A Large-Scale Dataset of 4G, NB-IoT, and 5G Non-Standalone Network Measurements

Konstantinos Kousias*, Mohammad Rajiullah†, Giuseppe Caso‡, Usman Ali¶, Ozgu Alay§, Anna Brunstrom†, Luca De Nardis¶, Marco Neri‖, and Maria-Gabriella Di Benedetto¶
*Simula Research Laboratory, Norway
†Karlstad University, Sweden
‡Ericsson Research, Sweden
§University of Oslo, Norway
¶Sapienza University of Rome, Italy
‖Rohde & Schwarz, Italy

Abstract

Mobile networks have become highly complex systems. In order to better understand how network features affect performance and suggest additional improvements, it is crucial to examine them from an empirical perspective. This paper presents a large-scale dataset of measurements collected over fourth generation (4G) and fifth generation (5G) operational networks, providing Long Term Evolution (LTE), Narrowband Internet of Things (NB-IoT) and 5G New Radio (NR) connectivity. We collected our dataset during a period of seven weeks in Rome, Italy, by performing several tests on the infrastructures of two major mobile network operators (MNOs). The open-sourced dataset has enabled multi-faceted analyses of network deployment, coverage, and end-user performance, and can be further used for designing and testing artificial intelligence (AI) and machine learning (ML) solutions for network optimization tasks.

Feature	Description
`Date`, `Time`, `UTC`	Temporal fields.
`Latitude`, `Longitude`, `Altitude`	Spatial fields; user equipment (UE) Global Positioning System (GPS) coordinates [in °].
`cellLatitude*`, `cellPosErrorLambda1*`, `cellLongitude*`, `cellPosErrorLambda2*`	Spatial fields; Cell GPS coordinates [in °] along with estimated positioning error margins [in m].
`Speed`	Moving speed [in km/h].
`Frequency`, `EARFCN`, `MNC`, `Band`	Mobile network information; Carrier frequency, E-UTRA Absolute Radio Frequency Channel Number (EARFCN), mobile network code (MNC), and band identifiers.
`PCI`, `cellIdentity*`, `eNodeB.ID*`, `SSBIdx`	Cell site information; Physical Cell ID (PCI), cell ID (CID), E-UTRAN Node B (eNB), and Synchronization Signal Block (SSB) beam identifiers.
`RSRP`$\hat{*}$, `RSRQ`$\hat{*}$, `SINR`$\hat{*}$, `Power`$\hat{*}$	Radio coverage information; Reference Signal Received Power (RSRP)$\hat{*}$ [in dBm], Reference Signal Received Quality (RSRQ)$\hat{*}$ [in dB], Signal to Interference and Noise Ratio (SINR)$\hat{*}$ [in dB], and received Power$\hat{*}$ [in dBm].
`scenario`, `campaign`	Campaign identifiers.
`n_CellIdentities*`	Number of cells per eNB.
`distance*`	Distance between the UE and a cell (assuming line of sight (LoS)) [in m].

* Only available for fourth generation (4G) and Narrowband Internet of Things (NB-IoT).
$\hat{*}$ The NB-IoT dataset provides such features for two antenna ports, referred to as `Tx1` and `Tx2`, while the fifth generation (5G) dataset covers different types of 5G signals, i.e., Secondary Synchronization Signal (SS), Demodulation Reference Signal (DMRS) Reference Signal (RS), Physical Broadcast Channel (PBCH), Primary Synchronization Signal (PSS), and SS-PBCH (e.g., `SS-RSRP` and `PSS-RSRP`).

TABLE I: Passive dataset features along with a short description.

