SHPIA: A Low-Cost Multi-purpose Smart Home Platform for Intelligent Applications

Florenc Demrozi and Graziano Pravadelli(✉)

Department of Computer Science, University of Verona, 37134 Verona, Italy
{florenc.demrozi,graziano.pravadelli}@univr.it

Abstract. Nowadays, smart devices have invaded the market and consequently our daily life. Their use in smart home contexts, to improve the quality of life, specially for elderly and people with special needs, is getting stronger and stronger. Therefore, many systems based on smart applications and intelligent devices have been developed, for example, to monitor people's environmental contexts, help in daily life activities, and analyze their health status. However, most of the existing solutions present disadvantages regarding accessibility, as they are costly, and applicability, due to lack of generality and interoperability.

This paper is intended to tackle such drawbacks by presenting SHPIA, a multi-purpose smart home platform for intelligent applications. It is based on the use of a low-cost Bluetooth Low Energy (BLE)-based devices, which "transforms" objects of daily life into smart objects. The devices allow collecting and automatically labelling different type of data to provide indoor monitoring and assistance. SHPIA is intended, in particular, to be adaptable to different home-based application scenarios, like for example, human activity recognition, coaching systems, and occupancy detection and counting.

The SHPIA platform is open source and freely available to the scientific and industrial community.

Keywords: Smart home platform · Automatic data annotation · Automatic data collection · Human activity recognition

1 Introduction

Nowadays, smart systems have invaded our daily lives with a plethora of devices, mainly from consumer electronics, like smartwatches, smartphones and personal assistants, and domestic appliances with intelligent capabilities to autonomously drive modern homes. These are generally based on wireless protocols, like WiFi, Bluetooth Low Energy (BLE) and ZibBee, low-cost sensing components, including Passive InfraRed (PIR) and Radio Frequency Identification (RFID)-based technologies, and several other kinds of environmental sensors. These devices communicate with each other without the need for human intervention, thus

contributing to the implementation of the Internet of Things (IoT) paradigm. Their number is continuously increasing, and their versatility enables several opportunities for different scenarios, ranging from simple environmental monitoring solutions to more complex autonomous control systems, in both private life (i.e., at home) and public contexts (i.e., social and working environments). Without claiming to be exhaustive, examples of applications can be found in healthcare [3,16] and elderly assistance [21,22], in smart building for Human Activity Recognition (HAR) [5], and energy management [11,20], as well as in smart industries [12] and smart cities [18].

In particular, concerning private life, in the last decade, the idea of a smart home has become of central interest, where its main aim concerns the recognition of the activities performed by the environment occupants (e.g., cooking, sitting down, sleeping, etc.), and the detection of changes in the environmental status due to such activities (e.g., temperature variation related to the opening or closing of a window) [2,7]. The concept of home, indeed, includes different connotations, and according to [13], it is characterized as a place for a) security and control, b) activity, c) relationships and continuity, and d) identity and values. Thus, to guarantee and promote these peculiarities, the design of a smart home cannot ignore the need of implementing the capability of recognizing human activities through HAR systems to provide real-time information about people's behaviors. HAR algorithms are based on pattern recognition models fed with data perceived by on-body sensors, environmental sensors, and daily life smart devices [2,9]. HAR algorithms and smart devices provide, then, the basics for implementing and integrating intelligent systems, which autonomously take decisions and support life activities in our homes.

1.1   Related Works

The literature concerning smart home platforms is extremely vast and differentiated, which makes impossible to exhaustively summarize them in a few lines. Among the cheapest solutions, we can cite the CASAS platform proposed in [6]. It integrates several ZigBee-based sensors for door, light, motion, and temperature sensing, with a total cost of $2,765. Based on the data collected by the platform, the authors were able to recognize ten different Activities of Daily Life (ADLs) executed by the environment occupants achieving, on average, approximately 60% accuracy.

In a similar way, in [19], the authors proposed a HAR model, fed with data collected through a smart home platform based on motion, door, temperature, light, water, and bummer sensors, to classify more than ten ADLs, achieving approximately 55% accuracy.

In [24], a study is presented where the authors installed a sensor network, composed of motion sensors, video cameras, and a bed sensor that measures sleep restlessness, pulse, and breathing levels, in 17 flats of an aged eldercare facility. They gathered data for 15 months on average (ranging from 3 months to 3 years). The collected information was used to prevent and detect falls and recognize ADLs by identifying anomalous patterns.

