Trustworthy AI for Next Generation Networks: the Fed-XAI innovative paradigm from the Hexa-X EU Flagship Project

Pietro Ducange[1,*], Francesco Marcelloni[1], Davide Micheli[3], Giovanni Nardini[1,2], Alessandro Renda[1], Dario Sabella[4], Giovanni Stea[1] and Antonio Virdis[1]

[1] Department of Information Engineering, University of Pisa, Largo Lucio Lazzarino 1, 56122 Pisa, Italy
[2] Center for Logistic Systems, University of Pisa, Via dei Pensieri 60, 57124 Livorno, Italy
[3] Telecom Italia S.p.a., Torino, Italy
[4] Intel Corporation Italia S.p.a, Viale Milano Fiori 4, 20057 Assago (MI), Italy

Abstract

This work presents the joint research activities on AI in and for 6G carried out by University of Pisa, Intel Corporation Italia s.p.a. and Telecom Italia s.p.a., within the Hexa-X EU project. Specifically, we focus on Federated Learning of Explainable Artificial Intelligence (Fed-XAI), which has been recently awarded as key innovation by the EU Innovation Radar. We present the main recent achievements, that can be summarised in algorithms for generating federated XAI models in a privacy-preserved environment, a communication framework for Federated-Learning-as-a-Service and orchestration algorithms of federated learning participants. Finally, we discuss a proof of concept, that showcases the aforementioned Fed-XAI components.

Keywords

Federate Learning, Explainable AI, Simu5G

1. Introduction

The fourth (4G) and fifth (5G) generations of cellular networks have not only allowed billions of people to communicate with each other, but they have also supported and boosted the digitalization of industries and public administrations, machine-to-machine communications, distributed wireless computing and several advanced services tailored for a wide range of customers. As 5G becomes increasingly established commercially and technologically, the academic and industrial world has already begun to lay the groundwork for sixth-generation (6G) cellular networks [1]. Indeed, it is expected that 6G will be deployed starting from 2030.

In this context, in 2020 European Commission funded the Hexa-X project[1] within the Horizon 2020 programme[2]. Hexa-X provides a flagship vision targeted to lay the foundations of Beyond 5G (B5G) and 6G systems [2], considering *trustworthiness*, *inclusiveness* and *sustainability*, as its three core values of next generation communication networks [3]. Trustworthiness requirements play a key role in Hexa-X as B5G/6G networks are expected to support a plethora of applications based on artificial intelligence (AI). Indeed, most of these applications will be highly pervasive in the daily processes of individuals, companies and institutions, and also used in critical domains. Therefore, native B5G/6G applications, based on AI models, pose increasing privacy, security and trust issues.

Within the Hexa-X project, the working group composed by University of Pisa, Intel Corporation Italia s.p.a. and Telecom Italia s.p.a. (TIM), collaborates in research activities in the context of federated learning (FL) of explainable AI (XAI) models, denoted as Fed-XAI[4] in the following. Fed-XAI can be regarded as an enabler for several families of use cases envisioned for B5G/6G and is proposed as a machine learning (ML) paradigm compliant for building trusted and ethical next generation networks. Indeed, on one hand it respects data privacy and on the other hand it ensures the transparency of the AI models [5]. Notably, Fed-XAI has been recently awarded as *key innovation* by the EU Innovation Radar[3].

The research activities of the working group include the development of: i) ad hoc FL strategies for learning XAI models, ii) a communication framework supporting FL as a service (FLaaS) on Multi-access Edge Computing

Ital-IA 2023: 3rd National Conference on Artificial Intelligence, organized by CINI, May 29–31, 2023, Pisa, Italy

*Corresponding author.
pietro.ducange@unipi.it (P. Ducange)
0000-0003-4510-1350 (P. Ducange); 0000-0002-5895-876X (F. Marcelloni); 0000-0002-7851-0532 (D. Micheli); 0000-0001-9796-6378 (G. Nardini); 0000-0002-0482-5048 (A. Renda); 0000-0002-8723-7726 (D. Sabella); 0000-0001-5310-6763 (G. Stea); 0000-0002-0629-1078 (A. Virdis)

[1] Hexa-X website: www.hexa-x.eu
[2] Horizon 2020 - G.A.: 101015956
[3] https://www.innoradar.eu/innovation/45988

architectures, and iii) orchestration algorithms for FL. Moreover, a proof-of-concept will be delivered: a real Fed-XAI application will be deployed on an edge computing node of an emulated B5G network. Specifically, we consider a realistic scenario Vehicle-to-Everything (V2X) [4], in which users collaborates to build an XAI model, in a FL fashion, for performing Quality of Experience (QoE) prediction for video streaming [5].

