# Beyond Throughput: The Next Generation a 5G Dataset with Channel and Context Metrics

Darijo Raca
Faculty of Electrical Engineering, University of Sarajevo, Sarajevo, BiH
draca@etf.unsa.ba

Dylan Leahy, Cormac J. Sreenan
Jason J. Quinlan
School of Computer Science and Information Technology, University College Cork, Ireland
{cjs,j.quinlan}@cs.ucc.ie

## ABSTRACT

In this paper, we present a 5G trace dataset collected from a major Irish mobile operator. The dataset is generated from two mobility patterns (static and car), and across two application patterns (video streaming and file download). The dataset is composed of client-side cellular key performance indicators (KPIs) comprised of channel-related metrics, context-related metrics, cell-related metrics and throughput information. These metrics are generated from a well-known non-rooted Android network monitoring application, G-NetTrack Pro. To the best of our knowledge, this is the first publicly available dataset that contains throughput, channel and context information for 5G networks. To supplement our real-time 5G production network dataset, we also provide a 5G large scale multi-cell ns-3 simulation framework. The availability of the 5G/mmwave module for the ns-3 mmwave network simulator provides an opportunity to improve our understanding of the dynamic reasoning for adaptive clients in 5G multi-cell wireless scenarios. The purpose of our framework is to provide additional information (such as competing metrics for users connected to the same cell), thus providing otherwise unavailable information about the eNodeB environment and scheduling principle, to end user. Our framework, permits other researchers to investigate this interaction through the generation of their own synthetic datasets.

## CCS CONCEPTS

• **Information systems → Multimedia streaming**; • **Networks** → Public Internet; **Wireless access networks**; **Network measurement**.

## KEYWORDS

Dataset, 5G, NR, Mobility, throughput, context information, adaptive video streaming, mmwave

**ACM Reference Format:**
Darijo Raca, Dylan Leahy, Cormac J. Sreenan, and Jason J. Quinlan. 2020. Beyond Throughput: The Next Generation a 5G Dataset with Channel and Context Metrics. In *MMSys'20 June 08–11, 2020, Istanbul, Turkey*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn



*MMSys'20, June 08–11, 2020, Istanbul, Turkey*

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

## 1 INTRODUCTION

From the first-generation, voice-only, mobile cellular communication systems to the current fourth generation system (4G - Long Term Evolution - LTE), development has progressed at a steady pace. While this evolution is typically fuelled by services and applications utilising the network, the current bandwidth demands on the 4G network was never anticipated [1]. Applications utilising social media, gaming, and recent advances in Augment/Virtual Reality, has accelerated the demands for the next, fifth generation (*5G*), of the cellular communication standard. 5G hold the promise of ubiquitous connection with vastly improved connectivity: high data rates (10x increase compared to “traditional” 4G network) and low latency (10x lower compared to a 4G network). In addition to the high rates (around 1Gbps) and low latency (1ms), 5G provides connectivity for tens of thousands of devices in order to support future IoT and Internet of Vehicle (*IoV*) paradigms. These enhancements require novel solutions in core network architecture and radio interface design [6].

The two most significant factors driving the development of next generation cellular standards is the rapid increase in the number of connected devices and the unrivalled rise in multimedia traffic, and as a direct result their increased throughput demands. Predictions for the number of connected devices by 2025 vary. The largest prediction being that the number of connected Internet of Things(*IoT*) devices is expected to reach 75.44 billion [2]. At the heart of this growth in throughput demand is video traffic, carried through different applications, from Video on Demand (*VoD*), live streaming and 360-degree video. Current streaming platforms utilise the HTTP adaptive streaming (HAS) technique [17] for video delivery. HAS allows graceful adaptation of video quality during the playback through the segmentation of video content. New video compression standards (H.265/HEVC) and ultra-high definition resolutions (e.g., 8K) have high bandwidth requirements [13]. These requirements are further exacerbated in 360-degree videos. For example, 24K 360-degree video with 120 frame-per-second can consume several Gbps. However, high bandwidth demand is not the only constraint. User can change field-of-view at any time. For a user to not experience motion sickness, latency during the transition needs to be less than 20ms [18]. While 5G can sustain these demands, it is yet to be proven that high-rate low-latency can be constantly supported in real networks. To support analysis of video performances, new datasets containing bandwidth and latency information, collected in production 5G networks, are needed.

