Mobility-Predictive Feature Extraction from 5GC, RAN, and Integrated Frameworks Using NetSim

Gwangun Yu, Gilhan Choi
Kongju National University
Cheonan 31080, Korea
{yugwangun, choikilhan12}@smail.kongju.ac.kr

Yonggang Kim*
Kongju National University
Cheonan 31080, Korea
ygkim@kongju.ac.kr

***Abstract*—Accurate user mobility prediction is a fundamental requirement for efficient resource management in 5G networks. However, the 5G Core (5GC) and the Radio Access Network (RAN) provide inherently distinct perspectives of mobility, with the former capturing coarse grained location and access behavior while the latter records radio level measurements and handover events. In this work, we construct an integrated dataset by systematically extracting and standardizing parameters from the 5GC, the RAN, and their joint integration, in accordance with 3GPP specifications (TS 23.288, TS 23.273, TS 38.331, and TS 38.305). The methodology employs Köln mobility traces as user input and simulates scenarios in NetSim, where logs are generated separately from each domain and then combined. The resulting dataset preserves 3GPP compliant semantics, enables cross-domain validation of mobility events, and provides a unified foundation for AI-driven mobility prediction and analytics.**

***Index Terms*—5GC, RAN, LCS, UE mobility prediction, NetSim, Köln trace,**

I. Introduction

The fifth-generation (5G) mobile network is designed to support enhanced Mobile Broadband (eMBB), Ultra-Reliable Low-Latency Communications (URLLC), and massive Machine-Type Communications (mMTC), as defined in the Third Generation Partnership Project (3GPP) specifications beginning with Release 15. With the increasing service demands and corresponding proliferation of heterogeneous base stations and network infrastructure, along with the adoption of diverse frequency bands such as A6G, precise prediction of user equipment (UE) mobility has emerged as a critical factor for improving performance across various 5G usage scenarios. For example, mobility management directly affects handover performance, session continuity, and Quality of Service (QoS). Inefficient or delayed handovers may lead to intermittent connection failure, increased latency, or degraded throughput, which are critical issues in applications such as autonomous driving, augmented reality, and mission-critical communications.

Various studies have highlighted the role of Machine Learning (ML) in improving mobility prediction. For example, Rydén *et al.* demonstrated that ML-based prediction of mobility, traffic, and radio channel conditions at the Radio Access Network (RAN) level can significantly enhance network performance [1]. Jeong *et al.* focused on the 5G Core (5GC) domain and showed that ML-assisted prediction frameworks are able to reduce signaling overhead by up to 75%, underlining the benefits of predictive analytics in session and control management [2]. Beyond these domain-specific efforts, Rago *et al.* proposed a spatiotemporal modeling of UE dynamics for Cloud-RAN environments, where prediction of user mobility and channel quality can enable more efficient resource allocation [3]. To effectively capture network topology and user inter-dependencies, Bermudez *et al* employed Graph Neural Networks (GNNs) for mobility management in O-RAN. The authors demonstrated that integrated-domain data can be effectively used to enable proactive handover decisions and intelligent link prediction. [4]

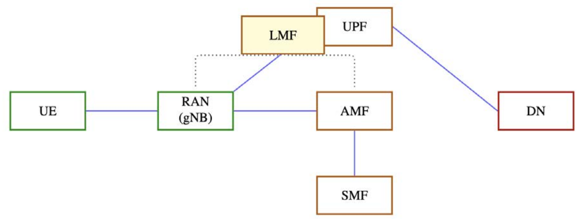

Fig. 1: Overview of 5G architecture.

While these studies provide strong evidence of the benefits of mobility prediction within either the RAN or 5GC domains, they are often limited to domain-specific perspectives. RAN-level metrics alone cannot capture session-level delays caused in the 5GC, while 5GC-domain analysis ignores the dynamic behavior of the radio environment. To overcome this limitation, this paper advocates an integrated domain perspective that combines RAN and 5GC-level data, in line with 3GPP's emphasis on end-to-end mobility optimization in Releases 17 and 18 [5]. By correlating radio layer measurements with 5GC network signaling and transport metrics, the integrated domain offers a holistic view of mobility. For effective mobility analysis, we utilize NetSim which we further modify to generate representative datasets and enable a methodology for extracting and analyzing RAN, 5GC, and integrated domain data. This work lays the groundwork for future AI-enhanced mobility prediction and adaptive mobility management in 5G and beyond.

