Location-based analytics in 5G and beyond

Stefania Bartoletti, Luca Chiaraviglio, Sergio Fortes, Takai Eddine Kennouche, Gürkan Solmaz, Giacomo Bernini, Domenico Giustiniano, Joerg Widmer, Raquel Barco, Giuseppe Siracusano, Andrea Conti, and Nicola Blefari Melazzi

***Abstract*—Location-based analytics will enable a plethora of new services for 5G verticals and will empower the optimized use of network resources. Such analytics build on enhanced 5G positioning obtained through new positioning signals and procedures defined within 5G standardization, together with the integration of heterogeneous technologies for achieving sub-meter accuracy. This paper proposes an end-to-end architecture integrated within the 5G network infrastructure to provide location-based analytics as a service. We also present an overview of cutting-edge applications in 5G and beyond, focusing on *people-centric* and *network-centric* location-based analytics.**

***Index Terms*—5G, localization, analytics, network management, machine learning.**

I. Introduction

The 5G cellular network and its long-term evolution are targeting the design of pervasive cyber-physical systems, whose operation and control depend on accurate location information. This is addressed by enhancing 5G networks and devices (under discussion in 3GPP Rel. 17) towards localization functionalities, thus elevating location information to a first class network service [1], [2]. Besides the localization of users, there is a growing interest in location-based analytics, i.e., the analysis of the location and behaviour of people and things in public areas, roads, and buildings, through dedicated infrastructures or by relying on user devices [3]–[10]. While closely related, location-based analytics are not a mere extension of user equipment (UE) localization, but rather a new paradigm that enables a large variety of scenarios and applications, including security, transportation and smart cities, as well as opportunistic networking.

Location-based analytics can be classified as *people-centric* and *network-centric*. People-centric analytics refer to the ensemble of information related to people presence and movements in physical spaces (e.g., people counting, dynamic map creation and people flow tracking, fusion of spatiotemporal data with multimodal information, and anomalous behaviour detection) [3]–[7]. Network-centric analytics refer to the ensemble of information related to network operation (e.g., network planning, fault detection, resilience, location-aware diagnosis and troubleshooting) [8]–[10]. On the one hand, the ability to operate 5G networks in both sub-6 GHz and Millimeter Wave (mmWave) frequency bands and the use of massive antenna arrays significantly extend the capabilities of people-centric localization. On the other hand, such new 5G features, including beamforming, multi-connectivity, and the adoption of new spectrum portions pose new challenges for autonomous network management. As a consequence, the exploitation of location information related to user equipment can boost the efficiency of network management for the provisioning of dynamic services.

The provision of location-based analytics relies on complex features and mobility patterns extracted from raw location data within physical and network events. This calls for an extension of the network functions (e.g., the scheduler) to interface with location data in a multi-layer and flexible architecture that i) facilitates secure sharing and re-use of accurate location and context data for diverse localization services, and ii) combines the different network functions for the extraction of location-based analytics. Moreover, network functions should be served on top of the 5G infrastructure and the pure 5G connectivity. There is a unique opportunity for network providers to make location-based analytics a network-native service in 5G and beyond, which will be pivotal to creating new disruptive services and to optimize network performance.

This paper proposes a full-stack architecture integrated with the 5G network infrastructure (Sec. II) to serve a plethora of services requiring location-based analytics. Such analytics rely on enhanced positioning provided in 5G, also integrating hetero-