In this paper, we present two datasets: the first collected from real 5G production network and the second synthetic dataset generated from a large-scale multi-cell 5G/mmwave ns-3 [3] framework. Our production dataset is similar to our previously collected dataset in 4G networks [15]. We collected traces from a major Irish operator with two mobility patterns, driving and static. Furthermore, we extend our download strategy beyond file download, running the same scenarios while streaming video content using Amazon Prime and Netflix streaming services. In addition to throughput values, the dataset contains information about latency, channel conditions (e.g., signal strength), user location (i.e., GPS coordinates) and more (see Section 4 for details). These metrics allow multi-purpose analysis including a comparison of different HAS approaches, handover prediction, coverage analysis, mobility prediction etc. To the best of our knowledge, this is the first publicly available dataset that contains throughput, channel and context information for 5G networks.

The remainder of this paper is organised as follows. Section 2 describes related work. The 5G production dataset collection and captured metrics are explained in Section 3, while Section 4 explores statistical traits of the production dataset for different mobility patterns. In Section 5, we present our 5G/ns-3 simulation framework and offer details on configuration, structure and illustrate sample outputs of the synthetic dataset. In Section 6 we outline possible use cases, while Section 7 concludes the paper.

## 2 RELATED WORK

Prior dataset in this area, were collected using 3G and 4G network technologies [4, 10, 15, 16, 19, 20]. These datasets typically contain throughput information logged in a range of timescales (from one to several seconds), and across a multitude of mobility patterns, including static, pedestrian, bus, train, ferry and car. Throughput information can be beneficial when evaluating the performance at the application layer, such as required by HAS algorithms in video streaming during the optimisation of video delivery in *rate-based* [9], *buffer-based* [7], and *hybrid* [5] schemes. Some of these datasets [4] were collected multiple times over the same route to get statistically significant results (as the network throughput can vary significantly over the same route). In addition to throughput performance [10, 15, 20], some datasets contain information about the channel (e.g., signal strength), context (e.g., GPS of the device, device velocity, eNBs ID), which can be beneficial when evaluating mobility patterns during handover. There are also many video streaming approaches that leverage information beyond throughput to make more intelligent decisions for the next chunk quality improving video QoE [14, 21, 22]. Recently, Narayanan et al. [12] conducted the first 5G measurement study of commercial 5G network in the U.S (Verizon). The authors only collected throughput and latency information and compared their results with the Verizon 4G network. Their log dataset which was not released, consists of UE location and IP, and, eNB ID and signal strength.

## 3 PRODUCTION DATASET GENERATION

For the collection of the 5G production dataset we utilise version 18.7 of the G-NetTrack Pro mobile network monitoring tool[1] for Android devices. The application is installed on a Samsung S10 5G Android device. G-NetTrack Pro permits the collection of multiple channel-related metrics, context-related metrics, cell-related metrics and throughput information (uplink and downlink) using the standard Android library. G-NetTrack Pro works across a range of Android devices and does not require rooted privileges. Some of the limitations include minimum one-second granularity for the channel metrics (this limitation comes from the Android API itself) and a non-unified capability of measuring all the metrics across the different mobile "system on a chip" (*SoC*) chipsets manufacturers. Implementation of callback methods for reporting channel values depends on the SoC manufacturer. Luckily, Samsung S10 5G (with *Exynos* chipset) provides a means of capturing all the 5G channel metrics (note: at the time of collection, we were limited by the choice of 5G supported mobile devices supported by the Irish mobile operator used for the dataset generation).

Our production dataset[2], consists of 83 traces, with a total duration of 3142 minutes. The mobile plan offered by the Irish mobile operator includes a fair use limit of 80GB data per month, before the download rate is reduced. The strategy for collecting the 5G data is as follows; for each combination of application (file download, Netflix, Amazon Prime) and mobility pattern (static, driving), we run experiments until all data is consumed per month. This leads to a limited number of traces. For example, we only capture four bandwidth traces in a static scenario with large file download. However, the total number of minutes for the static scenario is 160 minutes. This is intuitive, as file download case produces the highest throughput values and consumes data very quickly. These large duration traces can be split into shorter duration's, depending on the needs of an experimentation setup (typically, most of the video-related experiments considered in the literature utilise up to five minutes of bandwidth traces).