The rest of the paper is organized as follows. Section 2 provides some details on the activities carried out in the Hexa-X project by the research group. In Section 3, we briefly introduce our proof-of-concept. Section 4 argues on the impact that Fed-XAI may have in the market. Finally, Section 5 draws concluding remarks.

2. Activities on Fed-XAI in Hexa-X

The enabling of Fed-XAI within the Hexa-X project required to carry out the following four different and complementary activities:

- developing ad-hoc *strategies* for FL of XAI models, specifically rule-based models;
- developing a *Federated-Learning-as-a-Service (FLaaS)* communication framework;
- developing *orchestration strategies* for the FLaaS architecture;
- gathering, analyzing and pre-elaborating real data from live TIM mobile network.

The results of these activities have been embodied in the Fed-XAI demo, which is described in Section 4. In the following, we will describe the activities in detail.

2.1. Developing strategies for FL of XAI models

The objective of Fed-XAI is to devise algorithmic solutions to enable the collaborative training of AI systems with an adequate degree of explainability, according to the FL paradigm.

In *standard FL* a centralised topology is considered: a central entity fulfils the role of orchestrating the process and aggregating local model contributions, coming from the data owners, into a global updated model. In such a setting, FL algorithms mainly aim to optimize a global differentiable objective function, e.g., through variants of stochastic gradient descent (SGD) such as Federated Averaging (FedAvg) and Federated SGD. Consequently, models like Neural Networks, typically optimized through SGD, are immediately suitable for this type of collaborative optimization and have been widely investigated in the FL literature. However, they are generally considered as opaque, i.e. hard to interpret [6].

In certain scenarios other classes of models, such as Decision Trees (DTs) and Rule-based Systems (RBS), may be advisable due to their inherent interpretability and competitive performance, especially when data are represented in tabular form. However, they require ad-hoc training strategies for the federated setting since they are not typically learned through the optimization of a differentiable global objective function. Recently, a lot of research effort in the field of Fed-XAI aims to address this challenge. In the framework of FL of transparent models, Takagi–Sugeno-Kang Fuzzy Rule-Based systems (TSK-FRBSs) have been recently investigated [7]. These regression systems simply consist of collections of rules in the form "IF *antecedent* THEN *consequent*", where the parameters (in both the antecedent and consequent parts) are learned from the training data: the antecedent of a rule identifies a specific region of the attribute space through highly interpretable linguistic labels, whereas the corresponding consequent allows generating the predicted output within such a region as a linear combination of the input variables. In one of our recent works [8] we proposed a novel approach for FL of TSK-FRBS under the orchestration of a central entity. A high level illustration of the approach is depicted in Fig. 1.

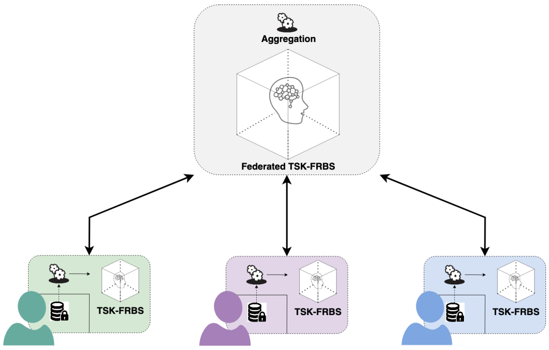

Figure 1: High level illustration of FL of TSK-FRBS.

In a nutshell, our approach involves an one-shot communication scheme: each participant learns a TSK-FRBS based on its local and private data, and shares the rule base with the central server. Then, the aggregation procedure consists in the juxtaposition of the rule bases received from the involved participants and in the resolution of rule conflicts. A conflict emerges when two or more rules from different participants have the same antecedent and different consequents. The aggregation strategy in this case consists in combining the conflicting rules into a single one with the same antecedent: the coefficients of the linear model of the new consequent are evaluated as the average of the coefficients of the original rules, weighted according to the support (how much a rule is activated) and confidence (how good the rule is in modelling data) of each rule as measured on the training data. More details on the proposed FL scheme for TSK-FRBS along with some preliminary results can be found in [8]

