Building Realistic Experimentation Environments for AI-enhanced Management and Orchestration (MANO) of 5G and beyond V2X systems

Nina Slamnik-Kriještorac*, Miguel Camelo Botero*, Luca Cominardi[†], Steven Latré*, Johann M. Marquez-Barja* * University of Antwerp - imec, IDLab - Faculty of Applied Engineering, Belgium [†] ADLINK Technology, France

Abstract—The plethora of heterogeneous and diversified services in 5G and beyond requires from networks to be flexible, adaptable, and programmable, i.e., to be able to correspondingly adapt to changes. As human intervention might significantly increase delays in MANagement and Orchestration (MANO) operations, automation and intelligence become imperative for orchestrating services and resources, especially the ones with stringent requirements for latency and capacity, such as Vehicle-to-Everything (V2X) services. As virtualization and Artificial Intelligence (AI) promise to mitigate those challenges towards enabling true automation in MANO operations, in this paper we present our effort towards building and fully utilizing the real-life testbeds, such as Smart Highway and Virtual Wall, located in Belgium, to conduct realistic experimentation and validation of distributed orchestration intelligence in a dynamic network such as V2X system.

Index Terms-management and orchestration, NFV, AI, ML, zenoh, testbeds, experimentation

I. INTRODUCTION AND MOTIVATION

With the arrival of contemporary technologies such as Software Defined Networking (SDN), Network Function Virtualization (NFV), and Multi-Access Edge Computing (MEC), 5G and beyond mobile communication systems are becoming able to enhance existing use cases and business models, and to foster new ones [1]. As networks are usually consisted of a complex set of broad and heterogeneous devices and resources that must be integrated to provide seamless service, the traditional, inherently manual, network MANagement and Orchestration (MANO) becomes impossible to maintain [2]. Given the heterogeneous and diversified nature of services in 5G and beyond, with complex and potentially conflicting demands, networks need to be flexible, adaptable, and programmable, to be able to swiftly adapt to changes [3]. These requirements need to be supported by technologies, such as virtualization [4] and Artificial Intelligence (AI) [2], in particular Machine Learning (ML), to enable automation and intelligence in the MANO of services, thereby coping with the network complexity.

The potential of integrating AI/ML techniques to NFV MANO systems, with the goal to enforce operations and automate them, is well recognized and there are some important research efforts invested to study this topic [3, 5, 6, 7]. However, there is still a gap in research when it comes to experimentation and testing the true impact of AI/ML on the optimization of NFV MANO operations. To this end, in this paper we present our ongoing work towards building and fully utilizing the potential of high-performance real-life testbeds, such as Smart Highway¹ [8] and Virtual Wall², to pursue testing and validation of distributed intelligence in a dynamic

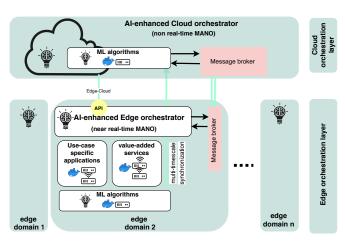


Fig. 1: The architecture of multi-domain AI-enhanced management and orchestration system for V2X use cases.

network such as Vehicle-to-Everything (V2X) system. We present the AI-enhanced MANO system for V2X services in Fig. 1, with cloud and edge orchestration layers, which are enabled to autonomously operate, but also to collaborate and balance their operations towards achieving desired Key Performance Indicators (KPIs).

Some of the reasons why an automated MANO in 5G and beyond is necessary are:

- a demand for new use cases for industry 5.0, Industrial Internet of Things (IIoT), and self-driving vehicles, which require significantly extensive broadband, efficient, resilient, and reliable connectivity, and network availability of five-nines [5];
- improving operational efficiency in NFV systems, as network complexity significantly increases with heterogeneous and distributed resources and services [7];
- dynamic changes in KPIs happen due to fluctuations in demands from verticals (e.g., IIoT, vehicular systems), user mobility patterns, etc., thus, network needs to improve its operation by learning from the environment and optimizing itself towards the desired and promised KPIs [5];
- complexity and heterogeneity, brought by combining different technologies and verticals in 5G and beyond, require NFV MANO systems that enable an intelligent interplay between edge and cloud [5];
- NFV MANO operations (e.g., service placement, migration, fault recovery, scaling) need to be further optimized and automated in 5G ecosystems with the help from AI/ML, as traditional optimization techniques are com-

¹Smart Highway: https://www.fed4fire.eu/testbeds/smart-highway/

²Virtual Wall: https://www.fed4fire.eu/testbeds/virtual-wall/

plex and lengthy, and heuristics are near-optimal, which might make them both ineffective in swift response to dynamic changes in the network [7]; and

• management and orchestration of computational resources in 5G and beyond becomes a major challenge, due to the practices of cloudification and virtualization of core network functions, and partially radio access network functions [3].