In [28], the authors used an application to continually record raw data from a mobile device by exploiting the microphone, the WiFi scan module, the device heading orientation, the light proximity, the step detector, the accelerometer, the gyroscope, the magnetometer, and other built-in sensors. Then, time-series sensor fusion and techniques such as audio processing, WiFi indoor positioning, and proximity sensing localization were used to determine ADLs with a high level of accuracy.

When developing a strategy to deploy technology for discreet in-home health monitoring, several questions arise concerning, for example, the types of sensors that should be used, their location, and the kind of data that should be collected. In [26], the authors deeply studied such issues, pointing out that no clear answer can be identified, but the perceived data must be accurately evaluated to provide insights into such questions.

Recently, relevant pilot projects have been developed and presented, such as HomeSense [25,27], to demonstrate seniors' benefits and adherence response to the designed smart home architecture. HomeSense exposes the visualization of activity trends over time, periodic reporting for case management, custom real-time notifications for abnormal events, and advanced health status analytics. HomeSense includes magnetic contact, passive infrared motion, energy, pressure, water, and environmental sensors.

However, all these innovative systems and applications frequently present disadvantages in terms of accessibility and applicability. In several cases, they are based on ad-hoc and costly devices (e.g., cameras) which are not accessible to everyone. In fact, as shown in [14], among 844 revised works, the system cost is the principal reason for the failure of projects concerning the design of smart health/home systems.

Some projects targeting the definition of low-cost solutions have been also proposed, but they are generally devoted to monitoring or recognizing single activities and/or specific use-case scenarios, thus lacking generality, or requiring the final users to install several non inter-operable solutions in their homes. For example, in [4] a set of very low-cost projects focusing on solutions for helping visually-impaired people are presented. Less effort has been spent, instead, designing solutions that use low-cost objects of daily life (ODLs) to monitor and recognize people's activity in general.

Finally, a further limitation of the existing smart home environments regards the necessity of annotating the collected data based on the video registration of the environment for training the pattern recognition algorithms, which is of central importance for the implementation of efficient HAR algorithms. Unfortunately, the annotation process is generally a very time-consuming and manual activity, while only a few prototypical automatic approaches are currently available in the literature [8].

1.2 Paper Contribution

According to the previous considerations and the limitation of existing solutions, this paper presents SHPIA, an open-source multi-purpose smart home platform

for intelligent applications. It is based on the use of low-cost Bluetooth Low Energy (BLE)-based devices, which "transform" objects of daily life into smart objects. The devices allow collecting and automatically labeling different types of data to provide intelligent services in smart homes, like, as example, indoor monitoring and assistance. Its architecture relies on the integration among a mobile Android application and low-cost BLE devices, which "transform" objects of daily life into smart objects. By exploiting these devices, SHPIA flexibly collects and annotates datasets that capture the interaction between humans and the environment they live in, and then the related behaviors. These datasets represent the basics for implementing and training HAR-based systems and developing intelligent applications for different indoor monitoring and coaching scenarios. SHPIA can be set up with less than \$200 and operates in a ubiquitous and not invasive manner (i.e., no camera is required). SHPIA's software is available to the scientific community through a public GitHub repository [1].

1.3   Paper Organization

The rest of the paper is organized as follows. Section 2 introduces preliminary information concerning the devices and the communication protocols used in SHPIA. Section 3 details the SHPIA architecture and describes how it enables data collection and labeling. Section 4 showcases and discusses the experimental results and application scenarios. Finally, Sect. 5 concludes the paper with final remarks.

2   Preliminaries

SHPIA is based on the use of the Nordic Thingy 52 device shown in Fig. 1. The choice of using such a device to implement the SHPIA platform has been made on the basis of its versatile and complete set of characteristics, which are summarized below in this section. However, this device can be easily replaced, without affecting SHPIA functionalities, with many other BLE-based inertial measurement units available on the market, provided that they allow to collect a similar set of data through their sensors.