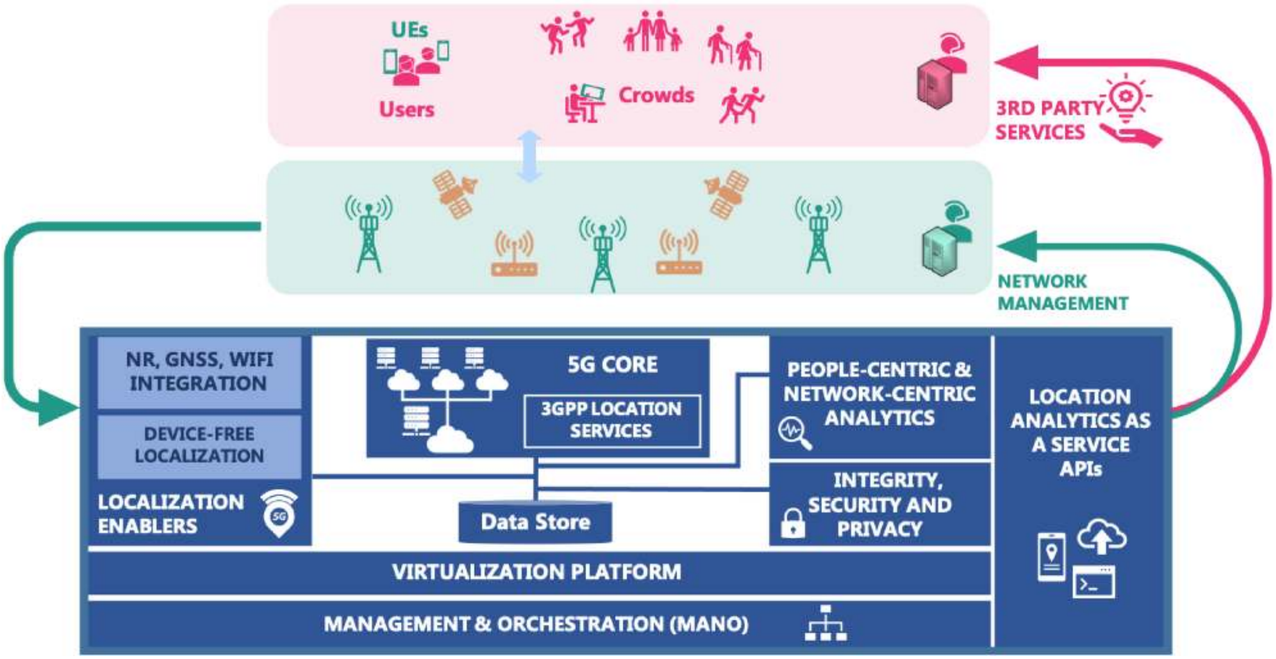

Fig. 1: System architecture for localization analytics as a service.

geneous data and device-free localization (Sec. III). Then, we detail a set of case studies on people-centric and network-centric analytics (Sec. IV).

II. End-to-end Architecture

We propose new system functionalities integrated in the 5G network infrastructure (spanning network edges and data centers) to allow operators and service providers to expose location-based analytics as a service. Such functionalities leverage 5G network information combined with heterogeneous data from other radio access technologies [2] (see Sec. III).

A. Localization and analytics functions

We propose the use of virtualization techniques to run the localization and analytics functions as virtual functions, i.e., providing an augmentation of the 5G architecture by leveraging on the ETSI network function virtualization framework, which represents the 3GPP standard for operators to deploy 5G network functions in virtualized infrastructures. This augmentation of the 5G architecture offers operators and service providers the possibility to expose new location-based analytics to third parties and exploit location data for smart network management applications.

Fig. 1 presents a comprehensive view of the proposed system architecture and includes details of how the location-related functions coexist on top of a virtualized infrastructure, for their on-demand deployment in the form of localization services. The proposed system is compliant with the 3GPP 5G Core architecture and makes use of a service based architecture that integrates with the 5G network functions and augments it with atomized and independent location functions. Specifically, the system is aligned with the enhanced 3GPP Location Service (eLCS) architecture, which specifies 5G network functions, interfaces, and workflows for location-related functionalities [1]. Here, the location management function (LMF) coordinates and calculates the user position for location-based services requested by external or internal eLCS clients, including other network functions.

In our proposed system architecture, the localization enablers provide the other system functions with user device location data (e.g., coordinates, velocity, direction). In particular, localization enablers implement two type of LMFs (see Sec. III) deployed on-demand to fulfill specific positioning requirements: i) integration of New Radio (NR), global navigation satellite system (GNSS), WiFi; and ii) device-free localization. Such LMFs provide location data to the location data analytics functions (LDAFs) for the provision of location-based analytics. Such LDAFs are deployed on-demand

and combined according to the localization service requirements and objectives. LDAFs can be considered as LCS clients and use positioning data from localization enablers LMFs. LDAFs for people-centric and network-centric apply descriptive, predictive, prescriptive, and diagnostic algorithms to respectively perform statistical analysis on location and network data, assess future possible conditions, search for actions to be taken, and determine the causes for specific conditions.

