POLITECNICO DI TORINO
Repository ISTITUZIONALE

Mobile RF Scenario Design for Massive-Scale Wireless Channel Emulators

Original
Mobile RF Scenario Design for Massive-Scale Wireless Channel Emulators / Rusca, R., Raviglione, F., Casetti, C.E., Giaccone, P., Restuccia, F.. - ELETTRONICO. - (2023), pp. 675-680. (European Conference on Networks and Communications (2023 EuCNC) Gothenburg (Sweden) 6-9 June 2023) [10.1109/EuCNC/6GSummit58263.2023.10188319].

Availability:
This version is available at: 11583/2977750 since: 2023-08-01T18:27:03Z

Publisher:
IEEE

Published
DOI:10.1109/EuCNC/6GSummit58263.2023.10188319

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

16 November 2026

Mobile RF Scenario Design for Massive-Scale Wireless Channel Emulators

Riccardo Rusca[†], Francesco Raviglione[*], Claudio Casetti[†], Paolo Giaccone[*], Francesco Restuccia[‡]

[†] Department of Control and Computer Engineering, Politecnico di Torino, Italy
[*] Department of Electronics and Telecommunications, Politecnico di Torino, Italy
[‡] Institute for the Wireless Internet of Things, Northeastern University, United States

***Abstract*—Large-scale wireless emulation is gaining momentum nowadays, thanks to its potential in the development and deployment of advanced use cases for next-generation wireless networks. Several novel use cases are indeed emerging, including massive MIMO, millimeter wave beamforming and AI-based Vehicle-to-Everything (V2X) optimized communication. The development and testing of a wireless application, especially at a large scale and when dealing with mobile nodes, faces several challenges that cannot be solved by simulation frameworks alone. Thus, massive-scale channel emulators are emerging, enabling the emulation of realistic scenarios which leverage real hardware and radio signals. However, this is a complex task due to the lack of realistic scenarios based on real datasets. We thus propose a novel framework for the design and generation of channel emulation scenarios starting from real mobility traces, either generated by means of dedicated tools, or collected on the field. Our framework provides a practical way of generating mobility scenarios with vehicles, pedestrians, drones and other mobile entities. We detail all the steps foreseen by our framework, from the provision of the traces and radio parameters, to the generation of a matrix describing the delay and IQ samples for each time instant and node in the scenario. We also showcase the potentiality of our proposal by designing and creating a vehicular 5G scenario with 13 vehicles, starting from a recently-disclosed open dataset. This scenario is then validated on the Colosseum channel emulator, proving how our framework can provide an effective tool for large-scale wireless networking evaluation.**

I. Introduction

Wireless channel emulators are becoming a fundamental enabler for the next generation of mobile networks. With the commercial global release of 5G and the active research on the new 6G network generation, the focus of researchers is increasingly shifting to the use of Artificial Intelligence (AI) and new Deep Machine Learning techniques to support wireless solutions.

Massive MIMO, wireless signal beamforming, Vehicle-to-Everything (V2X) communications are only some of the new applications that are arising. All these AI-driven applications require very complex testbeds to investigate and evaluate the performance under different conditions, and with different input parameters. Given the strict requirements imposed by the flagship communities for B5G / 6G development, like Hexa-X [1], the creation of a synthetic testbed with simulated datasets is no longer a possibility.

However, the development, deployment and on-field testing of applications is a complex process, which requires a non-negligible amount of time and effort. Being able to reliably simulate and emulate networks with real data has thus become crucial, especially during the pre-deployment phase and when dealing with large-scale scenarios.

While simulation frameworks can provide a viable mean of evaluating wireless networks, especially concerning mobile nodes, they are based on mathematical models for the underlying access technologies, which focus only on the most important aspects of a real environment. Indeed, they often compute the radio propagation and packet loss based on probabilistic considerations, and thus cannot model all the variables involved in complex environments.

Therefore, a further step towards emulation of large scale Radio Frequency (RF) scenarios is required. This kind of emulation involves real hardware and physical radio signals, which are exchanged in a dedicated testbed environment thanks to Software Defined Radios (SDRs). These signals are processed by channel emulators to generate a realistic path loss and phase shift, as if the SDRs were placed inside the actual physical environment. The setup for the evaluation of scenarios through RF channel emulation can be, however, fairly complex, due to the dynamics of the nodes and the lack of open datasets.