For file download trial, we use a large file (> 200MB) to allow the TCP sending window to ramp up to the maximum size. As stated, every sample is logged with one-second granularity. For Netflix and Amazon Prime we stream animated (circa 200m) and live-action (circa 400m) video content, while running G-NetTrack Pro application in the background collecting bandwidth and channel samples.

The following outlines the metrics within our production dataset:

- *Timestamp:* timestamp of sample
- *Longitude and Latitude:* GPS coordinates of mobile device
- *Velocity:* velocity in kph of mobile device
- *Operatorname:* cellular operator name (anonymised)
- *CellId:* Serving cell for mobile device
- *NetworkMode:* mobile communication standard (2G/3G/4G/5G)
- *DL_bitrate:* download rate measured at the device (application layer) (kbps)
- *UL_bitrate:* uplink rate measured at the device (application layer) (kbps)
- *State:* state of the download process. It has two values, either I (idle, not downloading) or D (downloading)
- $Ping_{avg}$, $Ping_{min}$, $Ping_{max}$, $Ping_{std}$, $Ping_{loss}$: ping statistics (average, minimum, maximum, standard deviation and loss)
- *RSRQ:* value for RSRQ. RSRQ Represents a ratio between RSRP and Received Signal Strength Indicator (RSSI). Signal strength (signal quality) is measured across all resource elements (RE), including interference from all sources (dB).
- *SNR*: value for signal-to-noise ratio (dB).

[1] http://www.gyokovsolutions.com/

[2] http://cs1dev.ucc.ie/misl/5Gframework/5G-production-dataset.zip

- *RSRP:* value for RSRP. RSRP Represents an average power over cell-specific reference symbols carried inside distinct RE. RSRP is used for measuring cell signal strength/coverage and therefore cell selection (dBm).
- *RSSI:* value for RSSI. RSSI represents a received power (wideband) including a serving cell and interference and noise from other sources. RSRQ, RSRP and RSSI are used for measuring cell strength/coverage and therefore cell selection (handover) (dBm).
- *CQI:* value for CQI of a mobile device. CQI is a feedback provided by UE to eNodeB. It indicates data rate that could be transmitted over a channel (highest MCS with a BLER probability less than 10%), as the function of SINR and UE's receiver characteristics. Based on UE's prediction of the channel, eNodeB selects an appropriate modulation scheme and coding rate.
- *NRxRSRQ & NRxRSRP:* RSRQ and RSRP values for the neighbouring cell.

Table 1 summarises the mobility patterns used to generate the dataset:

**Table 1: Mobility Patterns**

| Type | Summary |
|---|---|
| Static | Static trials (indoor and in car scenarios) |
| Car | Trials include urban and suburban scenarios |

In conjunction with the different mobility patterns, different download approaches were taken, and these are summarised in Table 2:

**Table 2: Application Patterns**

| Type | Summary |
|---|---|
| File Download | Continuous large file download |
| Netflix | Netflix service provider streamed video content |
| Amazon Prime | Amazon Prime service provider streaming video content |

# 4 PRODUCTION DATASET OVERVIEW

Due to page limitations, this section gives a short overview of the production dataset. The dataset is composed from two mobility patterns, *static* and *car*. Static traces were collected indoors and in-car, with the mobile device in a stationary position. Unlike the *static* car scenario, the *car* mobile scenario included driving through city and the suburban areas. The majority of the cases were collected during the morning and evening hours and can be further classified as commute traces.

## 4.1 4G vs. 5G

We start by comparing the throughput of traces collected over the 4G and 5G technologies. For 4G, we use our previously collected dataset [15]. To offer a fair comparison, we only compare traces from the same mobile operator and with the same mobility patterns. Furthermore, only the scenario with file download is compared across two mobile technologies. We use average throughput and variation range as performance metrics. Variation range is a percentile-wise measure of variation. Let's define $R$ as application throughput during time interval the $(t, t+1)$. Then variation range is defined as the interval $[R^L, R^H]$, where $R^L$ represents a $10^{th}$ percentile of $R$, and analogously $R^H$ a $90^{th}$ percentile of $R$ [8]. This range defines boundaries where 80% of measured throughput lies.