The Nordic Thingy 52 is a compact, power-optimized, multi-sensor device designed for collecting data of various type based on the nRF52832 System on Chip (SoC), built over a 32-bit ARM CortexTM-M4F CPU. The nRF52832 is fully multiprotocol, capable of supporting Bluetooth 5, Bluetooth mesh, BLE, Thread, Zigbee, 802.15.4, ANT, and 2.4 GHz proprietary stacks. Furthermore, the nRF52832 uses a sophisticated on-chip adaptive power management system achieving exceptionally low energy consumption. This device integrates two types of sensors: i) environmental and ii) inertial. Environmental concern temperature, humidity, air pressure, light intensity, and air quality sensors (i.e., CO_2 level). Instead, inertial concerns accelerometer, gyroscope, and compass sensors. Besides the data directly measured by the integrated sensors, the Thingy computes over the edge the following information: quaternion, rotation matrix, pitch,

roll, yaw, and step counter. Concerning the communication capabilities, Thingy 52 instantiates a two-side BLE communication with the data aggregator device, unlike BLE beacons. The communication between Thingy 52 and the data aggregator occurs at a frequency that goes from 0.1 Hz to 133 Hz, making SHPIA adaptable to applications scenarios were high sampling frequencies are required. Moreover, since the BLE provides the possibility to send more than just one value into every single transmitted package, the sensor's sampling frequency is not limited 133 Hz (i.e., the maximal frequency of the BLE communication), but it enables the sensor to sample at higher frequencies.
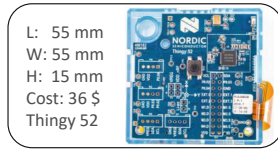

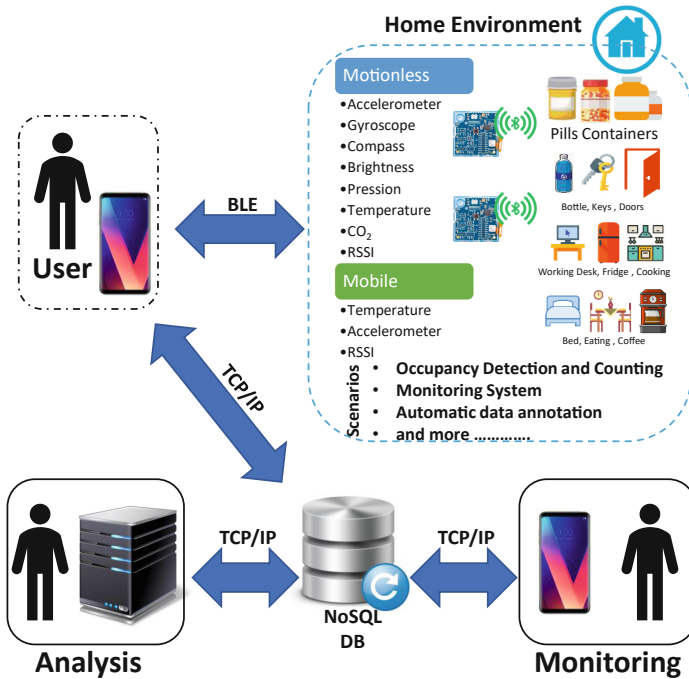**Fig. 1.** SHPIA compatible device: Nordic Thingy 52.

Fig. 2. Schematic view of the SHPIA platform.

3   SHPIA Architecture

This section introduces the core of the SHPIA platform, which is shown in Fig. 2. In particular, it describes the principal agents composing its architecture, and how a smart home environment can be defined and configured for enabling data collection and annotation.

3.1   Agents

The agents involved in the SHPIA architecture are classified as *abstract agents* and *real agents*. The abstract agents are necessary to analyze the status of the environment (in terms of included objects and environmental conditions) and the status of people living in it (in terms of presence, quantity, movements, and accomplished actions). On the other side, real agents are represented by the people occupying the environment, their smartphones, and the Thingy 52 devices. SHPIA enables communication capabilities among real agents as described below in the paper.

3.2   Environment Definition

Concerning the home environment, SHPIA defines it as a set composed of the home itself, people inside it, and available ODLs. In particular, ODLs enclose mobile objects (bottles, pills container, keys, etc.) and motionless objects (e.g., doors, desks, coffee machine, etc.) present inside the environment, as those shown in Fig. 2.

