Journal Title 0 (2023) 1–0
IOS Press

Remote Detection of Social Interactions in Indoor Environments through Bluetooth Low Energy Beacons

Paolo Baronti [a], Paolo Barsocchi [a], Stefano Chessa [b], Antonino Crivello [a], Michele Girolami [a], Fabio Mavilia [a] and Filippo Palumbo [a]

[a] *Italian National Council of Research, ISTI-CNR, Pisa, Italy*
E-mail: name.surname@isti.cnr.it
[b] *Department of Computer Science, University of Pisa, Pisa, Italy*
E-mail: stefano.chessa@unipi.it

Abstract. The way people interact in daily life is a challenging phenomenon to be captured and studied without altering the natural rhythm of the interactions. We investigate the development of automated tools that may provide information to the researchers that analyse interactions among humans. One important requirement of these tools is that should not interfere with the subjects under observation, in order to avoid any alteration in the subject's normal behaviour. Our approach is based on the detection of proximity among groups of people that is obtained using commercial wearable wireless tags based on Bluetooth Low Energy (BLE) and a novel algorithm called Remote Detection of Human Proximity (ReD-HuP) that analyses the wireless signal of tags and produce the proximity information. The algorithm, which has been validated against the ground truth of an experimental dataset, achieves an accuracy of 95.91% and an F-Score of 95.79%.

Keywords: Proximity, Bluetooth Low Energy, Social Interactions

1. Introduction

The analysis of social interactions, meant as the tendency of humans to interact with their own kind over time is a complex task being social interactions generally not easy to understand, capture and study with automated tools. However, this can be achieved indirectly by analyzing the features (sociological markers) that characterise the human interactions [1]. Among these markers, it is worth mentioning the co-location (or proximity) of people and their explicit interactions (if they talk to each other, their gestures, etc.). The combination of these markers with their duration over time, their intimacy, and their emotional intensity determine the strength of such interaction, as reported in [2]. In particular, proximity is a mandatory condition to identify ties among humans in the real world, since two humans need to be in the same place at the same time to have a face-to-face meeting. For this reason, the information about proximity can reveal non-trivial social dynamics that, in turn, can be used in many applications, like the detection of crowded areas and the optimisation of the work [3–5], the presence detection in indoor environments [6–9], or the forecasting of the spread of infection diseases [10]. The conventional methods for the detection of proximity rely on questionnaires or direct observations of the subjects under study [11, 12]. However, these methods are feasible for small groups of people, and they may also introduce biases since they may give the users a sense of control and of observation that, in the end, may alter their interactions.

For this reason, we aim at the development of a platform for the automatic detection of proximity among people in indoor environments. This platform is unobtrusive and low-cost, based on off-the-shelves devices (on wristbands and small fixed stations) and on a novel algorithm, called Remote Detection of Human Proximity (ReD-HuP). The platform architecture is modular and follows the paradigm of remote positioning, where the wristbands act as beacon emitters (beacons are short 30 bytes messages sent in broadcast) and the fixed stations, forming the infrastructure, receive the beacons and analyse them. In our specific case, the wristband are Bluetooth Low Energy (BLE) commercial tags, which have been already validated for this purpose in [13–16]. The wristbands are configured with appropriate frequency and power of emission of the beacons, so to limit the number and range of the beacons sent. The stations, in turn, collect all the beacons received, analyse, and aggregate them to identify who among the users are in proximity with each other. The ReD-HuP algorithm has been validated with a dataset produced in a data collection campaign in which we physically simulated real-world interactions and also collected the ground truth.

With respect to the previous work in the literature, the novelty of our approach is in the use of cheap, off-the-shelf commercial devices that are not specifically designed for the detection of proximity, in the use of BLE technology that is now available in most smart devices (like smartphones, wristbands and smart watches), and in the use of the novel ReD-HuP algorithm that runs over a distributed platform of fixed stations and that is based on a voting-based mechanism to aggregate the results computed from all the stations[1].

According to the experiments conducted with our dataset, the proposed algorithm achieves a 95% accuracy and an F-Score of up to 95%. The dataset, which includes about 100.000 beacons, has been constructed with a group of people in a real indoor environment of about $90m^2$ with 4 stations deployed.

The rest of the paper is organised as follows: Section 2 reports the state of the art in this area, sections 3 and 4 present the ReD-HuP algorithm and the experimental settings, respectively. Finally, Section 5 discusses the results obtained and Section 6 presents the conclusions and the work planned for the future.

[1]This work extends our previous work [17] by: a new extensive analysis of state-of-the-art algorithms and techniques in the field of social interactions detection; a novel in-depth analysis of the collected dataset used in the experimentation with a focus on the statistics of the collected RSSI that characterise the functionalities of the distributed system at node level both during the actual social interaction and when no interaction is ongoing; a brand new thoroughly performance analysis of the ReD-HuP algorithm with a new metric (Cohen's kappa coefficient), with a focus on the reliability and scalability of the system. The new analysis is based on the evaluation of the performance (accuracy, F-score, and Cohen's kappa coefficient) when decreasing the number of used fixed stations (up to a single station).

2. Related Work

Automatic proximity detection and localisation [18] has been extensively studied with different technologies. The SocioPatterns[2] platform [19, 20] is an interesting project which aims at studying dynamics of human interactions with real-world experiments. Differently from our approach, the platform is based on custom hardware. Specifically, it is composed by wearable badges equipped with RFID emitters and a number of RFID receivers deployed in the environment. Badges are configured to periodically emit low-power signals, while receivers record the signals and store them for further analysis. Badges are generally worn around the neck and receivers, typically, are deployed on the ceiling, these settings lead to an optimal scenario for detecting proximity among people. A similar approach is represented by the Sociometric Badge [3, 21]. Authors exploit a custom hardware to detect proximity among people and their voice activity. The badge not only emits and records RFID signals, but it also analyses data collected from voice activity, with the goal of detecting when people are in touch and when they are talking to each other. A recent improvement of the Sociometric Badge is the Rhythm platform presented in [4]. Rhythm is also designed for the assessment of human organisations, it is based on a Bluetooth badge acting as emitter and on mobile and stationary clients able to record and to store the data collected. SocioPatterns and the Sociometric Badge represent state-of-the-art platforms for studying dynamics of human interactions. However, they both rely on a custom hardware design which makes it difficult to think to a large-scale experiment. Furthermore, both of the badges are based on the RFID technology not always available on commercial smart phones or wearable smart devices.

A different approach is described in [22]. Authors propose a custom wristband based on BLE beacons to detect social interactions. The wristbands are based on the Android Wear OS and Tizen smartwatches, they emit, collect, and store beacons locally. The data analysis is achieved with a post-processing tool based on a

[2]http://www.sociopatterns.org/

binary classifier. Also in this case, the hardware platform described is not a ready-to-use commercial product. It is worth to notice another interesting project named the Copenhagen Networks Study [23]. Such study involved approximately 800 students from University of Denmark for a period of two years. The study aims to collect different kind of data through a mobile app. Among the data collected, also the proximity among students is detected by exploiting Bluetooth periodic scans and WiFi signals as reported in [24]. Such work increases the dimension of the experiments with a large dataset collected over the years.

Finally, a BLE-based approach for proximity detection have been proposed in [25]. Authors show a system able to recognise social interactions by utilizing commercial smartphone devices. Through the evaluation of Bluetooth RSSI values as input of a machine learning model, authors classified the proximity between two devices into three interaction zones.

Similar approaches based on radio signals involve Wi-Fi information are described in [26]: authors exhibits a system based on commercial smartphones able to recognize spatial settings between subjects and to utilise the built-in accelerometer for speech activity identification. The idea to utilise commercial smartphone has been also proposed in [27]. Authors present a sensor-driven social sharing application within the working environment of a research institution. Their study involves conversation monitoring and interaction with physical objects but authors report findings regarding privacy and user experience issues in terms of acceptability of such services by the users.

It is worth mentioning that some of the above-described technologies can be used as basis for a strictly-related application as indoor positioning and localisation or more generic [28–32] location-based services [33–37].