Table 3 shows performance metrics for the 4G and 5G. As expected, 5G allows higher rates, with a 50% increase for the average throughput for the static scenario. This observation is further supported by variation range, where the upper limit for throughput is 202Mbps for the 5G, almost 3x higher than that of 4G. However, this difference is less evident in the case of the car scenario. The average throughput increased by 27% for 5G compared to 4G. The main reason for "minor" improvement is lack of 5G base stations across all driving routes, forcing the device to use 4G. However, even with this limit, upper limit for variation range is still almost 2x higher for 5G than 4G. While the mentioned metrics gave the expected performance, the interesting values are peak rates that were observed during the collection. In the case of the static scenario, the maximum observed 5G throughput is 333Mbps. This is an increase of 3x times compared to the same 4G scenario (peak rate 97Mbps). The difference is even more evident for the driving scenario, where the 5G-supported device achieved a rate of 532Mbps, 5x times larger than 4G (peak rate 108Mbps).

## 4.2 File download vs. streaming

Table 4 shows a comparison between performance metrics for different application types. Intuitively, continuous file download has the highest average throughput and variation range. Netflix and Amazon Prime consume significantly less bandwidth, as seen in Figure 1, and is a consequence of application behaviour. Figure 1 depicts a boxplot illustrates the relationship between CQI and application throughput and shows the range of throughput values for each CQI separately. Overall, we observe an increasing trend in throughput proportional to CQI. However, the range of throughput values oscillates significantly for each CQI. For Amazon Prime, lower bitrates result in similar throughput rates across all CQI values. Streaming services download segments only during the ON phase (buffer filling). Also, bandwidth demand is limited by the maximum quality of encoded video content. Overall, Netflix consumes significantly more bandwidth than Amazon Prime for both mobility patterns, as a result of the higher encoding quality and thus larger segment sizes. Next, we analyse the collected traces latency performance. For the static and driving scenarios, the average latency is 75 and 90ms, respectively. This performance is much higher than the targeted 1ms, which is expected to be achieved as the technology matures.

# 5 5G/MMWAVE SIMULATION FRAMEWORK

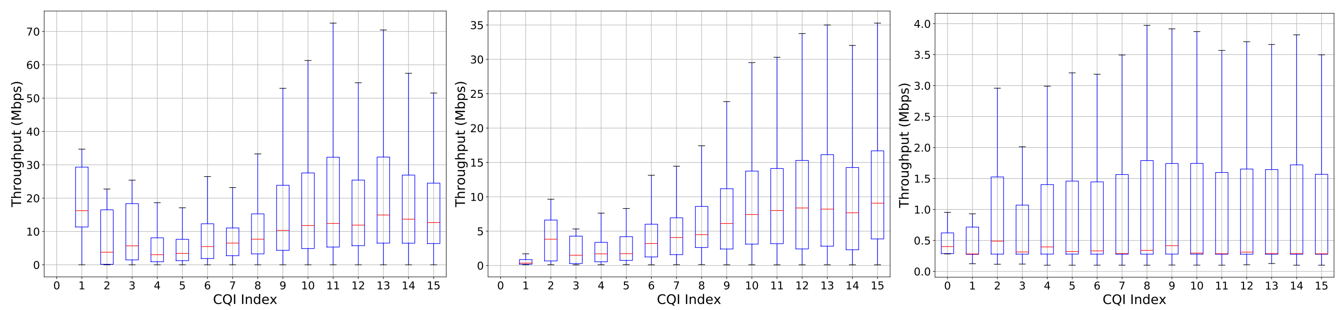
In [15], we presented our previous work which contained both a production and synthetic 4G trace dataset composed of client-side cellular key performance indicators (KPIs). The synthetic dataset was generated from a large-scale 4G ns-3 simulation that includes one hundred users randomly scattered across a seven-cell cluster. This synthetic dataset was beneficial in that it provides additional information (such as competing metrics for users connected to the same cell), thus providing otherwise unavailable information about the eNodeB environment and scheduling principle, to end user. While prior work utilised the open-source NS-3 LENA project [3],

**Table 3: Average/Variation Range of Application Throughput (Mbps) across different mobility patterns and network technologies (file download scenario only)**

| | *Mobility Patterns* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Network Technology** | **Static** | | | | **Car** | | | | | |
| | *Avg.* | *Var. Range* | *# Traces* | *Trace Dur. (m)* | *Avg.* | *Var. Range* | *# Traces* | *Trace Dur. (m)* | | |
| 5G | 66.9 | (22.0, 202.5) | 5 | 260 | 28.5 | (3.0, 88.5) | 16 | 459 | | |
| 4G | 42.6 | (21.3, 77.2) | 5 | 39 | 22.3 | (3.2, 49.1) | 12 | 290 | | |

