Part I

Edge Intelligence with 5G/6G Networks

1

Edge Networking Technology Drivers for Next-generation Internet of Things in the TERMINET Project

Athanasios Liatifis[1], Dimitrios Pliatsios[1], Panagiotis Radoglou-Grammatikis[1], Thomas Lagkas[2], Vasileios Vitsas[2], Nikolaos Katertsidis[1], Ioannis Moscholios[3], Sotirios Goudos[4], and Panagiotis Sarigiannidis[1]

[1]University of Western Macedonia, Greece
[2]International Hellenic University, Greece
[3]University of Peloponnese, Greece
[4]Aristotle University of Thessaloniki, Greece
E-mail: aliatifis@uowm.gr; dpliatsios@uowm.gr; pradoglou@uowm.gr; tlagkas@cs.ihu.gr; vitsas@it.teithe.gr; n.katertsidis@uowm.gr; idm@uop.gr; sgoudo@physics.auth.gr; psarigiannidis@uowm.gr

Abstract

The rapid growth of the Internet of Things has shaped the design and deployment of mobile networks to accommodate the need for ubiquitous connectivity and novel applications. The recent advancements in wireless communications and computing technologies allow the connection of a wider range and number of devices and systems, thereby enabling the design and development of next-generation Internet of Things applications. However, current networking technologies cannot accommodate the increasing number of IoT devices as well as satisfy the heterogeneous and stringent requirements in terms of bandwidth, connectivity, latency, and reliability. Motivated by these remarks, this chapter aims to provide an overview of key novel technologies, which are expected to be integral components of future networks and can effectively address the challenges associated with satisfying the aforementioned requirements.

Keywords: Next-generation Internet of Things, software-defined networking, network function virtualization mobile edge computing, digital twins, radio access network.

1.1 Introduction

The past couple of years have seen an increased interest in the Internet of Things (IoT) [1]. IoT is one of the fastest evolving technologies and it is being increasingly adopted and shaped by various industries and organizations to push their vision. As a result, IoT is expected to be a core component of the future Internet and has received much attention from both industry and academia due to its great potential to deliver customer services in many aspects of modern life. IoT enables the interconnection of various appliances and devices to the internet, enabling them to communicate and exchange data. This interconnected network of devices is expected to introduce substantial changes to the ways people live and work.

The main advantage of IoT is the ability to collect, aggregate, and analyze large volumes of data, enabling the automation of various processes or generating useful insights that assist in the decision-making process. IoT is being integrated into various application verticals, including smart healthcare, smart industry, autonomous vehicles, smart agriculture, and smart cities.

The latest advancements in wireless communications and computing technologies integrate enhanced connectivity, increased bandwidth and data rates, and ultra-low-latency communications, making it feasible to connect a wider range of systems and devices and enabling the realization of next-generation IoT (NG-IoT) applications [2]. To this end, researchers have identified a number of key challenges that have to be addressed:

1. **Large data volumes and number of devices:** A distinct characteristic of IoT is the dense deployment of massive numbers of devices. These devices generate a large volume of data that have to be efficiently transferred and processed. The Big Data concept is concerned with how these data are collected, stored, and processed.
2. **Ubiquitous wireless connectivity:** Mobile networks provide ubiquitous wireless connectivity, enabling reliable communications between humans. Consequently, they are promising candidates for the

communications infrastructure. However, the traffic generated by IoT devices features some special characteristics and has considerable differences compared to traffic generated by human-to-human communications. Therefore, these attributes have to be considered during the design and deployment of future mobile networks.

3. **Interoperability:** Interoperability refers to the ability of devices and applications from different vendors to work together seamlessly. It is considered one of the most important aspects of the IoT, as it enables the communication and sharing of data among devices, regardless of the manufacturer or technology they use. Key challenges in achieving interoperability in the IoT domain include the diversity of devices and protocols, as well as the lack of standardization in the field.
4. **Energy efficiency:** Energy efficiency is concerned with the minimization of energy consumption while ensuring the provisioning of a minimum quality of service (QoS). It is a critical factor in IoT, as most of the devices have limited energy reserves. Therefore, the reduction of the consumed energy assists in extending their operating time. Moreover, achieving a high energy efficiency level can effectively reduce the total network energy consumption.
5. **Cybersecurity considerations:** Cybersecurity considerations include the measures adopted to protect the devices, data, and networks from unauthorized access and cyberattacks, such as man-in-the-middle attacks, device spoofing, and denial-of-service attacks. Cybersecurity has become a major concern due to the increasing number of IoT devices. Furthermore, the limited processing capabilities of IoT devices make the deployment of advanced cybersecurity countermeasures challenging.

