The 5G-IANA platform: Bringing far-edge resources and ML potential to the disposal of automotive third parties

Francesca Moscatelli*, Thanos Xirofotos†, Amr Rizk‡, Nehal Baganal-Krishna‡, Edoardo Bonetto§
Eirini Liotou¶, Angelos Amditis¶
*Nextworks, Pisa, Italy
†UBITECH, Athens, Greece
‡University of Duisburg-Essen, Essen, Germany
§LINKS Foundation, Turin, Italy
¶Institute of Communications and Computer Systems, Athens, Greece

***Abstract*—The advent of 5G has been a game-changer in the Automotive Vertical sector. 5G acts as an enabler of advanced networking architectures which, in turn, allow the development of novel services, by attracting and engaging third parties to experiment using available 5G infrastructure and connectivity. 5G-IANA is a Horizon 2020 ICT-41 project that targets to create such an advanced open experimentation platform, bringing powerful novelty to third parties through the integration and provisioning of far-edge resources for network orchestration purposes, and through the offering of Machine Learning knowledge as an add-on service to them. The scope of this paper is, therefore, to present the 5G-IANA project overall goals, by emphasizing the most important expected innovations from the project regarding far-edge resources' orchestration and ML as a service provisioning.**

***Index Terms*—5G, automotive vertical, edge resources, orchestration, distributed machine learning**

I. Introduction

5G-based Automotive-related services (i.e., Connected and Automated Mobility services) are a broad range of digital services in and around vehicles including both safety-related and other commercial services provided, enabled, or supported by 5G networks. With the roll-out of 5G, mobile networks will offer a broad range of connectivity performances including gigabit speeds and mission critical reliability. Most importantly, 5G will be a unified multi-service platform, serving not only the traditional mobile broadband market but also enabling digital transformation in a number of vertical industries, which will result in the creation of unprecedented opportunities for innovation and economic growth.

In parallel to these opportunities that 5G brings, there is a clear momentum both from the European Commission and the actual market needs to deploy and provide integrated, open and cooperative experimentation infrastructures (i.e., platforms) on top of which third party experimenters tailored to specific vertical sectors will test their applications. Such third parties mainly entail Small and Medium Enterprises (SMEs), which, by themselves, do not have the means or the capabilities to enter the 5G services market without any support from larger 5G players. On top of that, the need to provide predefined 5G open source repositories that include, for instance, already available Virtual Network Functions (VNFs) or even chains of such VNFs in the forms of Network Applications (”NetApps”), is also critical. In this way, SMEs can relatively easily and quickly deploy their own services, applications or algorithms on top of reliable and accessible 5G experimentation platforms.

In this context 5G-IANA project's objective is to create and offer an open accessible environment in which third party developers can easily experiment and test their automotive-related applications. To do so, the project designs the Automotive Open Experimentation Platform (AOEP) as a multi-layered platform that extends from the end user (application) layer to the infrastructure layer and optimally combines context and network infrastructure-aware functionalities for supporting the deployment of advanced services represented as linked chains of virtualised Application Functions (AFs) and Network Functions (NFs). 5G-IANA offers also a catalogue of ready-to-use NetApps containing already chained AFs and/or NFs that a Vertical may want to re-use and eventually modify/extend to deliver 5G-enabled services. Then, the project provides a whole set of hardware and software resources that includes the computation and communication/transport infrastructure as well as the management and orchestration components, coupled with a NetApp Catalogue mostly tailored to the Automotive sector. More specifically, the target is to:

1) deliver an Automotive experimentation platform that operates on top of a 5G infrastructure (extending to vehicles) and provides a NetApp Catalogue and an AFs/NFs Repository open to third parties. In particular, the NetApp Catalogue will host a set of Starter Kits, i.e., baseline NetApps grouped per service category (e.g., vehicle movement, infotainment etc.), that third parties can use/include in new service chains to develop and design their own Automotive service,
2) provide functionalities for easing the design and chain-

ing of new Automotive services as well as their provisioning, validation and and bench-marking. These functionalities will be exposed to third parties through an enhanced Graphical User Interface (GUI) that will enable also the specification of service-related high-level Quality of Service (QoS) requirements,

3) deploy and orchestrate Automotive services and related components (i.e., NetApps and AFs) from both the application and the networking point of view and to monitor and dynamically adapt them at run-time. Emphasis is given to the "lightweight" orchestration on top of On-board Units (OBUs)/Road-side Units (RSUs) for offering a more flexible and scalable management of the related resources,
4) facilitate the definition and re-usability of ML models by any Vertical service that needs to train models in its business logic, e.g., Federated Learning iterative process,
5) abstract the infrastructure and technology complexity and provide uniform domain-agnostic deployments, enabling cross-domain and cross-platform interoperability.