Therefore, given a set of mobile ODLs (M) and a set of motionless ODLs (Ml), the environment (E) is formally defined as:

$$E = \{\{M \cup Ml\}, f_s, f_l, T, D\}, \text{with}$$
$$f_s : M \longrightarrow T$$
$$f_l : Ml \longrightarrow T$$

where, T is a set of Thingy 52 devices, D is a data aggregator node, while f_s and f_l are functions that associate, respectively, a Thingy 52 device to each mobile (M) and motionless (Ml) ODL. The data aggregator D identifies the device that collects the data perceived by the Thingy 52 devices, behaving as a gateway towards a Cloud database. SHPIA uses an Android smartphone as data aggregator.

3.3   Environment Configuration

To handle the definition of the environment, we designed the Android application shown in Fig. 3. It allows the users to create one or more environments and associate a single specific Thingy 52 to each ODL of interest. In addition, this application allows real-time visualization of the perceived data and enables the smartphone to operate as the data aggregator.

Figure 4 presents the steps that the user has to perform in order to configure the smart environment by means of the mobile application. They work as follows.
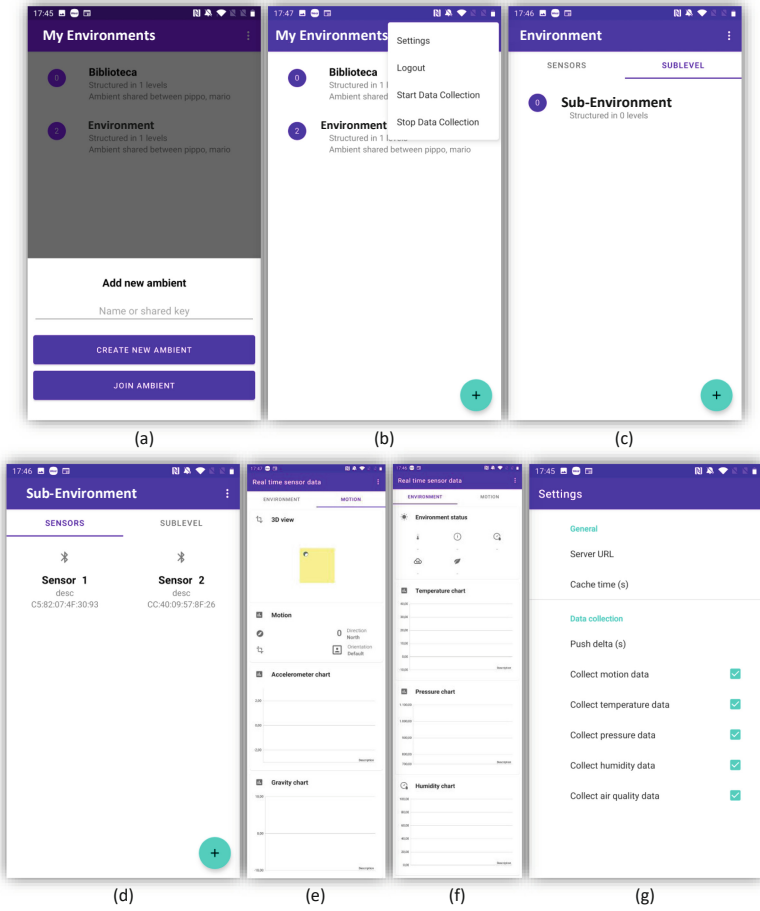
Fig. 3. SHPIA Android mobile application.

Fig. 4. SHPIA environmental configuration work-flow.

User Account Creation: In this first step, the SHPIA application allows, if not already existing, the creation of a user profile to associate the collected data. Once the user is verified, she/he can set the IP address of a Cloud-based NoSQL database from the setting page, to which the data will be transmitted. We want to emphasize that the transmitted data can be saved at any NoSQL database deployed on such IP. Users, for analysis purposes, need only to know the format that the SHPIA mobile application uses to save the data.

Create Environment: Once authenticated, the user can create one or more environments, as introduced in Sect. 3.2, by defining its name (i.e., *env_id*) and geographical address (i.e., *address*). Alternatively, if the environment already exists, the user can share it with other SHPIA users (Fig. 3(a) and Fig. 3(b)).