II. 5G Specifications and Architecture

According to 3GPP specifications, mobility management in the 5G system requires tight coordination between the

TABLE I: UE Mobility information collected from 5GC (3GPP TS 23.288, Rel. 18)

Information	Source	Description
UE ID	AMF	SUPI used for UE identification
UE location(s)	AMF	UE position (TA/cell entered) with timestamp
Fine-granularity location	LCS	Coordinates (GAD shape), timestamp, motion events, QoS accuracy
Linear distance threshold	NWDAF	Distance travelled before triggering new reports
Type Allocation Code (TAC)	AMF	Vendor/model info; helps detect abnormal mobility
Frequent mobility updates	AMF	Detects abnormal "ping-pong" mobility from cell reselections
UE access behaviour	AMF	Metrics on UE state transitions (idle, connected, handover)
UE location trends	AMF	Long-term statistics of UE location changes

5GC network and the RAN, as shown in figure 1. The 5GC architecture is well organized in TS 23.501 and the associated procedures in TS 23.502 define the roles of the Access and Mobility Management Function (AMF), Session Management Function (SMF), User Plane Function (UPF), and the Data Network (DN). AMF handles UE registration, connection, and mobility anchoring. The SMF manages Packet Data Unit (PDU) session establishment and policy-driven path control, while the UPF ensures user plane forwarding toward the Data Network (DN). United with these functions maintains session continuity as the UE moves across RANs. In the RAN, TS 38.300 describes the overall NR architecture and TS 38.331 specifies Radio Resource Control (RRC) signaling procedures. UEs perform measurements such as RSRP, RSRQ, and SINR, which are reported to the gNB. Based on these reports, the gNB initiates or coordinates handover execution with the AMF. Enhancements introduced in Release 16 and Release 17, including dual connectivity, conditional handover, and mobility robustness optimization, are documented in TR 38.863. These procedures ensure mobility performance in dense and high-mobility environments.

In addition to these 5GC and RAN functions, the Location Management Function (LMF) is employed to support 5G Location Services (LCS) as specified in TS 23.273 and TS 38.305. The LMF collects measurement information from the gNB (via NRPPa) and the UE (via LPP), and interacts with the AMF to support positioning and mobility related procedures. By bridging radio-level measurements and 5GC level contexts, the LMF enables location aware mobility management.

The integration of UE-related mobility data from both domains, augmented with LMF outputs, naturally advances the development of a unified framework. Such integration allows cross-validation of location and mobility patterns, enhances prediction accuracy, and establishes the foundation for the subsequent discussion on integrated domains. In this context, the inclusion of LCS-derived location data from 5GC, as specified in the 3GPP standards, becomes essential for enabling standardized and reliable mobility prediction.

III. 5GC, RAN, and Integrated Domains

As discussed in the preceding section, the 5GC and the RAN are responsible for fundamentally different aspects of the 5G system and therefore provide distinct types of information. According to the relevant 3GPP specifications, it is possible to identify the categories of data that must be extracted from each domain in order to enable accurate mobility analysis and prediction.

A. *Mobility Management Information in the 5GC*

In the 5GC domain, mobility-related information primarily refers to the broad area of user activity and its corresponding estimated location. This domain focuses on identifying the UE, tracking its access and registration status, and collecting location information at varying levels of granularity. The AMF provides UE identity and coarse location updates, while the LCS is responsible for fine-granularity positioning. In addition, the Network Data Analytics Function (NWDAF), standardized in 3GPP TS 23.288, enhances mobility management by deriving statistical and predictive insights such as abnormal mobility detection or distance-based reporting thresholds. Together, these functional entities form the principal sources of 5GC-domain mobility information, as summarized in Table I.

B. *Mobility-Related Information in RAN*

In contrast to the 5GC, the RAN domain provides fine-grained and almost real-time information derived from direct radio measurements. This domain emphasizes short-term dynamics of UE behavior, such as mobility history, handover performance, and radio link quality. According to 3GPP TS 38.300 and TS 38.331, the gNB and associated nodes are responsible for delivering such detailed mobility information, which can be leveraged for AI/ML-based prediction and optimization, as summarized in Table II.

C. *Unified Framework for 5GC and RAN Mobility Information*

While the 5GC and RAN individually provide valuable but distinct types of mobility information, their integration yields a more complete and actionable view of UE behavior. The 5GC offers broad and often estimated mobility data, such as location updates and access behavior, whereas the RAN contributes instance measurements, including radio link quality and handover events. When these two data sources are combined, the resulting dataset enables cross-domain correlation that can improve mobility prediction accuracy, anomaly detection, and resource optimization. This integrated perspective is particularly aligned with recent 3GPP directions, where cross-domain data analytics is increasingly emphasized to support AI/ML-driven mobility management.