**Table 4: Average/Variation Range of Application Throughput (Mbps) across different mobility patterns and application types**

| | *Mobility Patterns* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Application** | **Static** | | | | **Car** | | | |
| | *Avg.* | *Var. Range* | *# Traces* | *Trace Dur. (m)* | *Avg.* | *Var. Range* | *# Traces* | *Trace Dur. (m)* |
| File download | 66.9 | (22.0, 202.5) | 5 | 260 | 28.5 | (3.0, 88.5) | 16 | 459 |
| Netflix | 13.7 | (0.5, 31.1) | 10 | 576 | 7.5 | (0.4, 19.9) | 23 | 637 |
| Amazon Prime | 6.9 | (0.3, 11.2) | 8 | 582 | 1.3 | (0.3, 2.7) | 21 | 628 |




**(a) File Download** **(b) Netflix** **(c) Amazon Prime**

**Figure 1: Boxplot of CQI vs application throughput for across different application types (driving)**

recent advancements offers a full-stack simulation infrastructure of the ns-3 mmwave module [11].

The focus of the simulation 5G framework in this submission is to build upon said 4G simulation model and create a flexible and highly customisable 5G/mmwave [11] ns-3 simulation framework, which generates a trace dataset of 5G key performance indicators (KPIs) across numerous 5G clients (UEs) and base stations (eNBs). Example output includes time-series channel quality indicators: CQI/SNR/RSRP/RSRQ as well as throughput rates for the different evaluation scenarios. The framework produces data for a defined number of 5G base stations and clients, typically in a Line Of Sight (LOS) level environment. The dataset provides a unique mechanism to view the relationship of the channel quality indicators between the network(s) and the client(s) in a large-scale 5G simulation. All code, build and usage instructions for our 5G/mmwave ns-3 simulation framework are available online[3].

## 5.1 Framework Configuration

Table 5 illustrates the simulation input configuration fields, categorised by where in the code these fields are configured, and a default value and description for each. To provide ease of use, a python2 script, *start_mmwave.py*, is provided to initiate execution.

[3]http://cs1dev.ucc.ie/misl/5Gframework/5G-framework.zip

This script handles user input and redirects ns-3 commandline output to a dedicated logging file system for subsequent parsing once the simulation completes.

*start_mmwave.py* uses several configurable flags, as shown in Listing 1, to take user input directly from the command line. These inputs can be displayed when executing the script with the -h flag. The user input, as defined in Table 5, is configurable through this script. This script also configures the simulation environment by removing old log files, that may interfere with the simulation, and outputs the user-provided variables to the CLI before calling the execution. While the simulation is executing, this script will repeatedly scan to determine when the simulator process begins logging and reads current simTime from a special log file, timelog.txt, and outputs the current simTime. Additionally, the script calls auxiliary scripts appropriate to handle said logging.

**Listing 1: Python Test Template**

```
# python start_mmwave.py −ue %s −enb %s −t %s −src %s
      −log %s −x %s −y %s −z %s
      −xVel %s −yVel %s −zVel %s −i %s
```

In order to generate throughput and end-to-end latency information, a remote host is introduced to the simulation. This remote node acts as the destination for packets generated by the UE and the source of packets that each UE receives. The intention for using this remote node is to behave as a pseudo-internet location.

