Data in Brief 32 (2020) 106102

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

A bluetooth low energy dataset for the analysis of social interactions with commercial devices☆

Michele Girolami[a,*], Fabio Mavilia[a], Franca Delmastro[b]

[a] Institute of Information Science and Technologies, National Research Council, Pisa, Italy
[b] Institute of Informatics and Telematics, National Research Council, Pisa, Italy

ARTICLE INFO

Article history:
Received 29 June 2020
Revised 23 July 2020
Accepted 23 July 2020
Available online 30 July 2020

Keywords:
Social interactions
Bluetooth low energy
Co-location
Social sensing

ABSTRACT

This paper describes a data collection campaign and a dataset of BLE beacons for detecting and analysing human social interactions. The dataset has been collected by involving 15 volunteers that interacted in indoor environments for a total of 11 hours of activity. The dataset is released as a collection of CSV files with a timestamp, RSSI (Received Signal Strength Indicator) and a unique identifier of the emitting and of the receiving devices. Volunteers wear a wristband equipped with BLE tags emitting beacons at a fixed rate, and a mobile application able to collect and to store beacons. We organized 6 interaction sessions, designed to reproduce the three common stages of an interaction (Non Interaction, Approaching and Interaction). Moreover, we reproduced interactions by varying the volunteer's posture as well as the position of the receiving device. The dataset is released with a ground truth annotation reporting the exact time intervals during which volunteers actually interacted. The combination of such factors, provides a rich dataset useful to experiment algorithms for detecting interactions and for analyzing dynamics of interactions in a real-world setting. We present in detail the dataset and its evaluation in “Sensing Social In-

☆ This work is co-submitted with the paper entitled: “Sensing Social Interactions through BLE Beacons and Commercial Mobile Devices”
DOI of original article: 10.1016/j.pmcj.2020.101198

* Corresponding author.
E-mail addresses: michele.girolami@isti.cnr.it (M. Girolami), fabio.mavilia@isti.cnr.it (F. Mavilia), franca.delmastro@iit.cnr.it (F. Delmastro).

https://doi.org/10.1016/j.dib.2020.106102

teractions through BLE Beacons and Commercial Mobile Devices", in which we focus on two orthogonal analysis: quality of the dataset and RSSI symmetry of the channel during the interaction stage between pairs of users.

Specifications table

Subject	Computer Science
Specific subject area	Computational Social Science and Mobile Social Sensing for Detecting Human Social Interactions.
Type of data	Comma Separated Files (CSV) reporting the fields described in Table 1 for Data and Ground Truth annotation.
How data were acquired	BLE commercial tags emitting beacons, Android mobile application.
Data format	Raw data with Ground Truth annotation as separate files.
Parameters for data collection	The experimental sessions have been designed to reproduce three stages of a social interaction: Non Interaction (i.e., subjects are resting and they are not interacting), Approaching (i.e., subjects are moving close to each other) and Interaction (i.e., subjects are resting and they are interacting). Each of the three stages is denoted by a distance in meters.
Description of data collection	BLE beacons are emitted with a frequency of 5 Hz and power of emission set to -6 dBm. The mobile application is designed for Android OS and it records beacons in best-effort mode. Data collected are stored in the internal storage unit of the device. The data are collected during 6 sessions each designed to reproduce different conditions of the social interaction. Each session consists of as set of tests replicated for a number of runs.
Data source location	High School, Italy
Data accessibility	https://doi.org/10.17632/v7nc4zypgy.1
Related research article	M. Girolami, F. Mavilia, F.Delmastro, Sensing social interactions through ble beacons and commercial mobile devices, Pervasive and Mobile Computing (2020), 101198, (2020), doi: https://doi.org/10.1016/j.pmcj.2020.101198.170.