Create Sub-environment: SHPIA users can create as many environments as needed, and an environment is typically composed of other sub-environments (Fig. 3(c)). For example, an apartment consists of a lounge, a kitchen, two bedrooms, and two bathrooms. SHPIA does not present any limit in the depth of nested sub-environments (i.e., it can configure sub-environments of a sub-environment of . . . of an environment). This specific feature has been developed by considering the possibility of adopting SHPIA also in industrial, scholastic, or smart city scenarios. Moreover, SHPIA can be used also for not environmental-related contexts. For example, the users can adopt it for implementing a wireless body area network by associating the Thingy 52 devices to body parts instead of environments or ODLs, as in [15]. Overall, environments and sub-environments are described as shown in the example of Listing 1.1. Besides, *env_id* and *address*, the environments and sub-environments are identified by *owner*, *creation_time*, list of *sub_environments* and a brief *description*.

```
{
    "env_id": "Home_1",
    "description": "master bedroom",
    "owner": "florenc.demrozi@univr.it",
    "address": "Strada le Grazie 15, Verona, Italy",
    "creation_time":'27/12/2021 15:13:52.085',
    "sub_environments": ["Sub_Home_1","Sub_Home_2","
    Sub_Home_3","Sub_Home_4"]
}
```

Listing 1.1. Example of environment description in SHPIA.

Create Smart ODLs: Once the environment has been created, the user can finally attach a Thingy 52 device to each mobile or motionless ODL of interest to transform it into a smart ODL. At this point, the user is required to move his/her smartphone close (<10 cm) to the ODL equipped with the Thingy 52 to allow SHPIA to recognize it. SHPIA automatically associates the ODL with the nearest Thingy 52 device by exploiting the Received Signal Strength Indicator

(RSSI) measurement (Fig. 3(d)). RSSI, often used in Radio Frequency (RF)–based communication systems, is related to the power perceived by a receiver. In particular, it provides an indication of the power level at which the data frames are received. The rationale is that the higher the RSSI value, the stronger the signal and the closer the receiver and the emitter. The RSSI is used to reduce possible wrong associations in the presence of a high number of BLE devices distributed in the environment.

Data Collection: After the environment definition and the association of ODLs to Thingy 52 devices, the SHPIA mobile app will start collecting data from them (Fig. 3(e) and Fig. 3(f)). The data perceived by the smartphone are internally stored as JSON documents. As soon as an Internet connection is available, all data are saved on the remote NoSQL database (Fig. 3(g)). Listing 1.2 shows an example of the data perceived by the Thingy 52 device.

```
{
    "deviceID": "00:00:5e:00:53:af",
    "on_device_time":'27/12/2021 15:13:52.085',
    "on_aggreg_time":'27/12/2021 15:13:52.155',
    "parent_env_id": "Sub_Home_1"
    "ODL":       "Description",
    "temp":      23.8,
    "light":     43,
    "pressure": 101.325,
    "CO2": 56,
    "accel": [{
            "X":   0.0226898,
            "Y":  -0.382233,
            "Z":   9.54773,
        }],
    "gyro": [{
            "X":   15.022791,
            "Y":  -12.233382,
            "Z":    3.73547,
        }],
    "comp": [{
            "X":   0.6898022,
            "Y":  -2.233382,
            "Z":   5.385477,
        }],
    "quaternion": [{
            "a":   0.02,
            "b":  -0.38,
            "c":   9.54,
            "d":   5.47,
        }],
    "other": [{
            "pitch": 8.2,
            "roll": -3.8,
```

```
        "yaw":     9.54773,
      }],
    "rssi":     -65
}
```

Listing 1.2. Data format of Thingy 52 device.

Attributes *deviceID*, *on_aggreg_time* and *on_device_time* uniquely identify the document; the rest represents the data perceived by the device's sensors and the RSSI value measured by the smartphone. The *on_aggreg_time* variable represents the timestamp when the aggregator receives the data. Instead, *on_device_time* represents the timestamp when the data is perceived on the Thingy 52 device. Finally, *parent_env_id* is used to identify the (sub-)environment to which the device is collected. We want to emphasize that the transmitted data can be saved at any NoSQL database deployed on the target IP (Fig. 3(g)). Users, for analysis purposes, need only to know the data format (i.e., Listings 1.2) that the SHPIA mobile application adopts to save the data.

4   SHPIA Evaluation

This section deals with the evaluation of the performance of the SHPIA data aggregator to show the lightweight of the Android application in terms of power consumption and use of resources. In addition, it illustrates four application scenarios where SHPIA can operate. Such scenarios do not require any modification of the SHPIA platform, thus proving its versatility.