Finally, the integrity, security, and privacy functions provide authentication and advanced cryptographic techniques on the localization and analytics data to be exposed towards external applications, secure conditional sharing techniques and data management policies (e.g., anonymization, obfuscation).

B. Localization analytics as service APIs

The localization analytics are exposed as services through application programming interfaces (APIs) consisting of pipelines of functions, which are linked together to provide the application with the requested output. This process relies on a workflow execution engine, i.e. the management and network orchestration, which translates the service request into a number of functional steps, involving one or more localization functions.

The localization analytics output is passed to the application via service APIs, either by exposing the output as an on-demand RESTful service or by exposing it as a continuous data stream. This is done via dedicated access control functions within the API layer. In summary, the overall proposed approach has the main goal to enable a flexible and composable platform where the various localization functions can be combined while facilitating sharing and re-use of some of the key functionalities (e.g., those for the localization enablers or data security and privacy) across different localization services.

III. 5G Positioning

This section presents the ongoing 3GPP standardization activities and the research in the area of 5G localization, to better define the eLCS involved within the proposed architecture. We also provide an overview on the main technologies that can be fused together with 5G location data for enhanced localization, with a special focus towards device-free localization.

A. 5G standardization, metrics and use cases

Positioning in 5G was introduced in Rel. 15 for non-standalone operation (5G networks aided by existing 4G infrastructure) and continued in Rel. 16 with standalone NR operation, with further enhancements expected in Rel. 17. From a theoretical standpoint, positioning in 5G relies on single value estimation, where each measurement used for localization corresponds to the estimate of a single-value metric. In particular, 3GPP Rel. 16 compliant solutions mainly rely on downlink time difference of arrival (DL-TDoA) and beamforming angle of arrival (AoA). Depending on the use case, some received signal strength indicators (RSSIs) such as the reference signal received power (RSRP) and the reference signal received quality (RSRQ) can be also used for positioning. For instance, for people monitoring and flow control, fingerprinting-based solutions for 3GPP reference signals using machine learning can deliver accurate positioning results with online data.

Richer information increases the accuracy, especially in challenging environments, and can be extended to use soft information [11] to improve the positioning accuracy for 5G use cases. New compelling applications relying on high-precision positioning technology in autonomous applications, high integrity and reliability in addition to high accuracy and low latency become a necessity.

Integrity is the measure of trust that can be placed in the correctness of information supplied by a navigation system. It includes the ability of a system to provide timely warnings to user devices in case of failure. In 3GPP Rel. 17, integrity has started to be considered mainly for GNSS localization to support well-known automotive and railway use cases. Nevertheless, integrity solutions are becoming increasingly important also for 5G positioning and for providing accurate, timely, and reliable positioning data for people-centric and network-centric services.

B. Heterogeneous location data fusion

The fusion of radio access technology (RAT)-dependent and RAT-independent location data in a hybrid fashion is beneficial for the demanding positioning requirements on accuracy, latency and integrity level for 5G use cases. Although 3GPP already provides support for RAT-independent methods, new technology trends suggest that we are moving from having several independent chipsets in smartphones, towards the integration in a single chipset, turning the vision of heterogeneous location data fusion into reality in the coming years.

The GNSS is almost fully supported in 3GPP for both 4G and 5G. The combination of 5G cellular positioning and GNSS is needed for many use cases in which one technology is not fully operating or has limited coverage, such as in tunnels or urban canyon scenarios. Studies show that use of even only one high-accuracy 5G round-trip-time observable can remarkably improve the horizontal positioning accuracy with respect to GNSS stand-alone solutions by relaxing the positioning problem and improving the geometry of the solution [12].

Concerning the integration of other RAT-independent positioning methods, the combination of ranging measurements for a UE from multiple WiFi access points (APs) and 5G NR cells, for both indoor and outdoor scenarios, is envisaged to accomplish high-accuracy positioning. However, in 5G networks, the location server may not have the information about the WiFi APs exact locations; this limits the usefulness of WiFi data at the location server. In such cases, for instance, smartphone movements can be estimated using WiFi Fine Time Measurement ranging measurements [13] without relying on the knowledge of the AP position. These data can be integrated in a network-based location system defined in 3GPP, where the network collects round-trip-time measurements sent from the UE.