The integration of AI/ML into NFV MANO systems for 5G and beyond is expected to mitigate most of the challenges listed above, as it is now mature enough to provide efficient solutions for complex optimization and prediction problems [3]. If used in a suitable manner (e.g., very fine-grained AI algorithms carefully chosen for operation in a specific network domain), AI/ML can enable i) optimized service instantiation, ii) learning utilization patterns for computational resources of virtualized network services, iii) prediction models for proactive resource allocation/relocation, and iv) optimized service migration. Besides the aforementioned benefits, AI/ML techniques impose some additional challenges that need to be taken into account. Some of them are vulnerabilities in terms of i) security, scalability, and transferability [7], which limit the full potential of applying AI/ML to NFV MANO in 5G, ii) high computation power that might not be available in resource constrained edge nodes, and iii) need for quality data to train the ML algorithms, as their performance on making decisions (e.g., predicting, classification, taking actions) will depend on how close was the training data to the actual data used in production environments.

The AI-enhanced MANO system for V2X services that we present in this paper, and illustrate in Fig. 1, consists of two layers, i.e., cloud and edge. The system enables autonomous MANO operations in each of the domains, but enforces an interplay between them for offloading orchestration decisions, or for retrieving data from distributed data engineering pipelines available in all edge domains. Such a system orchestrates both services and applications developed for various use cases, but also Network Intelligent Functions (NIFs) that are represented by adopted and integrated AI/ML models. Despite the emerging popularity of bringing intelligence to network management and orchestration functions in 5G and beyond, most of the works on validating the impact of AI/ML on MANO are based on simulations. There is a gap between using synthetic data and real data when it comes to training and validating/testing AI/ML models, as real setups can create more realistic traces for training, with higher probability of good performance when deployed in production environments. However, building realistic Proof-of-Concepts (PoCs) is usually time-consuming and expensive, while the number of scenarios that can be covered is limited. On the other hand, simulators bring that flexibility but mostly at the cost of not capturing all dynamics of real environments. Thus, the real setups are fundamental to create hybrid approaches that ensure that the performance of AI/ML algorithms is not negatively impacted once they are dealing with real data. One of the attempts to pursue testing of AI/ML on the lifecycle management operation of scaling service functions is presented by Baranda et al. [6], where a scaling operation of virtual Content Delivery Network (vCDN) service is triggered by AI/ML algorithms, thereby integrating AI/ML into management platform of 5G-Transformer³. Thus, in this paper we present and illustrate a realistic experimentation environment that extends the scope of aforementioned PoC,

and enables studying and experimenting with AI-enhanced operations of proactive placement, scaling, migration, and termination, of challenging V2X services, towards understanding and resolving challenges imposed by AI/ML to overcome them and improve those MANO operations.

II. ARCHITECTURE OF AI-ENHANCED MANAGEMENT AND ORCHESTRATION SYSTEM

The architecture of multi-domain MANO system presented in Fig. 1 is applicable to all distributed and heterogeneous softwarized networks whose operation stretches from edge to the cloud, where services and applications are usually deployed with microservice-based approach, and connectivity ensured via different wireless technologies including 5G and beyond. As such networks are usually characterized by distributed resources belonging to different edge domains, which might belong to different Mobile Network Operators (MNOs), we follow the split between cloud (i.e., centralized) and edge orchestrators, which are deployed in a relationship m : n, $m < n, m, n \in \mathbf{N}$.