Besides radio-based technologies, in the literature are present video-based and/or audio-based approaches. In [38], authors present a system able to recognise human proximity with the use of audio sensing and a combination of the smartphone's accelerometer and microphone. The system proposed by the authors performs the voice profiling of the users and is able to detect real-time conversation. Unfortunately, their approach requires a continuous audio recording from each user's smartphone and, consequently, it has a big impact on the battery energy consumption. On the other hand, video-based techniques have been proposed in [39, 40]. In [40], authors propose a system able to detect social events using a ground-truth manually annotated. In [39], authors show a video-based approach able to track people or a single user showing interesting performances on a dataset collected on an outdoor scenario. However, video-based and audio-based approaches are resource-hungry. Furthermore, it is important to highlight that these approaches also bring more sensitive issues about privacy. Consequently, systems based on radio technologies are the most promising for detecting social interactions in indoor environment.

2.1. *Bluetooth Low Energy working principles*

The Bluetooth Low Energy (BLE) technology has emerged as a communication protocol enabling the Internet of Things paradigm due to its low energy consumption and to its wide support by end user devices. Basically, a BLE-equipped device advertises its presence to other devices. The protocol establishes three functional modes: idle, device discovery, and connection.

During the idle mode the device does not transmit or receive packets. The discovery mode allow three states: initiating, scanning and advertising. During the initiating state, the BLE device tries to initiate a connection with other devices. In scanning state, the device actively looks for advertisers. Basically, the device periodically scans the advertising channels and listens to advertising information sent by other devices. In the advertising state, the device periodically sends advertisement packets through a broadcasting operation. Finally, during the connection state, the BLE device can serve in slave or master role. If the previous state of the device was initiating, the device assumes the master role and set the transmission timing. A master device can be connected to multiple slaves. Otherwise, the device assumes the slave mode and it can be connected to one master device only. In advertising mode, a device transmits advertising packets. The advertising packets may contain a data payload and can be forwarded towards a specific device. Generally, the advertising events can be undirected and directed. An undirected advertising is used for detecting unknown devices and allows different responses. A directed advertising is used for establishing connections with already known devices.

3. Proximity detection through voting strategy

This proposed solution addresses the discovery of interactions among people inside an indoor environment. The implemented algorithm, called Remote Detection of Human Proximity (ReD-HuP), is able to detect an intentional proximity event among human beings. This section describes the algorithm and shows the proposed sensing architecture in order to build a dataset of BLE measurements for testing and validation purposes.

3.1. Human proximity detection through the ReD-Hup algorithm

A social interaction happens if a dyad of people stay at the same point at a short distance. We first set the lower bound of the distance ranging from 0.5 to 1.5 meters [41]. If the distance exceed the range we consider the proximity event as unintentional.

The proximity event is evaluated through the analysis of the BLE beacons. In this paper, the beacons are emitted by wearable devices (tags). In particular, BLE tags are off-the-shelf devices able to transmit and receive BLE beacons. A BLE device first set the transmission frequency and the power intensity and then it send beacons periodically. These beacons, in BLE technology they are also called advertisements, can be listened and received from nearby devices. A receiver device which collects a beacon estimates the Received Signal Strength Indicator (RSSI) in a decibel scale (dbm) and, potentially, it can be able to estimate the distance between the transmitter and itself [42–44].

In our experimental setting, we deploy transmitter stations in which the algorithm runs to collect and to analyse the beacons sent by tags. ReD-Hup collects beacons and RSSI values from tags worn by two users that are sending beacons at the same time.

As an example, Figure 1 shows data collected during a proximity event among two users. During this event, the users have a face-to-face meeting 4 minutes long. Successively, the meeting ends for 4 minutes. The red continuous line in Figure 1 represents the meeting ground-truth. The figure also shows the raw RSSI measurements gathered by each station deployed into the environment under different conditions. In fact, the graph on the top shows raw data collected from the closest station to the point in which the meeting occurs. In this case, the distribution of the RSSI values has a clear pattern in which the values increases during the meeting and decreases when the meetings ends and the users leave the meeting point. The meeting detection is not a trivial task, in fact, the second case of the figure shows a meeting that produces patterns not clearly recognizable. In fact, the RSSI measurements collected have a similar distribution. It is worth noting that the performances are strongly related to the availability of the raw data. In fact, in a real scenario, the quality of the beacon measurements depends on factors (e.g., users' body orientation, multipath effects). The ReD-HuP algorithm proposed in this paper take into account all these factors.

ReD-HuP is able to to estimate the time intervals during which a couple of people are in proximity. The main idea behind the algorithm is to perform a voting system considering the information retrieved from each station deployed into the environment. The voting process is performed through two steps.

Initially, the algorithm analyses the RSSI measurements received from all the station. Basically, each station votes 1 if a proximity events is detected, -1 if a no proximity events is detected and 0 otherwise. These output are produced according to the analysis of the RSSI values sent by a couple of tags during a time window of duration τ. Into the algorithm two threshold are set, namely σ_{RSSI} and Δ_{RSSI}. The threshold σ_{RSSI} limits the analysis to pairs of users that are both close enough to the fixed station that is analysing their RSSI. This value is set because we empirically observed the performances improves if data from two users that are far away from the fixed station are not considered. In other word, σ_{RSSI} set the limit for considering two users close enough to the fixed station. The value of Δ_{RSSI} is the difference between the RSSI received from both users; if Δ_{RSSI} is small enough means, the possibility that a meeting is ongoing is high.

The second phase of the algorithm mixes together the votes expressed by each station and produces the final output for a specific time interval. Basically, if the majority of the stations votes proximity, the sum of all votes are greater than 0. Then, the algorithm returns proximity for the couple of users. If the sum is less than 0, the ReD-HuP algorithm returns non-proximity. The sum is equal to 0 may mean that the stations are not able to provide a vote or that some stations voted proximity and others non proximity. In this case, the algorithm gives priority to the station that recorded with higher accuracy beacons emitted by the couple of users. Basically, each station evaluates and stores the RSSI mean value for each couple of users. If the sum is equal to 0, the algorithm provides the same output of the station in which the mean value is max. Differently, ReD-HuP is not able to detect proximity for the current time interval, and it returns the previous valid response (if available). More details on the implemented algorithm can be found in [17].

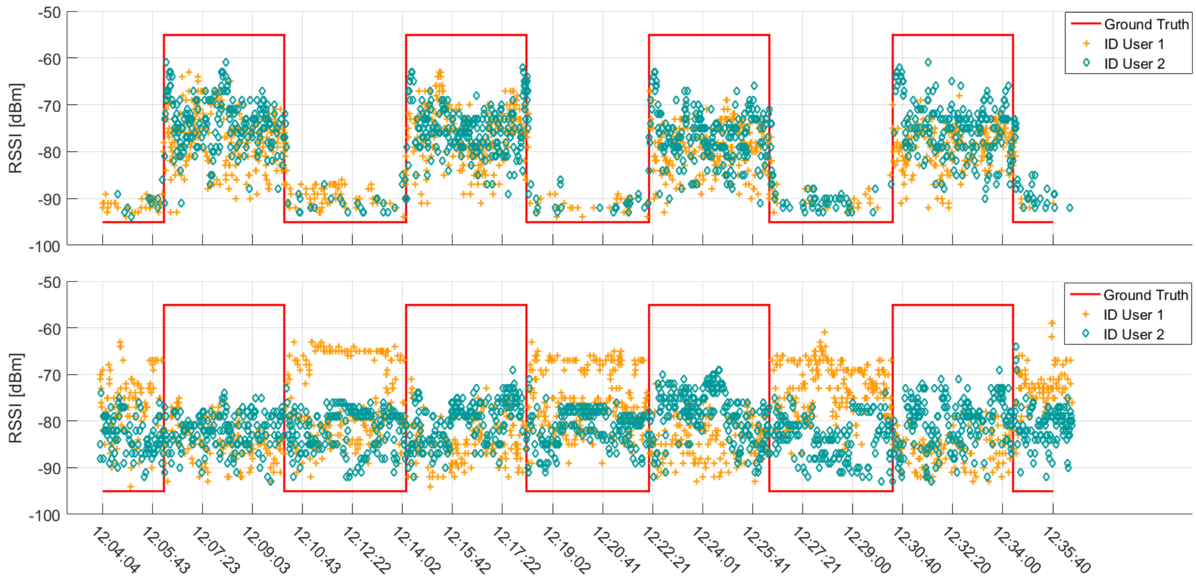

Fig. 1. Examples of data collected during a proximity event among two users (dyads) gathered by each station deployed into the environment under different conditions (on top: raw data collected from the closest station to the point in which the meeting occurs; on the bottom: a meeting that produces patterns not clearly recognizable).