We thus propose for the first time a framework for the practical creation of RF scenarios, *starting from real collected data*. Furthermore, we prove the effectiveness of our framework describing the creation a full-fledged mobility scenario for the emulation of vehicular networks, with the aim of implementing it in Colosseum [2]. We also outline all the steps needed to design V2X RF scenarios in Colosseum. The proposed scenario is based on a recently-disclosed open dataset [3], containing the traces of 19 vehicles travelling in an urban

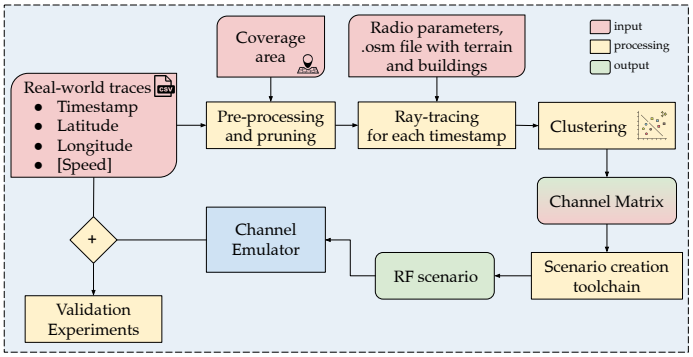

Fig. 1: High-level scheme of our data-driven framework for the creation and validation of an RF scenario.

environment. The framework aims at providing researchers working with massive-scale wireless channel emulators a fundamental tool for the development and testing of innovative applications using realistic wireless environments.

The rest of the paper is organized as follows. Section II presents the advantages brought by a massive-scale wireless channel emulator, while Section III describes our framework for the generation of scenarios starting from real traces. Then, Section IV presents a centralized 5G V2X scenario generated through our framework, whose details are reported in Section V, following our framework's processing pipeline. Our cellular-based scenario is then validated in Section VI. Finally, Section VII draws the conclusions.

II. Advantages of massive-scale channel emulators and Colosseum

The need for realistic emulation of radio frequency signals has increasingly grown over the years, motivated by the emergence of novel technologies in different fields, which include vehicular networks. The authors of [4] developed a first FPGA-based network channel emulator, allowing researchers to emulate the interaction of different RF nodes in real-time. Up to 12 nodes are supported, experiencing a realistic channel impulse response and proper interference model.

When the need for large-scale experiments emerged, several challenges started to arise, including how to properly emulate a large number of signals without easily exhausting the available resources in modern computers. Furthermore, the cost of simulating and emulating large-scale scenarios may be too high for most research groups, due to the necessity of buying and configuring dedicated hardware.

These challenges, together with the demand for accessible, effective large-scale radio emulation frameworks, led DARPA to develop the world's most powerful, one-of-kind wireless network emulator, called **Colosseum** [2]. This framework was then made available to industry, universities and research groups by Northeastern University. Colosseum combines 128 Standard Radio Nodes (SRNs) with a Massive digital Channel Emulator (MCHEM) backed by an extensive FPGA routing fabric, and it is considered the world's largest RF emulator. It enables the emulation of up to 256 independent radio nodes, 65,535 wireless channels in multiple operational environments with an emulation area up to $1\,\mathrm{km}^2$. RF and traffic scenarios are at the core of Colosseum, emulating not only the channel between transmitters and receivers but also the typical effects of the wireless propagation environment.

In the last few years, several researches have leveraged Colosseum to test their applications, thanks to a number of synthetic RF scenarios which were made publicly available. Recently, Villa *et al.* in [5], presented a synthetic V2X scenario for the Colosseum emulator, referred to an area around Tampa, FL. They leverage the tap and channel approximation framework previously presented in [6], which is also the basis for our work. The scenario proposed by Villa *et al.* consists of one Road Side Unit (RSU) and three On Board Units (OBUs, i.e., vehicles) moving on a straight path at a constant speed. The routes and trajectories of vehicles have been simulated by leveraging a commercial ray-tracing software namely, Wireless InSite® [7].