TABLE II: RAN-domain information relevant to mobility (3GPP TS 38.300/38.331; ref. TR 37.817 for AI/ML inputs)

Category	**Information (examples aligned with TS 38.331 RRC measurements/events)**	**Source**
Input	UE location context (serving cell ID, PCI/beam info), UE speed/velocity estimate	UE / gNB
Input	Radio measurements for serving/neighbor cells (RSRP, RSRQ, SINR), beam quality, CSI	UE (RRC MeasReport)
Input	Measurement events (e.g., A1–A6, conditional HO triggers), time-to-trigger, offsets	UE $\rightarrow$ gNB (RRC)
Input	UE mobility history (recent serving cells / TA updates observed at RAN side)	gNB
Input	Neighbor-cell context and historical KPIs (HO success/failure, block error, delay)	Neighbor gNBs
Input	Position/QoS/performance of historically handovered UEs around the target area	Neighbor gNBs
Input	Current/predicted RAN load (PRB utilization, buffer, scheduling backlog)	gNB / OAM
Input	UE handover logs (success/rollback/failure causes, radio link failure counters)	gNB
Input	Trajectory estimates (short-term path, likely next cell/beam)	Local analytics
Input	Predicted resource status (RB availability per cell/beam, expected interference)	Local analytics
Input	Current/predicted UE traffic profile (flow count/bitrate/burstiness)	Local analytics

IV. METHODOLOGY

The UE mobility traces and base station information were obtained from Köln dataset [6], which provides timestamped UE trajectories in the form `(time, ue_id, x, y, speed)` and base station data in the form `(bs_id, x, y)`. Each UE entry specifies the 2 dimensional position and instantaneous speed of an UE at a given simulation time, enabling accurate reconstruction of large scale mobility patterns within the city of Köln. In the simulation, gNBs are configured with unique identifiers `bs_id` and fixed deployment coordinates `x,y`, while UEs follow time varying trajectories directly derived from the trace dataset. This ensures that UE positions evolve realistically over time, rather than relying on synthetic mobility models.we set up the simulation environment according to the 5G standalone specifications. Each UE is equipped with four antennas, two for transmission and two for reception, with a transmission power of 23 dBm and height of UE 1.5m. On the other hand gNB is configured with height of 10m and a transmission power of 40 dBm. It employs a total of 12 antennas, comprising eight transmit and four receive antennas, and utilizes an omnidirectional antenna configuration. The simulation is carried out in NetSim, which supports 5GC functions such as AMF, SMF, UPF. However, NetSim does not support LCS and its related function LMF, which led to build our own.

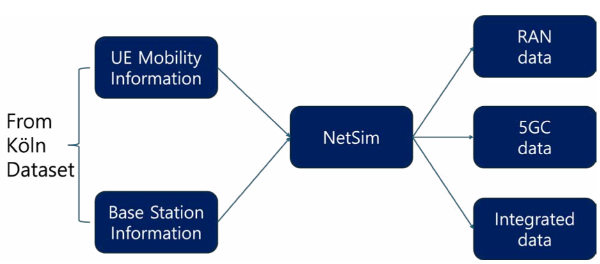

Fig. 2: Schematic representation of the proposed workflow.

From this setup, logs are collected from three perspectives: (i) RAN domain logs including RSRP, cell associations, and handover triggers, (ii) 5GC domain logs including registration, session management, and UE state transitions, and (iii) an integrated dataset that contains RAN and core domains characteristics. Figure 2 represents the integrated process flow, where BS and UE mobility traces from the Köln dataset are applied to NetSim. The resulting simulation generates logs corresponding to the RAN, the 5GC, and the integrated domain. The geometric distribution of the 246 gNB sites is represented through a Voronoi tessellation, as shown in Figure 3a. In this construction, each polygon corresponds to the coverage area of a gNB, determined by the locus of points that are closer to that gNB than to any other. This representation provides an abstraction of cell boundaries under ideal isotropic propagation conditions. Although it does not account for realistic channel effects such as fading or shadowing, it serves as a useful baseline for mobility and handover analysis, and can be cross-validated against NetSim-derived coverage logs.