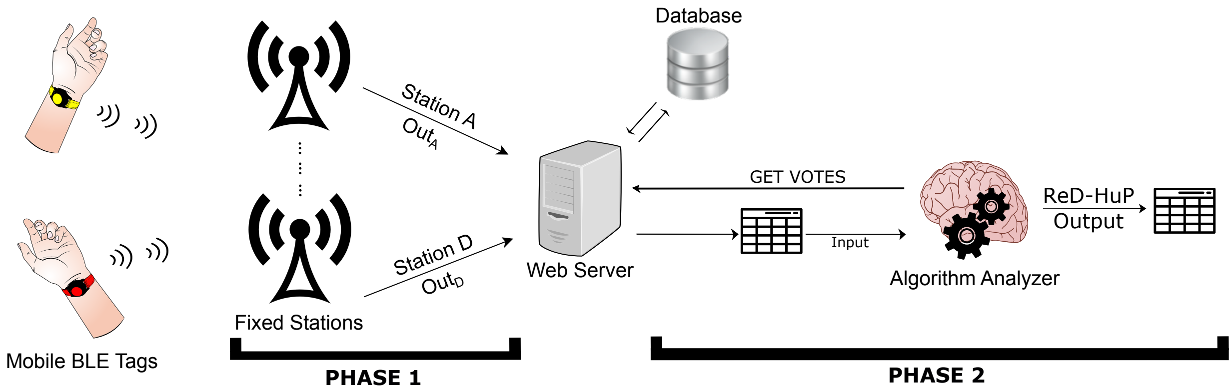

Fig. 2. The sensing architecture: users wear wristbands emitting BLE beacons at a given frequency and power; beacons are collected by fixed stations deployed in the environment; stations periodically produce a vote for every pair of users according to the ReD-Hup algorithm.

3.2. Sensing Architecture for Proximity Detection

In order to collect and analyse the RSSI values of BLE tags, we designed a distributed architecture as reported in Figure 2. Users wear BLE wristbands that only emit BLE beacons at a given frequency and power of emission. Beacons are collected by a number of fixed stations deployed in the environment. Stations periodically produce a vote for every pair of users according to the ReD-Hup algorithm. The available votes are: proximity (value 1); no proximity (value -1); no answer (value 0) for a pair of users. The Web Server runs the phase 2 of the ReD-HuP algorithm: it queries all the stations to retrieve the votes generated and it takes a final decision based on the ReD-Hup algorithm.

Table 1

Data format of each BLE signal received by stations and stored in the database.

Epoch-Time	Date-Time	Type	ID-Receiver	ID-Sender	RSSI

Note that stations store in their database only a limited number of beacon information, Table 1 shows the fields currently stored in the local databases.

4. The Experimental Dataset

We tested the ReD-HuP algorithm with a real-world dataset. The dataset mimics real-world face-to-face interactions among people in our working environment. This section firstly describes how we organise the data collection campaign (Section 4.1) and then it provides an analysis of the data we collect with the goal of assessing the quality of the dataset (Section 4.2).

4.1. The Testing Environment

The dataset has been collected in our research institute, namely ISTI-CNR[3] located in Pisa, Italy. We selected one of the wings of ISTI-CNR, whose layout is reported in Figure 3. The map covers approximately 130 m^2 with 4 offices (referred to as A to D) opening on the right and left side of the main corridor. The offices are approximately 25 m^2 and they are equipped with common working furniture such as wood-based desks, metal closets and chairs. The floor is made of tiles of 60x60 cm^2, Figure 3 shows a super-imposed grid.

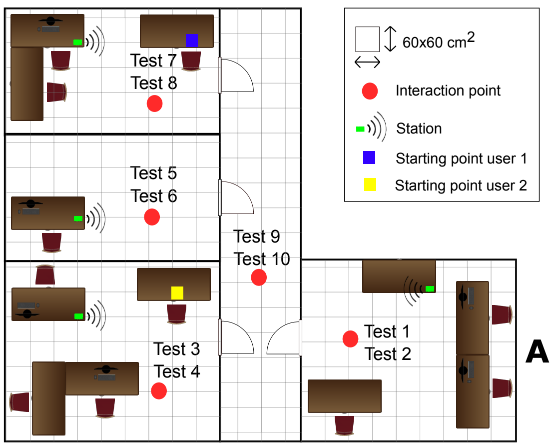

Fig. 3. Map of the testing area: approximately 130 m^2 with 4 offices (referred to as A to D) approximately 25 m^2 equipped with common working furniture.

Interactions are reproduced with a pair of volunteers (the dyad) conducting several tests in 5 different locations shown as red dots in Figure 3. In each of the locations, the dyad executes 2 tests for a total of 10 tests. Each test consists of 2 stages: 4 minutes or Non Interaction followed by 4 minutes of interaction. During the Non Interaction stage, volunteers rest in their starting point (shown as blue and yellow markers in Figure 3), while during the Interaction stage volunteers move close to each other at a distance of approximately 1.5 meters in one of the 5 locations.

During the tests, volunteers are asked to wear a commercial BLE wristband produced by Global Tag[4]. The wristband is equipped with a easy-to-configure and fully compliant Bluetooth v4 chipset designed to advertise BLE beacons in broadcast. Every wristband supports several settings. In particular it is possible to configure the beacon protocol (Eddystone or iBeacon), the beacon advertisement rate (from 1Hz to 10Hz) and the transmission power (from -23 dbm to 4 dbm). We configure the wristbands with the iBeacon protocol set at 2Hz and -6 dBm.

The environment is also equipped with several fixed stations, shown as green markers in Figure 3. The stations are powered with Raspberry PI boards provisioned with an USB Bluetooth dongle, namely the BLED112 from Bluegiga[5]. Stations are connected to a local network via Ethernet wired cable. In particular, the life-cycle of the Java code is designed to perform one single operation: stations listen for BLE beacons sent by the wristbands. Beacons are uploaded remotely with a REST API that we also coded. The API allows to upload a BLE beacon as a JSON object and to store the object in a NO-SQL database. To this end, we configured a MongoDB collection.

Figure 2 reports the whole monitoring infrastructure used for our tests. We deploy 4 stations in our environment, one in each office. Stations are generally deployed over desks and they are connected to the power line.

Table 2 reports the distance (in meters) between each of the 4 stations and the starting point of the two volunteers. We record a maximum distance of 8.4 meters (S_A - User 1) and a minimum distance of 2.4 meters (S_B - User 2).

We also asked to volunteers to log in diary the starting and ending time of each of the interactions that they had, as well as any important remarks during the tests. Such log represents our ground-truth that we

[3] lat: 43.718302, lon: 10.422085

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

[5] https://www.bluegiga.com/

Table 2
Distances in meters between stations and each of the volunteer's starting points.

User ID	Distance from starting point [m]			
	S_A	S_B	S_C	S_D
1	8.42	5.47	3.01	2.47
2	7.25	2.40	3.10	6.02

used to assess the performance of the ReD-HuP algorithm.

4.2. Analysis of the Dataset

We now analyse the quality of the dataset. The analysis first focuses on the amount of beacons collected and lost during the tests, and then it provides a detailed analysis of the RSSI distribution reported from all the stations both during the Non Interaction and Interactions stages.