2.2. Developing the FLaaS Communication Framework

The implemented communication framework for FLaaS allows B5G/6G networks to offer support for FL applications to their users. The objective is to offer users (i.e., User Equipments or UEs) the ability to: i) discover FL services and applications that the network may support, ii) join them and therefore use XAI models built in a federated fashion, and iii) contribute to the FL process without sharing their data. The proposed FLaaS is meant to accommodate UEs with high computing power (e.g., cars), that can both acquire data and participate in the FL process, as well as UEs with low computing power (e.g., IoT devices), which can sense but would not be able to participate in the FL process. This last type of devices can leverage Multi-access Edge Computing (MEC) [9], and delegate the computationally onerous task of training and exchanging models to MEC Apps. In our FLaaS, network operators or third parties define onboard FL services. UEs, on the other hand, have an interface to discover the FL services offered by their network, to join/leave running instances of these services, to participate in the FL training process, or to just receive model updates. FLaaS can run simultaneously several instances of the same FL service, called *FL processes*, which are useful to handle different sets of UEs in different geographical areas. This also allows the network to *orchestrate* FL processes according to some criteria, for instance the type of UEs, the type of data that they possess, the geographical location, the computation capabilities, etc.

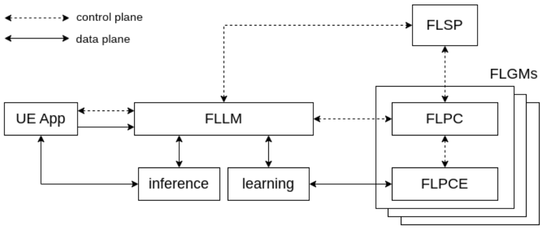

Figure 2: Overview of the entities involved in the FLaaS framework

Figure 2 presents a high-level view of the FLaaS framework. The FL Service Provider (FLSP) maintains a library of available FL services and the list of running FL processes. An FL Local Manager (FLLM) is associated with each data owner (i.e., UE), and there is one FL Global Manager (FLGM) per FL process, that is responsible for model aggregation. As already stated, FLLMs may reside directly on the UE or run as MEC Apps. They include the logic to interface with the FLGM, and they may include a learning module, for allowing their owner to participate in the FL process, and an inference module, for making prediction using an available XAI model. The FLGM comprises two modules: the FL process controller (FLPC) and the FL process computation engine (FLPCE). The former manages control-plane interactions with the FLSP (e.g., authorization grants) and the FLLMs, and the latter is the entity that actually builds the global AI model, by exchanging local and global AI model updates with the learning modules of the FLLMs.

We implemented the above FLaaS framework within Simu5G [10, 11], a *model library* for the OMNeT++ discrete-event simulation framework[4], which provides the modules for the simulation of the data plane of 4G/5G mobile networks, as well as for ETSI MEC. In the latter, FL entities are developed as ETSI MEC applications running on a MEC host, with the exception of the FLLM, that can instead be deployed on the UE according to its needs.

2.3. Developing Orchestration Strategies for the FLaaS architecture

AI/ML-driven processes are expected to be crucial in managing several aspects of B5G/6G networks, and therefore are being natively included both as elements to be managed by the infrastructure and as elements aiding the infrastructure management. One of the Hexa-X activities has focused on this aspect by defining an orchestration architecture to support the lifecycle of AI/ML-driven processes, enabling their dynamic creation, configuration and migration, but also supporting other orchestration entities to make use of above mentioned processes [12]. As anticipated in the previous section, the proposed FLaaS framework support this view thanks to its ability to execute several instances of the same FL processes, simultaneously and independently, thus paving the way for enforcing complex orchestration policies.

Orchestration of AI/ML processes is indeed a critical feature to ensure the feasibility of the proposed Fed-XAI, as an FL process may involve a large number of participants. From a system perspective it is not desirable to include all of them in the same FL instance. Within this context, several orchestration policies can be considered [13]. A first type of orchestration strategies may regard the selection of participants to the FL process. Several performance KPIs, such as model accuracy, model explainability and communication-resource consumption, may be considered in the selection process. The

[4] https://omnetpp.org

strategies should take into consideration the different trade-offs between the different KPIs, with the aim of meeting the defined requirements (both functional and non-functional). A second type of orchestration strategies may address the problem of selecting the timing that governs FL interactions, such as how often agents should report to the federation server or how often the server should send updates to federated agents. The orchestration policies discussed above can only be effectively addressed if the management of FL tasks is included in a holistic orchestration system that can i) enforce coordination among multiple FL tasks to avoid overloading the underlying communication/computation resources; ii) schedule and reserve communication and computation resources, possibly working proactively; and iii) monitor the various performance indicators of each FL task, checking whether the service requirements are met or not.