Value of the data

- This dataset provides a detailed set of experimental sessions reproducing social interactions. The sessions differentiate with respect to the posture of the subjects (Standing or Sitting) and the way the recording device is hold (in a hand, front pocket, back pocket). Such combinations reproduce realistic and non-trivial conditions of the social interactions. Moreover, the dataset is collected by using commercial devices (BLE tags and a mobile app installed on 10 different models). All the interactions are annotated with a Ground Truth to allow the performance assessment of novel algorithms designed to detect social interactions.
- Researchers and start-up players can use this dataset in order to analyze and automatically detect proximity among people in order to design social-aware services.
- The dataset provides a valuable starting point for the automatic detection of the different stages of an interaction. Algorithms based on the recognition of the interactions can be used for further experiments: (i) to study dynamics of the interactions in a natural context (e.g. students, collaborative groups), (ii) to detect cases of *social isolation* of fragile subjects (e.g. older adults or subjects affect by diseases) or (iii) to study and predict the diffusion of diseases in the crowd.
- The dataset presents also additional values: (i) a detailed Ground Truth of the social interactions, (ii) subjects involved are not trained for the experiments. Data are gathered with commercial devices, as a result the dataset provides non-biased traces of the RSSI and (iii) the experiments are designed to consider heterogeneous condition of the subjects involved.

1. Data description

The dataset is released in the form of a set of CSV files containing a time series of the Bluetooth Low Energy (BLE) beacons recorded by devices during the experimental sessions. Volunteers wear a commercial BLE tag advertising beacons at a fixed rate, while an Android-based app has been used to collect and store BLE beacons broadcasted by other devices. The BLE tags are produced by Global Tag[1], as shown in Fig. 1.

The mobile app logs the following information: `timestamp`, `receiver ID`, `sender ID` and `RSSI`, as described in Table 1. Such fields are stored in a CSV file whose name identifies uniquely the session, the tuple and test executed (filename: < `session identifier`, `tuple identifier`, `test identifier`>; content: <`timestamp`, `receiverID`, `senderID`, `RSSI`>).

As an example, `1_2_3.csv` represents the file containing data from session 1, tuple 2, test 3, as reported in Fig. 2.

The dataset is organized in 6 experimental sessions. Each session consists of a set of tests repeated for a number of runs. The tests differentiate with respect to how subjects behave, i.e. body posture or the way they hold the device. The released files reflect such structure.

Each of the experimental sessions is identified with a unique identifier. A session is a collection of t tests repeated for r runs. Runs are repetitions of the same test. For example, session S_1 comprises $t = 3$ tests, each of them repeated for $r = 5$ times, for a total of 15 different runs. During a test, subjects can reproduce one or more of the 3 stages for an interaction: Interaction, Approaching and Non Interaction, as shown in Fig. 3. Each stage is characterize by a distance of the subjects according to the definition of [2]. In particular, the Interaction stage spans up to 1.2 m, the Approaching stage ranges between 1.2 and 2.5 m and the Non Interaction ranges between 3 and 3.5 m.

We split the dataset in sessions in order to reproduce several scenarios typical of the social interactions. In particular, we consider 3 main types of sessions:

Fig. 1. BLE Tags used during the experimental sessions.

[1] http://www.global-tag.com

Table 1
Data format of the CSV files.

Timestamp	ID Receiver	ID Sender	RSSI

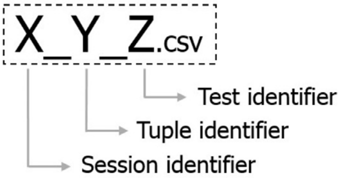

Fig. 2. Structure of the file names.