Table 3 provides an overview of the tests. The table shows for each interaction point (Room A, B, C, D and Corridor): the tests performed in such point, the distance (in meters) from the stations, the number of BLE beacons collected and the duration of the tests.

Our dataset comprises a total of about 100.000 beacons recorded from 5 stations, with an average of 19.878 beacons recorded at each interaction point. We experience that the 5 stations do not record the same amount of beacons in each interaction point. In particular, tests in Room B report the minimum number of BLE beacons (17.970), while tests in Room C report the highest value (22.572). Such differences can be caused by multiple factors. As for example, the presence of obstacles between a station and a wristband, the posture of the volunteers during the interactions and any wireless interference are all factors that might decrease the amount of beacons that a station records.

We further investigate the ratio between the amount of beacon recorded with respect to the amount of beacons expected, we refer to such ratio as the beacon loss rate. Table 4 reports the beacon loss rate computed for all the stations and for all the 10 tests conducted. The table also reports the average beacon loss rate for each station during the 10 tests (last column of Table 4). The amount of beacons expected is determined by the advertisement rate of the wristband set to 2Hz for our tests. As expected, the beacon loss rate increases with the distance between emitter and receiver. As a general trend, the higher the distance the higher the beacon loss rate. However, this trend is not always confirmed. As a meaningful example, we observe that the beacon loss rate of Station A for tests 3 and 4 in Room B is higher than that of tests 5 and 6 in Room C (94.2 dbm - 94.1 dbm for tests 3 - 4, 80.1 dbm - 79.4 dbm for tests 5 -6) even if the distance between Station A and Room B is lower than that of Room C (6.79, 7.75 respectively). Stations report a beacon loss rate ranging from 59.8% (Station B) to 84.7% (Station A) with an average of 67.9% for all the stations.

We also compute the beacon loss rate during each test, it ranges from 63.9% in test 5 to 73.7% in test 3 with an average of approximately 67%.

Furthermore, we study how stations located in different locations record beacons during the same test. In particular, we analyse the RSSI variations of beacons recorded from 4 stations at different distances from the same interaction point during the Interaction and Non Interaction stages. To this purpose, we decided to consider a meaningful example able to highlight some interesting features of RSSI variation among the station. We select test 7 located in Room D, and we analyse the RSSI values from Station A to D, as shown in Figure 4 and Figure 5. Figure 4 shows two overlapping time series, one for each of the volunteer joining test 7. The figure also shows as a blue-star line the ground truth, namely the time intervals during which the dyad is interacting. Station C is located in Room C and it is placed at 2.9 meters away from Room D. The time series recorded are similar most of the time since Station C is not able to clearly distinguish between the Interaction and Non Interaction. Differently, Station D is located at 2.9 meters away from Room D and it records RSSI values that change more clearly during the time. Station C and D represent two opposite cases. In particular, Station C represents our worst case in which the beacons collected are not so helpful to distinguish between Interaction and Non Interaction. Station D represents our an optimal case, not only for the amount of beacons collected but also for the evident fluctuation of the values of RSSI during Interaction and Non Interaction. In this last case, it is easy to distinguish between Interaction (increase of the value of the RSSI) and Non Interaction (decrease of the value of the RSSI). Stations A and B are also interesting cases. Station A is located 7.94 meters away from Room D, it collects few beacons with apparently few fluctuations during the two stages. Station B is located 5.2 meters away from Room D. User 1 remains in Room D for all the time, since the interaction point and the starting point are on Room D. Differently, User 2 moves from

Table 3

Dataset overview.

Interaction point	Tests	Distance from interaction point [m]				Collected Beacons	Duration of each test [min]
		S_A	S_B	S_C	S_D		
Room A	1,2	3.34	6.92	7.25	9.57	19225	32.2
Room B	3,4	6.79	3.13	4.17	7.08	17970	32.5
Room C	5,6	7.75	3.29	1.52	3.67	22572	32.2
Room D	7,8	7.94	5.27	2.91	2.9	19445	32.4
Corridor	9,10	5.3	4.95	3.95	5.72	20149	32.3

Table 4

Ratio between the amount of beacon recorded with respect to the amount of beacons expected, we refer to such ratio as the Beacon Loss Rate.

Station	User ID	Test										Total
		1	2	3	4	5	6	7	8	9	10	
A	1	73.196	71.912	94.285	94.177	80.196	79.406	95.092	87.513	83.196	80.963	**84.772**
	2	71.186	73.928	93.132	89.61	84.54	81.835	93.43	94.221	89.038	84.183	
B	1	58.015	57.855	64.172	60.159	58.195	57.674	63.522	61.455	65.98	70.067	**59.852**
	2	62.268	57.339	55.305	53.617	57.316	55.866	59.202	63.416	57.874	57.702	
C	1	65.103	64.393	70.323	62.724	59.928	59.302	60.583	61.429	61.58	62.983	**64.097**
	2	76.443	72.946	71.912	63.751	59.074	58.579	66.028	64.19	60.911	59.686	
D	1	69.794	68.786	67.581	59.03	52.74	55.504	54.908	52.58	59.701	60.046	**63.38**
	2	80.747	79.948	74.27	65.162	58.842	57.597	62.091	60.939	61.297	66.023	
Total		**69.59**	**68.47**	**73.72**	**68.38**	**63.97**	**63.32**	**69.10**	**68.25**	**67.40**	**67.69**	

its starting point located in Room B to the interaction point located in Room D. The two time series obtained from Station B well reflects such mobility. In particular, the RSSI values of User 2 increase while the dyad is *not* interacting and they decrease while they are interacting.

We finally analyse the probability density function of the RSSI in order to better understand the contribution of RSSI during the Interaction and Non Interaction stages, as shown in Figure 6. The figure shows, for each station, the probability density of User 1 (red curve) compared to User 2 (blue curve) during Interaction and Non Interaction by considering all the tests. Each density also shows a super-imposed line useful to better understand the trend, the line is obtained with a Kernel Density Estimation (KDE) interpolation. As previously discussed, some stations cannot clearly distinguish between beacons emitted during Interaction and Non Interaction stages. This is the case of Stations C in which the densities during Interaction and Non Interaction are very similar when compared to the ones obtained from Station D. It is worth to notice that also Station A provides useful values of RSSI. In this last case, even if the amount of beacon is limited, the densities are different during the two stages.

Out of the 10 tests described in Table 4, Test 7 represents a good benchmark for our experimental analysis for two main reasons:

– Stations A to D exhibit remarkable different trends of the RSSI collected during the Interaction and Non Interaction stages. Such trends will be further analysed in the following section. In particular, we will show the optimal case of Station D that clearly distinguish between the two stages, and the worst case of Station C that does not provide clearly such a distinction. Differently, stations A and B also provide some interesting corners-cases. Such consideration are better reflected with test 7 with respect to the other ones.