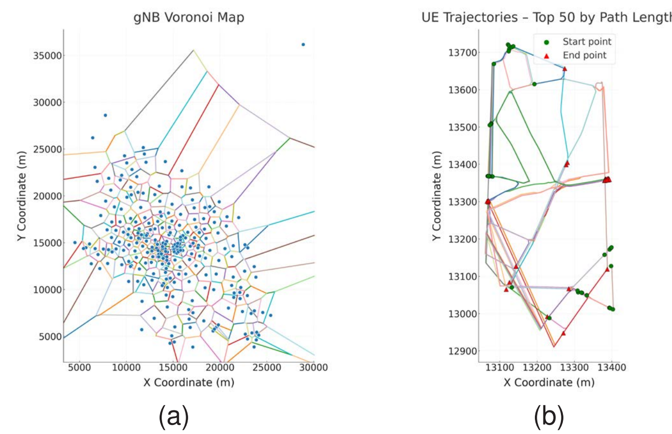

Fig. 3: (a) Voronoi tessellation of gNB sites, (b) UE trajectories within the simulation area.

As illustrated in Fig. 3b, the trajectories of UE are distributed within a simulation area of $800\,\mathrm{m} \times 500\,\mathrm{m}$. Each UE follows a file-based mobility setup from Köln dataset. Among all simulated UEs, the top 50 devices with the longest travel distances are highlighted, demonstrating their higher mobility and frequent transitions across multiple cells. These highly mobile UEs are particularly important since they generate more frequent handovers and contribute significantly to network signaling loads. Figure 4 illustrates the NetSim graphical user interface (GUI) where the simulation scenario is executed.

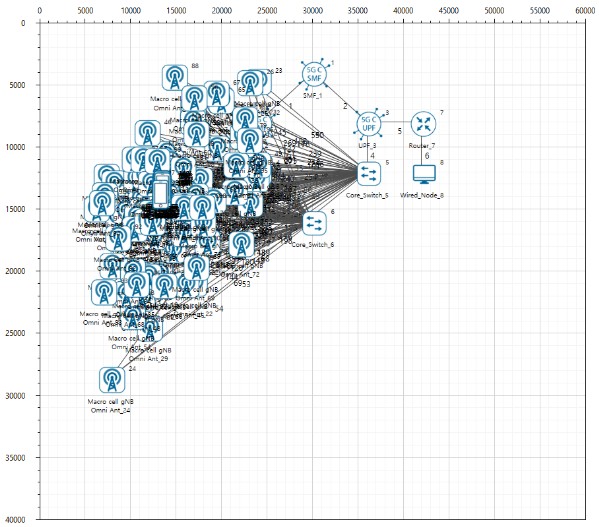

Fig. 4: NetSim GUI screen with 500 active nodes.

A large number of gNBs and UEs are deployed by importing XML-based configuration files, enabling flexible control over node placement and simulation parameters. The example shown in the figure contains 500 active nodes including 246 base stations. Figure 5 illustrates a representative handover log produced by NetSim. Each entry captures the UE identifier, serving gNB, target gNB, and the event time at which the handover is triggered. The log also incorporates the Time-to-Trigger (TTT) parameter, defined in 3GPP TS 38.331, which specifies the minimum duration that the handover condition (e.g., RSRP difference exceeding the event A3 offset plus hysteresis) must be satisfied before execution. This prevents unnecessary "ping-pong" handovers caused by short-term fading. In addition, the log captures the Handover Interruption Time (HIT), which represents the service disruption experienced by the UE during the transition from the serving to the target gNB. HIT is a critical performance indicator since it directly affects latency-sensitive applications. By logging both TTT and HIT, the dataset aligns with standard-compliant handover definitions and provides quantitative insight into mobility robustness. These logs form the foundation for later integration with 5GC and LCS data, enabling cross-domain mobility prediction and analytics. Figure 6 example event trace obtained from NetSim, showing system-level logging of protocol activities across different network entities.

V. Conclusion

This work demonstrated that the extraction and analysis of RAN, 5GC, and integrated domain data through NetSim constitute a comprehensive framework for evaluating UE mobility. Unlike domain-specific studies, the integrated-domain perspective captures the interplay between radio conditions, session management, and end-to-end QoS. Our results suggest that this holistic approach not only aligns with the goals of 3GPP Releases 17 and 18, but also offers practical advantages for designing more resilient mobility management strategies. The significance of this study lies in demonstrating that multi-