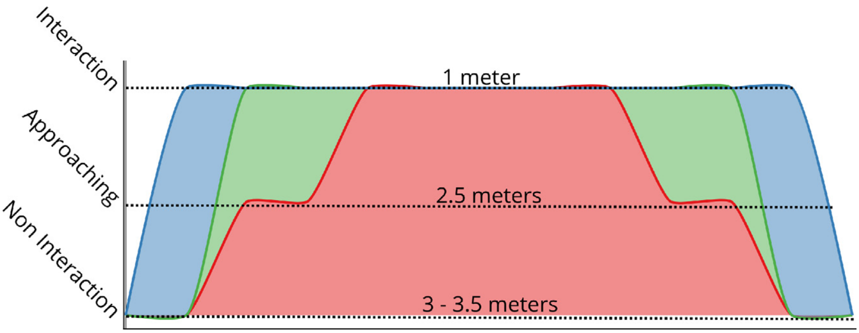

Fig. 3. Stages of a face-to-face interaction.

- T_1: the goal is to collect beacons only for the calibration purpose. In particular, we collect RSSI values while volunteers are involved only in the Interaction stage.
- T_2: the goal is to collect beacons during two stages: Non Interaction and Interaction. Volunteers first do not interact, and later they move in proximity and start interacting.
- T_3: the goal is to collect beacons during 3 stages: Non Interaction, Approaching and Interaction. Volunteers first do not interact, then they move in the Approaching stage and finally they interact.

Table 2 describes the settings for each of the 6 experimented sessions. The table provides the following information:

- Session ID and Test ID;
- Interaction Time/Distance: duration of the Interaction stage expressed in minutes and distance between subjects expressed in meters;
- Non interaction Time/Distance: duration of the Non Interaction stage expressed in minutes and distance among subjects expressed in meters;
- Approaching Time/Distance: duration of the Approaching stage expressed in minutes and distance among subjects expressed in meters;
- Runs: number of runs for the same test;
- Tuples: the layout of the subjects: dyad (2), trio (3), foursome (4), groups of five (5);
- Posture: subjects can stand (ST) or sit (SI) during the tests;
- Receiver position: receiver can be hold in hand (H) or put in the front pocket (FP) or in the back pocket (BP). For some of the tests, subjects hold the device differently. For example 2 subjects can put the device in the front pocket, while 2 other subjects put the device in the back pocket. In these cases, we adopt the naming convention: 2 x FP, 2 x BP.

Table 2
Settings for all the Sessions. ST/SI stand for standing/sitting postures, FP/BP stand for front/back pocket.

Session	Test	Type	Interaction Time/Distance	Non interaction Time/Distance	Approaching Time/Distance	Runs	Tuples	Posture	Receiver Position
1	1	T_1	5′/1m	–	–	5	2	ST	FP
	2		5′/1m	–	–	5	2	SI	FP
	3		5′/1m	–	–	5	2	ST	BP
2	1	T_2	3′/1m	1′/3m	–	5	3	ST	FP
	2		3′/1m	1′/3m	–	5	3	SI	FP
	3		3′/1m	1′/3m	–	5	3	ST	BP
3	1	T_2	4′/1m	2′/3m	–	5	5	ST	FP
	2		4′/1m	2′/3m	–	5	5	SI	FP
	3		4′/1m	2′/3m	–	5	5	ST	BP
4	1	T_3	4′/1m	2′/3.5m	2′/2.5m	4	4	ST	FP
	2		4′/1m	2′/3.5m	2′/2.5m	4	4	SI	FP
	3		4′/1m	2′/3.5m	2′/2.5m	4	4	ST	BP
	4		4′/1m	2′/3.5m	2′/2.5m	4	4	ST	H
5	1	T_3	4’/1m	2’/3.5m	2’/2.5m	4	4	ST	2 x H 1 x FP 1 x BP 2 x FP 2 x H
	2		4′/1m	2′/3.5m	2′/2.5m	4	4	ST	2 x BP 2 x FP
	3		4’/1m	2’/3.5m	2’/2.5m	4	4	ST	2 x BP 1 x FP 1 x H l 2 x FP 1 x BP 1 x H
	4		4′/1m	2′/3.5m	2′/2.5m	4	4	ST	2 x FP 2 x H
6	1	T_3	4′/1m	2′/3.5m	2′/2.5m	3	2	ST	1 x FP 1 x BP
	2		4′/1m	2′/3.5m	2′/2.5m	3	2	ST	1 x BP 1 x H
	3		4′/1m	2′/3.5m	2′/2.5m	3	2	ST	1 x FP 1 x H

For the majority of the tests, tuples had the same posture and they hold the device in the same way (see Sessions: 1, 2, 3, 4). For example, Session 1 - test 1 the 2 tuple were **ST** with device in FP, while in Session 1 – test 2 the 2 tuple were **SI** with device in FP.