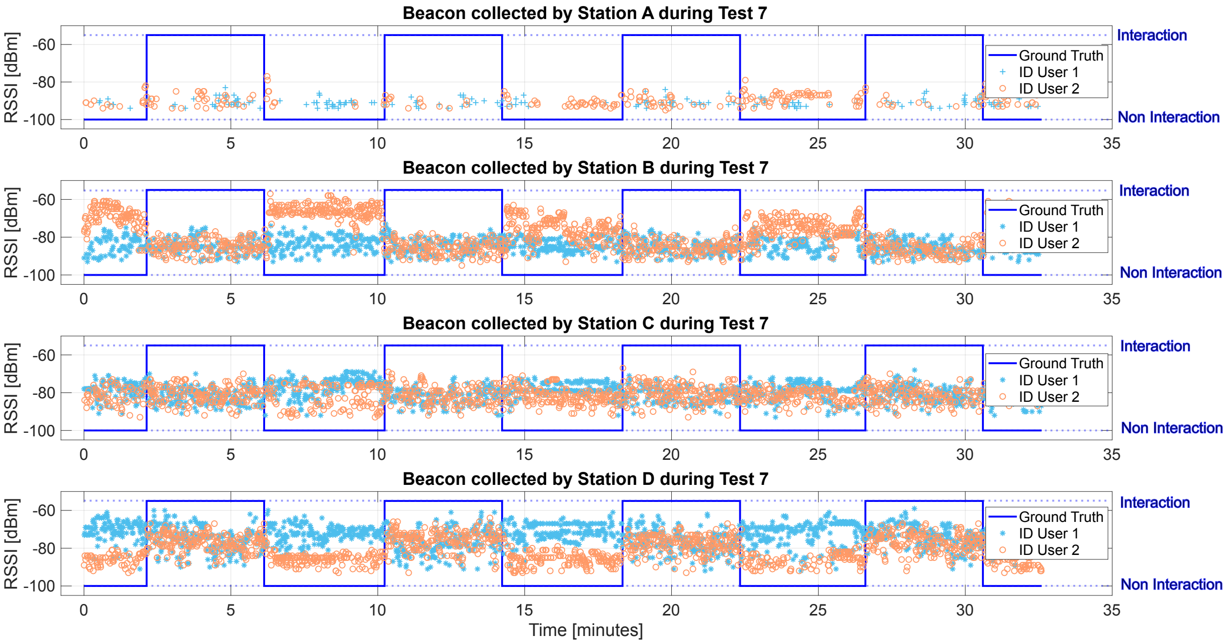

Fig. 4. RSSI variations of beacons recorded from 4 stations at different distances from the same interaction point during the Interaction and Non Interaction stages.

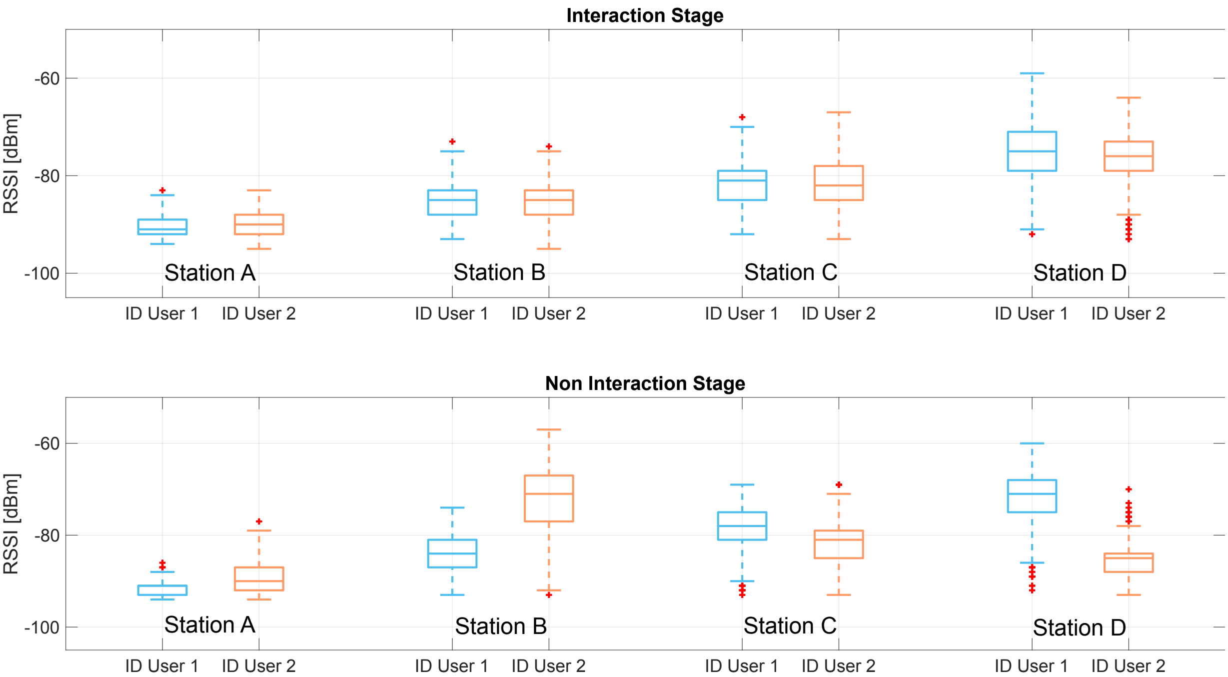

Fig. 5. Variation of the RSSI recorded from different stations as box plots.

– We selected a test with a considerable percentage of beacon loss so that to discuss features of the RSSI in a typical scenario and not in an ideal condition. To this purpose and based on the pre-

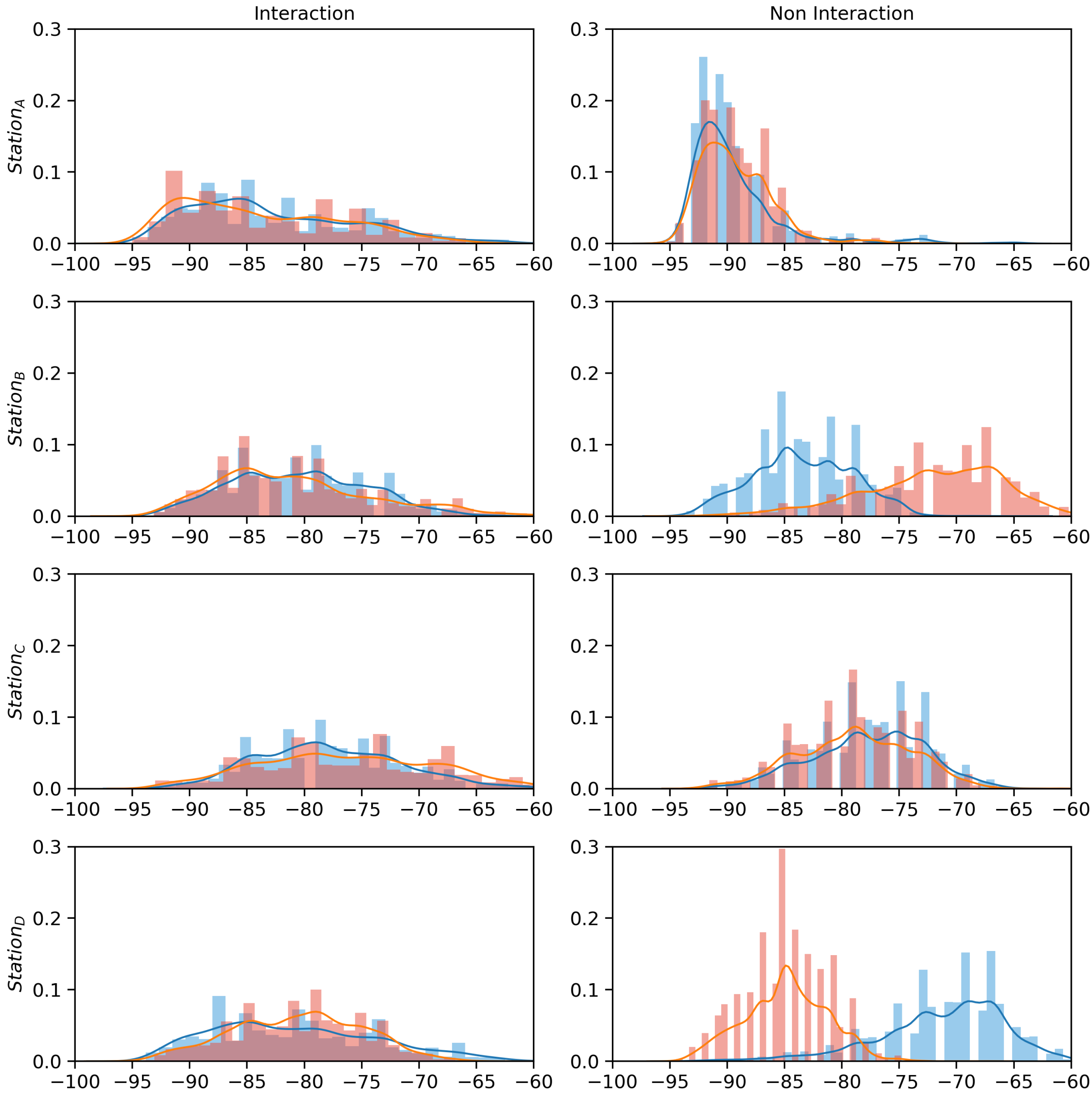

Fig. 6. Probability density function of the RSSI during Interaction and Non Interaction stages. For each station, the probability density of User 1 (red curve) is compared to User 2 (blue curve) during Interaction and Non Interaction by considering all the tests.

vious consideration, we selected test 7 which results with 69% of beacon loss rate (that is, the 80th percentile of the beacon loss percentages we measured).

