Data in Brief 52 (2024) 109954

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Experimental mmWave WiGig-based backhaul network dataset

Tânia Ferreira[a,b], Duarte Raposo[a,*], Alexandre Figueiredo[a], Eurico Dias[a,b], Pedro Rito[a], Miguel Luís[a,c], Susana Sargento[b]

[a] *Instituto de Telecomunicações, Aveiro 3810-193, Portugal*
[b] *Departamento de Electrónica, Telecomunicações e Informática, Universidade de Aveiro, Aveiro 3810-193, Portugal*
[c] *Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, Lisboa 1049-001, Portugal*

ARTICLE INFO

Article history:
Received 28 October 2023
Revised 4 December 2023
Accepted 6 December 2023
Available online 12 December 2023

Dataset link: Millimeter-Wave Dataset (Original data)

Keywords:
Outdoor network
Blockage
802.11ad
WiGig
Backhaul network
5G

ABSTRACT

The wireless backhaul has emerged as an attractive alternative to traditional fiber backhaul for 5G technology, offering greater flexibility and cost-effectiveness thanks to the availability of high bandwidths capable of achieving fiber-like data rates. However, the millimeter-wave-based (mmWave) protocols, namely IEEE 802.11ad and later IEEE 802.11ay, suffer from a high susceptibility to obstruction, which only allows correct operation under Line-of-Sight conditions (LOS). Any sudden obstructions can significantly reduce the maximum achievable throughput, leading to delays exceeding acceptable limits for critical applications, and may even culminate in link failure in certain circumstances. Therefore, it is essential to assess how different types and durations of obstructions impact different network OSI layers to determine the feasibility of mmWave. WiGig-based technologies for wireless backhaul scenarios. This article describes a dataset collected from an experimental IEEE 802.11ad backhaul network, mmWave-based mesh network at 60 GHz, deployed in an outdoor environment. The data contains multi-layer information, including MAC, PHY, and network data, which provides valuable insights into the WiGig network behavior under three distinct scenarios. These scenarios include normal operation, long-term blocked scenario, and short-term blocked scenario, based on the type and duration of the

---

[*] Corresponding author.
E-mail address: dmgraposo@av.it.pt (D. Raposo).

https://doi.org/10.1016/j.dib.2023.109954

blockage event crossing the LOS path. The dataset presents an extensive PHY, MAC and transport layer measurement campaign for an outdoor WiGig network, and thus it is a valuable resource for researchers and professionals interested in understanding the behavior and performance of real-life mmWave-based WiGig networks aimed for 5G backhauling.

Specifications Table

Subject	Computer Networks and Communications
Specific subject area	Wireless Networks
Type of data	Table
Data format	Filtered
How the data were acquired	The test data was collected using a custom automation script, under various blockage conditions, with iperf3 client and a server set up at each end link. The following scenarios were used to collect physical and network metrics: Normal scenario) data was collected for each experiment (MCSs: 1, 3, 5, 7, 9, and Auto) during 15 minutes with continuous LOS between each pair of STAs; Long-Term Blockage scenario) using a metallic obstacle to simulate continuous blockage between each STA pair tested for 15 minutes; Short-Term blockage) for the automatic MCS mode, a temporary LOS throughout during 5 minutes, by having people moving freely across the path formed between the two STAs. The collected data was then uploaded to an accessible database, and recorded timestamps were used to retrieve metrics for further analysis. Hardware used: (i) Three CCS Metnet 60 GHz nodes [1]; (ii) two accelerated processing units (APUs) were placed at each node end to inject traffic into the mmWave connections.
Data Collection	All data was uploaded to a database for further analysis. Data is sent between an iperf3 server located at the PBSS central point (PCP) node sending data to an iperf3 client located at two remote nodes, node A and B. These remote nodes are connected one hop away from PCP and work according to the IEEE 802.11ad norm.
Data Source Location	City/Town/Region: Aveiro Country: Portugal
Related Research Article	Ferreira, T. et al. “Millimeter-wave feasibility in 5g backhaul: A cross-layer analysis of blockage impact”, in IEEE Access, volume 11, pp. 5178–5192, 2023, doi:10.1109/ACCESS.2023.3236100. Ferreira, T. et al. “Improving mmwave backhaul reliability: A machine-learning based approach”, in Ad Hoc Networks, volume 140, 2023, doi:10.1016/j.adhoc.2022.103050 (2023).
Data Accessibility	Repository name: SNOB5G Data Direct link to data: https://github.com/nap-it/SNOB5GData DOI: https://doi.org/10.5281/zenodo.10050012