Differently, in Session 5 and 6 we introduced some changes on the postures and on the receiver positions. In particular, Session 5 - tests 1 and 3, the 2 foursome (Q_1 and Q_2) had different layouts. In test 1, 2 subjects in Q_1 hold: the device in hand (2xH), 1 in the front pocket (1xFP) and 1 in the back pocket (1xBP), while for Q_2 2 subjects hold: the device in front pocket and 2 in hand. In test 3, 2 subjects in Q_1 hold the device in back pocket (2xBP), 1 in the front pocket (1xFP) and 1 in hand (1xH), while for Q_2 2 subjects hold the device in front pocket, 1 in the back pocket and 1 in hand. Similar considerations apply also for Session 6.

We report in Tables 3 and 4 a detailed description of Session 1, 2, 3 and 4 and of Session 5, 6 respectively so that to better understand the different settings. The two tables report the session ID, the tuple ID, the IDs of the subjects joining the tuple, the device model, the Version of Android OS and a summary of Posture and Receiver Position in all the session's tests. Moreover, the two tables refer also the CSV file containing the tests described.

In order to clarify the layout of the experimental sessions, we report in Fig. 4 a visual representation of the tests. We show the posture of the volunteers (Standing or Sitting) and the position of the smartphone (front pocket or back pocket). As previously described, volunteers wear the BLE tag on one of their arms.

Table 3
Details of Sessions 1, 2, 3 and 4. ST/SI stand for standing/sitting postures, FP/BP stand for front/back pocket.

Session	Tuple ID	User ID	Device Model	Android	Posture/Receiver Position				Dataset filename			
				Version	Test 1	Test 2	Test 3	Test 4	Test 1	Test 2	Test 3	Test 4
1	1	15	Samsung Galaxy S7	7.0	ST/FP	SI/FP	ST/BP	–	1_1_1	1_1_2	1_1_3	–
		36	Nexus 6	7.0								
	2	23	Nexus 6	7.0					1_2_1	1_2_2	1_2_3	–
		26	LG G4	6.0								
	3	20	Huawei P9 Lite	6.0					1_3_1	1_3_2	1_3_3	–
		22	Nexus 6	7.0								
	4	41	Nexus 6	7.0					1_4_1	1_4_2	1_4_3	–
		42	Nexus 6	7.0								
2	1	4	Honor 8	7.0	ST/FP	SI/FP	ST/BP	-	2_1_1	2_1_2	2_1_3	–
		23	Nexus 6	7.0								
		26	LG G4	6.0								
	2	44	Nexus 6	7.0					2_2_1	2_2_2	2_2_3	–
		45	Samsung Galaxy S5	6.0.1								
		47	Samsung Galaxy J3	5.1.1								
	3	10	Nexus 6	7.0					2_2_1	2_3_2	2_3_3	–
		28	Nexus 6	7.0								
		46	Huawei P9 Lite	7.0								
3	1	23	Nexus 6	7.0	ST/FP	SI/FP	ST/BP	-	3_1_1	3_1_2	3_1_3	–
		26	LG G4	6.0								
		46	Huawei P9 Lite	7.0								
		49	Nexus 6	7.0								
		50	Nexus 6	7.0								
	2	22	Nexus 6	7.0					3_2_1	3_2_2	3_2_3	–
		27	Nexus 6	6.0								
		29	Nexus 6	7.0								
		48	Samsung Galaxy S5	6.0.1								
		51	Huawei P9 Lite	7.0								
4	1	4	Honor 8	7.1	ST/FP	SI/FP	ST/BP	ST/H	4_1_1	4_1_2	4_1_3	4_1_4
		23	Nexus 6	7.0								
		52	Samsung Galaxy S8	7.0								
		53	Nexus 6	7.0								
	2	15	Samsung Galaxy S7	7.0					4_2_1	4_2_2	4_2_3	4_2_4
		17	Samsung Galaxy S7 Edge	7.0								
		27	Nexus 6	7.0								
		28	Nexus 6	7.0								