5. Experimental Settings and Results

In order to assess the performance of the solution proposed, we compare the results of ReD-HuP against the ground truth. As a meaningful example, we show in Figure 7 a graphical representation of the performance of ReD-HuP in Test 7. Figure 7 shows a time series of the outputs computed by ReD-HuP with red dots. We also report as a star-line the ground truth, namely the time intervals during which the two volunteers had a social interaction. The figure shows qualitatively that ReD-HuP is able to detect correctly all the interactions and non interaction during the 4 runs of Test 7.

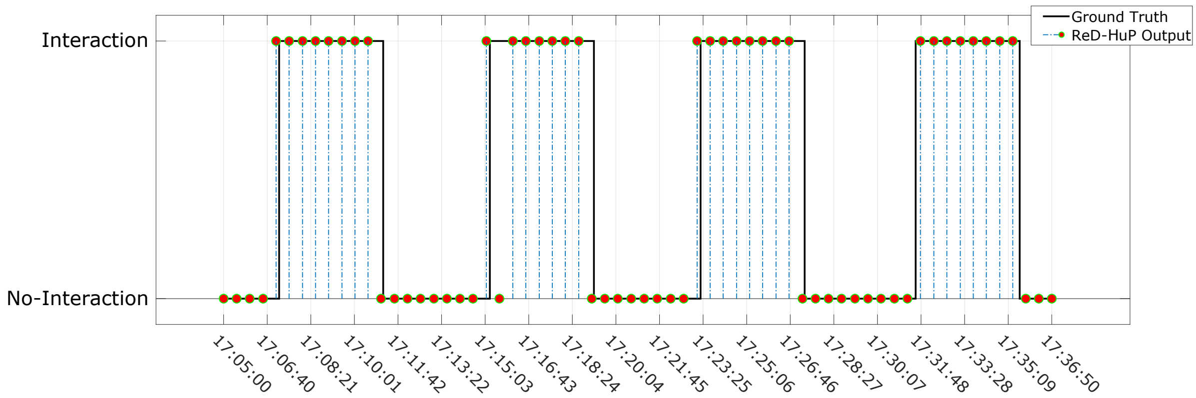

Fig. 7. Time series of the classification result of the ReD-HuP with respect to the ground truth in Test 7: the outputs computed by ReD-HuP (red dots); ground truth (blue star dots). ReD-HuP is able to detect correctly all the interactions and non interaction during the 4 runs of Test 7.

In order to measure quantitatively the performance of ReD-HuP, we build the confusion matrix. It measures the following quantities: true positive (TP), true negative (TN), false positive (FP) and false negative (FN) answers. These metrics assess the number of right/wrong answers given by ReD-HuP, with respect to the events reported by the ground truth diary (i.e. the existence or not of an interaction for the dyad).

Given the confusion matrix, we consider the following 3 evaluation metrics, namely Accuracy, F-Score and Kappa. Accuracy is given by:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

it measures the proportion of correct answers of the algorithm with respect to the total amount of observations. F-Score is defined as:

$$F - Score = \frac{2 * P * R}{P + R}$$

and it combines both precision $P = TP/(TP + FP)$ and recall $R = TP/(TP + FN)$. When $F - Score = 1$, the algorithm obtains perfect precision ($P = 1$) and perfect recall ($R = 1$). Finally, we consider Kappa coefficient defined as:

$$Kappa = \frac{Accuracy - RA}{1 - RA}$$

where Random Accuracy (RA) is the proportion of agreements expected purely by chance. RA is defined as $RA = \frac{(TN+FP)*(TN+FN)+(FN+TP)*(FP+TP)}{N}$, with N being the total number of occurrences. Kappa coefficient measures the inter-rate reliability of the two classes and how much homogeneity exists between them. This coefficient is important in the case that one of the classes to be identified is much more prevalent with respect to the others. In this case, a generic classification algorithm might result with a high value of accuracy, but with a a low value of Kappa coefficient.

We first consider the two parameters of our algorithm, namely σ_{RSSI} and Δ_{RSSI}. We evaluate the overall accuracy, F-Score and Kappa for all the tests, by varying such parameters from 0.5 to 10 (with a step of 0.5) and from -83dbm to -87dBm for Δ_{RSSI} and σ_{RSSI} respectively.

Figure 8 shows a graphical representation of accuracy and F-score metrics. From the figure we observe that the two graphs have a similar trend: they increase as Δ_{RSSI} increases up to the maximum value of $\Delta_{RSSI} = 2.5$. Moreover we observe that the trend of the two metrics decrease with values of $\Delta_{RSSI} \geqslant 5$. It is worth to notice that values of Δ_{RSSI} and σ_{RSSI} outside the ranges reported in Figure 8 do not affect significantly the accuracy and F-Score.

From such results, we can derive the optimal setting of ReD-HuP:

$$\Delta_{RSSI} = 2.5, \sigma_{RSSI} = -85dBm \qquad (1)$$

With such settings, we obtain an overall accuracy of 95.91% and F-Score of 95.79%. We further analyse the performance of ReD-HuP by analysing the Kappa coefficient with the optimal setting, as shown in Figure 9. At this conditions, we obtain a value of Kappa of 92% that confirms the effectiveness of ReD-HuP in detecting social interactions. Finally, we report on Table 5 the cumulative confusion matrix obtained by consider-

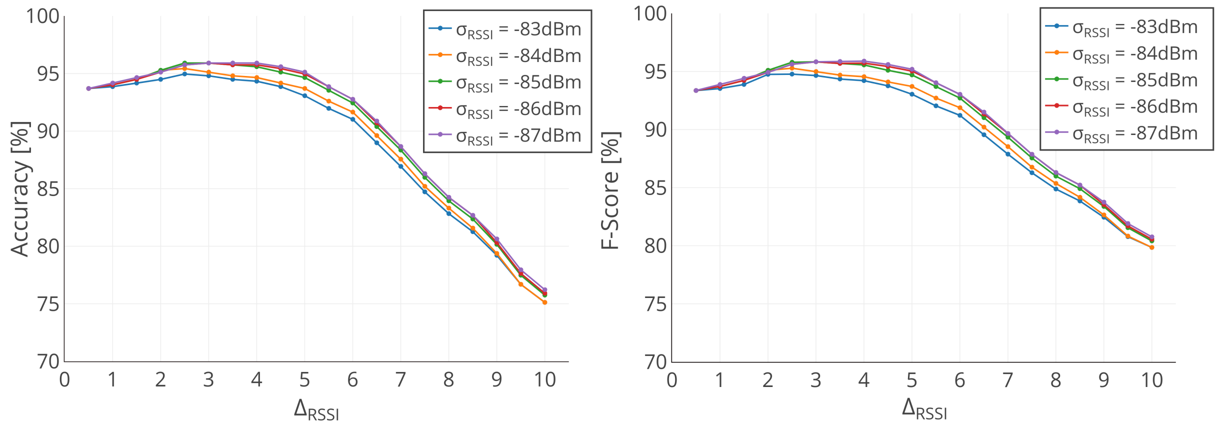

Fig. 8. Variation of Accuracy and F-Score with σ_{RSSI} and Δ_{RSSI}. The two graphs increase as Δ_{RSSI} increases up to the maximum value of $\Delta_{RSSI} = 2.5$, while the trend of the two metrics decrease with values of $\Delta_{RSSI} \geqslant 5$.