1. Value of The Data

- Importance: The data provided shows the behavior of an experimental IEEE 802.11ad backhaul network deployed in an outdoor environment in different real-life use cases. To this author’s knowledge, it is the first of its kind as it includes extensive MAC, PHY, and network-layer metrics, providing insights into mmWave-based WiGig network behavior under normal operation, long-term blocked, and short-term blocked scenarios. Previous measurement campaigns on mmWave-based WiGig networks mostly focused on indoor settings and measured only a few metrics like TCP throughput, PHY Rate and RSSI. They also only analyzed static blockages.
- Target audience: The dataset provided is a valuable asset to researchers in the field of millimeter-wave communications, especially those lacking the resources or infrastructure to conduct extensive outdoor IEEE 802.11ad measurements with specialized outdoor equipment.

It is also a great tool for those who wish to get a deep understanding of the impact of the different types and durations of blockage on the different OSI layers of an outdoor mmWave-based WiGig network. Note that this dataset includes a broader set of metrics (particularly from a lower perspective) and a wider range of blockage scenarios compared to other related works.

- Future use: The dataset provided allows defining the impact of both long-term and short-term obstructions on mmWave-based WiGig networks, as explored in [2]. It has also supported the development of a proactive machine learning framework [3] in an outdoor setting [4], leveraging real-life data to enhance the reliability and resilience of WiGig-based networks that are prone to blockages. This includes a link quality classifier that can differentiate normal operation from long-term and short-term blockages, as well as a deep learning forecasting model that accurately predicts relevant link metrics ahead of time. Additionally, this data can be used to further evaluate the feasibility of IEEE 802.11ad solutions in outdoor environments suffering from static and transient blockages [5].

2. Data Description

The dataset contains the complete data from experiments conducted during July and August 2021 in an outdoor IEEE 802.11ad-compliant network deployed on the rooftop of the Institute of Telecommunications in Aveiro, Portugal. The data includes PHY, MAC and network metrics for each network scenario (normal, long-term blockage, and short-term blockage), and for each possible MCS combination and link that could be formed using the different radios at each unit. All the data was recorded in a database in separate .cvs files, which were later used to generate a single file containing all the metrics for all the scenarios, MCS combinations tested, and tested radio paths. This dataset contains two files that include PHY, MAC and network data for all the experiments:

- “dataset_complete.csv”: This file is the complete version of the dataset, i.e., it has all the metrics for three scenarios for all MCS modes tested (1, 3, 5, 7, 9 and auto), and for each path (different radio combinations) tested to form the two backhaul connections (PCP and A, and PCP and B). It contains 124980 entries, each collected with a 1s periodicity, where each row is composed of a set of test-specific identifiers (Link, Time, scenario, and MCS mode) and the multi-layer metrics (Beam Index RX, Beam Index TX, Beam Index TX (Min), Beam Index RX (Min), Beam Index TX (Max) and Beam Index RX (Max), MCS RX, MCS TX, Mbps TX, Mbps TX (Max), Packet Error(%), RCPI (dBm), SNR (dB), Packet Loss (%), Sent (Mbps), Received (Mbps), Retransmits, RTT (Mean), RTT (Max), RTT (Min)). Table 1 describes all the file’s attributes along with their description and missing values.
- “dataset_P_A_auto.csv”: This file is a shorter version of the data presented in dataset_complete.csv. Specifically, this file has only 3979 data entries and only contains the information regarding the backhaul connection between nodes PCP and A in the automatic MCS mode for the three scenarios tested. Similarly to the complete version, each row is composed of a set of test-specific identifiers (Link, Time, scenario, and MCS mode) and the multi-layer metrics (Beam Index RX, Beam Index TX, MCS RX, MCS TX, Mbps TX, Packet Error(%), RCPI (dBm), MCS mode, Sent (Mbps), Received (Mbps), Retransmits, RTT (Mean), RTT (Max), RTT (Min)).

