Secure and Agile 6G Networking

DOI:
10.1109/ICTON59386.2023.10207418

Document status and date:
Published: 08/08/2023

Document Version:
Accepted manuscript including changes made at the peer-review stage

Secure and Agile 6G Networking – Quantum and AI Enabling Technologies

Carlos Rubio García[1], Oumayma Bouchmal[1], Catalina Stan[1], Panagiotis Giannakopoulos[1], Bruno Cimoli [1], Juan Jose Vegas Olmos[2], Simon Rommel[1] and Idelfonso Tafur Monroy[1]
[1]*Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands*
[2]*Software Architecture, NVIDIA Corporation, Yokneam, Israel*
e-mail: c.rubio.garcia@tue.nl

ABSTRACT This paper proposes a novel architecture for enabling ultra-fast and ultra-safe 6G networks that can support complex and challenging real-time applications based on four key enabling technologies: 1) performance prediction, 2) AI-enabled task offloading, 3) quantum machine learning, and 4) quantum-resistant communication. With the emergence of 6G applications where the real-time quality of experience is prioritized, AI-enabled task offloading leverages the benefits of edge computing. Moreover, the execution time of complex applications can be reduced by using quantum computers at the edge or in the cloud. In addition, by incorporating quantum key distribution and post-quantum cryptography, we can ensure the safety of mobile networks in the quantum computing era. Collectively, these technologies will provide ultra-fast and ultra-safe 6G networks, meeting the requirements of challenging real-time applications that were not supported in the previous generations, thus advancing the state of the art of mobile communication networks.
Keywords: 6G, edge computing, performance prediction, task offloading, QML, QKD, PQC

1. INTRODUCTION

As the world moves towards an era of billions of interconnected devices, the requirements of mobile networks are shifting from being human-centric to equally considering machine-to-machine type communications. Billions of humans, robots and smart devices will deal with far more complex and challenging applications, generating Zettabytes of digital information whose processing, transmission and security requirements will be higher than ever. An example of such applications are Machine to everything (M2X) communications, augmented and virtual reality (AR/VR), Telepresence and Massive twinning [1]. To enable the aforementioned use cases in mobile networks, our proposed architecture is built on top of the current state of the art 6g architecture proposed by [1], taking into account performance prediction, AI-enabled task offloading, quantum machine learning and quantum-resistant communications. In our view, these technologies are the key drivers for leveraging the benefits of complex AI algorithms by utilizing the computational resources at the edge, far edge and cloud, ultimately facilitating a quantum-resistant, data-driven approach for next-generation mobile networks.

2. 6G ENABLING TECHNOLOGIES

2.1 Performance prediction

Computing and communication systems experience performance variability even when executing the same task [2] and their behavior can change during different periods of time [3]. Such variability can cause fluctuating performance, reduced reliability, increased risks of errors or failures, and potential impacts on quality, and customer satisfaction. Predicting the performance of the system plays a crucial role for addressing the aforementioned issues and delivering fast and reliable connections. The term performance prediction refers to the ability to anticipate and estimate how a network will behave under various conditions, such as the number of connected devices, the type of applications being used, and the physical environment. Predictions are made by generating performance models through analytical modeling, simulation, benchmarking, profiling, as well as, machine learning techniques which capture the behavior and performance characteristics of the system. Such models are generated based on observations or collected historical data from the actual system performance. Accurate performance predictions enable network operators to make informed decisions on optimizing their networks for the best possible user experience. The main benefits of using performance prediction in the context of 6G networks is improved resource allocation and reliability. With accurate performance predictions, network operators can allocate resources where they are needed most, reducing unnecessary traffic and improving overall network efficiency. Furthermore, by predicting potential problems and identifying areas of weakness, network operators can take steps to prevent outages and minimize downtime.

2.2 AI-enabled task offloading

In future 6G networks with demanding use cases and stringent quality of service (QoS) and connectivity requirements, efficient real-time processing cannot depend only on traditional mobile cloud computing (MCC) due to the large propagation delay encountered in the data transmission path. In this context, distributing the computing, networking, and storage resources at the far-edge and edge, through technologies such as multi-access edge computing (MEC), is essential to build a computing continuum to support ultra-low latency applications. Since edge computing is set to be deployed closer to the end users, the path between data source and the target destination is shortened, therefore encouraging more frequent interaction between the edge nodes and the users and at the same time balancing the load on the backhaul network by alleviating the volume of data flows sent to the cloud. In this distributed computing environment envisioned for 6G, large data sets can be collected and pre-processed at the edge, therefore creating an ideal environment for executing AI/QML algorithms that aim to enhance performance by making ultra-fast decisions within the network and bring connected intelligence into both current and future deployments. Furthermore, equipping the network with native and connected AI [4] addresses the need to solve joint optimization problems coming from the time-varying network behavior, e.g., to predict and detect future anomalies and network states [5], as described in Section 2.1.