Time(ms)	UE Id	Serving Cell	Target Cell	Serving SSB SNR(dB)	Target SSB SNR(dB)	Remarks
171	UE_193366	MACRO CELL GNB OMNI ANT_9	MACRO CELL	4.626742	24.557729	Handover Initiated
171	UE_1242881	MACRO CELL GNB OMNI ANT_9	MACRO CELL	17.250621	22.637664	Handover Initiated
171	UE_1559770	MACRO CELL GNB OMNI ANT_9	MACRO CELL	15.630406	24.901041	Handover Initiated
171	UE_1583497	MACRO CELL GNB OMNI ANT_9	MACRO CELL	10.117265	16.717494	Handover Initiated
171	UE_1308	MACRO CELL GNB OMNI ANT_28	MACRO CELL	19.400608	22.599605	Handover Initiated
171	UE_1016964	MACRO CELL GNB OMNI ANT_28	MACRO CELL	19.090596	28.974148	Handover Initiated
171	UE_1030954	MACRO CELL GNB OMNI ANT_33	MACRO CELL	3.979584	8.798052	Handover Initiated
171	UE_1494584	MACRO CELL GNB OMNI ANT_36	MACRO CELL	7.607872	14.724864	Handover Initiated
171	UE_1508533	MACRO CELL GNB OMNI ANT_36	MACRO CELL	8.82388	26.55115	Handover Initiated
171	UE_1398234	MACRO CELL GNB OMNI ANT_56	MACRO CELL	10.933702	18.870576	Handover Initiated
171	UE_11959	MACRO CELL GNB OMNI ANT_59	MACRO CELL	4.629557	13.098483	Handover Initiated
171	UE_203273	MACRO CELL GNB OMNI ANT_75	MACRO CELL	21.587562	24.746342	Handover Initiated
171	UE_1075695	MACRO CELL GNB OMNI ANT_75	MACRO CELL	15.726121	25.711707	Handover Initiated
171	UE_1167440	MACRO CELL GNB OMNI ANT_75	MACRO CELL	21.912581	32.217371	Handover Initiated
171	UE_1207508	MACRO CELL GNB OMNI ANT_75	MACRO CELL	17.094848	20.352804	Handover Initiated
171	UE_1376892	MACRO CELL GNB OMNI ANT_75	MACRO CELL	13.944662	22.362785	Handover Initiated
171	UE_1509319	MACRO CELL GNB OMNI ANT_75	MACRO CELL	20.238579	24.275922	Handover Initiated
206	UE_1509319	MACRO CELL GNB OMNI ANT_28	MACRO CELL	15.847923	20.238579	Handover Initiated
206	UE_1559770	MACRO CELL GNB OMNI ANT_28	MACRO CELL	8.61283	18.292857	Handover Initiated
206	UE_1308	MACRO CELL GNB OMNI ANT_36	MACRO CELL	13.259751	18.08434	Handover Initiated
206	UE_203273	MACRO CELL GNB OMNI ANT_36	MACRO CELL	-2.352197	19.212029	Handover Initiated
206	UE_1376892	MACRO CELL GNB OMNI ANT_36	MACRO CELL	17.860178	26.662153	Handover Initiated
206	UE_1398234	MACRO CELL GNB OMNI ANT_76	MACRO CELL	-6.291654	0.392691	Handover Initiated
206	UE_1030954	MACRO CELL GNB OMNI ANT_82	MACRO CELL	-13.146829	2.55234	Handover Initiated
206	UE_11959	MACRO CELL GNB OMNI ANT_97	MACRO CELL	10.621795	20.815473	Handover Initiated

Fig. 5: Mobility-predictive data: NetSim handover log.