Instead, our work, for the first time, proposes a new framework that can be used for the creation of scenarios with real mobility data and, in general, real datasets. To the best of our knowledge, this work proposes for the first time a full-fledged framework for the creation RF scenarios *starting from real mobility traces*, thus adopting a data-driven approach. We also develop, leveraging our framework, the first V2X scenario developed for Colosseum with real mobility traces, using state-of-the-art ray-tracing algorithm to compute the channel path loss.

III. Framework

Our framework proposes for the first time a pipeline for the practical generation of RF scenarios for massive-scale emulators, starting from real mobility data, either generated by means of dedicated tools, or collected on the field.

All the input, processing and output steps foreseen by our solution are depicted in Figure 1. Furthermore, an open source MATLAB implementation of the whole framework has been made available to the research community[1], to enable easy, practical generation of new RF scenarios starting from data collected in the real-world.

As can be seen, our framework requires three main inputs: (i) first, a set of real-world traces in CSV format, containing at least a millisecond-accurate timestamp, and the position of each node over time, as latitude and longitude; (ii) second, a coverage area of a given size; (iii) finally, the set of radio parameters for the generation of the scenario and an OpenSteeetMap (`.osm`) file including information on terrain and buildings within the coverage area. The radio parameters include the User Equipment (UE) and base station (eNB or gNB) transmission power, or the mobile node and Road Side Unit one in case of Wi-Fi-based scenarios, the operating frequency (e.g., 1 GHz for generic cellular-based scenarios, or 5.9 GHz for a scenario based on V2X Dedicated Short-Range Communications – DSRC), the position of the base station, if any, and the antenna height. The remainder of this Section will detail all the steps foreseen by our framework:

Step 1. Pre-processing: the first step is related to the pre-processing of the set of traces (which will be referred to as *dataset* from now on), and pruning of the mobile nodes which never enter the coverage area in the desired time instants. It should be noted how the user is responsible for choosing a meaningful coverage area, e.g., where most nodes are located most of the time. Therefore, the pre-processing module also provides the user with a way to understand where nodes are located in time, helping the process of selection of a geographical area.

Step 2. Ray-tracing: starting from the provided radio parameters and depending on whether a cellular-based or DSRC-based scenario is selected, ray-tracing is performed to compute the path loss gain and phase shift, which represent a required input for the clustering step. Optionally, the building material and terrain material can be specified, together with the maximum number of reflections. If not specified, the latter is set equal to the third order, as, according to our tests, it enables the generation of realistic scenarios without taking

[1] https://github.com/riccardo-rusca/Mobile_RF_Scenario_Design

into account too many low-power reflections. A noise level can also be configured to prune some of the less significant rays.

The ray-tracing step is one of the most important ones. Indeed, in wireless communication systems, wireless signals are transmitted via multiple propagation paths. Ground and buildings cause reflections, diffraction and, in general, unpredictable penetration patterns in these paths.

Realistically determining the behaviour of radio waves in mobile scenarios can represent a significant challenge, and leveraging simple path loss models (like the ones proposed by 3GPP [8]) may not be enough to reach the desired emulation accuracy. A ray-tracing algorithm can thus be used to determine the path loss and phase shift of each ray (i.e., of each component at the receiver) using electromagnetic analysis, including tracing the horizontal and vertical polarization of a signal. The latter is performed by analyzing the propagation path of the signals transmitted from the mobile nodes (transmitters) to the receivers, e.g., a fixed eNodeB (for 4G) or gNodeB (for 5G), or other moving devices.

Step 3. Clustering: the ray-tracer output model often includes tens of multi-path components. Due to the high computational complexity required to generate scenarios, and to the large space needed to store them, even the most powerful channel emulators accept up to four non-zero-valued channel taps. In order to reduce the number of components returned by the ray-tracer to the maximum value allowed, we followed the two-step procedure presented in [6], which is included in our framework as part of the **clustering** step. Specifically, the first clustering sub-step consists in applying a Machine Learning clustering algorithm. The goal is to find a certain number of centroids, one for each channel tap, that approximate the characteristics of all the components within the cluster.

To perform clustering, we applied a K-means algorithm leveraging as distance function the Multi-path Component Distance (MCD). As demonstrated by the authors of [9], the MCD distance can effectively improve the clustering performance of channel data over the classical Euclidean distances (i.e., Squared Euclidean Distance – SED – and Joint SED [6]). MCD takes into consideration the distance between the time of arrival and the angles of arrival/departure of the multiple paths.