To address the aforementioned challenges, several technologies have emerged. This chapter aims to provide an overview of these technologies, describe their key features, and outline potential applications. An illustration of the technology drivers for NG-IoT is presented in Figure 1.1.

1.2 Technology Drivers

1.2.1 Software defined networking and network function virtualization

Software defined networking (SDN) and network function virtualization (NFV) have revolutionized the way networks are designed, deployed, and

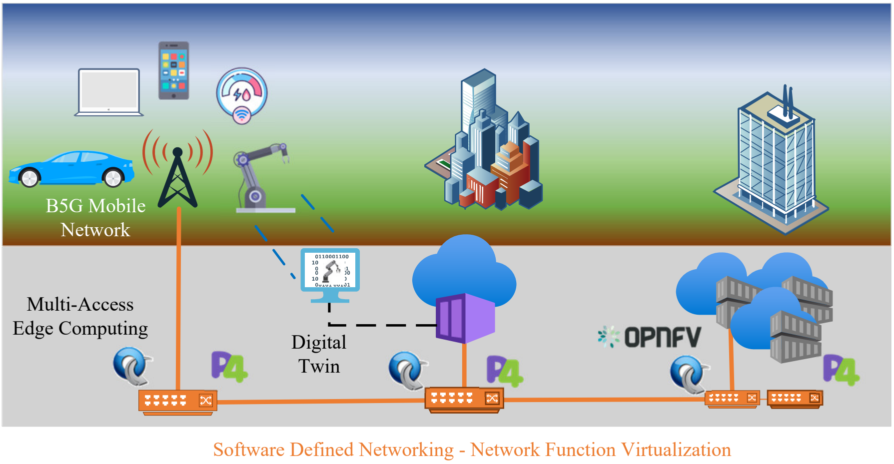

Figure 1.1 NG-IoT technology drivers.

operated [3]. Traditionally, data forwarding and control mechanisms were intertwined in every forwarding device (e.g., switch, router, firewall, etc.) resulting in limited flexibility in terms of new functionality and freedom of choice of hardware solutions. Each vendor offered specific tools and frameworks rendering multi-vendor deployment a hard task for large organizations (operational expenditure – OPEX). SDN, and, notably, OpenFlow [4], are promising technologies that can address the challenges associated with managing and operating complex networks. Consequently, service providers can focus on developing novel and customized solutions suitable for the size, needs, and customer profiles while maximizing the usage of networking infrastructure.

Networking functionalities, such as firewalls and load-balancing processes, were implemented by dedicated and expensive hardware solutions. Also, network scalability was not particularly considered, resulting in higher capital expenditures [5]. Moreover, the support for new protocol demands required the replacement of existing infrastructure and time-consuming processes. By leveraging the advancements of virtualization technologies, NFV offers dynamic network scaling and allows for placing network functionalities in the appropriate location, thereby minimizing network downtime.

The combination of SDN and NFV allows for realizing a dynamic, efficient, and customized networking infrastructure. Network services can be deployed easily in any part of the network and combined to form a chain of services [6]. Furthermore, SDN complements NFV offering a standardized mechanism to manipulate the behavior of forwarding devices through well-defined interfaces. The control plane can identify abnormal behavior in the data plane and re-adjust the path flows traverse to avoid cascading effects.

Despite the fact that SDN and NFV have acted as catalysts for many network services, current solutions are accompanied by several limitations. The next generation of computer networks, also referred to as next-generation SDN (NG-SDN), is characterized by fully programmable data planes offering packet-based monitoring solutions, whereas SDN and NFV will be tightly coupled further complementing each other. Programming languages and frameworks like P4, eBPF/XDP, and DPDK enable end-to-end programmability and experimentation with new protocols following a continuous integration/continuous development (CI/CD) approach [7]. Additionally, network functions can be, fully or partially, offloaded to programmable targets alleviating the control plane from continuous monitoring of the data plane state. Advances to artificial intelligence and machine learning (AI/ML) are also reflected in the control plane through intelligent applications that utilize the enriched statistics collected from the data plane [8], while existing SDN controllers are updated or new ones are developed to take full advantage of what programmable targets have to offer [9]. Frameworks like open programmable infrastructure (OPI) aim to unify programmability across infrastructure components (computing, AI/ML, networking, storage, etc.) granting developers standardized mechanisms of programming the entire infrastructure [10].