The main project novelties lie, therefore, on two main aspects. The first one concerns the integration and provisioning of far-edge resources for network management and orchestration purposes, which will allow third party developers to take full advantage of the capabilities offered by even the far-edge nodes of the infrastructure (i.e., OBUs attached on vehicles and RSUs deployed on the roads). In this direction, programmable OBUs/RSUs ready to host containerized services are envisioned by the platform. The second novel aspect is the availability of a Distributed Machine Learning (DML) framework as part of the 5G-IANA platform and of the correlate collected information as a service to the Over The Top (OTT)/Original Equipment Manufacturer (OEM) third party. As an example, network monitoring with respect to data traffic will yield QoS predictions that support the service provisioning decisions of such third parties.

The remainder of the paper is structured as follows: Section II describes main edge-related technical challenges that the project addresses, which concern the support of the requirements of automotive-related services facilitated by the introduction of edge computing, as well as the challenges to support the Life Cycle Management of far-edge devices. Then, Sections III and IV present the two main project novelties, which concern respectively the orchestration of far-edge resources and the support of Machine Learning (ML) as a service. Next, Section V presents the design overview (in terms of components and functions) of the 5G-IANA platform, which will enable the deployment of both innovations previously described, followed by the Conclusions section.

II. Technical Challenges

A. Automotive-related

Edge computing is one of the most important technical aspects to be considered when rolling out Vehicle-to-Everything (V2X) use cases and services, especially considering strict performance requirements in terms of low latency and high reliability. The goal is to move computing and storage out of the central cloud and closer to end-users and data sources to meet the desired QoS requirements, e.g., latency in case of ultra Reliable Low Latency Communication (uRLLC) services. Multi-access Edge Computing (MEC) is considered as a fundamental enabler for supporting the roll-out of new Automotive use cases and services while leveraging network slicing for tailored network deployments that fulfil customers' requirements and demands. V2X use cases include a variety of services with different requirements to be considered depending on their specific scope/objectives, spawning from the vehicle movement and hazard notification service areas, where applications are usually latency demanding, till infotainment and AR/VR services which require high Uplink (UL)/Downlink (DL) throughput. One of the major challenges is related to MEC applications design for achieving the expected performances service developers need to identify the functionalities that require processing at the edge with respect to others that do not require a near-real-time exchange of information and can be deployed at the central/remote cloud locations. Therefore, the design of the desired service-chain should support distributed processing, synchronisation of contexts, and multi-level load-balancing [1]. The Automotive market highly influences edge computing architectures, targeting scenarios with multiple vehicles, from different manufacturers, and many other devices (e.g. sensors, actuators, smartphones etc.) connected to the infrastructure, e.g., RSUs, and the mobile network.

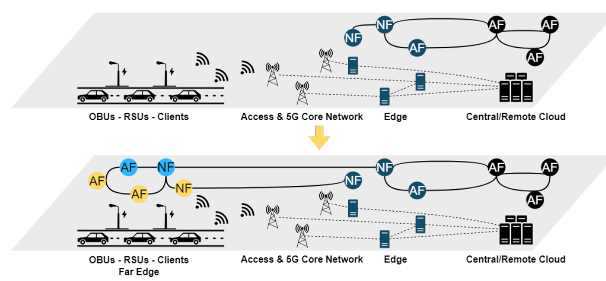

Fig. 1. Example of a service-chain extended to the far edge starting from a three-tier usual deployment

The introduction of edge computing brings a shift towards three-tier deployments, where UEs (i.e., client-side components) are expected to communicate with service server-side components potentially distributed across the edge and central cloud locations. In beyond-5G networks, a further challenge consists in extending the deployment of server-side components to the UE side, which, as depicted in Fig. 1, in the Automotive segment mostly means relying on orchestration-enabled OBUs and RSUs. In [2], ETSI analyses MEC-related requirements in terms of deployments and interoperability for a set of use cases. With respect to the possibility of executing service workloads on top of OBUs acting as MEC hosts, on one hand, the main identified benefits concern an enhanced contextual awareness by exploiting radio network, location, and/or other information relevant to the changing environment during journey time. On the other hand, the fluctuation of connectivity conditions an in-vehicle MEC host may experience, together with its possibly limited compute resources (i.e., processing, memory and storage capabilities) may impact its efficient management.