In this context, the large bandwidth of mmWave networks not only provides very high accuracy positioning, but enables simultaneous localization and mapping (SLAM) though AoA information. SLAM in mmWave networks relies on anchor location estimation, device localization, and environment mapping for both physical and virtual anchors. A low complexity SLAM algorithm fully integrated with a mmWave communication system is feasible with median error smaller than 0.3-0.5 m [14].

C. Device-free localization

Device-free localization addresses the identification and analysis of signals backscattered by single and multiple device-free targets (persons, vehicles, etc.) and relies on sensor radar networks.Such networks sense the wireless environment to infer the location of targets from signal reflections and obstructions and can take advantage of any modulated signal at any frequency of operation.

The ultra-low latency connectivity and a finer radar range resolution enabled by 5G are paving the way to the use of 5G NR waveforms for radar networks. As an example application, device-free localization has been proposed in [15] as an integrated radar service for future vehicle networks. In this context, the use of mmWave technology is particularly relevant since the reduced wavelength at mmWave allows the use of massive arrays with electronic steering capabilities, thus enhancing the directionality properties for detection and tracking of device-free targets.

IV. From Localization to Analytics

The ubiquity of sensors in the 5G ecosystem provides unprecedented opportunities for obtaining large-scale mobility attributes as well as for understanding human mobility patterns. Such ubiquity also enables to exploit the correlation between human activities and network events. This section presents a set of case studies for location-based people-centric and network-centric analytics, with examples that can be implemented as LDAFs within the proposed system architecture.

A. People-centric location-based analytics

People-centric analytics provide insights and empower domains such as smart cities and transportation and can enable a number of 5G services in such contexts.

1) Group detection and people counting: There is a growing interest in designing crowd-centric device-free [5] and device-based [6] methods for group detection and people counting that infer the number of targets directly from the measured data without estimating their locations.

This use case considers people counting and group detection based on *wireless activity of mobile devices* using wireless scanners. City-scale measurements [7] are conducted to analyze crowd mobility and Group-In method [6] is developed for group inference from wireless traces. In particular, we consider the scalability of group detection in the city-scale pilot study conducted in Gold Coast, Australia. We applied graph algorithms in [6] at a

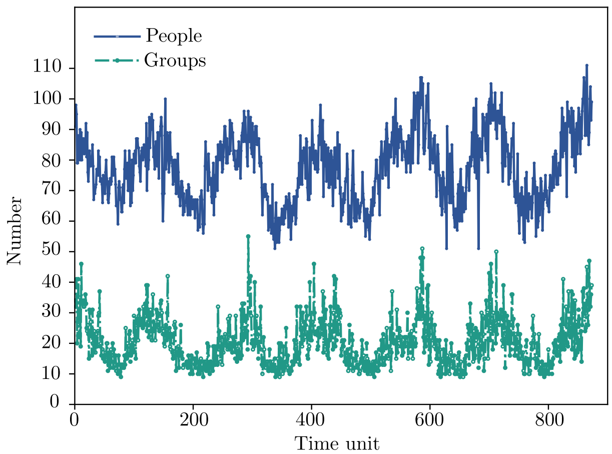

Fig. 2: People counts and number of groups observed every 10 minutes based on wireless activities in the Gold Coast city. The results show equal time units in the one week period.

large scale using WiFi and 5G datasets for verifying the feasibility of group detection in large areas. The parameters are trained in controlled laboratory environments, after which the obtained models are applied to the larger-scale data. In the centralized computing phase of Group-In, we apply the wireless fingerprint match algorithm based on RSSIs. In the group-inference phase, the highly-connected subgraphs algorithm is applied.

Fig. 2 shows the observations from Group-In application to a one-week period (10 min. time interval, 30 sec. sampling time). We observe a positive correlation between the number of groups and the number of people. Data have a daily trend with a peak value (up to 110 people) every day. Initial results indicate that Group-In can be applied to analyze larger datasets such as city-scale data for long periods, and it can be used to create general group behaviour insights. Accurate localization through 5G will lead to more granular insights for people counting and group behaviour identification. Moreover, similar insights can be produced in real time for much larger regions through wide deployment and availability of the 5G infrastructure.