1.2.2 Beyond 5G mobile networks

5G mobile networks facilitate the design and development of applications with highly heterogeneous communication requirements by defining specific application classes such as enhanced mobile broadband (eMBB), ultra-reliable and low-latency communications (URLLC), and massive machine-type communications (mMTC). Nevertheless, the rapid growth of intelligent and autonomous IoT networks is likely to exceed the capabilities of 5G mobile networks.

As a result, researchers from academia and industry are focusing on the beyond 5G (B5G) mobile networks and their accompanying technological advances in order to prepare the way for NG-IoT development [11]. Due to its superior features over previous network generations, such as extremely high throughput, ultra-low-latency communications, and intelligent network capabilities, B5G networks are expected to deliver a new level of service and user experience in IoT networks.

- **Support for massive IoT:** B5G networks will enable the deployment of a massive number of IoT devices, with the ability to deploy millions of devices in a square kilometer. This will allow IoT applications to be deployed in domains such as smart cities, industrial automation, and agriculture [12].
- **Ultra-reliability and low-latency communications:** B5G networks facilitate the realization of IoT-based application scenarios that require ultra-reliability and low latency. Such application scenarios include remote surgery, virtual and augmented reality applications, autonomous vehicles, and intelligent industrial robotics [13].
- **Seamless wireless protocol integration:** The integration of various wireless protocols will assist in achieving the stringent requirements of NG-IoT. For example, the integration of Wi-Fi and B5G protocols will enable the leverage of the high-capacity links of Wi-Fi and the enhanced coverage of mobile networks. Moreover, additional wireless protocols can be integrated, with each one introducing different advantages, such as Bluetooth low energy (BLE), LoRa, etc.
- **High-frequency communication:** High-frequency communication refers to the use of higher frequency bands in wireless communications, such as millimeter wave (mmWave) and terahertz (THz) frequencies [14], [15]. The mmWave range includes frequencies in the range of $30-300$ GHz, while the THz range includes frequencies between 100 GHz and 10 THz. These frequencies offer a much larger bandwidth compared to traditional mobile network frequencies. As a result, larger numbers of devices can be accommodated and very high data rates can be achieved. However, these signals are more susceptible to obstacles such as trees and buildings, making the deployment of such high-frequency networks more challenging.
- **Novel radio access schemes:** Radio access schemes are a key component of wireless communication systems, as they manage the use of the limited radio frequency (RF) spectrum. Due to the massive connectivity and low-latency requirements, several radio access schemes have emerged, such as the non-orthogonal multiple access (NOMA) and grant-free (GF) access schemes. NOMA schemes aim to schedule multiple devices over the same radio resources, while GF schemes allow

devices to transmit their data without having to request radio access beforehand, effectively reducing latency [16].

- **Reconfigurable intelligent surfaces:** Reconfigurable intelligent surfaces (RIS) enable the manipulation of electromagnetic signals through a large number of passive, programmable, and low-cost elements [17]. These elements are able to control the absorption, refraction, and reflection of the signals, therefore reducing the signal fading, reducing the required transmission power, and enhancing the spatial spectrum reuse. Moreover, RIS reduce network expansion costs, as they mitigate the need to deploy more base stations for providing coverage to devices located at the network edge or behind obstacles.

1.2.3 Digital twin

The digital twin concept enables the digital representation of a physical object or system. A digital twin is generated by accumulating data from sensors on the physical device or system and utilizing that data to construct a virtual model of the real device or system [18]. The digital system coexists with the physical one as the two systems are interconnected through reliable and low-latency links. Real-time data transfers between physical and digital systems provide synchronized and coherent operation of the physical and virtual counterparts [19].

This enables the implementation of various simulation scenarios and analyses using this model for evaluating the possible outcomes, prior to their actual application in the real world. Consequently, digital twins can be used to construct digital representations of IoT devices and/or networks, in order to monitor, control, and assess the performance of the device/system in real time. Digital twins can be used in a variety of applications and scenarios including the following:

- **Virtual prototyping:** The digital twin of a device can be used for testing and assessing various design options, leading to an optimal design of the final product before its construction. This can effectively reduce the overall manufacturing time, as well as the associated testing costs.
- **Predictive maintenance:** By monitoring the performance of a device or system over time, the digital twin can be used to predict when maintenance is needed and to schedule maintenance at the most optimal time. Moreover, "what-if" scenarios can be simulated in order to find the best course of action with respect to the object's maintenance.

- **Building automation:** Typically, buildings consist of various components from different domains, including ventilation, heating, energy, mechanics, and plumbing. The digital twin of a building, a building complex, or a whole city can effectively facilitate building/city management and mitigate the environmental footprint (i.e., realizing a green building).