3. Experimental Design, Materials and Methods

The datasets discussed in this document were collected to study the impact of long-term and short-term obstruction on experimental mmWave-based WiGig networks. Metal obstacles can have significant impacts on mmWave signal propagation, regarding first order reflections, leading to signal attenuation and multipath propagation, resulting in destructive interface. Thus,

Table 1
Datasets attributes description.

Attributes	Description	Missing Value
Link	The communication path identifier (e.g. P/A/2/1). The first two letters identify the transmitting node and the receiving node, respectively, and the next two numbers identify the radio used by the receiving node and the radio used by the transmitting node, respectively. Note that the uplink and downlink between each node pair have a different tag associated	A/B/x/x and B/A/x/x
Time	The date of the specific measurement time. This metric has a second precision	
Beam Index RX (Min), Beam Index RX (Max), Beam Index RX	The ID of the maximum, minimum and average beam used for reception	
Beam Index RX (Min), Beam Index RX (Max), Beam Index RX	The ID of the maximum, minimum, and average beam used for transmission	
MCS RX	Modulation and Coding Scheme (0-9) used during transmission in one direction (uplink or downlink)	2,4,6 and 8
MCS TX	Modulation and Coding Scheme (0-9) used during transmission in the other direction	
Mbps TX (Max) and Mbps TX	The peak and average rate measured at the MAC layer in Mbps	
Packet Error (%)	Packet Error Rate, i.e., the ratio (in percentage) between the number of packets with errors after FEC and the tota number of transmitted packets.	
RCPI (dBm)	Received Power Indicator in dBm	
SNR (dB)	Signal to Noise Ratio in dB. This metric is used by WiGig protocols to select the optimal RX/TX communication beam pair	
Packet Loss	Packet losses at the transport layer in percentage	
Sent (Mbps) and Received (Mbps)	TCP transmit and receive rates in Mbps	
Retransmists	The total number of TCP retransmissions	
RTT (Min), RTT (Max) and RTT (Mean)	The minimum, maximum and average Round-Trip Time (in ms)	

in our experiment, a metal obstacle was selected as obstruction. The testbed represents a typical 5G backhaul network, deployed on a city-scale. For this purpose, an IEEE 802.11ad network composed of three CCS Metnet 60 GHz [1] nodes was deployed on the rooftop of Instituto de Telecomunidações building -at the height of about 1.5 m and with a 50-m distance between the nodes-, to ensure that the tests were performed under a closed and controlled outdoor environment. Note that there were no other devices operating in this band on this environment, thus there is no co-channel nor adjacent channel interference. Fig. 1 shows the complete infrastructure used for the data collection under the various test scenarios.

The network nodes employ the standardized IEEE 802.11ad (WiGig) technology that operates between the 57 GHz and 66 GHz unlicensed band, which supports channels 1 to 4 each with a bandwidth of 2.16 GHz. Each node has four radio modules covering a 90º range, overlapping to cover a 300º horizontal field of view. Each radio employs a 19 dBi beamforming steerable antenna (a 16×2 element array with 37 dBm EIRP) that enables the establishment of directional links between STAs, ultimately allowing it to cope with the propagation losses associated with mmWave channels. The 300º wide range covered by each node is further subdivided into 64 discrete sectors (with 5º horizontal beamwidth), which can be used to send and receive the signal toward a specific direction.

While the standardized IEEE 802.11ad defines a point to multi-point network architecture that requires LOS - which typically covers distances usually up to 300 m with phased-arrays antennas -, the used system deploys a multi-point to multi-point topology and can cover distances up to 500 m, depending on the location, availability, and available capacity.