One particular technology that stands to benefit from both MEC and AI is task offloading which aims to provide enhanced quality of experience (QoE) for end users running applications from the non-exhaustive list that includes mixed reality, immersive games or AI-assisted V2X [1]. Although offloading application tasks to the edge nodes for faster execution can provide an improved experience for the users, it also comes with a set of challenges. First, the decision of whether offloading tasks to the edge rather than processing them locally is not trivial since the communication time across dynamic wireless channels [6] and optical transport network are employed in the process. Moreover, the energy consumption required for task information transmission and computation is another factor weighting in the offloading decision. Second, with the distributed deployment of the edge nodes and the limited available computing resources, where to offload the task while avoiding network congestion is another challenge. To tackle these issues, the network can employ AI to optimize task offloading by leveraging the benefits of a data-driven 6G network where data sets collected from the radio, core and transport domains are used to make optimal offloading decisions.

2.3 Quantum computing & quantum machine learning

Quantum computing (QC) is an area that focuses on building processors based on the principles of quantum physics [7]. Since QC obeys the laws of quantum physics, it can support the parallel processing of multi-dimensional data on large scales, which is known as quantum parallelism. On the other hand, Quantum machine learning (QML) emerged as a combination of ML techniques and QC. The main idea behind QML is to encode classical data into a compliant quantum language so that it can be executed on a quantum computer, by using different encoding methods such as amplitude encoding and Qsample encoding [8].

The range of applications and problems that can be addressed using QML techniques depends heavily on the specific quantum hardware employed in the process. Nowadays, there are three types of quantum computers: 1) Quantum annealers, specialized for optimization problems. 2) Analog quantum simulators, capable of more complex problems and expected to achieve quantum speedup with 50-100 qubits. 3) Universal quantum computers, offering exponential speedup and versatility, and projected to have above thousands of qubits [9]. Quantum annealers and analog quantum simulators are better suited for edge deployment as their design and architectural requirements are less complex and demanding compared to the universal ones. Their specialized capabilities and lower qubit counts make them more feasible for integration into edge devices. This approach which can be named "QML at the edge", will enable the fast pre-processing of the large data sets collected at the edge (as mentioned in Section 2.2), and provide resolution of network optimization problems such as resource allocation and traffic routing. On the other hand, universal quantum computers can be accessed via the cloud as it is very challenging deploying them at the edge due to their complexity, error correction, and scalability requirements. These devices can tackle a wider range of applications and more complex problems including the execution of all the types of advanced quantum algorithms, and providing a sophisticated QML capable of offering a significant speedup over classical techniques, improving various aspects of 6G networks such as data processing and compression, and enhancing various complex AI applications in 6G such as autonomous systems and advanced natural language processing. The integration of Quantum processors in the network edge and having them located also in the cloud will enable more advanced applications that were not supported or realized in the previous generation such as enhanced AR/VR, next-generation healthcare, and massive IoT.

2.4 Quantum key distribution and post-quantum cryptography

Quantum key distribution (QKD) and post-quantum cryptography (PQC) are two technologies that enable quantum-resistant communications. While classical and post-quantum cryptography algorithms rely on the use of mathematical operations easy to compute in one direction, but infeasible to reverse within a reasonable amount of time, QKD guarantees rely on two theorems of quantum physics: The no-cloning theorem and Heisenberg's uncertainty principle. QKD allows two communicating parties to distill a symmetric secret key that is the result of successfully transmitting quantum signals between two authorized partners connected through