Compared to the pseudocode in [6] (not reported here due to the limited space), when we loop over all centroids and average all paths assigned to a given cluster, we enhanced the code by introducing a random permutation of the centroid position. This is performed, specifically, during the initialization of the centroids, in case zero paths are assigned to a centroid. This results in having at least one path assigned to each centroid, where each centroid represents the approximated position of one channel tap. Then, in order to reconstruct the approximated taps, each centroid gain is computed as the sum of all multi-path components within the respective cluster. It should be noted how this procedure is repeated for each pair of transmit-receive node and for each timestamp available in the set of traces.

Finally, an approximated taps re-sampling step is foreseen, during which each centroid is aligned to the specific Finite Impulse Response (FIR) tap of the channel emulator, as described in [6]. The goal of the re-sampling step is to match the approximated centroid time of arrival with the channel emulator FIR filter indexes.

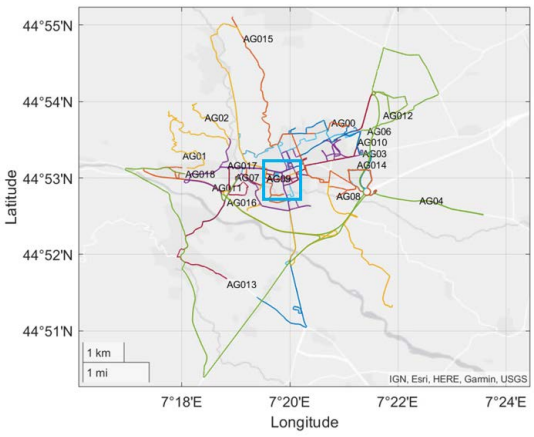

Fig. 2: SAMARCANDA traces with 1 km^2 area highlighted in blue. The picture showing the traces on a map is taken from [3].

Step 4. Generation of the Channel Matrix: the output of the clustering algorithm is a Channel Matrix of size *Number of nodes* $\times$ *Number of nodes* $\times$ *Number of timestamps*, which includes, for each link and timestamp, the values of the FIR filter delay and IQ coefficients for the path gain and phase shift of each signal.

Step 5. Scenario creation: the Channel Matrix is then used as input to the specific scenario creation toolchain available for the target channel emulator, producing as final output the RF scenario that can be installed in the emulator.

A relevant example of toolchain, dedicated to the Colosseum emulator, is represented by the Channel emulation generator and Sounder Toolchain (*CaST*), proposed in [5]. *CaST* receives as input the Channel Matrix described earlier, providing as output a full-fledged Colosseum scenario.

After the generation of the RF scenario, the latter can be leveraged to perform realistic experiments with connected mobile nodes (e.g., Wi-Fi clients or UEs), thanks to the channel emulator. A use case scenario, leveraging real V2X mobility traces and the Colosseum emulator is detailed in the next Section.

IV. V2X USE CASE SCENARIO: DATASET AND PROTOCOLS

Following the definition of our framework, we leveraged the proposed pipeline to create a V2X Colosseum scenario starting from a recently-disclosed high-accuracy open dataset [3]. This shows on the one hand how our framework provides an effective toolchain for the generation of data-driven channel emulation scenarios, and, on the other hand, provides the research community with a scenario for the evaluation of innovative V2X services within a realistic V2X mobility environment.

The core input of our scenario is represented by an open vehicular dataset, comprising the traces of 19 vehicles travelling in an urban and sub-urban area. The dataset is called Synthetic Accurate Multi-Agent RealistiC Assisted-gNss DatAset, in short *SAMARCANDA*, and was made available to the research

community by the authors of [3]. The traces include the dynamic data of 19 vehicles (herein called *agents*) travelling in the area around Pinerolo, a large town near Turin, in Italy.

The data for each vehicle is stored in CSV files and includes position (in terms of latitude and longitude), heading, speed, acceleration and an accurate timestamp for each of these values [3]. On average, an update every 100 ms is available for each of the 19 vehicles. Therefore, the minimum time granularity which can be emulated within our scenario is equal to 100 ms. This value is enough to emulate the great majority of V2X use cases, with frequent dynamic data updates.