2.4. Data gathering, analysis and pre-elaboration from TIM live mobile network

In order to test the whole Fed-XAI framework with realistic B5G/6G data, we gathered live measurements from TIM's mobile 5G network, such as base stations positions and user data volume. Moreover, we also acquired geolocated data from live TIM's Radio Area Network (RAN) by exploiting the Minimization of Drive Test (MDT) functionality. After a preliminary analysis of the extracted data, they have been pre-elaborated, and aggregated in order to use them for feeding Simu5G for generating realistic network scenarios for some envisioned B5G/6G applications, such as Tele-operated Driving (TOD). The resulting generated datasets are publicly available[5], and can be used for building AI models to predict the future quality of video-streams. The advantage of our data-driven approach adopted in Fed-XAI is twofold: first, privacy preservation is guaranteed by leveraging FL during collaborative training of XAI models, especially suitable in heterogeneous B5G/6G scenarios; second, an adequate degree of explainability of the models themselves is ensured, with benefits for industrial customers in terms of high dependability, and for end-users in terms of trustworthiness.

3. Fed-XAI Proof-of-Concept

Results of the Fed-XAI activities have been also made apparent in a Proof of Concept (PoC) within the Hexa-X project. The PoC shows how XAI models, trained in a federated fashion, are able to predict the quality of a video stream in a vehicular network context.

We consider a ToD scenario, where connected cars send video streams to a remote driver at the edge of the network, representing the pilot-seat view. ToD is only safe if this video stream plays smoothly and with good enough quality. This makes it important to be able to *predict* loss of quality in advance, in order to allow both the network and the remote driver to take the appropriate countermeasures. Individual cars participate in the collaborative training of an XAI model for predicting the quality of the video streaming, under the Fed-XAI paradigm and using the FLaaS framework discussed above. The XAI model enables to identify root causes of quality degradation, which is useful for network/system (re)configuration, service level agreement (SLA) verification, and appropriate online reaction on part of the remote driver. For this reason, the Fed-XAI PoC provides a dashboard showing the predictions and root causes in real time.

In an initial (offline) *training phase* based on the FL learning scheme for TSK-FRBS models discussed in Section 2.1, we exploit Simu5G to generate a training dataset including a large set of Quality of Service (QoS)-related data produced by video-streaming sessions. In a subsequent (online) *prediction phase*, Quality of Experience (QoE) of the video streaming is predicted and the predictions, along with their explanations, are shown on the dashboard in real time. The actual video streaming and the prediction of its QoE are realized by emulating in real time a portion of 5G network, which covers roads with semaphore-regulated intersections where connected cars move. These cars locally run the sender side of the video-streaming application, whereas the receiver side (that represents the remote driver) runs as an application on a MEC host. The traffic scenario is set based on data extracted from TIM's live network, extrapolated to model a future 6G traffic configuration. Moreover, the position of Base Stations (BSs) in the emulation mirrors the actual one in the city of Turin.

Figure 3 shows the scheme of the PoC testbed. It includes the real-time emulated network: the cars as connected to it as 5G UEs by means of 5G base stations called gNodeBs (gNBs). The network is emulated in real time on a desktop machine running Simu5G [14] and also the MEC system is included in the emulation. The video source and the video player are hosted on two different laptops, connected via Ethernet interfaces to the machine that runs Simu5G. Packets generated by / addressed to external interfaces are swapped in real time with Simu5G messages, so that they experience the same network treatment that they would in a mobile network.

The Fed-XAI application supports the collaborative learning of XAI models and their exploitation for inference purposes and is deployed on a dedicated server. The application includes the following modules: the Fed-XAI Local Manager can be considered as a UE-related control