Thus, each edge domain that consists of one or multiple edge nodes (i.e., MEC hosts) is governed by one edge orchestrator, which is, following ETSI NFV MANO framework, in charge of lifecycle management (e.g., instantiation, scaling, termination) of all underlying services, i.e., i) use case-related services, ii) value-added services, and iii) NIFs that embody AI/ML models. On the other hand, cloud orchestrator is rather in charge of global optimization in the system, thereby making less-granular decisions depending on the e.g., locations and density of vehicles on the roads for our particular real-life use case. One particular example of these decisions is service migration from one edge domain to another, triggered by higher density of vehicles (i.e., edge service consumers) in one edge domain, or by need for optimization of energy consumption in MEC hosts across edge domains.

Two MANO layers communicate with each other in the two following ways: i) via Edge-Cloud reference point, which is used to either offload decision-making tasks between two orchestrators or to pass the already taken decision, and ii) via message brokers, which exchange data in a controlled way depending on the type of AI/ML technique that has been applied in the system, thereby using that data to either perform training or model adjustments and online learning. Thus, depending on the time-scales of optimization (global or local, i.e., edge-specific), it is required that MEC hosts can connect data to AI/ML models in a transparent and efficient way (e.g., using Zenoh framework introduced in Section III). In case of federated learning, which is suitable for distributing intelligence across edge nodes, thus deploying AI/ML agents in edge nodes, we consider that each edge orchestrator trains the local model based on the data collected from its own domain. On the other hand, if security in data sharing between two message brokers laying in two orchestration layers can be preserved, multi-agent reinforcement learning may use data collected from other edge domains to optimize policies.

Some examples of V2X services that might benefit from such AI-enhanced NFV MANO system are:

- *infotainment services*, such as vCDN, where cloud orchestrator optimizes distributed vCDN deployment across edge domains, based on the locations of vehicle (e.g., retrieved from the location service) and computational resource utilization in each of the edges (e.g., collected from message brokers in edges) [8, 9];
- emergency services, such as lane clearance for emergency vehicles, where edge orchestrators optimize service oper-

³5G-Transformer - the project on 5G Mobile Transport Platform for Verticals: http://5g-transformer.eu/

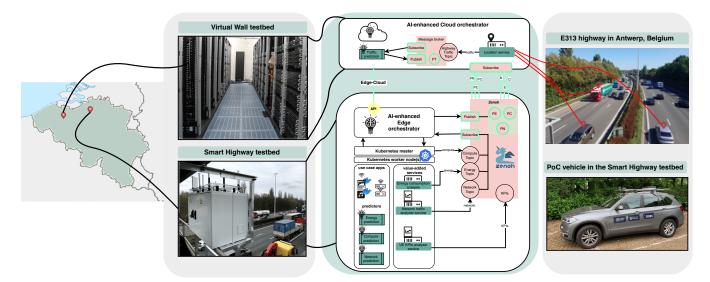


Fig. 2: The AI-enhanced management and orchestration system mapped to the real-life testbed environment (PE - pre-processed/predicted energy, etc.).

ation on the edge level, while cloud orchestrator makes sure that service is deployed in all edge domains that are on the route affected by emergency situation [8, 10];

• *maneuver recommendation services*, which are services that provide vehicles with recommendations about in which lane to drive, when to merge or exit the lane, etc., taking into account location/speed/destination of vehicles (e.g., retrieved from cloud orchestrator) [8, 11].

III. PROOF-OF-CONCEPT

In Fig. 2, we map the testbed components to the elements of AI-enhanced MANO framework presented in Section II. Starting from the edge, we provide the NFV infrastructure in MEC hosts by virtualizing computational resources in Roadside Units (RSUs), which are deployed along the E313 highway in Antwerp, Belgium, as a part of the Smart Highway testbed [8]. We presented the collocation of MEC platforms with RSU units in [9], and used it in the demo setup for emergency V2X services in [10]. To make use of the computational resources for performing lifecycle management of edge V2X services, we deploy Kubernetes (K8s), where edge orchestrator embodies the role of K8s master and extends it to i) support cross domain operations, i.e., edge-cloud interaction, and ii) receive dynamic triggers from AI/ML models deployed in NIFs for optimizing MANO operations. Such K8s master with extended and enhanced operation deploys services and applications on designated worker nodes. In the PoC, both master and worker nodes can be deployed on the bare metal, as well as in Linux containers (LXC), which is a more suitable practice for shared experimentation environments as testbeds.