As the dataset is composed of real traces, they model with high accuracy the behaviour of real vehicles, allowing researchers to test V2X applications without the need of resorting to less accurate mathematical mobility models. The traces were collected with a high-end Global Navigation Satellite System (GNSS) device with 10 Hz update rate and Real-Time Kinematic (RTK) corrections, including an Inertial Measurement Unit (IMU) for gathering additional dynamic information (e.g., acceleration) [3]. This kind of GNSS receiver is increasingly becoming the receiver of choice for testing and deploying innovative V2X applications as, thanks to RTK, it enhances the positioning accuracy up to a centimeter-level, enabling lane-level localisation.

The emulated scenario is a urban scenario around the city center of Pinerolo, representative of a mid-size city with a more crowded city center. As Colosseum supports the definition of an area up to $1\,\text{km}^2$, we selected as coverage area the most significant location in Pinerolo, i.e., its city center, where most vehicles are located most of the time. Furthermore, the choice of this central area enables a more accurate emulation of the effect of buildings in an urban scenario, with respect to other, more rural areas covered by SAMARCANDA.

The main area covered by the SAMARCANDA traces, together with the chosen $1\,\text{km}^2$ coverage area is depicted in Figure 2.

As mentioned earlier, the proposed framework requires to define whether to emulate a cellular-based or DSRC-based scenario, and set the input radio parameters accordingly. Given the low latency and high throughput requirements of most innovative V2X services, we focus on 5G connectivity through a central base station (i.e., a central gNB). Centralized approaches for automated and connected vehicles, such as centralized Federated Learning (FL) and highly automated maneuver management, are currently being investigated by several European Projects, which proved their effectiveness when combined with a reliable 5G network [10]. Furthermore, we believe these approaches are of particular interest for the research community, as centralized solutions typically provide a better and more accurate "view" of the road than purely decentralized algorithms [11]. Considering the reasonable range for a 5G base station, when focusing on centralized V2X scenarios, a $1\,\text{km}^2$ square appears to be technically sound as maximum emulation area, out of which vehicles can be considered out of coverage.

V. Creation of a V2X Scenario

This Section describes how we generated a Colosseum V2X scenario starting from our framework, and from the SAMARCANDA dataset. As mentioned earlier, our scenario models a 5G cellular-based communication between vehicles and a 5G base station.

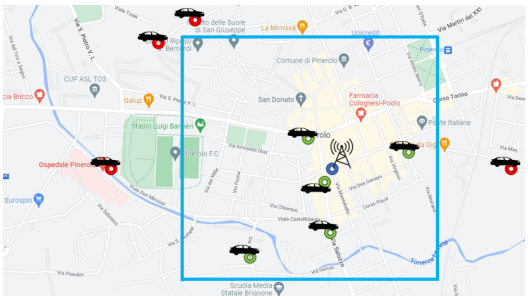

Fig. 3: Map of the selected area for the Colosseum scenario, with vehicles and the 1 km^2 zone highlighted in blue.

A. Vehicular scenario, pre-processing and pruning

To derive the Colosseum scenario, we leveraged the mobility traces extracted from the SAMARCANDA dataset. As Colosseum can, at the time of writing, only emulate scenarios in geographical areas up to $1\,\text{km}^2$, we selected an area of such a size for dataset pruning. The area has been selected following the rationale described in the previous Section, and it is represented by the blue square in Figure 2. After the pre-processing step, we obtained a pruned set of traces with 13 vehicles and one fixed antenna.

The pruning step also comprises the deletion of all timestamps of periods when no vehicles are travelling inside the area of interest. Even though they correspond to a very low percentage over all timestamps, in this specific case, this allowed us to provide a more compact representation of the dataset before the ray-tracing step. The final pruned dataset thus comprises 13 cars, and 5,200 timestamps for each vehicle, corresponding to a rough total of 8.7 minutes of total emulated time.