B. Applicability of ETSI MEC in far-edge device

The far-edge devices are typically characterized by a limited amount of storage, memory and computing resources. These characteristics of the far-edge environment have a deep impact on the most suitable solutions to be adopted for the virtualization, the life management, and the orchestration of the applications to be run on far-edge devices. Lightweight solutions are envisaged so that the limited available resources are not exhausted.

Mobile far-edges, such as the OBU located on vehicles, can also experience disruption in the mobile network communication in geographical areas where network coverage is limited. In these situations, the far-edge orchestration should be self-dependent, and it should not be relying on edge or network support.

Another limitation is the limited availability of information on the far-edge devices with respect to the MEC server case. This can limit the ETSI MEC API that can be offered on far-edge devices. For instance, the ETSI GS MEC 013 “Location API” provides position information about UE terminals in the area covered by the MEC server exploiting network-based information. The far-edge devices are unlikely to have this information available from the network as they are UE themselves. The “Location API” could then be hardly implemented at the far-edge devices in full compliance with the standards. Some adaptations should be introduced such as retrieving UEs position information from the MEC server or from other sources.

In Section III, the 5G-IANA approach for the dynamic orchestration of far-edge resources (i.e., OBUs and RSUs) is presented, while in Section V the AOEP proposed design is described, introducing the architecture components and functionalities proposed to address the far-edge orchestration challenge.

III. Orchestration of far-edge resources

A. State of the art on orchestration solutions

A thorough analysis of the products offered by the major players in the V2X arena highlighted that virtualization and orchestration solutions are not mentioned explicitly when illustrating the commercially available OBU and RSU devices. These devices are equipped with an ITS communication stack software with related APIs. Third party software applications can be hosted on the OBUs or developed using a Software Development Kit (SDK).

In the research field, several proposals of virtualization and orchestration approaches for OBU/RSU devices have been introduced. In [3], different virtualization technologies have been evaluated for improving the safety and security of vehicular EBUs. A similar approach, which focused on OBU virtualization, was also introduced in [4]. In both studies, container virtualization was not considered in the analysis. A container-based approach was considered in [5]. The solution introduced was considered a self-standing environment where an internal orchestration is performed. No interaction with a central orchestrator was devised.

Other scientific works instead focused on a more centralized approach for the virtualization of OBU applications. One approach was to virtualize the entire OBU and have a virtual OBU at the edge server that is a replica of the real OBU [6], [7]. This approach can offload the OBU from the processing, but it would require a continuous exchange of information between the real OBU and the virtual OBU. This could prevent the real OBU to be self-sufficient when the mobile connection is absent. A similar approach has been proposed in [8]. In this case, it is not considered to have a complete virtualized copy of the real OBU, but only specific services in form of docker containers are migrated to the Edge server.

The most similar approach to the one proposed by 5G-IANA was introduced by the 5GinFIRE project [9]. In this project, several vertical industries were targeted, among them, the automotive one. Indeed, the 5GinFIRE project made available the IT-Av Automotive Environment where third party developers can test their application on real OBU and RSU hardware by deploying them as Virtual Network Functions using virtual machines and exploiting the orchestration features provided by a multi-site orchestrator based on OSM MANO. A docker-based deployment has been also introduced as new functionality to the 5GinFIRE platform.