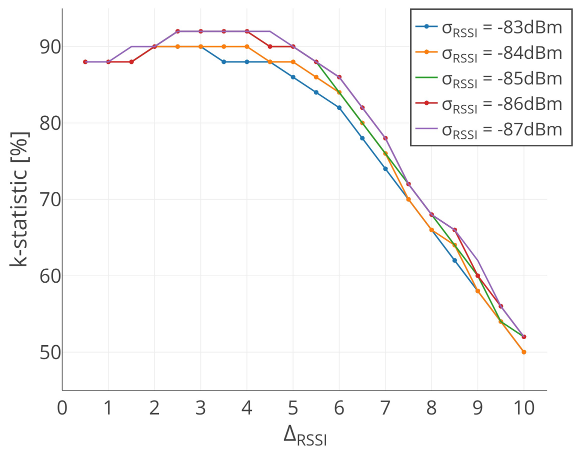

Fig. 9. Variation of Kappa coefficient with σ_{RSSI} and Δ_{RSSI}. We obtain a Kappa coefficient with the optimal setting of 92% that confirms the effectiveness of ReD-HuP in detecting social interactions.

ing the optimal configuration of σ_{RSSI} and Δ_{RSSI} reported in (1).

The results presented in Figure 8 and 9 have been obtained by exploiting the 4 available voting stations A to D, as reported in Figure 3. We finally analyse the robustness of ReD-HuP by reducing progressively the number of stations moving from 4 to 1 single station. The process we follow consists in computing the accuracy, F-Score and Kappa with 4, 3, 2 and 1 station and by considering all the available combinations. As a result, we tested ReD-HuP with 14 different settings, as shown in Table 6. From the table we observe that ReD-HuP is able to provide high performance with at least one combination of stations. In particular, we report below the optimal values obtained with the number of voting stations and their layout:

- with 4 stations ReD-HuP obtains 95.91 %, 95.79 % and 92% of accuracy, F-score and Kappa respectively;
- with 3 stations and layout S_A S_B S_D ReD-HuP obtains 87.23%, 87.07% and 74.45% of accuracy, F-score and Kappa respectively;
- with 2 stations and layout S_A S_D ReD-HuP obtains 89.72%, 89.94% and 79.44% of accuracy, F-score and Kappa respectively;
- with 1 stations and layout S_D ReD-HuP obtains 85.67%, 85.11% and 71.33% of accuracy, F-score and Kappa respectively;

6. Conclusions

The automatic detection of proximity among humans is a useful tool to support the analysis of human social interactions and the studies in social networking. We have presented a platform that makes use of low-cost, off-the-shelf commercial BLE wristbands and a small network of fixed stations that receive the BLE beacons emitted by the wristbands and execute ReD-HuP, a novel algorithm for proximity detection. This approach opens a new perspective on the detection of social interaction, as, in the future, BLE beacons can even be easily embedded into clothes, thus further reducing the impact for the users.

In our future work we plan to further pursue this approach, by further developing the ReD-HuP algorithm.

Table 5
Cumulative confusion matrix.

	True conditions: Interaction	True conditions: No-Interaction	Accuracy [%]	F-Score [%]	K-statistic [%]
ReD-HuP: Interaction	296	3	95.91	95.79	92.00
ReD-HuP: No-Interaction	23	313			

Table 6
Performance of ReD-HuP algorithm by varying the number of voting stations.

Number of stations	Stations Layout	Accuracy [%]	F-Score [%]	K-statistic [%]
4	S_A S_B S_C S_D	95.91	95.79	92.00
3	S_A S_B S_C	68.22	63.83	36.40
	S_A S_B S_D	87.23	87.07	74.45
	S_A S_C S_D	75.23	74.15	50.45
	S_B S_C S_D	85.20	85.54	70.41
2	S_A S_B	84.58	84.98	69.16
	S_A S_C	55.30	51.44	10.55
	S_A S_D	89.72	89.94	79.44
	S_B S_C	68.22	67.41	36.44
	S_B S_D	84.58	83.95	69.15
	S_C S_D	71.03	65.56	42.00
1	S_A	54.52	53.80	9.03
	S_B	83.84	83.82	67.26
	S_C	57.17	55.72	14.31
	S_D	85.67	85.11	71.33

To this purpose, we also plan to extend the experimental campaign to collect a larger dataset, also with settings including more than two people, in order to optimise the two parameters σ_{RSSI} and Δ_{RSSI} of ReD-HuP. We will also aim at the automatic reconfiguration of these two parameters in order to avoid the additional calibration of the system when it is deployed in a new environment.

Acknowledgements

This work has been partially funded by the EU H2020 project "Novel Empowering Solutions and Technologies for Older people to Retain Everyday life activities" (NESTORE), GA769643.

References

[1] M. Gupte and T. Eliassi-Rad, Measuring tie strength in implicit social networks, *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12* (2012).

[2] M. Granovetter, The Strength of Weak Ties: A Network Theory Revisited, *Sociological Theory* **1** (1983), 201–233. http://www.jstor.org/stable/202051.

[3] J.-P. Onnela, B.N. Waber, A. Pentland, S. Schnorf and D. Lazer, Using sociometers to quantify social interaction patterns, *Scientific Reports* **4** (2014), 5604 EP -.

[4] O. Lederman, A. Mohan, D. Calacci and A.S. Pentland, Rhythm: A Unified Measurement Platform for Human Organizations, *IEEE Multimedia* **25**(1) (2018), 26–38. doi:10.1109/MMUL.2018.112135958.

[5] A. Crivello, F. Mavilia, P. Barsocchi, E. Ferro and F. Palumbo, Detecting occupancy and social interaction via energy and environmental monitoring, *International Journal of Sensor Networks* **27**(1) (2018), 61–69.

[6] P. Barsocchi, A. Crivello, M. Girolami, F. Mavilia and E. Ferro, Are You in or Out ? Monitoring the Human Behavior Through an Occupancy Strategy, in: *MoCS*, 2016. ISSN 15301346. ISBN 9781509006793. doi:10.1109/ISCC.2016.7543732.

[7] D. Bacciu, S. Chessa, E. Ferro, L. Fortunati, C. Gallicchio, D. La Rosa, M. Llorente, A. Micheli, F. Palumbo, O. Parodi et al., Detecting socialization events in ageing people: the experience of the DOREMI project, in: *2016 12th International Conference on Intelligent Environments (IE)*, IEEE, 2016, pp. 132–135.

[8] C. Gallicchio, A. Micheli, P. Barsocchi and S. Chessa, User movements forecasting by reservoir computing using signal streams produced by mote-class sensors, in: *International Conference on Mobile Lightweight Wireless Systems*, Springer, 2011, pp. 151–168.

[9] D. Bacciu, C. Gallicchio, A. Micheli, S. Chessa and P. Barsocchi, Predicting user movements in heterogeneous indoor environments by reservoir computing, in: *Proc. of the IJCAI Workshop on Space, Time and Ambient Intelligence (STAMI), Barcellona, Spain*, 2011, pp. 1–6.

[10] R. Mastrandrea, J. Fournet and A. Barrat, Contact patterns in a high school: A comparison between data collected using wearable sensors, contact diaries and friendship surveys, *PLoS ONE* **10**(9) (2015), 1–26. ISBN 1932-6203 (Electronic)\r1932-6203 (Linking). doi:10.1371/journal.pone.0136497.

[11] A. Martınez, Y. Dimitriadis, B. Rubia, E. Gómez and P. De La Fuente, Combining qualitative evaluation and social network analysis for the study of classroom social interactions, *Computers & Education* **41**(4) (2003), 353–368.

[12] L. Bazzani, M. Cristani, D. Tosato, M. Farenzena, G. Paggetti, G. Menegaz and V. Murino, Social interactions by visual focus of attention in a three-dimensional environment, *Expert Systems* **30**(2) (2013), 115–127.

[13] M. Girolami, F. Mavilia, F. Delmastro and E. Distefano, Detecting Social Interactions through Commercial Mobile Devices, in: *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2018, pp. 125–130. doi:10.1109/PERCOMW.2018.8480397.