During the conduction of this study, the channel impulse response (CIR) measurement was performed by computing the autocorrelation function at the receiver, - using a continuous

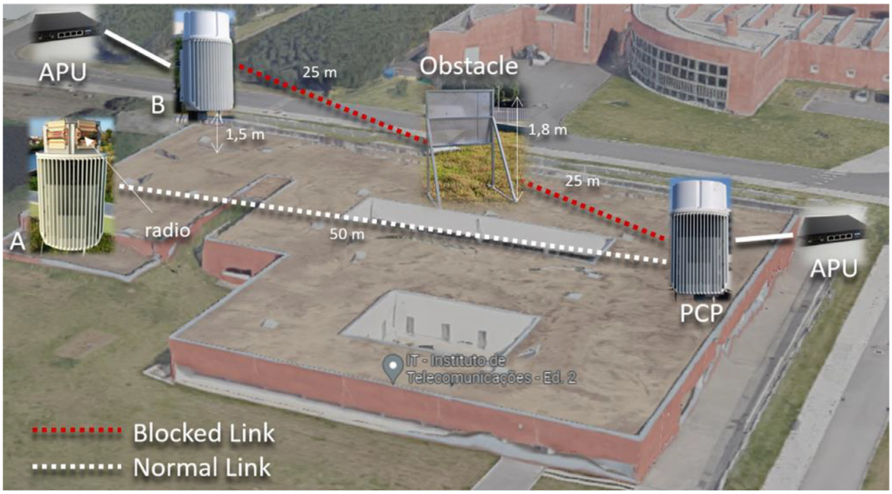


Fig. 1. Nodes and topology of the deployed outdoor testbed, presenting the obstacle used for the long-term blockage shortage [2].

Table 2
Metnet mesh version 1 node specification [1].

Frequency Band	60 GHz mmWave unlicensed 60 GHz mm Wave unlicensed 57 GHz to 66 GHz band
Topologies	Point to MultiPoint MultiPoint to MultiPoint (Mesh)
Capacity	10 Gbps per node
Radio Access	Metnet SON utilizing TDMA Dynamic TDD
Antenna	Beamforming Phased array 19 dBi gain
Beam Angle	Horizontal 300° electronic steerable 5° beamwidth, Vertical 20° fixed
Modulation and encoding	10 levels of adaptive encoding MCS 0–9 up to 2.5 Gbps
Range	Up to 500m* Dependant on location, availability, and capacity
Interfaces	Up to 4 Ethernet interfaces 2 x Fixed RJ45 100/1000 Base-T 2 x optional 10G or 1G SFP optical

Pseudo-Random Binary Sequence (PRBS) as the probe signal -, to have an accurate description of the channels' propagation conditions in both the uplink and downlink directions. The PRBS signal's impulse-like autocorrelation function minimizes its influence on CIR estimation [6].

When a communication channel has reflections, multiple signal copies will arrive with different delays, resulting in additional peaks in the estimated CIR. The obtained CIRs indicate the existence of small NLOS components caused by reflections on the ground or metal edges surrounding the roof; however, these can be neglected compared to the dominant LOS component, as shown by Fig. 2.

As shown in Fig. 3, the nodes' beamforming capabilities allow the sender node to concentrate the signal in the best direction it has found towards the receiver during the IEEE 802.11ad beamforming training. More details on the node specs can be found in Table 2. To inject the traffic within the mmWave mesh, two iperf3 hosts were connected to both start and end nodes, as can be seen by the introduction of the Accelerated Processing Units (APUs) 1 and 2 in the network diagram of Fig. 1. The deployed network consists of both wired and wireless 60 GHz mesh nodes, and if a node is not connected to the Point of Presence, it acts as a remote node (e.g. node A and B), with access to the core network provided through a wired node.