[5] http://docenti.ing.unipi.it/g.nardini/ai6g_qoe_dataset.html

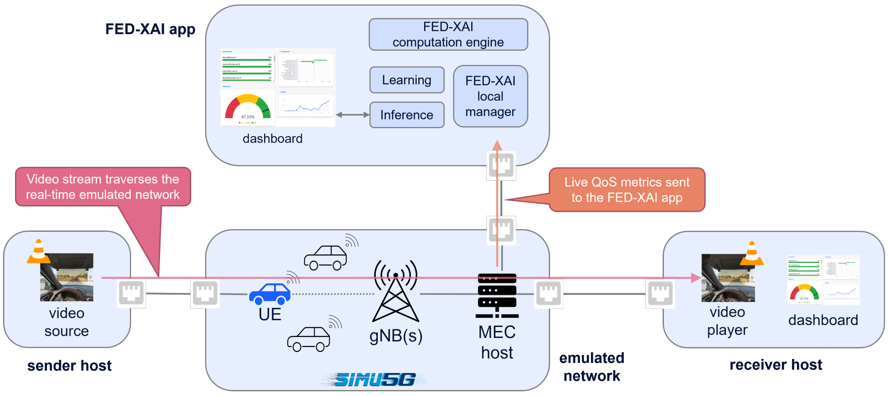

Figure 3: Scheme of the testbed of the Fed-XAI PoC

entity of the FL process; the Learning module performs the training stage of local XAI models; the Fed-XAI Computation Engine aggregates the local models and produces a federated model; the Inference module allows the end-user to leverage an XAI model for performing inference on local data; the Fed-XAI Dashboard allows users to visualize prediction and to leverage explainability to uncover model operation. All Fed-XAI application modules have been designed and implemented to be compliant with an edge computing environment: we exploit a fully virtualized architecture and deploy each module inside a container of the Docker ecosystem[6] so as to allow portability regardless of the underlying hardware and software infrastructure and for easier migration in real-world mobility scenarios [15]. The actual Fed-XAI process exploits the Intel OpenFL[7] library, purposely extended to support FL of inherently interpretable models, such as TSK-FRBS. Finally, the Fed-XAI dashboard has been implemented as a web application.

4. Impact of the proposed Fed-XAI approach

A benefit of using the Fed-XAI approach is the increase of trust in AI for 6G-enabled services. This has an immediate business impact for 6G stakeholders. In fact, the V2X applications described above are typical cases where a collaboration (and related business agreement) is needed between mobile network operators (MNOs), possibly in partnership with edge-computing service providers and car original equipment manufacturers (OEMs). Both car OEMs and MNOs can benefit from explanations about XAI models predictions and any consequent decision making: MNOs can provide a more explainable set of 6G functionalities (e.g., FL agents enabling QoS predictions) and expose them to their customers (including car OEMs, but also application developers and system integrators); OEMs can also benefit from more information on network predictions, exploitable to improve the automated driving features offered to their end-customers (i.e., the actual drivers). More in detail, from a business perspective, the goal of OEMs is to provide V2X services and advanced features related to automated driving that can leverage the functionalities of communication networks. For such advanced functionalities, the prediction of the levels of QoS and QoE are critical to assess the reliability of the network and of the offered functionalities themselves. On the other hand, MNOs need to provide the required performance in a trusted environment, so that an agreement with the OEMs can be found. The boundary between MNOs and OEMs worlds is typically governed by a set of Service Level Requirements (SLRs), defining indeed the terms and conditions of the agreement between these two stakeholders (see 5GAA reports for the V2X cases[8]). These SLRs are service-specific, and may include minimum throughput, maximum delay, but also availability and reliability of the guaranteed KPIs. With this in mind, in the transition to 6G, accurate, timely, and explainable predictions are critical to provide advanced and very challenging use cases with a time horizon ranging from extremely short to long periods. Thus, it is evident how Fed-XAI is critical to improve mutual understanding and trust among stakeholders in the 6G ecosystem.

5. Conclusions

Federated Learning of eXplainable AI models (Fed-XAI) can be regarded as a key enabler towards trustworthy AI in several application domains, including next gen-

[6] https://www.docker.com/products/container-runtime/
[7] https://github.com/intel/openfl
[8] https://bit.ly/3JT1GXV

eration wireless networks. In fact, it ensures *privacy preservation* during the model training stage and *explainability* through the adoption of highly interpretable models. The study of Fed-XAI is in its early stages: in this paper we have reported on three dimensions which are being investigated in the context of the Hexa-X EU flagship project: the design of ad-hoc strategies for FL of XAI models; the proposal of a communication framework supporting FLaaS on MEC architectures; the design of orchestration algorithms for FL. A Proof-of-Concept is also described, aimed at showcasing the adoption of Fed-XAI technologies in a vehicular networking application. Finally, the potential business impact of the adoption of such paradigm is discussed with reference to the same vehicular scenario.