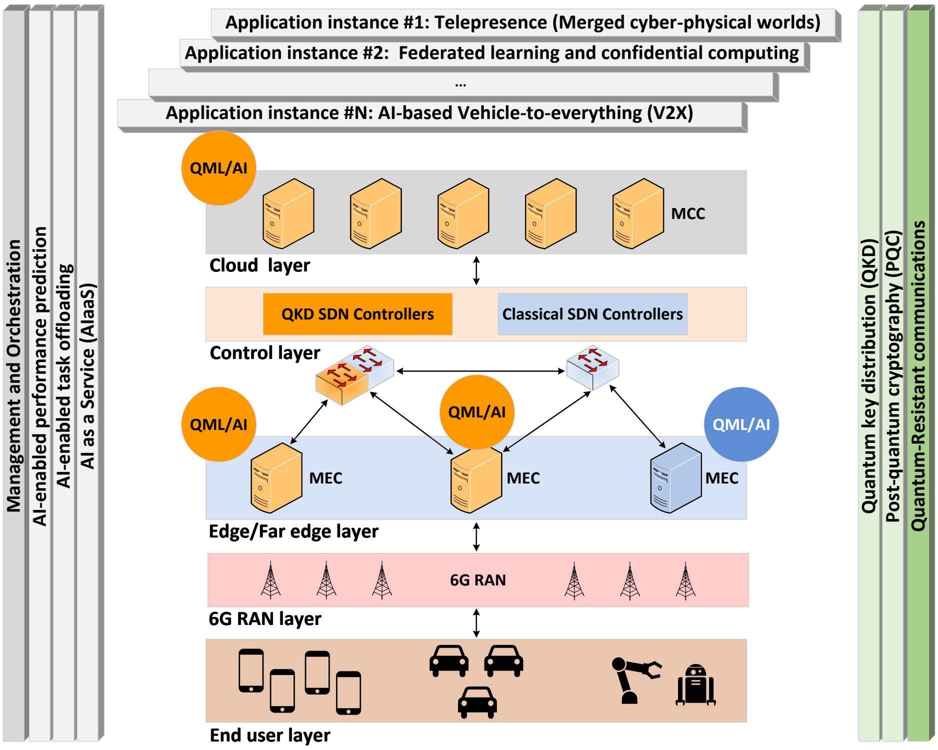

Figure 1: Secure and Agile 6G Networking – Quantum and AI Enabling Architecture

a quantum channel [10]. Any attempt to intercept the communications on the quantum channel will inevitably disturb the quantum signals, changing their original quantum states, indicating the presence of an eavesdropper on the quantum channel. If no eavesdropper is detected, communications are guaranteed to be secure. PQC refers to a collection of public key cryptography (PKC) algorithms that are believed to be secure against the attacks of a quantum computer. The quantum-resilience is achieved by mathematical operations such as lattice-based or hash-based cryptography, for which quantum algorithms offer little to no advantage in speed [11]. PQC cryptography imitates the behavior of classical PKC algorithms, allowing operations such as key-pair generation, digital signatures, key-establishment and key encapsulation/decapsulation.

QKD and PQC can complement each other to provide quantum-resistant end-to-end communications. While QKD provides better security guarantees than PQC, its implementation is harder due to its hardware requirements, requiring a quantum channel through which the quantum signals can be transmitted. For this reason, it is vital that QKD can be integrated with already existing telecommunications infrastructure. Software defined networking (SDN) plays a big role in this regard, as it provides QKD networks with the flexibility needed for an efficient and dynamic allocation of resources, as well as coexistence with current infrastructure. In the parts of the networks where optical communications infrastructure exist, QKD could be implemented for securing the communications. Moreover, QKD could even be combined with PQC to provide a quantum-resistant communication system whose security is guaranteed via two different quantum-resistant assumptions. In the parts of the network where optical fiber is not available, PQC can be employed to ensure quantum-resistant communications. To coordinate this, SDN could guarantee the interoperability among QKD and PQC solutions, and provide quantum-resistant communications for every user, allowing for adaptive and scalable quantum-resistant networks.

3. 6G ARCHITECTURE VISION

The introduction of quantum and AI-based techniques (QC, QML, QKD, PQC, AI-enabled task offloading, etc.) in 6G systems impose the definition of a new architecture with different entities and functionalities. In this regard, in this section, we present our vision for a 6G architecture. The architecture depicted in Fig. 1 is described across five layers: end user, 6G radio access network (RAN), edge/far edge, control, and cloud layer.

In the long-term approach, it is reasonable to consider a geographically distributed architecture where complex AI and QML algorithms, as well as efficient real-time processing, must be offloaded to the network edge, where powerful servers and analog quantum computers could be deployed. In this case, the service provider can choose to offload some of the computation tasks to the network edge. Moreover, in ultra-fast networks, we could imagine a scenario where SDN controllers need to make real-time optimized routing decisions. In this situation, the quantum processors (analog and annealers) can be used in order to execute the task faster than the classical approaches. In other scenarios, where complex applications demand fast processing, a different offloading decision can be made. In the case of edge devices that still do not provide enough performance for real-time functionality, universal quantum computers in the cloud could be considered a solution as mentioned in section 2.3. By calculating the offloading time of either going to the edge or to the cloud, as well as predicting the processing time that a complex task might require either with classical equipment/analog quantum simulators at the edge or universal quantum computers at the cloud, an offloading algorithm could determine that the fastest option is to offload a very complex task to a universal quantum computer on the cloud and solve it using QML algorithms. Although the offloading time itself might increase due to greater geographic distance, the processing time in a fault-tolerance quantum computer could highly outperform the time in a classical machine, enabling real-time processing of very complex applications. In this case, QML can be used for complex computation and heavy learning processes that cannot be supported by conventional techniques or analog quantum computers.