1.2.4 Multiple-access edge computing

Multi-access edge computing (MEC) is an emerging paradigm that attempts to converge telecommunication and IT services by delivering a cloud computing platform at the network edge [20]. MEC provides storage and computing resources at the network's edge, therefore lowering latency for mobile end users and maximizing the utilization of mobile backhaul and core networks. Consequently, MEC enables the development of a wide range of novel applications and services and the introduction of new business models. Key applications enabled by the MEC include the following:

- **Computation offloading:** Computationally intensive applications such as augmented and virtual reality applications require large data volumes to be transmitted and processed in real time. As a result, the conventional approach of transmitting and routing the data to datacenters with ample computing resources cannot ensure the real-time constraint. Computation offloading enables resource-constrained devices to partially or fully offload computationally intensive tasks to the computation resources offered by MEC, thereby reducing the processing time, increasing battery lifetime, and enhancing the network's energy efficiency [21], [22]. Additionally, as the processing takes place at the network edge, the communication and latency overheads are mitigated.
- **Next-generation Internet of Things:** Traditional IoT application scenarios involve aggregating large data volumes and forwarding them to a cloud environment for further processing. However, emerging next-generation IoT applications have increased constraints in terms of latency and reliability (e.g., autonomous vehicle scenarios). MEC can facilitate the deployment of storage and computing resources in close proximity to the devices, ensuring both redundancy and fast responses to device requests [23]. For instance, MEC is a key enabler for vehicle-to-infrastructure and vehicle-to-vehicle communications. Vehicles connected through the distributed MEC nodes can send and receive

real-time information, such as traffic congestion and warnings from other vehicles.

- **Content delivery and caching:** Image and video sharing constitute a large portion of mobile traffic. Such content is stored in datacenters and is distributed to the users through content delivery networks (CDNs). Nevertheless, there may be a lack of content in the proximity of the users, resulting in increased buffering time and lower quality of experience (QoE). To this end, MEC can effectively assist in realizing a distributed CDN by also taking into consideration information and context about the users in the proximity [24], [25].

1.3 Conclusion

The rapid expansion of the IoT considerably affects the planning of networks in order to satisfy the need for ubiquitous connectivity, increased data rates, and low-latency communications. NG-IoT is considered an evolution of the conventional IoT paradigm that is enabled by the latest advancements in wireless communications and computing technologies. NG-IoT allows the design and development of innovative services and applications.

NG-IoT applications feature heterogeneous and stringent requirements in terms of latency, reliability, and connectivity. To this end, various technologies and frameworks have emerged, aiming to address these requirements and the associated challenges. This chapter provided an overview of key technology drivers, namely the NG-SDN, NFV, B5G mobile networks, digital twins, and MEC, by presenting the main principles and outlining several application scenarios. These technologies are expected to be core components of future networks and facilitate the development of novel applications and services.

Acknowledgements

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

References

[1] H. Cao, M. Wachowicz, C. Renso, and E. Carlini, "Analytics everywhere: Generating insights from the internet of things," *IEEE Access*, vol. 7, pp. 71749–71769, 2019, doi: 10.1109/ACCESS.2019.2919514.

[2] NG-IoT Consortium, "Building a roadmap for the next generation internet of things: Research, innovation and implementation 2021-2027." Sep. 2019.

[3] D. Kreutz, F. M. V. Ramos, P. E. Veríssimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-defined networking: A comprehensive survey," *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2015, doi: 10.1109/JPROC.2014.2371999.

[4] N. McKeown *et al.*, "OpenFlow: Enabling innovation in campus networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69–74, Mar. 2008, doi: 10.1145/1355734.1355746.

[5] Z. Latif, K. Sharif, F. Li, M. M. Karim, S. Biswas, and Y. Wang, "A comprehensive survey of interface protocols for software defined networks," *Journal of Network and Computer Applications*, vol. 156, p. 102563, 2020, doi: https://doi.org/10.1016/j.jnca.2020.102563.

[6] K. Kaur, V. Mangat, and K. Kumar, "A comprehensive survey of service function chain provisioning approaches in SDN and NFV architecture," *Computer Science Review*, vol. 38, p. 100298, 2020, doi: https://doi.org/10.1016/j.cosrev.2020.100298.

[7] A. Liatifis, P. Sarigiannidis, V. Argyriou, and T. Lagkas, "Advancing SDN: From OpenFlow to P4, a survey," *ACM Computing Surveys*, Aug. 2022, doi: 10.1145/3556973.