B. Advances in 5G-IANA – strategies

In 5G-IANA, the designed solution for the orchestration of virtual resources on top of OBUs and RSUs foresee the integration of Kubernetes as lightweight orchestration framework. Each OBU/RSU hosts a Kubernetes cluster that may acts as an independent orchestration platform that executes the Life Cycle Management (LCM) of the deployed services or can be coordinated by the upper centralized orchestration layers for facilitating the deployment of more extensive service-chains across multiple edge and far-edge locations. As previously discussed, distributed service-chains, partially deployed on top of OBUs/RSUs pose an additional challenge related to the potential discontinuity of mobile connectivity. In 5G-IANA different mitigation solutions are currently under investigation: i) the execution of graceful roll-back procedures in case a LCM operation coordinated by the centralized layers fails due to a connectivity issue, ii) the possibility of notifying through the vehicle Human Machine Interface (HMI) the absence of mobile connectivity and the consequent need of taking over the control of the vehicle in case of vehicle movement related applications and iii) depending on the service nature, the possibility of designing the distributed service-chain in order to have potentially independent sub-chains running on top of OBUs/RSUs, for which is not mandatory to exchange data with service components running at the edge and/or central cloud. Furthermore, the AOEP performs a tailored selection of OBUs and RSUs. The platform will provide a subscription functionality to determine which OBUs and RSUs are registered and can be potentially used to execute services, while a continuous resource discovery will allow to understand which OBUs are currently available when a request for the provisioning of a specific Automotive service is issued. The selection of OBUs and RSUs can be performed taking into account various parameters, including the ones specified by the Service Provider, e.g., the coverage area, for which the expected vehicle trajectory should be considered, and/or host-specific capabilities like the availability of data-sets for training/profiling, availability of connected sensors/actuators etc.

IV. ML AS A SERVICE

A. State of the art on distributed ML in 5G

Traditional ML systems require centralized data collection and processing, consuming network resources and posing a potential risk regarding the privacy of the collected data. The wide range of connected nodes especially OBUs and RSUs in the context of 5G networks leads to new distributed ML services that leverage the cooperative and distributed data collection in addition to an unprecedented amount of data availability leading to improved statistical analysis. Note that Distributed ML (DML) dispenses with the high bandwidth and storage requirements of centralized ML architectures as well as the privacy concern. In combination with the growth of the processing capabilities of Far Edge nodes such as OBUs and RSUs this allows ML tasks to be carried out more efficiently at the Far Edge rather than on a central server [10].

In a nutshell, distributed or specifically federated ML frameworks [11] dispatch and ML model to the data source locations, in our case OBUs and RSUs, to be locally trained. This is achieved using a variation of the Stochastic Gradient Descent (SGD) algorithm. To obtain one final, coherent model the locally trained models are collected and averaged at an aggregation node. This procedure may be repeated until a certain stopping criterion is achieved. The orchestration and communication framework of the DML which is based on the selection of nodes to form a (typically) star topology with the model aggregation set in the center node (MEC) [12].

The framework above can on the one hand side be considered as a generic support technology for 5G Core's Network data analytic function [13]. However, in this paper we argue in the next section that this can also be offered as an OTT service by the 5G network.

The limitations of the state-of-the-art DML methods when blindly applied in 5G networks arise due to heterogeneity of the quality and availability of the data as well as the heterogeneity of the training devices themselves (OBU, RSU) as well as the network conditions [14]. This is clearly seen in a 5G mobile scenario where different OBUs observe qualitatively different data (e.g., network conditions or vehicle trajectories) and the fact that the OBUs themselves are composed of heterogeneous hardware that might not be permanently available for DML tasks. Finally, as parameter consistency can vary from synchronous configurations over stale-synchronous, asynchronous up to ensemble learning configurations we observe that in fact depending on the number of OBUs that contribute to a DML task the optimal choice of the configuration changes [15].

B. Advances in 5G-IANA

Based on the state-of-the-art above, the first direction of concern in 5G-IANA is the integration of the system design for DML in 5G and secondly, the optimization of DML training with federated concepts for 5G automotive services.

The Integration of DML orchestration in the 5G-IANA architecture is based on, first, DML-specific resource monitoring, and secondly, a 5G-specific distribution of ML models. The first aspect concerns the assessment and verification of the underlying ML assumptions with respect to the overall training resources and data availability in a distributed and inherently non-uniform environment. DML orchestration in the context of 5G requires key monitoring information regarding the overall availability of OBU and RSU resources and data, as well as the heterogeneous OBU devices. This is reflected in a continuous DML specific resource discovery and monitoring such as monitoring the OBU training resource availability or the OBU training data availability. The second aspect concerns the coupling of the DML decision-making mechanisms with the 5G-IANA service orchestration. Based on OBU and RSU monitoring data, the DML orchestration acts as a filter to decide on the selection of specific OBUs to run a DML service on. This filtering information is fed back to the slice management layer as a constraint for the deployment. Further, as mentioned in the previous subsection, the choice of an optimal DML configuration (synchronous, asynchronous, etc.) depends on the number of OBUs and RSUs selected for training. As the DML-specific filtering process runs on the DML orchestrator it also supplies this choice as a configuration parameter for the DML service.

