-by-design operation management principles, avoiding the disclosure of personal information and enabling only the exchange of knowledge [2].

B. Technological challenges

- It is proven that by utilising **novel bandwidth-saving approaches**, next generation wireless networks are now able to handle **huge traffic demands despite the resource constraints** [9].
- Efficient Edge to Cloud data localisation and cognitive processing offloading by incorporating **AI-based recourse management** should be considered for further research. The aim is to **increase the dynamicity and adaptability** of future frameworks based on devices' demand and preferences, and system performance. **Latency-aware dynamic resource allocation** is another aspect to be greatly considered.
- Design and implementation of **novel AI-powered Intrusion Detection and Prevention (IDPS) systems** [10] that rely on Machine Learning (ML) and Deep Learning (DL) techniques [11], [12], toward enhancing fault tolerance, privacy, and controllability of the distributed data to avoid misuse of data and secure the user's identity.
- Promotion of **decentralised Smart Contracts through Blockchain Technology** for automating the offering and acceptance of user data. Decentralised Blockchain technologies can be used for the exchange and verification of learning model updates.

C. Social and ethical challenges

- Given the potentially far-reaching impact of the NG-IoT, societies need to employ **self-regulation frameworks** able to detect and mitigate ethical and social risks by design. International collaborations or anonymous information channels could be formed to debate the nature of such issues and produce best practices.
- Another important and challenging research direction is the design and deployment of **unbiased AI systems**. Avoiding undesired discrimination of vulnerable groups is an important aspect of the NG-IoT digital societies.
- Considering the decrease in the amount of **greenhouse gas (GHG), new ways of charging batteries** and daily charging routines should be investigated, towards sustainable and energy-consistent NG-IoT infrastructures.

IV. Conclusions

Despite the existence of fundamental technical barriers, current technological advancements showcase that we are already one step closer to the realisation of Industry 5.0 and AI-enabled NG-IoT applications in the Cloud-Edge Continuoum.

Acknowledgment

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957406 (TERMINET).

References

[1] P. D. D. et. al, "Optimal design and orchestration of mobile edge computing with energy awareness," *IEEE Transactions on Sustainable Computing*, vol. 7, no. 2, pp. 456–470, 2022.

[2] I. S. et. al, "Unsupervised ethical equity evaluation of adversarial federated networks," in *Proceedings of the 16th International Conference on Availability, Reliability and Security*, 2021.

[3] "NexT GEneRation SMart INterconnectEd IoT," https://terminet-h2020.eu/, 2020, [Online; accessed 14-April-2023].

[4] G. K. et. al, "Towards climate smart farming—a reference architecture for integrated farming systems," *Telecom*, vol. 2, no. 1, p. 52–74, Feb 2021. [Online]. Available: http://dx.doi.org/10.3390/telecom2010005

[5] "Metaminds Innovations," https://metamind.gr/, 2021, [Online; accessed 14-April-2023].

[6] A. L. et. al, "Advancing sdn from openflow to p4: A survey," *ACM Comput. Surv.*, vol. 55, no. 9, jan 2023. [Online]. Available: https://doi.org/10.1145/3556973

[7] C. C. et. al, "Cherry tree disease detection dataset," 2022. [Online]. Available: https://dx.doi.org/10.21227/ehfm-9j20

[8] ——, "Peach tree disease detection dataset," 2022. [Online]. Available: https://dx.doi.org/10.21227/w67n-0q72

[9] V. P. et. al, "Hierarchical multiple access (hima) for fog-ran: Protocol design and resource allocation," *IEEE Transactions on Wireless Communications*, vol. 21, no. 2, pp. 960–975, 2022.

[10] V. K. et. al, "Ids for industrial applications: A federated learning approach with active personalization," *Sensors*, vol. 21, no. 20, p. 6743, Oct 2021. [Online]. Available: http://dx.doi.org/10.3390/s21206743

[11] P. R.-G. et. al, "Iec 60870-5-104 intrusion detection dataset," 2022. [Online]. Available: https://dx.doi.org/10.21227/fj7s-f281

[12] ——, "Dnp3 intrusion detection dataset," 2022. [Online]. Available: https://dx.doi.org/10.21227/s7h0-b081