Table 4
Details about Sessions 5 and 6. ST/SI stand for standing/sitting postures, FP/BP stand for front/back pocket.

Session	Tuple ID	User ID	Device Model	Android	Posture/Receiver Position				Dataset Filename			
				Version	Test 1	Test 2	Test 3	Test 4	Test 1	Test 2	Test 3	Test 4
5	1	4	Nexus 6	7.0	ST/H	ST/FP	ST/H	ST/H	5_1_1	5_1_2	5_1_3	5_1_4
		23	Nexus 6	7.0	ST/FP	ST/BP	ST/BP	ST/FP				
		26	LG G4	6.0	ST/H	ST/FP	ST/FP	ST/H				
		53	Nexus 6	7.0	ST/BP	ST/BP	ST/BP	ST/FP				
	2	10	OnePlus 5	8.0	ST/H	ST/FP	ST/FP	ST/H	5_2_1	5_2_2	5_2_3	5_2_4
		22	Nexus 6	7.0	ST/BP	ST/BP	ST/FP	ST/H				
		28	Nexus 6	7.0	ST/H	ST/FP	ST/H	ST/FP				
		29	Nexus 6	7.0	ST/BP	ST/BP	ST/BP	ST/FP				
6	1	3	Nexus 6	7.0	ST/BP	ST/H	ST/H	–	6_1_1	6_1_2	6_1_3	–
		4	Honor 8	7.1	ST/FP	ST/BP	ST/FP	–				
	2	15	Samsung Galaxy S7	7.0	ST/BP	ST/BP	ST/H	–	6_2_1	6_2_2	6_2_3	–
		29	Nexus 6	7.0	ST/FP	ST/H	ST/FP	–				
	3	10	OnePlus 5	8.0	ST/FP	ST/H	ST/H	–	6_3_1	6_3_2	6_3_3	–
		28	Nexus 6	7.0	ST/BP	ST/BP	ST/FP	–				

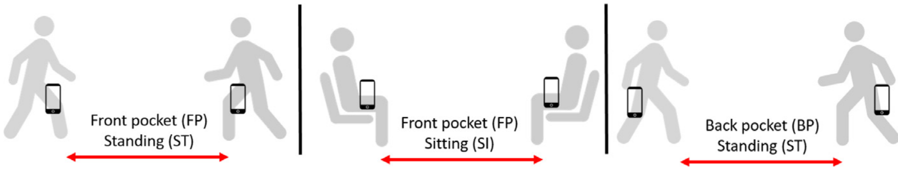

Fig. 4. Layout of the experimental sessions.

The dataset we release also provides a detailed Ground Truth (GT) of all the experiments conducted, as a separate set of CSV files. The name of the files containing the GT exactly matches with the name of the files of the experiments, as shown in Fig. 2. Such files report the start time of the meeting (Interaction) and the end time of the meeting (Non Interaction) according to the following format: `Start Meeting Timestamp, End Meeting Timestamp`.

2. Experimental design, materials, and methods

The experiments described in this work are the result of a data collection campaign designed to capture all the complexities behind the detection of interactions among humans through commercial mobile devices. We consider two main aspects during the design phase of this campaign:

- the use of commercial and standard devices for emitting and receiving beacons;
- the variability of the way subjects act during an interaction.

This section describes both of these aspects. In particular, we start with a description of the hardware and the software, and then we report some preliminary observations concerning the data collected.