An instantiation of the DML framework appears in the description of the following case study, where the focus is on monitoring the status of the application layer network.

C. Case study: Network Status Monitoring

In this case study, we monitor the application-layer network behavior and provide timely spatial QoS predictions for automotive verticals. As the quality of communication may vary spatially and temporally, and autonomous vehicles cannot afford to suddenly lose communication, having a foresight of future network behaviour will enhance the decision making capabilities of autonomous vehicles.

As depicted in Fig. 2, we deploy a network monitoring NetApp that uses passive probing to obtain network measurements. Here, passive probing means that the regular network traffic is monitored to study the network patterns for prediction. Results show that this method has less data overhead compared to active probing from a single host [16].

Here, we train a forecasting ML model with the the following goals: (1) to learn data traffic patterns for data traffic prediction, (2) to learn network condition models to provide QoS predictions, and (3) to learn to distinguish between normal and abnormal network behaviours to detect and predict faults. The forecasting model to be used here is a Long Short-Term Memory (LSTM) model which overcomes the vanishing gradient problem [17] by introducing a linear unit called Constant Error Carousel [18] and outperforms conventional benchmark models such as Support vector regression (SVR) and Autoregressive Distributed Lag Model (ADLM) in forecasting tasks [19], [20] and [21].

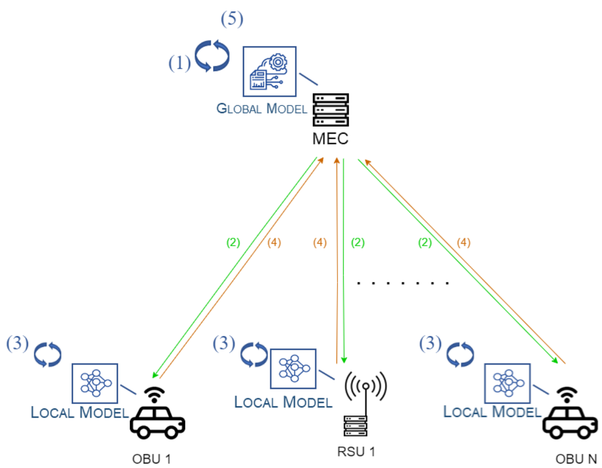

Fig. 2. Operational Flow of a DML NetApp including (1) client (OBU/RSU) selection; (2) model dispatch; (3) local training; (4) model delivery; (5) model aggregation

As the OBU local datasets are non-independent and identically distributed and unbalanced, since some OBUs might use certain services (based on their location, requirements and tasks) more than other OBUs, we explore DML algorithms beyond distributed averaging such as Federated Averaging (FedAvg) [11], FedProx [22] or FedOpt [23]. In a nutshell, to balance heterogeneous data and heterogeneous devices, random subsets of suitable training nodes (OBUs) are selected by the DML orchestrator for the individual DML tasks in combination with asynchronous SGD optimization at the aggregation on the edge.

The specific operational flow is as follow: The initial DML client (OBU/ RSU) selection by the DML orchestrator is performed based on node availability, node connectivity, node battery resources, ML resource availability, node location, data availability, data volume, and data features. These selection criteria reflect the required knowledge to perform DML tasks for automotive applications over a 5G mobile network as to control data quality and device heterogeneity.