[14] F. Palumbo, P. Barsocchi, S. Chessa and J.C. Augusto, A stigmergic approach to indoor localization using bluetooth low energy beacons, in: *2015 12th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, IEEE, 2015, pp. 1–6.

[15] P. Barsocchi, A. Crivello, M. Girolami, F. Mavilia and F. Palumbo, Occupancy detection by multi-power bluetooth low energy beaconing, in: *Indoor Positioning and Indoor Navigation (IPIN), 2017 International Conference on*, IEEE, 2017, pp. 1–6.

[16] P. Baronti, P. Barsocchi, S. Chessa, F. Mavilia and F. Palumbo, Indoor Bluetooth Low Energy Dataset for Localization, Tracking, Occupancy, and Social Interaction, *Sensors* **18**(12) (2018), 4462.

[17] F. Mavilia, F. Palumbo, P. Barsocchi, S. Chessa and M. Girolami, Remote Detection of Indoor Human Proximity using Bluetooth Low Energy Beacons, in: *2019 15th International Conference on Intelligent Environments (IE)*, IEEE, 2019.

[18] P. Cassará, F. Potortí, P. Barsocchi and M. Girolami, Choosing an RSS device-free localization algorithm for Ambient Assisted Living, in: *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2015, pp. 1–8. doi:10.1109/IPIN.2015.7346788.

[19] V. Sekara and S. Lehmann, The strength of friendship ties in proximity sensor data, *PLoS ONE* **9**(7) (2014), 1–14. doi:10.1371/journal.pone.0100915.

[20] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.F. Pinton, M. Quaggiotto, W. van den Broeck, C. Régis, B. Lina and P. Vanhems, High-resolution measurements of face-to-face contact patterns in a primary school, *PLoS ONE* **6**(8) (2011). ISBN 1932-6203 (Electronic)\r1932-6203 (Linking). doi:10.1371/journal.pone.0023176.

[21] T. Kim, E. McFee, D.O. Olguin, B. Waber and A. Pentland, Sociometric badges: Using sensor technology to capture new forms of collaboration, *Journal of Organizational Behavior* **33**(3) (2012), 412–427.

[22] A. Montanari, S. Nawaz, C. Mascolo and K. Sailer, A Study of Bluetooth Low Energy performance for human proximity detection in the workplace, in: *Pervasive Computing and Communications (PerCom), 2017 IEEE International Conference on*, IEEE, 2017, pp. 90–99.

[23] A. Stopczynski, V. Sekara, P. Sapiezynski, A. Cuttone, M.M. Madsen, J.E. Larsen and S. Lehmann, Measuring large-scale social networks with high resolution, *PLoS ONE* **9**(4) (2014). ISBN 1932-6203. doi:10.1371/journal.pone.0095978.

[24] P. Sapiezynski, A. Stopczynski, D.K. Wind, J. Leskovec and S. Lehmann, Inferring person-to-person proximity using WiFi signals, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **1**(2) (2017), 24.

[25] N. Palaghias, S.A. Hoseinitabatabaei, M. Nati, A. Gluhak and K. Moessner, Accurate detection of real-world social interactions with smartphones, in: *2015 IEEE International Conference on Communications (ICC)*, IEEE, 2015, pp. 579–585.

[26] A. Matic, V. Osmani, A. Maxhuni and O. Mayora, Multimodal mobile sensing of social interactions, in: *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, IEEE, 2012, pp. 105–114.

[27] C. Efstratiou, I. Leontiadis, M. Picone, K.K. Rachuri, C. Mascolo and J. Crowcroft, Sense and sensibility in a pervasive world, in: *International Conference on Pervasive Computing*, Springer, 2012, pp. 406–424.

[28] J. Neburka, Z. Tlamsa, V. Benes, L. Polak, O. Kaller, L. Bolecek, J. Sebesta and T. Kratochvil, Study of the performance of RSSI based Bluetooth Smart indoor positioning, in: *2016 26th International Conference Radioelektronika (RADIOELEKTRONIKA)*, IEEE, 2016, pp. 121–125.

[29] J. Torres-Sospedra, R. Montoliu, S. Trilles, Ó. Belmonte and J. Huerta, Comprehensive analysis of distance and similarity measures for Wi-Fi fingerprinting indoor positioning systems, *Expert Systems with Applications* **42**(23) (2015), 9263–9278.

[30] P. Barsocchi, A. Crivello, D. La Rosa and F. Palumbo, A multisource and multivariate dataset for indoor localization methods based on WLAN and geo-magnetic field fingerprinting, in: *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, IEEE, 2016, pp. 1–8.

[31] F. Palumbo and P. Barsocchi, Salt: Source-agnostic localization technique based on context data from binary sensor networks, in: *European Conference on Ambient Intelligence*, Springer, 2014, pp. 17–32.

[32] F. Potortì, P. Cassarà and F. Palumbo, Robust Device-Free Localisation with Few Anchors, in: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, ACM, 2018, pp. 1184–1189.

[33] O. Belmonte-Fernandez, A. Gasco-Compte, E. Sansano-Sansano, M. Quinde, G. Manuel, J. Gines and J.C. Augusto, Evaluation of crowdsourcing Wi-Fi radio map creation in a real scenario for AAL applications, in: *2019 15th International Conference on Intelligent Environments (IE)*, IEEE, 2019.

[34] D. Sora and J.C. Augusto, Managing multi-user smart environments through BLE based system, in: *Intelligent Environments 2018: Workshop Proceedings of the 14th International Conference on Intelligent Environments*, Vol. 23, P.N.O.A. Ioannis Chatzigiannakis Yoshito Tobe, ed., IOS Press, Amsterdam, The Netherlands, The Netherlands, 2018, pp. 234–243.

[35] Ó. Belmonte-Fernández, A. Puertas-Cabedo, J. Torres-Sospedra, R. Montoliu-Colás and S. Trilles-Oliver, An indoor positioning system based on wearables for ambient-assisted living, *Sensors* **17**(1) (2017), 36.

[36] F. Potortì and F. Palumbo, CEO: A context event only indoor localization technique for AAL, *Journal of Ambient Intelligence and Smart Environments* **7**(6) (2015), 745–760.

[37] P. Barsocchi, E. Ferro, L. Fortunati, F. Mavilia and F. Palumbo, EMS@CNR: an energy monitoring sensor network infrastructure for in-building location-based services, in: *2014 International Conference on High Performance Computing & Simulation (HPCS)*, IEEE, 2014, pp. 857–862.

[38] H. Zhang, W. Du, P. Zhou, M. Li and P. Mohapatra, DopEnc: acoustic-based encounter profiling using smartphones, in: *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, ACM, 2016, pp. 294–307.

[39] L. Bazzani, M. Cristani and V. Murino, Decentralized particle filter for joint individual-group tracking, in: *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2012, pp. 1886–1893.

[40] H. Hung and B. Kröse, Detecting f-formations as dominant sets, in: *Proceedings of the 13th international conference on multimodal interfaces*, ACM, 2011, pp. 231–238.

[41] E.T. Hall, *The hidden dimension / Edward T. Hall*, [1st ed.] edn, Doubleday Garden City, N.Y, 1966, pp. xii, 201 p., [6] leaves of plates :–.

[42] F. Potortì, A. Corucci, P. Nepa, F. Furfari, P. Barsocchi and A. Buffi, Accuracy limits of in-room localisation using RSSI, in: *2009 IEEE Antennas and Propagation Society International Symposium*, IEEE, 2009, pp. 1–4.

[43] P. Barsocchi, F. Furfari, P. Nepa and F. Potortì, RSSI localisation with sensors placed on the user, in: *2010 International Conference on Indoor Positioning and Indoor Navigation*, IEEE, 2010, pp. 1–6.

[44] P. Barsocchi, S. Chessa, E. Ferro, F. Furfari and F. Potorti, Context driven enhancement of RSS-based localization systems, in: *2011 IEEE Symposium on Computers and Communications (ISCC)*, IEEE, 2011, pp. 463–468.