**Table 5: ns-3 5G/mmWave Configurable Attributes and Simulation Default Values**

| Field | Script | Config Via. | Description |
|---|---|---|---|
| simTime | -t | Terminal | Default 1.0 seconds. Length of time to run the simulation. |
| numUe | -ue | Terminal | Default 1. The number of UE to simulate. |
| numEnb | -enb | Terminal | Default 1. The number of eNodeB to simulate. |
| maxX | -x | Terminal | Default 100. The size of the simulation space along the X-axis. |
| maxY | -y | Terminal | Default 100. The size of the simulation space along the Y-axis. |
| maxZ | -z | Terminal | Default 100. The size of the simulation space along the Z-axis. |
| maxXVel | -xVel | Terminal | Default 100. The maximum velocity of UE along the X-axis. |
| maxYVel | -yVel | Terminal | Default 100. The maximum velocity of UE along the Y-axis. |
| maxZVel | -zVel | Terminal | Default 0. The maximum velocity of UE along the Z-axis. |
| interval | -i | Terminal | The time interval in the completed datasets. |
| | -src | Terminal | The directory location of the NS3 waf executable. |
| | -log | Terminal | The directory to store logs. |
| minX | not defined | Source | Default 0. Assumed to be irrelevant due to the ability to configure maxX easily. |
| minY | not defined | Source | Default 0. Assumed to be irrelevant due to configure maxY easily. |
| minZ | not defined | Source | Default 0. MmWave module does not support underground simulation (yet). |
| Data Rate | not defined | Source | Default value conforms to the expected/promissory value of 5G/mmWave. Adjustment rarely required. |
| Link Delay | not defined | Source | Default value conforms to the expected/promissory value of 5G/mmWave. Adjustment rarely required. |

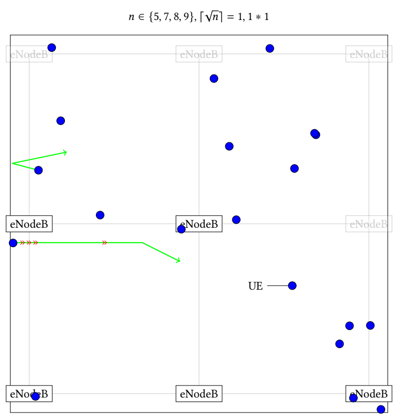



**Figure 2: 20 Scattered UEs and Mobility, eNodeB layout for $n \in \{5, 7, 8, 9\}$, Arrows Denoting Mobility**

## 5.2 Simulating Mobility and Handover

The simulation area hosts many UE nodes scattered randomly throughout. The number of UEs is configurable by the user, as is maximum velocity. The simulation randomly selects values between the minimum (no movement) and maximum velocity to assign to each UE. The travel path of the UEs follow the Gaussian Mobility model and will change direction when encountering the bounds of the simulation area. At random intervals, UE velocity will change to another randomly generated velocity. UEs initially connect to the closest eNodeB. From here, as they move throughout the simulation area, they will connect via handover to whichever eNodeB provides them with the best connection at any given time. This environment is visualised in Figure 2.

## 5.3 Simulation Evaluation and Dataset Structure

In order to test the functionality and configurations of our proposed framework, several test case simulations as detailed in Fig 3, are used to validate our design and implementation. These tests are executed multiple times over a single simulated second. Running over a single second with an interval of 0.1 gives 100 entries for the uplink and downlink of each UE and eNodeB. This setting produces adequate data for the sake of comparison while remaining within reasonable runtimes. The generated output logs are gathered and saved as numerous UE and eNB datasets in a folder called "mmwave_log". When compared these logs produce the same structure for all UEs and eNodeBs. Figure 4 illustrates an example of the sample output.

| simTime | numUe | numEnb | maxX | maxY | maxZ | maxXVel | maxYVel | maxZVel |
|---|---|---|---|---|---|---|---|---|
| 10 | 8 | 4 | 100 | 100 | 100 | 400 | 400 | 400 |
| 20 | 4 | 8 | 200 | 200 | 200 | 300 | 300 | 300 |
| 30 | 2 | 10 | 300 | 300 | 300 | 200 | 200 | 200 |
| 40 | 7 | 4 | 400 | 400 | 400 | 100 | 100 | 100 |
| 35 | 11 | 11 | 400 | 200 | 0 | 200 | 5 | 0 |
| 22 | 20 | 20 | 3 | 400 | 0 | 80 | 12 | 0 |

**Figure 3: Sample Test Cases**

To provide ease of use, we also offer an Ubuntu 19.10 VirtualBox VM containing all required dependencies and our 5G framework[4]. Username and password for the provided VM is "godashbed".

# 6 USE CASE

The production dataset and the output of the simulation framework are both exceptionally adaptive by design, allowing users to capture a variety of scenarios. Thus, there are a wide assortment of potential uses for the generated data. These potential uses extend to industries and fields such as machine learning, networking, research and development. Detailed examples are presented in Table 6.