The emitting devices are cheap, easy-to-configure and fully compliant with the most common beacon protocols, such as *iBeacon* and *Eddystone*. The advertisement rate of the beacons and the transmission power can be tuned, ranging from 1 Hz to 10 Hz and from −23dBm to 4dBm, respectively. For our experiments, we configure the BLE wristbands with a power of emission to −6dBm and an advertisement rate to 5Hz. We configure the wristbands to emit beacons with iBeacon payload. The payload only carries the ID of the wristband, represented as an integer value. Wristbands are based on silicon material with a small shell needed for plugging the beacon tag itself. The tag is of circular shape with a diameter of 3 cm, the System on Chip (SoC) includes antenna, a coin battery and the chip-set. Wristbands can be configured wirelessly through a mobile app.

For what concerns the receiving devices, we developed an Android application listening for beacon messages and providing a graphical user interface to allow the user to check the status of the experiment and the actions to be taken (e.g., Receiver ID, number of received beacons, start and end time of the encounter). In addition, the app logs the fields reported in Table 1 to the internal storage of the device. Before establishing the final version of the Android app, we investigated several features by exploiting the native functions of the operating systems, but experiencing several technical limitations, as detailed in [1]. The app can run both in foreground and in background and it implements a restarting mechanism in order to recover from crashes or involuntary stops.

Finally, we report some preliminary observations of the collected data with some significant cases. We provide a graphical representation of the RSSI fluctuation of the dyad 22 - 28 in 3 tests of Session 5. Figs. 5–7 show respectively:

- Session 5 - Test 1: user 22 puts the device in BP, user 28 holds the device in H;
- Session 5 - Test 2: user 22 puts device in BP, user 28 puts the device in FP;
- Session 5 - Test 3: user 22 puts device in FP, user 28 holds the device in H.

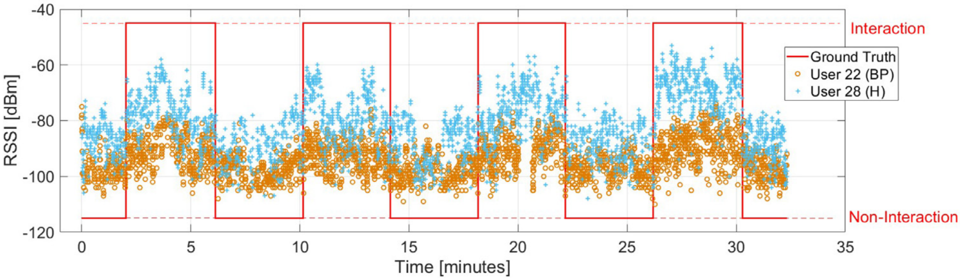

Fig. 5. Example of raw data of RSSI values from dyad 22 and 28 during Session 5, Test 1.

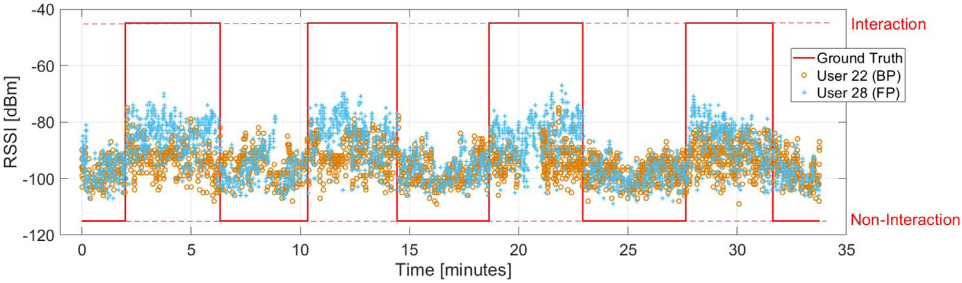

Fig. 6. Example of raw data of RSSI values from dyad 22 and 28 during Session 5, Test 2.