For each type of data that is collected, i.e., computational and network resource utilization, energy consumption, KPIs measured at users' side, and users' locations, we also deploy MEC value-added services, as per definition in ETSI MEC [12], which perform data retrieval and pre-processing before publishing them on Zenoh [13]. Given its minimal network overhead (as little as 5B), and its small footprint (around 60kB on Arduino board), Zenoh is adopted in our PoC as a framework for data engineering pipeline. In particular, Zenoh provides a minimal set of primitives to deal with data in motion (e.g., real-time stream of vehicles' location/speed/destination), data at rest (e.g., historic data for vehicles' and edge nodes' computational resource utilization and energy consumption) and remote computations (e.g., on-demand calculation of the best route and speed limit). Each edge and cloud orchestrator acts as a subscriber for various types of data that can be stored on edges, and used for training or online learning/optimization.

Furthermore, concerning the vehicle as a client, our current PoC includes one vehicle that is capable to communicate with the edge services via long range 4G (to be extended to 5G in future). Thus, the client application is installed in the Onboard Unit (OBU) of the vehicle, and it utilizes Uu link to exchange Cooperative Intelligent Transport System (C-ITS) messages with services, and inform them about its location, speed, heading, and destination.

Cloud orchestrator is running on the bare metal on top of the Virtual Wall testbed, located in Ghent, Belgium (Fig. 2). It is deployed as a web server (using Flask framework in python), which is capable of i) processing decision-offloading requests coming from the edge orchestrators, ii) location data processing and publishing on Zenoh, iii) injecting decisions on the north-bound interface of edge orchestrators to instruct them to proactively migrate/relocate services from one edge to another, and iv) receiving notifications from NIFs deployed on the cloud, which enhance their operations and help them make efficient decisions on managing underlying resources and edge orchestrators.

In Fig. 3 we show the result of average response time, and CPU utilization, of the vCDN server deployed on the MEC host in our PoC. To stress the load and increase the number of vehicles, we run Locust⁴ stress test inside the vehicle. We can see that the number of vehicles that are simultaneously requesting content from the same server affects the response time, and CPU utilization as well. In case NIF predicts the traffic demand, and the number of vehicles in this specific geographic region, they are expected to optimize the operation of an edge orchestrator, as it will perform horizontal scaling and additional deployments of vCDN server on other MEC hosts, so that users (vehicles) can still experience low response time. As the response time consists of communication latency

⁴Locust: https://docs.locust.io/

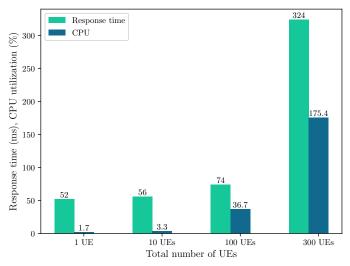


Fig. 3: Average response time and CPU utilization of vCDN server deployed in our PoC.

(uplink and downlink, impacted by network load), and computational latency (affected by CPU load), its increase is mainly affected by an increase in CPU utilization on edge nodes, which needs to be carefully monitored and optimized e.g., by corresponding NIFs. As a part of our future work, we are going to utilize this PoC to present the impact of various ML models on the operation of orchestration operations, measured at the client side in terms of response time, throughput, and other relevant KPIs.

IV. CONCLUSION

To mitigate the challenges in MANO operations imposed by human interventions (i.e., delayed operations, reactive approach) the automation and intelligence become an imperative for orchestrating services and resources, especially the ones with stringent requirements for latency and capacity, such as those in V2X systems. In this paper, we presented our ongoing work in building and utilizing the PoC on the real-life testbeds, to pursue realistic experimentation and validation of the impact that AI/ML have on the management and orchestration in distributed and heterogeneous networks such as V2X.

V. ACKNOWLEDGEMENT

This work has been performed in the framework of the European Union's Horizon 2020 project DAEMON co-funded by the EU under grant agreement No. 101017109, the Horizon 2020 Fed4FIRE+ project, Grant Agreement No. 723638, and the Horizon 2020 5G-Blueprint project under Agreement No. 952189.