What is also crucial to consider within a 6G distributed network architecture is the risk of having critical network information traveling from the user to the edge, or even to the cloud. In 6G and beyond 6G (B6G) networks, an assumption must be made that any attacker might have access to universal quantum computers and, thus, it should be a requirement that every communication channel within a 6G network architecture must be quantum-resistant, guaranteeing reliability, confidentially, privacy and integrity across different types of environments. Confidential computing and federated learning are examples of such cases in which information needs to be safely transported to either the edge or the cloud for faster and more efficient processing provided by AIaaS technologies. In this regard, the proposed architecture highly fits the security and performance requirements of such applications to be widely adopted in 6G networks, as it allows the secure offload of data either to the edge, far edge, nor to the cloud for further processing.

4. Conclusions

This paper investigates and proposes how performance prediction, AI-enabled task offloading, quantum machine learning, and quantum-resistant communication are key actors in enabling 6G ultra-safe, ultra-fast networks. We describe the aforementioned technologies and the benefits that they bring to the table. We further propose a distributed architecture that allows their coexistence and interoperability, with the end goal of developing a more advanced network architecture that can integrate, in a safe and reliable way, the potential of AI-enabled task offloading and QML for real-time processing of the complex and challenging applications of 6G networks.

Acknowledgment

This work was supported by the EC H2020 MSCA ITN-ETN IoTalentum (GA 953442), the ECSEL-JU project BRAINE (GA 876967), the Dutch Ministry of Economic Affairs and Climate Policy (EZK), as part of the Quantum Delta NL programme and the ADAPTOR project (NWO OTP-18651).

REFERENCES

[1] 5G PPP Architecture Working Group. The 6g architecture landscape. Technical report, 5GPPP, feb 2023.

[2] Aleksander Maricq, Dmitry Duplyakin, Ivo Jimenez, Carlos Maltzahn, Ryan Stutsman, and Robert Ricci. Taming performance variability. In *2018 13th USENIX Symposium on Operating Systems Design and Implementation*, OSDI'18, page 409–425, USA, 2018. USENIX Association.

[3] Dmitry Duplyakin, Alexandru Uta, Aleksander Maricq, and Robert Ricci. In datacenter performance, the only constant is change. In *2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID)*, pages 370–379, 2020.

[4] Carlos J. Bernardos and Mikko A. Uusitalo. European vision for the 6G network ecosystem, June 2021.

[5] R. Boutaba, Mohammad Salahuddin, Noura Limam, Sara Ayoubi, Nashid Shahriar, Felipe Estrada-Solano, and Oscar Caicedo Rendon. A Comprehensive Survey on Machine Learning for Networking: Evolution, Applications and Research Opportunities. *J. Internet Serv. Appl.*, 9, 05 2018.

[6] Catalina Stan, Simon Rommel, Ignacio de Miguel, Juan José Vegas Olmos, Ramón J. Durán, and Idelfonso Tafur Monroy. 5G Radio Resource Allocation for Communication and Computation Offloading. In *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2023.

[7] Eleanor Rieffel and Wolfgang Polak. An introduction to quantum computing for non-physicists. *ACM Computing Surveys (CSUR)*, 32(3):300–335, 2000.

[8] Tariq M Khan and Antonio Robles-Kelly. Machine learning: Quantum vs classical. *IEEE Access*, 8:219275–219294, 2020.

[9] Francesco Tacchino, Alessandro Chiesa, Stefano Carretta, and Dario Gerace. Quantum computers as universal quantum simulators: state-of-the-art and perspectives. *Advanced Quantum Technologies*, 3(3):1900052, 2020.

[10] Valerio Scarani, Helle Bechmann-Pasquinucci, N. Cerf, Miloslav Dusek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Review of Modern Physics*, 81:1301, 09 2009.

[11] Daniel J Bernstein. *Introduction to post-quantum cryptography*. Springer, 2009.