2) Mobility clustering: This use case investigates the mobility patterns in large-scale mobility datasets, which can be implemented within the proposed architecture using as input 5G LMFs. Such datasets exhibit challenges in terms of granularity, regularity, and accuracy, which motivate the use of modern deep learning techniques to be implemented as LDAFs. In this context, we investigated recurrent networks based sequence-to-sequence autoencoders [4] for human mobility analysis. We conducted unsupervised spatiotemporal clustering on the OpenPFLOW dataset [3], which represents walking, biking and commuting mobility in the city of Tokyo for 24h at regular 1 min. timesteps. The autoencoding model is formed by stacking layers of gated recurrent units in an encoder/decoder structure.

After training, spatiotemporal aspects of the mobility data are encoded in the latent space represented by the encoder output. There we apply principle component analysis, and then use K-Means method to detect clusters. Fig. 3 shows the process applied to walking trajectories from [3]. The visualization on the actual Tokyo map indicates potential trends such as regional, sub-regional, and cross-regional mobility concentration, as well as patterns of stationary and non-stationary behavior across different time periods. The fusion of 5G location data from heterogeneous technologies together with additional contextual information, as enabled by the proposed end-to-end architecture, will further improve such mobility analytics and enable, for example, dedicated network functions for anomaly detection of irregular trajectories and the interplay with other aspects of human activity.

B. Network-centric location-based analytics

Two use cases are now presented to show the use of location-based analytics for network management: *network optimization* for efficient service provisioning considering the dynamic changes in the network; and *location-aware diagnosis/troubleshooting* for the maintenance of the cellular network by identifying problems as well as ensuring the resilience of the network itself.

1) Network optimization: An example of promising techniques for location-aware network optimization is pencil beamforming based on the estimated UE position. We have performed a preliminary analysis for the impact of pencil beamforming on the QoS of 5G networks and the ElectroMagnetic field (EMF) exposure. To this aim, an open-source simulator has been developed [9] that is able to synthesize the traffic beams for each gNB, both in direction and beamwidth, by exploiting user equipment (UE) localization accuracy. Each beam is directed towards the centre of a circular area in which the UE is assumed to be, where the diameter

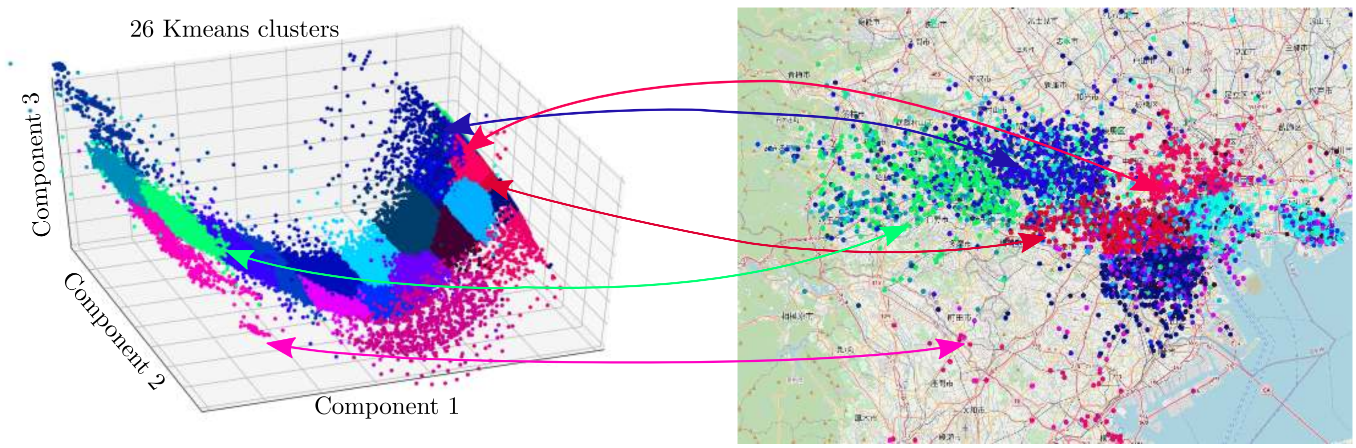

Fig. 3: Pedestrian mobility autoencoder-based clustering in Tokyo.

of this circular area indicates the uncertainty level for UE location estimate.

Table I presents the values of the average EMF [V/m] and the average throughput [Mbps] according to different location uncertainty levels. Results show that an increase of the location uncertainty level results in a higher EMF (due to possible overlap of the wider beams) and a lower throughput (the higher beam width lowers the beam's directivity). Therefore, accurate localization that reduces uncertainty levels helps to reduce the EMF exposure while increasing the throughput.