We decided to place one 5G antenna inside the $1\,\text{km}^2$ area, i.e., the blue point in the middle of the area in Figure 3. It is located to match the position of a real Vodafone LTE antenna as can be seen in [12]. The exact position of the antenna is $[44.88338°; 7.33152°]$. Furthermore, as Colosseum is optimized to work over a frequency band around 1 GHz, we selected it as reference frequency. However, it should be noted how, thanks to our framework, other frequency bands can be easily selected by tuning the input radio parameters. For instance, a new scenario emulating communication on the 5G N77 band, at 3.7 GHz, could be easily created by tuning only the transmitter frequency radio parameter.

B. Ray-Tracing with MATLAB

Concerning the ray-tracing step, among the currently available software suites, one of the most flexible is represented by the ray-tracing package within the MATLAB Communications Toolbox. This is thus the tool of choice for our scenario and for the open implementation of our framework. Specifically, we leveraged the `raytrace` function available in MATLAB [13], and we configured it with the radio parameters summarized in Table I. Specifically, we leveraged a transmitter frequency of 1 GHz, and a transmission power of 23 dBm, or 199.53 mW,

Fig. 4: Ray tracing with multiple propagation paths between the transmitter and the receiver.

according to [14]. We also left the maximum number of reflections up to the third order and set the buildings and terrain material to concrete.

The output of the `raytrace` function is a `Ray` object, representing all the rays found between the corresponding transmitter and receiver. For each ray a set of properties are specified, including: `PathLoss`, `PhaseShift`, `PropagationDelay`, `AngleOfArrival` and `AngleOfDeparture`. All this data was input in the clustering phase for the generation of the Channel Matrix.

Figure 4 depicts an example of the propagation rays and their power level (as shown in the legend on the left side), between a vehicle and the antenna.

C. Clustering

As described in Section III, we applied the optimized K-means algorithm, that uses the MCD as distance function, to the output of the ray-tracing step. Figure 5 depicts the results of the algorithm applied to the multi-path components shown in Figure 4. More in detail, the output are K clusters, with $K = 4$. The results are depicted by plotting each multi-path component in terms of path loss (in dB) versus the corresponding propagation delay (in seconds). The "x" symbols correspond instead to the centroids of each cluster, as determined by the K-means algorithm. After summing up the contribution of all multi-path components, we obtained a Channel Matrix with four delay and In-phase and Quadrature (IQ) values for each link towards the gNB and for each timestamp. As our aim is to model a centralized scenario, we modelled all the links from the vehicles to the base station, and from the base station back to the vehicles, leaving empty values for the vehicle-to-vehicle links. The Channel Matrix has been finally provided as input to the *CaST* toolchain for the generation of the RF scenario.

VI. Validation of the V2X Scenario

With the aim of validating the pipeline proposed as part of our framework and proving the validity and effectiveness of

TABLE I: MATLAB simulation parameters

Parameters	Value
Transmitter Power	23 dBm
Transmitter Frequency	1 GHz
Transmitter Antenna Height	1.5 m
Transmitter Antenna Gain	0 dBi
Receiver Antenna Height	30 m
Receiver Antenna Gain	0 dBi
Max Number of Reflections	3
Buildings and Terrain Material	Concrete
Ray-tracing method	SBR
Noise level	-89.1 dBm [6]

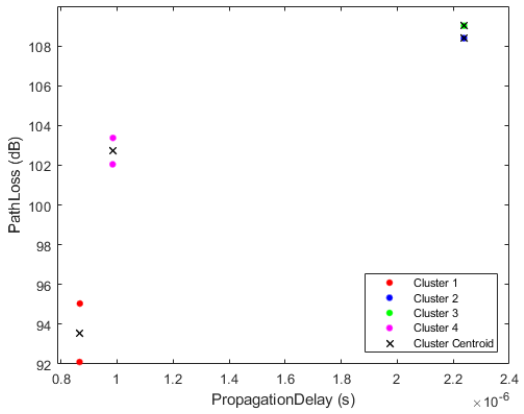

Fig. 5: Result of the K-means algorithm with Multi-path Component Distance (MCD).

the scenario described in the previous Section, we tested it inside Colosseum.

Specifically, each SRN has been configured to run SCOPE [15], as a full-fledged cellular-network system on both the base station and each of the 13 vehicles. SCOPE automatically manages the instantiation of the base station and Evolved Packet Core (EPC), and the association procedure for UEs (i.e., vehicles). SCOPE by default elects the lowest SRN id number as base station and the others as UEs. Since, in our scenario, the base station is the 14th SRN, we modified the default configuration so that the 14th node is elected as base station and all other nodes as UEs.