The CCS system features an EMS that provides a graphical overview of the state of the nodes and the mmWave links, displaying real-time MAC layer metrics (e.g., RCPI, SNR, MCS, PER). To

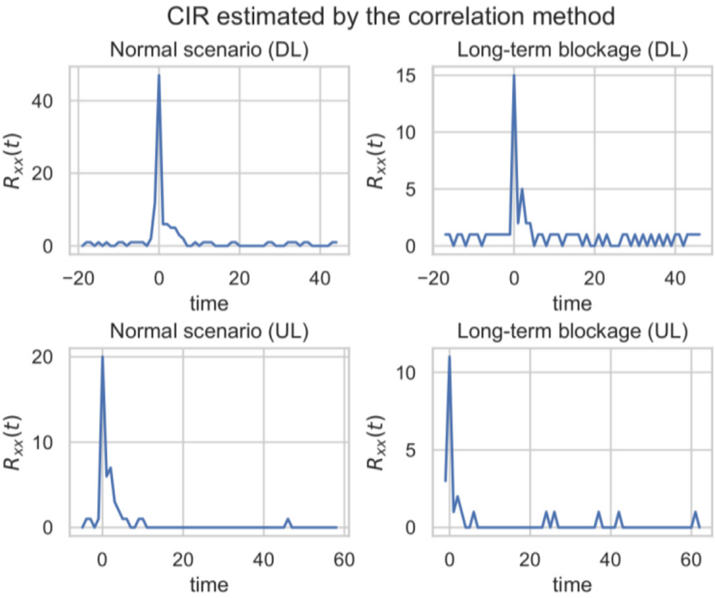
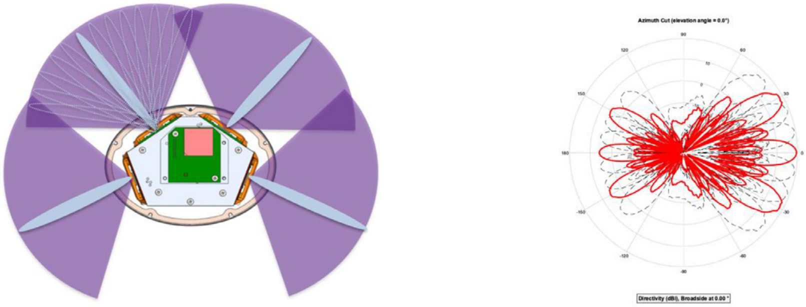


Fig. 2. CIR estimated by the autocorrelation function ($R_{xx}(t)$) at the receiver end using a continuous Pseudo-Random Binary Sequence (PRBS).



a) *CCS node's multisector beamforming phased arrays[4].*

b) Radiation pattern that can be used by a radio's node to establish 3 different directional beams.

Fig. 3. Example of establishing directional links with phased-array beamforming process.

access these link metrics outside the web page context, a python script was developed using HTTP requests to access the backend API. The metrics were then uploaded to a collection inside a newly deployed Mongo database hosted outside the EMS container. Additionally, a network probe using the iperf3 tool was developed to collect UDP and TCP-related metrics simultaneously. This probe comprises an application server and the respective client. For the client part, a python script was developed, which instantiates two parallel iperf3 clients with a bandwidth of 500 Mbps each to avoid link saturation) whose jobs are to collect UDP and TCP metrics. Moreover, the pingparser library was used to parse the output of the system's ping, so that the RTT metric could be obtained. The ping metrics were calculated by averaging the metrics measured for ten packet transmissions. The client is also responsible for loading the metrics at a given timestep into a new collection inside the deployed Mongo database. This database collection is composed of entries that display the instantaneous metrics values (described in Table 1) for each specific path. As for the server, it was implemented by simply running two iperf3 servers in the same ports defined by the client script.

The network-level probe was used to obtain data relating to each link in a sequential manner, with individual test runs of the network probe performed for each link. Note however that, while a UDP-capable probe was developed, it was found to stop loading metrics into the database due to a sporadic error. Thus, the UDP data was removed from the final dataset to avoid repeating these time-consuming experiments.

Since the firmware version used lacked a GUI functionality for selecting specific radios that establish communication in each link node, a JSON file, stored inside each remote node, had to be manually configured each time a specific path was to be tested. Because this was a time-consuming process, a bash script was developed to automate the selection process, allowing the appropriate parameters to be configured for testing a specific path. Besides configuring the appropriate parameters for testing a specific path (i.e., the radios used on each end), this script tested different paths sequentially and registered the exact time a specific path test began and ended.