Acknowledgments

This work was partially supported by: the EU Commission through the H2020 projects Hexa-X (Grant no. 101015956); by the PNRR - M4C2 - Investimento 1.3, Partenariato Esteso PE00000013 - “FAIR - Future Artificial Intelligence Research” - Spoke 1 “Human-centered AI”; by the PNRR “Tuscany Health Ecosystem” (THE) (Ecosistemi dell’Innovazione) - Spoke 6 - Precision Medicine & Personalized Healthcare (CUP I53C22000780001) under the NextGeneration EU programme; by the Italian Ministry of University and Research (MIUR), in the framework of the FoReLab project (Departments of Excellence); and by the Center for Logistic Systems of Livorno.

References

[1] M. A. Uusitalo, P. Rugeland, M. R. Boldi, E. C. Strinati, P. Demestichas, M. Ericson, G. P. Fettweis, M. C. Filippou, A. Gati, M.-H. Hamon, et al., 6G vision, value, use cases and technologies from european 6G flagship project Hexa-X, IEEE Access 9 (2021) 160004–160020.

[2] M. A. Uusitalo, M. Ericson, B. Richerzhagen, E. U. Soykan, P. Rugeland, G. Fettweis, D. Sabella, G. Wikström, M. Boldi, M.-H. Hamon, et al., Hexa-X the European 6G flagship project, in: 2021 Joint European Conf. on Networks and Communications & 6G Summit (EuCNC/6G Summit), IEEE, 2021, pp. 580–585.

[3] Hexa-X Deliverable D1.2 - Expanded 6G vision, use cases and societal values – including aspects of sustainability, security and spectrum, 2021. URL: https://hexa-x.eu/wp-content/uploads/2021/05/Hexa-X_D1.2.pdf.

[4] A. Renda, P. Ducange, F. Marcelloni, D. Sabella, M. C. Filippou, G. Nardini, G. Stea, A. Virdis, D. Micheli, D. Rapone, et al., Federated Learning of Explainable AI Models in 6G Systems: Towards Secure and Automated Vehicle Networking, Information 13 (2022) 395.

[5] J. L. C. Bárcena, M. Daole, P. Ducange, F. Marcelloni, A. Renda, F. Ruffini, A. Schiavo, Fed-XAI: Federated Learning of Explainable Artificial Intelligence Models, in: 3rd Italian Workshop on Explainable Artificial Intelligence (XAI.it 2022), 2022.

[6] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, et al., Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Information fusion 58 (2020) 82–115.

[7] X. Zhu, D. Wang, W. Pedrycz, Z. Li, Horizontal Federated Learning of Takagi–Sugeno Fuzzy Rule-Based Models, IEEE Transactions on Fuzzy Systems 30 (2022) 3537–3547.

[8] J. L. C. Bárcena, P. Ducange, A. Ercolani, F. Marcelloni, A. Renda, An Approach to Federated Learning of Explainable Fuzzy Regression Models, in: 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), IEEE, 2022, pp. 1–8.

[9] F. Giust, X. Costa-Perez, A. Reznik, Multi-access edge computing: An overview of ETSI MEC ISG, IEEE 5G Tech Focus 1 (2017) 4.

[10] G. Nardini, D. Sabella, G. Stea, P. Thakkar, A. Virdis, Simu5G–An OMNeT++ Library for End-to-End Performance Evaluation of 5G Networks, IEEE Access 8 (2020) 181176–181191.

[11] A. Noferi, G. Nardini, G. Stea, A. Virdis, Rapid prototyping and performance evaluation of etsi mec-based applications, Simulation Modelling Practice and Theory 123 (2023) 102700. doi:https://doi.org/10.1016/j.simpat.2022.102700.

[12] J. Péerez-Valero, A. Virdis, A. G. Sánchez, C. Ntogkas, P. Serrano, G. Landi, S. Kukliński, C. Morin, I. L. Pavón, B. Sayadi, AI-driven Orchestration for 6G Networking: the Hexa-X vision, in: 2022 IEEE Globecom Workshops (GC Wkshps), 2022, pp. 1335–1340.

[13] Hexa-X Deliverable D6.2 - Design of service management and orchestration functionalities, 2022. URL: hexa-x.eu/wp-content/uploads/2022/05/Hexa-X_D6.2_V1.1.pdf.

[14] G. Nardini, G. Stea, A. Virdis, Scalable Real-Time Emulation of 5G Networks With Simu5G, IEEE Access 9 (2021) 148504–148520.

[15] L. Ma, S. Yi, Q. Li, Efficient Service Handoff across Edge Servers via Docker Container Migration, in: Proceedings of the Second ACM/IEEE Symposium on Edge Computing, SEC ’17, Association for Computing Machinery, New York, NY, USA, 2017.