V. 5G-IANA ARCHITECTURE OVERVIEW

The centralized architecture, depicted in Fig. 3, is comprised by two main orchestration layers that reside on top of multiple Management and Orchestration (MANO) platforms in charge of orchestrating compute resources placed in different locations. In particular, in 5G-IANA, we distinguish among three possible locations, namely the central cloud, the edge and the far-edge, where the latest includes orchestration-enabled OBUs and RSUs. By design, the 5G-IANA platform is suitable for operating on top of distributed 5G infrastructures, enabling the deployment of Vertical services across non-federated Network Function Virtualisation Infrastructure (NFVI) Point of Presences (PoPs) while supporting the expected QoS depending on the specific 5G service profile. In addition, with the aim of facilitating Verticals in the provisioning of the desired services, the layered approach of the architecture follows the separation of concerns among the orchestration of the application components and the required network services that support them, as well as the modelling and selection of application-aware 5G Network Slices. The goal is to build an abstraction layer exposed to Verticals that hides the infrastructure and technology complexity and can lead to a uniform and optimal application design and execution, while the procedures for the selection of the most appropriate 5G Network Slice, along with the allocation of the compute resources across the distributed infrastructure are executed in a transparent manner from a platform end-user point of view.

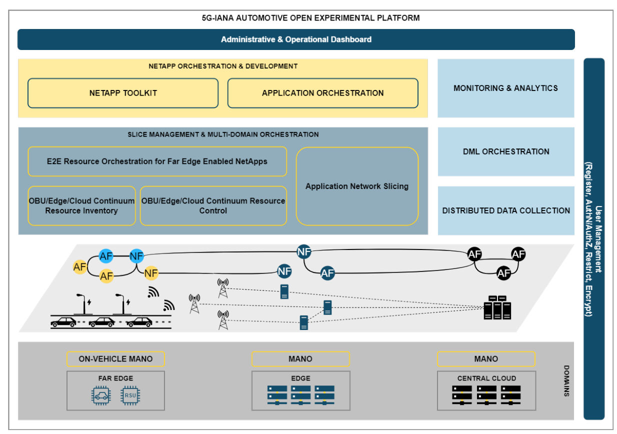

Fig. 3. 5G-IANA Automotive Open Experimental Platform System Design

The NetApp Orchestration and Development layer represents the entry point to the overall system for a subset of the identified stakeholders. This layer provides tailored functionalities to fulfil the requirements of Service Providers and NetApp Developers. The two main building blocks, namely the NetApp Toolkit and the Application Orchestrator, provide respectively the functionalities related to modelling/design and provisioning/orchestration phases of the Vertical service/NetApp lifecycle. On one hand, the NetApp Toolkit exposes functionalities for: i) the management of NetApp packages along with the included AFs/NFs packages, ii) the composition of NetApps and AFs in service chains and iii) the specification of high-level QoS parameters (e.g., number of expected vehicles, coverage areas, etc.) along with application-specific optimization policies. On the other hand, the Application Orchestrator handles: i) the Vertical Service deployment and LCM, ii) the mapping of the Vertical service provisioning request into a slice intent, and iii) the Vertical service run-time monitoring in conjunction with the its profiling and related policies executions.

The Slice Management and Multi-domain Orchestration layer is responsible for handling the intent-based request issued by the Application Orchestrator and translating it into the most appropriate 5G Network Slice profile to be selected to fulfill the specified high-level QoS requirements. This layer hosts also functionalities for coordinating a coherent provisioning of needed compute resources across the target locations (i.e., far-edge, edge and central cloud) where NetApps and related AFs/NFs are expected to be executed. Furthermore, to support the orchestration of service-chains on top of OBUs and RSUs, this layer provides also functionalities that realize the far-edge resource subscription as well as their continuous discovery, collecting periodically and/or on-demand the resources status along with information related to specific capabilities (e.g., availability of data-sets for training, power consumption values, etc.). To do so, this block leverages on an Information Localization service executed at the OBU/RSU to collect the needed information for properly handling the resources selection and allocation.

DML Orchestration, Monitoring and Analytics and distributed data collections are services that run on top of the previous two.

The DML block hosts the functional components in charge of handling the selection of the most appropriate resources (including the OBU and RSU) for supporting the operation of the desired DML Vertical Service. The telemetry and analytics functionalities, are coupled with the applications orchestration, as a software management layer for controlling and monitoring internal, service-to-service traffic in microservices-based applications. Both monitoring and analytics are achieved for the data and control plane by a set of intelligent proxies deployed alongside the application software components supporting the provision of support/backing services (e.g., service discovery, load balancing, health checking). This enables traceability and logging and reporting for vertical and network slice and infrastructural metrics.