After the data collection probes and the test automation tools described earlier were developed and thoroughly tested, the network data collection began one at a time for each of the three network scenarios (normal, long-term blockage, and short-term blockage). Each scenario was tested for each possible MCS combination and each different path that could be formed with the four radio units in each node for the links between node PCP and A, and node PCP and B. The tested scenarios were defined according to the type and duration of the blockage event crossing the LOS path, and can be described as:

- Normal scenario: metrics were collected for each path, for all MCS combinations (1, 3, 5, 7, 9, and automatic mode), during 15 min in a non-obstructed propagation environment with continuous LOS between each pair of tested STAs (PCP and A, PCP and B).
- Long-term blockage: metrics were collected for each path, for all MCS combinations (1, 3, 5,7, 9, and automatic mode) during 15 min under continuous LOS blockage as a metallic obstacle, shown in Fig. 1, was placed across the LOS path formed between each STA pair tested.
- Short-term blockage: metrics were collected for each path for the automatic MCS mode, during 5 minutes, in the presence of random temporary LOS blockages caused by having 1 and 2 people at a time moving across the path formed 5/7 between the two tested STAs. These movements were sometimes made in a completely random manner, and in other times they were performed in an alternating way following a horizontal, vertical, and diagonal line to the line where the link between the two STAs was estimated to exist. Note that, due to the nature of the short-term blockage scenario, it would be very time-consuming to collect data for all possible paths with all MCS modes for the same 15 min mark used in the other scenarios. Thus, the data collection had to be limited to a smaller set of paths and a shorter time frame.

After the data collection for each scenario was complete, the transport, network, and physical metrics were retrieved from the database using the recorded timestamps that marked the beginning of each experiment. The data was later analysed to study the consequences of long-term and short-term blockage on the multiple network layers [2].

Limitations

None.

Ethics Statement

The work here presented did not involve human subjects, animal experiments, or any data collected from social media platforms.

Data Availability

Millimeter-Wave Dataset (Original data) (GitHub)

CRediT Author Statement

Tânia Ferreira: Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing – original draft, Visualization; **Duarte Raposo:** Conceptualization, Methodology, Validation, Software, Resources, Writing – review & editing; **Alexandre Figueiredo:** Validation, Resources, Data curation; **Eurico Dias:** Software, Validation, Resources, Data curation; **Pedro Rito:** Conceptualization, Methodology, Validation, Writing – review & editing; **Miguel Luís:** Conceptualization, Methodology, Validation, Writing – review & editing; **Susana Sargento:** Conceptualization, Methodology, Validation, Writing – review & editing, Funding acquisition.

Acknowledgments

This work is supported by the European Union / Next Generation EU, through Programa de Recuperação e Resiliência (PRR) [Project Nr. 11: New Space Portugal].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Metnet 60G Unlicensed mmwave Mesh Datasheet. https://www.ccsl.com/v1/uploads/files/Metnet-60G-Mesh-datasheet.pdf (accessed September 20th 2023).
[2] T. Ferreira, et al., Millimeter-wave feasibility in 5g backhaul: a cross-layer analysis of blockage impact, IEEE Access 11 (2023) 5178–5192, doi:10.1109/ACCESS.2023.3236100.
[3] T. Ferreira, et al., Improving mmwave backhaul reliability: a machine-learning based approach, Ad Hoc Netw. 140 (2023), doi:10.1016/j.adhoc.2022.103050.
[4] Broadnet, A. R. & Brady, J. “Using self-organizing mmwave to deliver 5g services”, https://www.bcba.ca/application/files/5515/5812/9724/CCS_Metnet__Broadnet.pdf. (accessed September 20th 2023).
[5] Ferreira, T. et al. “Cross-layer data of WiGig links submitted to blockage in an outdoor environment”, https://github.com/nap-it/SNOB5GData (2022)
[6] I. Kochanska, J. Schmidt, A. Schmidt, Study of probe signal bandwidth influence on estimation of coherence bandwidth for underwater acoustic communication channel, Appl. Acoust. (2021) doi:183. 108331, doi:10.1016/j.apacoust.2021.108331.