4.1   Data Aggregator Performance Evaluation

The performances of three Android smartphones with different characteristics and prices have been evaluated while acting as data aggregators for the SHPIA platform. The characteristics of the tested smartphones are reported in Table 1. Instead, Table 2 presents the results of profiling the data aggregator nodes over a collection phase of 4/4/4 h, by using five Thingy 52 with sensors sampling data set 50 Hz, 100 Hz, 200 Hz. The data aggregator nodes were placed over a table at the height of 100 cm. Instead, the Thingy devices were associated with different desks. The distance between the data aggregator and the thingy nodes varied between 2 and 7 m. Overall, the average RAM use per hour was <116 Mbh, the storage memory use was <126 Mbh[1], and the battery usage was <680 mAh[2]. On average, CPU usage and data loss were respectively 37% and 0%.

It is worth noting that smartphones executing an Android version older than v11 can be connected simultaneously with up to seven Thingy 52. Instead, smartphones running Android v11 can support the simultaneous connection with up

---

[1] Cumulative, if Internet connection is missing, e.g., 1260 Mb in 10 h without connectivity.

[2] 270 mAh excluding the results of the Honor 7S smartphone.

Table 1. Characteristics of the tested data aggregator nodes.

Model	RAM (GB)	Storage (GB)	Battery (mAh)	Weight (g)	Price ($)	Android version
Honor 7S	2	16	3020	142	78	8.1
LG X Power 2	2	16	4500	164	89	8.1
Galaxy S9	4	64	3000	163	262	10.0

Table 2. Data aggregators profiling.

	Honor 7S			Galaxy S9 Edge			LG X Power 2		
Frequency (Hz)	50	100	200	50	100	200	50	100	200
RAM (MB/h)	93	93	92	127	132	140	100	100	110
Storage (MB/h)	50	100	196	48	99	198	104	127	219
CPU (%)	35	46	55	19	21	23	40	50	50
Battery (mAh)	1150	1208	1389	360	360	390	225	450	630
Data loss (%)	0	0	0	0	0	0	0	0	0

to eleven Thingy 52. To overcome this limitation, SHPIA implements a computation balancing module that allows different smartphones (thus, different users sharing the same environment) to automatically balance the number of Thingy 52 devices connected to them and save their information on the same dataset. Thus, in practice, SHPIA can handle more than 11 Thingy devices by jointly using more than one smartphone. Moreover, this balancing process helps to further reduce smartphone battery consumption.

These results show that the proposed platform works well on different data aggregator nodes, proving there is no need to buy costly top-level smartphones to run the SHPIA application. Concerning the Thingy 52, they can efficiently operate for more than three days without recharging the battery. In addition, a lower sampling frequency would further extend the battery life consistently for both smartphone and BLE nodes [17].

4.2   SHPIA Application Scenarios

In the following, we provide an overview of four different applications exploiting the SHPIA platform: a) environmental monitoring, b) occupancy detection and counting, c) automatic data annotation of ADLs, and finally, d) virtual coaching.

Environmental Monitoring: The primary use of SHPIA is that of collecting data concerning environmental conditions. For example, we used SHPIA to monitor a working office, shown in Fig. 5, shared by ten persons. We associated the Thingy 52 devices with 5 motionless nodes (indicated by red arrows in Fig. 5) and six mobile nodes (indicated by green arrows in Fig. 5) to perceive

the environment status. The Honor 7S smartphone, permanently connected to the electric current, described in Table 1 acted as a data collector. The Thinghy 52 associated with the motionless nodes were placed as follows: one at the office door, two on the windows, one on the desk at the office center, and one inside the locker. Instead, the mobile nodes were used to monitor different ODLs and the activity that employees performed on them (e.g., one Thingy 52 was attached to a bottle of water). The data collection process was conducted for two consecutive weeks.
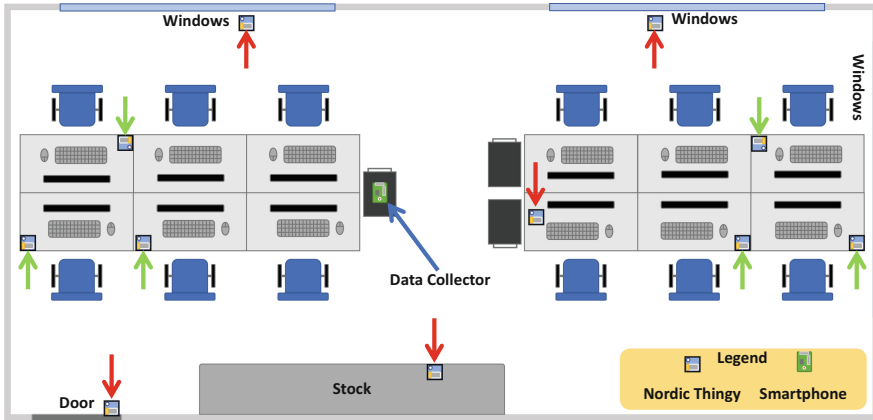
Fig. 5. Office 1.71. Motionless (red) and mobile (green) nodes and data collector (grey). (Color figure online)