[8] R. Amin, E. Rojas, A. Aqdus, S. Ramzan, D. Casillas-Perez, and J. M. Arco, "A survey on machine learning techniques for routing optimization in SDN," *IEEE Access*, vol. 9, pp. 104582–104611, 2021, doi: 10.1109/ACCESS.2021.3099092.

[9] R. Vilalta *et al.*, "Cloud-native SDN network management for beyond 5G networks with TeraFlow." virtual event, Jun. 2021. Available: https://zenodo.org/record/5089908

[10] Linux Foundation, "Open programmable infrastructure project." Sep. 2022. [Online]. Available: http://www.opiproject.org/

[11] Z. Zhang *et al.*, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Vehicular Technology Magazine*, vol. 14, no. 3, pp. 28–41, 2019, doi: 10.1109/MVT.2019.2921208.

[12] X. Chen, D. W. K. Ng, W. Yu, E. G. Larsson, N. Al-Dhahir, and R. Schober, "Massive access for 5G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 615–637, 2021, doi: 10.1109/JSAC.2020.3019724.

[13] B. S. Khan, S. Jangsher, A. Ahmed, and A. Al-Dweik, "URLLC and eMBB in 5G industrial IoT: A survey," *IEEE Open Journal*

of the Communications Society, vol. 3, pp. 1134–1163, 2022, doi: 10.1109/OJCOMS.2022.3189013.

[14] L. Zhang *et al.*, “A survey on 5G millimeter wave communications for UAV-assisted wireless networks,” *IEEE Access*, vol. 7, pp. 117460–117504, 2019, doi: 10.1109/ACCESS.2019.2929241.

[15] K. M. S. Huq, S. A. Busari, J. Rodriguez, V. Frascolla, W. Bazzi, and D. C. Sicker, “Terahertz-enabled wireless system for beyond-5G ultra-fast networks: A brief survey,” *IEEE Network*, vol. 33, no. 4, pp. 89–95, 2019, doi: 10.1109/MNET.2019.1800430.

[16] D. Pliatsios, A.-A. A. Boulogeorgos, T. Lagkas, V. Argyriou, I. D. Moscholios, and P. Sarigiannidis, “Semi-grant-free non-orthogonal multiple access for tactile internet of things,” in *2021 IEEE 32nd annual international symposium on personal, indoor and mobile radio communications (PIMRC)*, 2021, pp. 1389–1394. doi: 10.1109/PIMRC50174.2021.9569640.

[17] M. A. ElMossallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, “Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 3, pp. 990–1002, 2020, doi: 10.1109/TCCN.2020.2992604.

[18] R. Minerva, G. M. Lee, and N. Crespi, “Digital twin in the IoT context: A survey on technical features, scenarios, and architectural models,” *Proceedings of the IEEE*, vol. 108, no. 10, pp. 1785–1824, 2020, doi: 10.1109/JPROC.2020.2998530.

[19] S. Mihai *et al.*, “Digital twins: A survey on enabling technologies, challenges, trends and future prospects,” *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2255–2291, 2022, doi: 10.1109/COMST.2022.3208773.

[20] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, “On multi-access edge computing: A survey of the emerging 5G network edge cloud architecture and orchestration,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1657–1681, 2017, doi: 10.1109/COMST.2017.2705720.

[21] S. Thananjeyan, C. A. Chan, E. Wong, and A. Nirmalathas, “Mobility-aware energy optimization in hosts selection for computation offloading in multi-access edge computing,” *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1056–1065, 2020, doi: 10.1109/OJCOMS.2020.3008485.

[22] D. Pliatsios, P. Sarigiannidis, T. D. Lagkas, V. Argyriou, A.-A. A. Boulogeorgos, and P. Baziana, "Joint wireless resource and computation offloading optimization for energy efficient internet of vehicles," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 3, pp. 1468–1480, 2022, doi: 10.1109/TGCN.2022.3189413.

[23] S. Hamdan, M. Ayyash, and S. Almajali, "Edge-computing architectures for internet of things applications: A survey," *Sensors*, vol. 20, no. 22, p. 6441, Nov. 2020, doi: 10.3390/s20226441.

[24] E. F. Maleki, W. Ma, L. Mashayekhy, and H. La Roche, "QoS-aware 5G component selection for content delivery in multi-access edge computing," 2021. doi: 10.1145/3468737.3494101.

[25] X. Jiang, F. R. Yu, T. Song, and V. C. M. Leung, "A survey on multi-access edge computing applied to video streaming: Some research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 871–903, 2021, doi: 10.1109/COMST.2021.3065237.