# 7 CONCLUSION

In this paper, we present both a 5G trace dataset collected from a major Irish mobile operator, and a large-scale multi-cell 5G/wwave simulation framework. The 5G dataset is composed of client-side key performance indicators, and illustrates the variance in throughput demand in both client streaming and download scenarios. The

[4]http://cs1dev.ucc.ie/misl/5Gframework/5G-ns3-ubuntu19.10.zip

**Table 6: Proposed Use Cases**

| Industry/Field | Use | Explanation |
|---|---|---|
| Machine Learning | Predictive Models | Analysing datasets to reveal predicted network performance. |
| | Distribution Analysis | Revealing distributions and correlations between underlying low-level statistics. |
| | Generative Models | Analysis of dataset could be used to train a generative model to produce new data. |
| Networking | Prototyping | Prototyping the performance of proposed 5G/mmWave network topologies. |
| | Training | Educating would-be 5G/mmWave network engineers on the underlying concepts of 5G/mmWave. |
| Research | Experimentation | Providing a simulated environment in which to conduct experiments. |
| | Result Confirmation | Confirming published results of other researchers 5G/mmWave projects. |
| Development | Software/Hardware Testing | By introducing custom behaviour representative of new software/hardware into the simulation. |

| | time | frame | subF |
|---|---|---|---|
| DL | 0.0413333 | 41 | 3 |
| DL | 0.0513707 | 51 | 3 |
| DL | 0.0613749 | 61 | 3 |
| **1stSym** | **symbol#** | **cellId** | **rnti** |
| 5 | 3 | 1 | 3 |
| 14 | 3 | 1 | 3 |
| 15 | 3 | 1 | 3 |
| **ccId** | **tbSize** | **mcs** | **rv** |
| 0 | 1429 | 12 | 0 |
| 0 | 1429 | 12 | 0 |
| 0 | 1429 | 12 | 0 |
| **SINR(dB)** | **corrupt** | **TBler** | **Time** |
| 6.51983 | 0 | 6.05E-10 | 0.05 |
| 6.13488 | 0 | 0.000262906 | 0.06 |
| 6.36455 | 0 | 8.48E-13 | 0.07 |
| **UE** | **CurXPos** | **CurYPos** | **CurZPos** |
| 2 | 66.9202 | 45.6918 | 41.2589 |
| 2 | 67.0004 | 45.6647 | 41.2593 |
| 2 | 67.0807 | 45.6375 | 41.2597 |
| **CurXVel** | **CurYVel** | **CurZVel** | **CalcClosestENB** |
| 8.02687 | -2.71869 | 0.0435891 | 0 |
| 8.02687 | -2.71869 | 0.0435891 | 0 |
| 8.02687 | -2.71869 | 0.0435891 | 0 |
| **CalcClosestEnbDist** | **enbX** | **enbY** | **enbZ** |
| 90.9305 | 0 | 0 | 0 |
| 90.9761 | 0 | 0 | 0 |
| 91.0218 | 0 | 0 | 0 |
| **CQI** | *Cells* | | |
| 6 | *Intentionally* | | |
| 7 | *Left* | | |
| 7 | *Blank* | | |

**Figure 4: Sample 5G/mmwave output**

5G framework offers a mechanism to investigate large scale deployments of mobile devices in a multi-cell 5G environment. To the best of our knowledge, this is the first publicly available dataset that contains throughput, channel and context information for real-time analysis of a production 5G network. As the ns-3 cellular model evolves to include sub-6Ghz variations, in future work we would plan to extend our simulation framework to combine both mmwave and sub-6Ghz, as this combination is suggested in real-world next generation 5G networks.

## ACKNOWLEDGMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant 13/IA/1892, and is co-funded under the European Regional Development Fund under SFI Grant 13/RC/2077. We also thank Andrew Nash and Kieran Horgan for their invaluable assistance in evaluating the 5G/mmwave module.

## REFERENCES

[1] M. Agiwal et al. 2016. Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Communications Surveys Tutorials* 18, 3 (2016), 1617–1655.

[2] Tanweer Alam. 2018. A Reliable Communication Framework and Its Use in Internet of Things (IoT).

[3] Nicola Baldo et al. 2011. An Open Source Product-oriented LTE Network Simulator Based on Ns-3. In *Proceedings of the 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '11)*.