After starting the scenario, we measured several metrics, including the Round-Trip-Time (RTT) while the vehicles move around the coverage area, and the measured Downlink Signal-to-Noise-Ratio (SNR) in time. Then, these metrics have been compared to the position of each vehicle and to their distance from the gNB, leveraging the positions available in the SAMARCANDA dataset. For the sake of brevity, we report here the most important metrics, taking as reference one of the most interesting among the 13 vehicles.

Figure 6 shows the distance from the gNB as a function of time, comparing it with the RTT between the vehicle and the base station, measured thanks to the `ping` tool. The dotted lines delimit the time period in which the vehicle is located within the coverage area, while the vertical black lines delimit the moments in time included in the dataset after the pruning phase, i.e., the actual emulated timestamps.

As expected, communication appears to be possible only when the vehicle is located inside the coverage area, i.e., at a maximum distance of around 500 m from the gNB, being the latter placed around the center of the selected area. Furthermore, the average RTT seams to improve when the distance is reduced, thanks to the higher SNR and received power values. As a realistic urban scenario is modelled, several oscillations in both metrics can be observed, due to the effect of multi-path and buildings blocking and reflecting the signal.

Figure 7 reports instead the SNR as a function of time, compared with the distance from the gNB. We report the results as average values obtained for 5 experiments, with 99% confidence intervals. Consistently with the previous plot, it shows how a higher SNR is measured as the distance from the

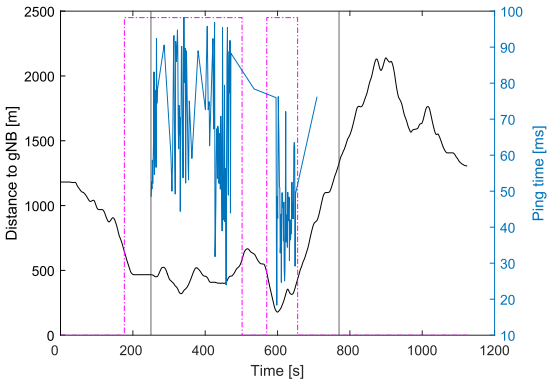

Fig. 6: RTT as a function of time, compared to the distance from the base station of the vehicle 9 in SAMARCANDA [3]

gNB decreases, with non-linearities and oscillations due to the effects of terrain and buildings, realistically modelled thanks to our framework and to the Colosseum emulator. Furthermore, the plot also shows how an SNR greater than, roughly, 3.0 is required to provide a stable-enough communication link between the reference vehicle and the gNB. Finally, it should be noted that the confidence intervals are very small, showing how the selected wireless emulation system can provide repeatable results. These results allowed us to validate both the proposed framework and the 5G scenario derived from it, showing the effectiveness of our approach.

VII. Conclusions

This paper has proposed for the first time a framework for the generation of large-scale RF scenarios for realistic channel emulators, starting from real traces and leveraging a data-driven approach. Building from the solution proposed in [5], we proposed a framework enabling the creation of arbitrary scenarios starting from pre-recorded real traces, detailing all the involved steps, and publicly releasing an open-source MATLAB implementation. Thanks to our framework, it becomes possible to easily create and deploy both cellular-based and Wi-Fi-based scenarios to large-scale emulators such as Colosseum [2], which has been used in our testbed.

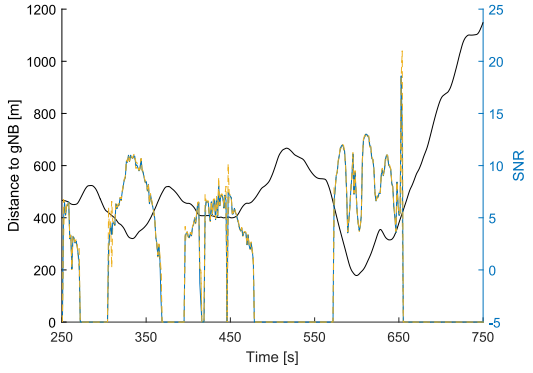

Fig. 7: SNR as a function of time, compared to the distance from the base station of the vehicle 9 in SAMARCANDA [3]. The blue line represents the average SNR over 5 experiments, with the 99% confidence intervals represented by orange dotted lines.