VI. Conclusions and Future Work

This paper has described the 5G-IANA project platform at a relatively high-level, with enough detail though to showcase the potential of its openness, accessibility and adaptability to enable third party experimenters in the Automotive vertical to develop, deploy and test their services. This open platform offers two main innovations to third parties that would be interested in experimenting on top of it, namely an available option for orchestration of far-edge resources that lie at the vehicle- or road-side, and an option for exploiting ML-awareness for provisioning automotive-related services in a more efficient and effective way, by utilising network-side predictions. Future work includes the software deployment of all architectural components described in this paper, their integration and validation, as well as the design of a proper third party engagement planning, service/algorithm/application integration and experimentation execution that will demonstrate and highlight the re-usability and potential impact of the 5G-IANA platform.

Acknowledgement

This research was performed in the context of the 5G-IANA project, co-funded by the European Commission under the Horizon 2020 Research and Innovation Programme (grant agreement No 101016427).

References

[1] 5GAA, “Mec for automotive in multi-operator scenarios.” 5GAA, 2021.

[2] E. G. MEC, “Etsi gs mec 002 v2.1.1 multi-access edge computing (mec); phase 2: Use cases and requirements.” ETSI, 2018.

[3] Pelzl and et al., “Virtualization technologies for cars: Solutions to increase safety and security of vehicular ecus,” *Automot. Saf. Secur*, vol. 2008, pp. 164–173, 2008.

[4] Zhu and et al., “Providing flexible services for heterogeneous vehicles: An nfv-based approach,” *IEEE network*, vol. 30, no. 3, pp. 64–71, 2016.

[5] Morabito and et al., “Lightweight virtualization as enabling technology for future smart cars,” in *2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*. IEEE, 2017, pp. 1238–1245.

[6] Santa and et al., “Surrogates: Virtual obus to foster 5g vehicular services,” *Electronics*, vol. 8, no. 2, p. 117, 2019.

[7] Afaq and et al., “Sensor virtualization and data orchestration in internet of vehicles (iov),” in *2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*. IEEE, 2021, pp. 998–1003.

[8] Campolo and et al., “Mec support for 5g-v2x use cases through docker containers,” in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–6.

[9] Aloizio and et al., “5ginfire: An end-to-end open5g vertical network function ecosystem,” *Ad Hoc Networks*, vol. 93, p. 101895, 2019.

[10] Nassef and et al., “A survey: Distributed machine learning for 5g and beyond,” *Computer Networks*, vol. 207, p. 108820, 2022.

[11] McMahan and et al., “Communication-efficient learning of deep networks from decentralized data,” in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.

[12] Nishio and et al., “Client selection for federated learning with heterogeneous resources in mobile edge,” in *IEEE international conference on communications (ICC)*. IEEE, 2019, pp. 1–7.

[13] Niknamand and et al., “Federated learning for wireless communications: Motivation, opportunities, and challenges,” *IEEE Communications Magazine*, vol. 58, no. 6, pp. 46–51, 2020.

[14] Dongdong and et al., “Federated learning in vehicular edge computing: A selective model aggregation approach,” *IEEE Access*, vol. 8, pp. 23 920–23 935, 2020.

[15] Hang and et al., “A free stale synchronous parallel strategy for distributed machine learning,” in *Proceedings of the 2019 International Conference on Big Data Engineering*, 2019, pp. 23–29.

[16] Haxhibeqiri and et al., “In-band network monitoring technique to support sdn-based wireless networks,” *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 627–641, 2021.

[17] Hochreiter, “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, no. 02, pp. 107–116, 1998.

[18] Han and et al., “A review of deep learning models for time series prediction,” *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7833–7848, 2021.

[19] Fischer and et al., “Deep learning with long short-term memory networks for financial market predictions,” *European Journal of Operational Research*, vol. 270, no. 2, pp. 654–669, 2018.

[20] Abbasimehr and et al., “An optimized model using lstm network for demand forecasting,” *Computers & industrial engineering*, vol. 143, p. 106435, 2020.

[21] Kulshrestha and et al., “Bayesian bilstm approach for tourism demand forecasting,” *Annals of tourism research*, vol. 83, p. 102925, 2020.

[22] Tian and et al., “Federated optimization in heterogeneous networks,” *Proceedings of Machine Learning and Systems*, vol. 2, pp. 429–450, 2020.

[23] Sashank and et al., “Adaptive federated optimization,” *arXiv preprint arXiv:2003.00295*, 2020.