[4] Ayub Bokani et al. 2016. Comprehensive Mobile Bandwidth Traces from Vehicular Networks. In *Proceedings of the 7th International Conference on Multimedia Systems (MMSys '16)*. ACM, Article 44, 6 pages.

[5] L. De Cicco et al. 2013. ELASTIC: A Client-Side Controller for Dynamic Adaptive Streaming over HTTP (DASH). In *IEEE International Packet Video Workshop*.

[6] M. Cosovic et al. 2017. 5G Mobile Cellular Networks: Enabling Distributed State Estimation for Smart Grids. *IEEE Communications Magazine* 55, 10 (Oct 2017), 62–69. https://doi.org/10.1109/MCOM.2017.1700155

[7] T.-Y. Huang et al. 2014. A Buffer-based Approach to Rate Adaptation: Evidence from a Large Video Streaming Service. In *SIGCOMM*. ACM.

[8] Manish Jain et al. [n.d.]. End-to-end Estimation of the Available Bandwidth Variation Range. In *Proceedings of the 2005 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '05)*. 12.

[9] J. Jiang et al. 2014. Improving Fairness, Efficiency, and Stability in HTTP-Based Adaptive Video Streaming With Festive. *IEEE/ACM Transactions on Networking* 22, 1 (Feb 2014).

[10] L. Li et al. 2015. A measurement study on TCP behaviors in HSPA+ networks on high-speed rails. In *2015 IEEE Conference on Computer Communications (INFOCOM)*. 2731–2739. https://doi.org/10.1109/INFOCOM.2015.7218665

[11] Marco Mezzavilla et al. 2018. *End-to-End Simulation of 5G mmWave Networks*. https://ieeexplore.ieee.org/document/8344116

[12] Arvind Narayanan et al. 2019. A First Measurement Study of Commercial mmWave 5G Performance on Smartphones. arXiv:cs.NI/1909.07532

[13] J. Nightingale et al. 2018. 5G-QoE: QoE Modelling for Ultra-HD Video Streaming in 5G Networks. *IEEE Transactions on Broadcasting* 64, 2 (June 2018), 621–634. https://doi.org/10.1109/TBC.2018.2816786

[14] Darijo Raca et al. 2019. Empowering Video Players in Cellular: Throughput Prediction from Radio Network Measurements. In *Proceedings of the 10th ACM Multimedia Systems Conference (MMSys '19)*. Association for Computing Machinery, New York, NY, USA, 201–212. https://doi.org/10.1145/3304109.3306233

[15] Darijo Raca, Jason J. Quinlan, Ahmed H. Zahran, and Cormac J. Sreenan. 2018. Beyond Throughput: A 4G LTE Dataset with Channel and Context Metrics. In *9th ACM Multimedia Systems Conference (MMSys '18)*. 460–465.

[16] Haakon Riiser et al. 2013. Commute Path Bandwidth Traces from 3G Networks: Analysis and Applications. In *Proceedings of the 4th ACM Multimedia Systems Conference (MMSys '13)*. ACM, 114–118.

[17] T. Stockhammer. 2011. Dynamic Adaptive Streaming over HTTP – Standards and Design Principles. In *MMSys '11 Proceedings of the second annual ACM conference on Multimedia systems*. New York, 133.

[18] Liyang Sun et al. 2018. Multi-Path Multi-Tier 360-Degree Video Streaming in 5G Networks. In *Proceedings of the 9th ACM Multimedia Systems Conference (MMSys '18)*. Association for Computing Machinery, New York, NY, USA, 162–173.

[19] J. van der Hooft et al. 2016. HTTP/2-Based Adaptive Streaming of HEVC Video Over 4G/LTE Networks. *IEEE Communications Letters* 20, 11 (2016), 2177–2180.

[20] Q. Xiao et al. 2014. TCP Performance over Mobile Networks in High-Speed Mobility Scenarios. In *2014 IEEE 22nd International Conference on Network Protocols*. 281–286. https://doi.org/10.1109/ICNP.2014.49

[21] Xiufeng Xie et al. 2015. piStream: Physical Layer Informed Adaptive Video Streaming over LTE. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, 413–425.

[22] C. Yue et al. 2017. LinkForecast: Cellular Link Bandwidth Prediction in LTE Networks. *IEEE Transactions on Mobile Computing* (2017).