Event_Id	Event_Type	Event_Time(초)	Device_Type	Device_Id	Interface_Id	Application_Id	Packet_Id	Segment_Id	Protocol_Name	Subevent_Type	Packet_Size(Bytes)	Prev_Event_Id
1	TIMER_EVENT	0	AMF	2	2	0	0	0	ETHERNET	ETH_IF_UP	0	0
2	TIMER_EVENT	0	UPF	3	2	0	0	0	ETHERNET	ETH_IF_UP	0	0
3	TIMER_EVENT	0	SWITCH	4	1	0	0	0	ETHERNET	ETH_IF_UP	0	0
4	TIMER_EVENT	0	SWITCH	4	2	0	0	0	ETHERNET	ETH_IF_UP	0	0
5	TIMER_EVENT	0	SWITCH	4	3	0	0	0	ETHERNET	ETH_IF_UP	0	0
6	TIMER_EVENT	0	SWITCH	4	4	0	0	0	ETHERNET	ETH_IF_UP	0	0
7	TIMER_EVENT	0	SWITCH	4	5	0	0	0	ETHERNET	ETH_IF_UP	0	0
8	TIMER_EVENT	0	SWITCH	4	6	0	0	0	ETHERNET	ETH_IF_UP	0	0
9	TIMER_EVENT	0	SWITCH	4	7	0	0	0	ETHERNET	ETH_IF_UP	0	0
10	TIMER_EVENT	0	SWITCH	4	8	0	0	0	ETHERNET	ETH_IF_UP	0	0
11	TIMER_EVENT	0	SWITCH	4	9	0	0	0	ETHERNET	ETH_IF_UP	0	0
12	TIMER_EVENT	0	SWITCH	4	10	0	0	0	ETHERNET	ETH_IF_UP	0	0
13	TIMER_EVENT	0	SWITCH	4	11	0	0	0	ETHERNET	ETH_IF_UP	0	0
14	TIMER_EVENT	0	SWITCH	4	12	0	0	0	ETHERNET	ETH_IF_UP	0	0
15	TIMER_EVENT	0	SWITCH	4	13	0	0	0	ETHERNET	ETH_IF_UP	0	0
16	TIMER_EVENT	0	SWITCH	4	14	0	0	0	ETHERNET	ETH_IF_UP	0	0
17	TIMER_EVENT	0	SWITCH	4	15	0	0	0	ETHERNET	ETH_IF_UP	0	0
18	TIMER_EVENT	0	SWITCH	4	16	0	0	0	ETHERNET	ETH_IF_UP	0	0
19	TIMER_EVENT	0	SWITCH	4	17	0	0	0	ETHERNET	ETH_IF_UP	0	0
20	TIMER_EVENT	0	SWITCH	4	18	0	0	0	ETHERNET	ETH_IF_UP	0	0
21	TIMER_EVENT	0	SWITCH	4	19	0	0	0	ETHERNET	ETH_IF_UP	0	0
22	TIMER_EVENT	0	SWITCH	4	20	0	0	0	ETHERNET	ETH_IF_UP	0	0
23	TIMER_EVENT	0	SWITCH	4	21	0	0	0	ETHERNET	ETH_IF_UP	0	0
24	TIMER_EVENT	0	SWITCH	4	22	0	0	0	ETHERNET	ETH_IF_UP	0	0
25	TIMER_EVENT	0	SWITCH	4	23	0	0	0	ETHERNET	ETH_IF_UP	0	0

Fig. 6: Mobility-predictive data: NetSim event trace log.

domain insights can transform the way mobility is managed in future networks. By revealing dependencies between RAN metrics and 5GC behavior, our methodology enables operators and researchers to identify bottlenecks that would remain hidden in isolated analyses. In the future, our focus will be on building AI-driven mobility prediction models that leverage the extracted datasets.

Acknowledgement

This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the National Program for Excellence in SW, supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation) in 2025 (2024-0-00073).

References

[1] H. Rydén, A. Palaios, L. Hévizi, D. Sandberg, T. Kvernvik, and H. Farhadi, "Mobility, traffic and radio channel prediction: 5G and beyond applications," *arXiv preprint arXiv:2203.08047*, 2022.

[2] J. Jeong, D. Roeland, J. Derehag, A. Johansson, V. Umaashankar, G. Sun, and G. Eriksson, "Mobility prediction for 5G core networks," *IEEE Communications Standards Magazine*, vol. 5, no. 1, pp. 56–61, 2021.

[3] A. Rago, P. Ventrella, G. Piro, G. Boggia, and P. Dini, "Towards an optimal management of the 5G cloud-RAN through a spatio-temporal prediction of users' dynamics," in *2020 Mediterranean Communication and Computer Networking Conference (MedComNet)*, pp. 1–4, 2020.

[4] A. G. Bermudez, M. Farreras, M. Groshev, J. A. Trujillo, I. de la Bandera, and R. Barco, "Graph neural networks for o-ran mobility management: A link prediction approach," *arXiv preprint arXiv:2502.02170*, 2025.

[5] X. Lin, "An overview of 5G advanced evolution in 3GPP Release 18," *IEEE Communications Standards Magazine*, vol. 6, no. 3, pp. 77–83, 2022.

[6] S. Uppoor, O. Trullols-Cruces, M. Fiore, and J. M. Barcelo-Ordinas, "Generation and analysis of a large-scale urban vehicular mobility dataset," *IEEE Transactions on Mobile Computing*, vol. 13, no. 5, pp. 1061–1075, 2014.