Feature	Description
Date, Time	Temporal fields.
GPS Long, GPS Lat	UE GPS coordinates [in °].
X PCI	Physical cell identity of the serving cell; X={5G, LTE}
(SS-)RSRP, (SS-)RSRQ, (SS-)SINR	Radio coverage information; Received power (RSRP* [in dBm]), Signal quality (RSRQ* [in dB]), and Signal-to-noise ratio (SINR* [in dB]) of the serving cell. The SS (Synchronization Signal) prefix is used in 5G.
RAT Info	UE Connectivity; 5G EN-DC: the UE is connected to a 5G cell, Long Term Evolution (LTE): the UE is connected to a 4G cell.
X Y Throughput	X throughput measured at Y carrier [kb/s]; X={5G, LTE}, Y={PDSCH, PUSCH}. PDSCH stands for Physical Downlink Shared Channel and PUSCH stands for Physical Uplink Shared Channel.
X Y Z Rate	Number of transmission time intervals (TTIs) that used Z as modulation divided by the total number of TTIs associated to the UE [as %]. This feature is available for Y and Z; X={5G, LTE}, Y={PDSCH, PUSCH}, Z={QPSK, 16QAM, 64QAM, 256QAM}. Please note that 256QAM is not used for LTE PUSCH.
5G X Y MCS	Indication of the Modulation and Coding Scheme (MCS) index (between 0 and 31) assigned to the UE during the observation period. This feature is available as X for Y; X={Min, Avg, Max}, Y={PDSCH, PUSCH}
LTE X MCS Level	MCS index (between 0 and 31) in X direction. Please note that for DL, MCS is taken at the first antenna; X={DL, UL}
5G Serving SSB Index	ID of the serving SSB beam (when the UE is connected to a 5G cell with SSB beamforming).
5G # SSB Beams	Number of detected SSB beams of the serving 5G cell.
5G X Num Y RB	X number of resource blocks (RB) in TTIs associated to the UE for Y carrier; X={Min, Avg, Max}, Y={PDSCH, PUSCH}
LTE X RB Y	Y number of RB used in the observation period for X carrier; X={PDSCH, PUSCH}, Y={Min, Avg, Max}
5G X Y TBS	X Transport Block Size (TBS) in the observation period [bytes] for Y; X={Min, Avg, Max}, Y={PDSCH, PUSCH}
LTE X Y TBS	X TBS transmitted in the TTIs associated to the UE over all Component Carriers [bits] for Y; X={Min, Avg, Max}, Y={PDSCH, PUSCH}
LTE-CA	Number of configured Secondary Component Carriers (SCCs). LTE-nCA denotes that n carriers (including the primary component carrier (PCC)) are configured.
LTE-CA Type	Type of carriers, for example, LTE or licensed assisted access (LAA).
DQA Result	Data Quality Analyser (DQA) data result which can be one of: Good Session; Blocked Session; Dropped Session; Application Error; No Service.
Current netw. X Y Throughput*	Current Y Speedtest throughput during a session [kb/s] for X; X ={DL, UL}, Y= {Instantaneous, Avg, Max}
Mean network X Throughput*	Mean Speedtest throughput [kb/s] for X; X ={DL, UL}
Num. Sent Packets*	Total number of packets sent during a session.
Num. Lost Packets*	Total number of missing packets (i.e., packets lost or exceeding the latency budget).
Channel QoS 3GPP*	Channel quality based on the packet error rate calculated from Num. Lost Packets and Num. Sent Packets [%].
X RTT*	X Round trip time (RTT) of the packets sent during a session [seconds]; X= {median, 10th percentile values}
X PDV*	X Packet delay variation (PDV) of the packets sent during a session evaluated according to RFC 5481 [seconds]; X= {median, 99th percentile values}
Interactivity score*	Interactivity score (*i-score*) evaluated as a function of RTT, PDV, and packet error rate [%] (see [**?**]).
UE Mode	UE setting during the measurement; 5G-enabled or 5G-disabled.
Scenario	Mobility scenarios - Indoor static (IS), outdoor walking (OW), or outdoor driving (OD)
Operator	Anonymized MNO; Op_1 or Op_2.
Campaign	Sub-campaign identifier

* quality of service (QoS) features for Speedtest.
* QoS features for real-time online gaming. The dataset contains values measured at the end and during each session. Feature attributes are prefixed with Cur. in the latter case.

TABLE II: Active dataset features along with a short description.