2) Network diagnosis: Location-aware network diagnosis can rely on *contextualized indicators*, i.e., time-series metrics combining location and cellular network measurement. Such indicators are extracted from the network measurements reported by the users in different areas of interest, including the cell coverage, center, and edge. This concept can be especially beneficial for 5G ultra-dense scenarios, characterized by a high dynamicity of users and an increased demand due to the reduced coverage areas and inter-site distances [10]. Supported by the high-accuracy localization provided by 5G systems, novel developments have aspired to complete automation of the definition of the areas of interest. This has led to an increased number of available contextualized indicators that can be used for diagnosis: each cell coverage area, center, edge, influencing area on other cells and area being influenced by each of their neighbours.

Figure 4 compares the performance of failure diagnosis mechanisms using classic metrics only with the use of both classic and contextualized metrics

TABLE I: Average EMF [V/m] and average throughput [Mbps] for different values of location uncertainty level.

	2 m	4 m	8 m	16 m	20 m
Avg EMF [V/m]	0.797	1.558	2.948	4.700	5.509
EMF C.I. [V/m]	0.795 0.798	1.555 1.561	2.944 2.952	4.694 4.706	5.502 5.516
Avg throughput [Mbps]	422.001	138.593	43.518	13.159	9.255
Thr. C.I. [V/m]	411.683 432.318	134.197 142.999	41.776 45.261	12.583 13.735	8.860 9.650

(fusion). This is done for the indoor ultra-dense scenario with 12 picocells and multiple modelled failures presented in [10]. Network diagnosis is performed based on three classifiers, namely k-nearest neighbors, discriminant analysis classification and multiclass error-correcting output codes classification. Results show how for the different classifiers the use of contextualized data considerably decreases the diagnosis error rate with respect to only using classical metrics, thus providing a powerful tool for 5G failure management. The availability of localization data for the generation of the location-enriched metrics allow the median diagnosis error rate for the three classifiers to be reduced significantly, going below 1% for disc and multiclass. This demonstrates the relevance of location-aware information for improving failure management of 5G networks.

V. Conclusion

This paper has proposed a new system architecture for the provision of location-based analytics as a service, which will enable a plethora of

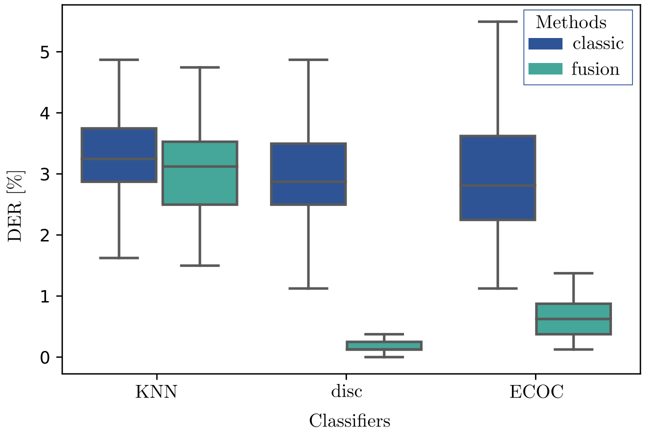

Fig. 4: Comparison between the diagnosis error rate (DER) obtained by classic and location-enriched contextualized metrics in an ultra-dense scenario, using k-nearest neighbors (KNN), discriminant analysis classification (disc) and multiclass error-correcting output codes classification (ECOC).

people-centric and network-centric applications for 5G verticals. The system architecture is proposed as an augmentation of the 5G architecture, where network and user data from heterogeneous technologies are combined to extract on-demand analytics that can serve third party applications or can be used to optimize the network performance. Example analytics for people grouping, mobility clustering, network optimization, as well as network diagnosis have been presented.

VI. Acknowledgment

This work was supported by the European Union's Horizon 2020 research and innovation programme under Grant no. 871249. The pilot study in Gold Coast is conducted with NEC Australia.

References

[1] 3GPP, "5G System (5GS) Location Services (LCS); Stage 2," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 23.273, 7, version 16.4.0.

[2] 3GPP, "Universal Terrestrial Radio Access (UTRA) and Evolved UTRA (E-UTRA); User Equipment (UE) performance requirements for Radio Access Technology (RAT) Independent Positioning Enhancements," 3GPP, Technical Specification (TS) 37.171, 10, version 16.1.0.