With the aim of showcasing our framework and providing the research community with a full-fledged RF scenario for development and prototyping of wireless applications, we also propose a V2X 5G centralized scenario for the Colosseum emulator, starting from the open SAMARCANDA dataset [3]. The creation and evaluation of such a scenario allowed us to validate our approach and show the effectiveness of our solution when creating, deploying and leveraging a realistic communication scenario with mobile nodes.

Acknowledgments and Disclaimer

This Work is partially funded by the European Commission through the H2020 projects Hexa-X (Grant Agreement no. 101015956), by the National Science Foundation under Grant CNS-2120447, CNS-2134973, OAC-2201536 and ECCS-2229472 as well as within the MOST – Sustainable Mobility National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033 17/06/2022, CN00000023). This manuscript reflects only the authors' views and opinions, and do not necessarily reflect the views of the National Science Foundation, neither the European Union nor the European Commission can be considered responsible for them.

References

[1] Hexa-X. [Online]. Available: https://hexa-x.eu/
[2] Leonardo Bonati et al., "Colosseum: Large-scale wireless experimentation through hardware-in-the-loop network emulation," in *Proc. of IEEE Intl. Symp. on Dynamic Spectrum Access Networks (DySPAN)*, Virtual Conference, December 2021.
[3] F. Raviglione, S. Zocca, A. Minetto, M. Malinverno, C. Casetti, C. Chiasserini, and F. Dovis, "From collaborative awareness to collaborative information enhancement in vehicular networks," *Vehicular Communications (Elsevier)*, vol. 36, 2022.
[4] J. Yackoski, B. Azimi-Sadjadi, A. Namazi, J. H. Li, Y. Sagduyu, and R. Levy, "Rfnest™: Radio frequency network emulator simulator tool," in *MILCOM*, 2011.
[5] D. Villa, M. Tehrani-Moayyed, P. Johari, S. Basagni, and T. Melodia, "CaST: A Toolchain for Creating and Characterizing Realistic Wireless Network Emulation Scenarios," in *ACM WiNTECH*, Sydney, Australia, October 2022.
[6] M. Tehrani-Moayyed, L. Bonati, P. Johari, T. Melodia, and S. Basagni, "Creating RF scenarios for large-scale, real-time wireless channel emulators," in *MedComNet*, 2021.
[7] Wireless EM Propagation Software - Wireless InSite - Remcom. [Online]. Available: https://www.remcom.com/wireless-insite-em-propagation-software
[8] ETSI, "ETSI TR 138 901 V15.0.0 (2018-07) - 5G; Study on channel model for frequencies from 0.5 to 100 GHz (3GPP TR 38.901 version 15.0.0 Release 15)," European Telecommunications Standards Institute, Standard ETSI TR 138 901 V15.0.0, 2018.
[9] N. Czink, P. Cera, J. Salo, E. Bonek, J. Nuutinen, and J. Ylitalo, "Improving clustering performance using multipath component distance," *Electronics Letters*, vol. 42, 2006.
[10] 5G-CARMEN. [Online]. Available: https://5gcarmen.eu/
[11] F. Raviglione, C. M. R. Carletti, C. Casetti, F. Stoffella, G. M. Yilma, and F. Visintainer, "S-LDM: Server Local Dynamic Map for vehicular enhanced collective perception," in *IEEE VTC2022-Spring*, 2022.
[12] Cell mapper. [Online]. Available: https://www.cellmapper.net/
[13] Reference manual for raytrace. [Online]. Available: https://www.mathworks.com/help/comm/ref/txsite.raytrace.html
[14] P. Joshi, F. Ghasemifard, D. Colombi, and C. Törnevik, "Actual output power levels of user equipment in 5G commercial networks and implications on realistic RF EMF exposure assessment," *IEEE Access*, vol. 8, 2020.
[15] L. Bonati, S. D'Oro, S. Basagni, and T. Melodia, "SCOPE: An Open and Softwarized Prototyping Platform for NextG Systems," in *ACM MobiSys*, Virtual Conference, June 2021.