REFERENCES

- D. Soldani and A. Manzalini, "Horizon 2020 and Beyond: On the 5G Operating System for a True Digital Society," *IEEE Vehicular Technology Magazine*, vol. 10, no. 1, pp. 32–42, 2015. doi: https://dx.doi.org/10.1109/ MVT.2014.2380581.
- [2] ETSI, "Autonomous Networks, supporting tomorrow's ICT business," ETSI White Paper No. 40, 2020. Online [Available]: https://www.etsi.org/images/files/ETSIWhitePapers/etsiwp-40-Autonomous-networks.pdf.

- [3] D. M. Gutierrez-Estevez, M. Gramaglia, A. D. Domenico, G. Dandachi, S. Khatibi, D. Tsolkas, I. Balan, A. Garcia-Saavedra, U. Elzur, and Y. Wang, "Artificial Intelligence for Elastic Management and Orchestration of 5G Networks," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 134–141, 2019. doi: https://dx.doi.org/10.1109/MWC.2019.1800498.
- [4] A. Bujari, C. E. Palazzi, D. Polonio, and M. Zanella, "Service Function Chaining: a lightweight containerbased management and orchestration plane," in 2019 16th IEEE Annual Consumer Communications Networking Conference (CCNC), pp. 1–4, 2019. doi: https://dx.doi. org/10.1109/CCNC.2019.8651862.
- [5] M. Bagaa, T. Taleb, J. Riekki, and J. Song, "Collaborative Cross System AI: Toward 5G System and Beyond," *IEEE Network*, vol. 35, no. 4, pp. 286–294, 2021. doi: https: //dx.doi.org/10.1109/MNET.011.2000607.
- [6] J. Baranda and et al., "On the Integration of AI/MLbased scaling operations in the 5Growth platform," in 2020 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), pp. 105–109, 2020. doi: https://dx.doi.org/10.1109/NFV-SDN50289.2020.9289863.
- [7] D. M. Manias and A. Shami, "The Need for Advanced Intelligence in NFV Management and Orchestration," *Netwrk. Mag. of Global Internetwkg.*, vol. 35, p. 365–371, Mar. 2021. doi: https://dx.doi.org/10.1109/ MNET.011.2000373.
- [8] J. Marquez-Barja, B. Lannoo, D. Naudts, B. Braem, V. Maglogiannis, C. Donato, S. Mercelis, R. Berkvens, P. Hellinckx, M. Weyn, *et al.*, "Smart Highway: ITS-G5 and C2VX based testbed for vehicular communications in real environments enhanced by edge/cloud technologies," in *EuCNC2019, the European Conference on Networks and Communications*, pp. 1–2, 2019. Available [Online]:https://biblio.ugent.be/publication/8642435.
- [9] N. Slamnik-Kriještorac and J. M. Marquez-Barja, "Unraveling Edge-based in-vehicle infotainment using the Smart Highway testbed," in 2021 IEEE 18th Annual Consumer Communications Networking Conference (CCNC), pp. 1–4, 2021. doi: https://dx.doi.org/10.1109/ CCNC49032.2021.9369622.
- [10] N. Slamnik-Kriještorac, G. M. Yilma, F. Zarrar Yousaf, M. Liebsch, and J. M. Marquez-Barja, "Multi-domain MEC orchestration platform for enhanced Back Situation Awareness," in *IEEE INFOCOM 2021 - IEEE Conference on Computer Communications Workshops (INFO-COM WKSHPS)*, pp. 1–2, 2021. doi: https://dx.doi.org/ 10.1109/INFOCOMWKSHPS51825.2021.9484632.
- [11] S. A. Ashraf, R. Blasco, H. Do, G. Fodor, C. Zhang, and W. Sun, "Supporting Vehicle-to-Everything Services by 5G New Radio Release-16 Systems," *IEEE Communications Standards Magazine*, vol. 4, no. 1, pp. 26–32, 2020. doi: https://dx.doi.org/10.1109/MCOMSTD.001. 1900047.
- [12] ETSI, "Multi-Access Edge Computing (MEC); Framework and Reference Architecture," ETSI ISG MEC, ETSI GS MEC 003 V2.1.1, 2019. Online [Available]: https://www.etsi.org/deliver/etsi_gs/MEC/ 001_099/003/02.01.01_60/gs_MEC003v020101p.pdf.
- [13] Eclipse zenoh, "Project website: Eclipse zenoh," 2020. Online [Available]: https://zenoh.io, Last accessed on 2021-9-20.