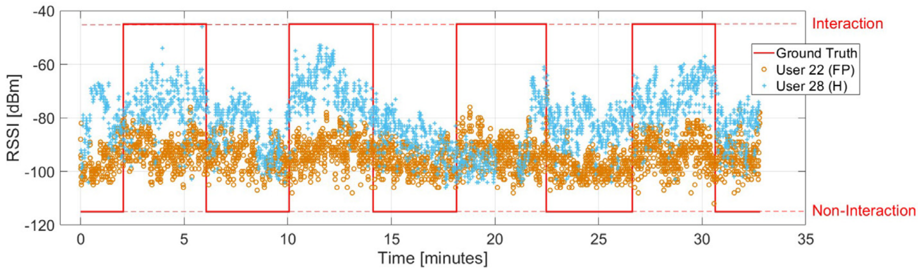

Fig. 7. Example of raw data of RSSI values from dyad 22 and 28 during Session 5, Test 3.

The Figs. 5, 6 and 7 show the RSSI values estimated by the device of user 22 (in orange dots) and of user 28 (in blue crosses) as well as the Ground Truth as a red line. A first observation is that when a device is hold in H, it generally provides high values the RSSI. This is the case of Test 1 and Test 3 for user 28 which hold the device in H. Differently, devices put in FP or BP record low values of RSSI. Such expected behaviour is particular evident in Test 1 and 2 for user 22 holding the device in BP.

A second observation concerns the amount of recorded beacons. The posture of the subjects not only affects the RSSI of the beacons, but also the amount of beacons recorded. We experienced with tests in which the amount of the collected beacons is far lower than the expected value. As a meaningful example, Test 2 provides a worst case in which the amount and the RSSI value of the recorded beacons decrease remarkably. Among the influencing factors we mention: the body posture, the model of the device and the location of the Bluetooth antennas of the recording device.

We finally report in Table 5 all the anomalies we faced during the dataset collection. Specifically, session by session, we list in detail data lacks due to hardware failures or software locks.

Table 5
Anomalies detected during the data collection.

Session	Test	Anomalies
1	2	Run 4: no beacons transmitted by User 23
	3	Runs 3, 4 and 5 not performed
2	all	All Runs: no beacons transmitted by User 23
3	1	All Runs: no beacons transmitted by User 49
	1	Runs 1, 2 and 3: no beacons transmitted by User 22
	2	Runs 3, 4 and 5: no beacons collected by User 51
5	1	All Runs: no beacons transmitted by Users 4 and 53
	2	Run 3: half beacons collected by User 10
	2	Run 4: no beacons collected by User 10
	3	Runs 3 and 4: no beacons collected by User 10
	3	Run 4: no beacons collected by User 26
6	3	Runs 1 and 3: very few beacons collected by User 3
	3	Runs 2: no beacons collected by User 3
	3	Runs 2 and 3: no beacons collected by User 28
	3	Runs 2 and 3: no beacons collected by User 28

Such observations only provide some qualitatively and preliminary insights of the features of the RSSI collected with commercial device during face-to-face interaction. We refer to [1] for a detailed analysis of the presented dataset.

3. Ethics statement

This dataset has been collected with the support of a number of subjects that accepted to join our experiments, after an appropriate information and training phase and by signing an Informed Consent form.

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.

Acknowledgments

This work is partially supported by the Socialize ME project funded by “Fondazione Cassa di Risparmio di Lucca”, a collaboration between CNR-ISTI, CNR-IIT and Polo Scientifico Tecnico Professionale E. Fermi G. Giorgi.

References

[1] M. Girolami, F. Mavilia, F. Delmastro, Sensing social interactions through ble beacons and commercial mobile devices, Pervasive Mob. Comput. 67 (2020) 101198, doi:10.1016/j.pmcj.2020.101198.
[2] E. Hall, C.P.C.L. of Congress, The Hidden Dimension, Anchor books, Doubleday, 1966.