[3] T. Kashiyama *et al.*, "Open PFLOW: Creation and evaluation of an open dataset for typical people mass movement in urban areas," vol. 85, pp. 249–267, 2017.

[4] I. Sutskever *et al.*, "Sequence to sequence learning with neural networks," in *Adv. in Neural Inf. Processing Sys.*, 2014, pp. 3104–3112.

[5] S. Bartoletti *et al.*, "Device-free counting via wideband signals," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1163–1174, May 2017.

[6] G. Solmaz *et al.*, "Group-in: Group inference from wireless traces of mobile devices," in *Proc. ACM/IEEE Int. Conf. on Inf. Process. in Sensor Netw. (IPSN)*, 2020, pp. 157–168.

[7] G. Solmaz *et al.*, "Toward understanding crowd mobility in smart cities through the internet of things," *IEEE Commun. Mag.*, vol. 57, no. 4, pp. 40–46, 2019.

[8] J. Ferragut *et al.*, "Traffic and mobility management in networks of femtocells," *Mob. Networks Appl.*, vol. 17, no. 5, pp. 662–673, 2012.

[9] L. Chiaraviglio *et al.*, "Pencil beamforming increases human exposure to electromagnetic fields: True or false?" in *arXiv preprint arXiv:2010.16288*, Nov. 2020.

[10] S. Fortes *et al.*, "Contextualized indicators for online failure diagnosis in cellular networks," *Computer Netw.*, vol. 82, pp. 96 – 113, April 2015.

[11] A. Conti *et al.*, "Soft information for localization-of-things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.

[12] J. del Peral-Rosado *et al.*, "Exploitation of 3D city maps for hybrid 5G RTT and GNSS positioning simulations," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, Barcelona, Spain, 2020.

[13] M. Rea, D. Giustiniano, and J. Widmer, "Virtual inertial sensors with fine time measurements," in *Proc. of IEEE Mass*, 2020.

[14] J. Palacios *et al.*, "Communication-driven localization and mapping for millimeter wave networks," in *Proc. IEEE Conf. on Computer Commun.* IEEE, 2018, pp. 2402–2410.

[15] R. S. Thoma *et al.*, "Cooperative passive coherent location: A promising 5G service to support road safety," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 86–92, Sep. 2019.

Stefania Bartoletti is Researcher at the National Research Council of Italy (IEIIT-CNR) and Research Associate at CNIT, working on location-aware wireless networks.

Luca Chiaraviglio is Associate Professor at University of Rome Tor Vergata and Research Fellow at CNIT, working on health risks assessment of 5G/B5G networks and optimization techniques.

Sergio Fortes is Researcher at Universidad de Málaga (UMA). His work focused on the development of machine learning and context-awareness approaches for the management of cellular networks.

Takai Eddine Kennouche is a Researcher at VIAVI Solutions France. His current research interests include 5G localization and analytics, urban computing and machine learning.

Gürkan Solmaz is Senior Researcher at NEC Laboratories Europe, Germany. His research interests are primarily in IoT and machine learning for smart cities and human mobility.

Giacomo Bernini is R&D Technical Coordinator at Nextworks. His research interests include SDN, NFV, MEC, cloud computing, 5G service orchestration and network slicing.

Domenico Giustiniano is Research Associate Professor at IMDEA Networks Institute, Madrid, Spain. His research interests are in wireless systems, including LiFi, spectrum sensing and 5G localization.

Joerg Widmer is Research Director as well as Research Professor at IMDEA Networks Institute, Madrid, Spain. His research focuses on wireless networks and millimeter-wave communications.

Raquel Barco is Full Professor in Telecommunication Engineering at the University of Malaga, specialized in mobile communication networks and smart-cities.

Giuseppe Siracusano is Senior Researcher at NEC Laboratories Europe, Germany. His research interests are computer systems and networks, with a focus on machine learning applied to systems.

Andrea Conti is a Professor at the University of Ferrara and a CNIT affiliate. His research interests include localization and tracking, wireless communications, and quantum information science.

Nicola Blefari Melazzi is Professor of telecommunications with the University of Roma Tor Vergata and is currently the Director of CNIT, a non-profit Consortium among 37 Italian Universities.