Table 3 shows an overview of the collected data. The first column introduces the used sensors. The second and third columns show the sensor sampling frequency and the measurement unit. Column four shows the number of samples collected by the system during the two weeks (i.e., 1209600 s). Column five identifies the number of data sources (BLE Thingy 52 nodes). Finally, the last column shows the memory space required to store the sensed data. The last row concerns the collection of RSSI data, since the data collector extracts and associates the reception timestamp, the RSSI measure, and the emitter identity to each received BLE packet.

Once collected through SHPIA, such data were successfully used by a HAR-based analyzer to perform environmental monitoring, recognition of people's actions (e.g., drinking), and localization of ODLs and people in the environment. Because of the adopted low sampling frequency, the mobile and motionless BLE nodes perfectly worked for the overall duration of the experiments (2 weeks) without being recharged.

Occupancy Detection and Counting: Occupancy detection and counting represent a fundamental knowledge for implementing smart energy management

Table 3. Results of data collection.

Sensor type	Frequency (Hz)	Measure unit	# of samples	# of BLE devices	Stored data
Temperature	0.33	Celsius	2016000	5	≈ 15Mb
Pressure	0.33	Bar	2016000	5	≈ 15Mb
Brightness	0.33	Lux	2016000	5	≈ 15Mb
Acceleration	10	m/s^2	133056000	11	≈ 1Gb
RSSI	10	dBm	133056000	11	≈ 1Gb

Table 4. Distance estimation results based on RSSI measurements captured 60 Hz.

| Regression model	5 m				3 m				2 m			
	Raw data		Features		Raw data		Features		Raw data		Features	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Gradient Boosting	51	31	57	33	17	12	30	15	12	9	30	15
Random Forest	51	31	55	28	17	12	25	8	12	9	25	8
Linear	74	58	104	79	27	22	80	60	22	16	80	60
Ridge	74	58	69	45	27	22	60	39	22	16	60	39
RANSAC	84	54	113	86	28	21	85	59	43	26	82	59
Bayesian	74	58	233	158	27	22	229	171	22	16	229	171
TheilSen	78	54	96	70	29	21	89	65	24	18	89	65

systems, as well as solutions for security and safety purposes [23]. Existing techniques for occupancy detection and counting can be categorized as a) not device free [29] and b) device free [30]. The SHPIA platform provides the capability to implement both categories.

Concerning the former, SHPIA can detect a user inside an environment by estimating the distance between the user's smartphone and a Thingy 52 device associated with the environment itself, based on the RSSI measurement. To test this scenario, we evaluate the accuracy of the distance estimation between the user's smartphone and ODLs equipped with Thingy 52 node. The evaluation has been performed on three different distance ranges: a) 0–5 m, b) 0–3 m, and c) 0–2 m by using two opposite setups: i) the smartphone in the user's hand and the ODL in a fixed position, and ii) the ODL on the user's hand and the smartphone in a fixed position.

Table 4 presents the results obtained by seven different regression models trained on RSSI data perceived 60 Hz. The models were trained in two ways: by using the raw data, and by using features extracted from one-second RSSI time windows. The quality of the achieved results is shown in terms of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) while estimating the distance in centimeters between the emitter and the receiver. The Random Forest model achieved the lowest RMSE/MAE value in both the raw and the feature-based data representation. The second most performing model was Gradient Boosting. Overall, we achieved an RMSE on raw RSSI data of 51 cm in the range 0–5 m, 17 cm in the range 0–3 m, and 12 cm in the range 0–2 m. Moreover, in

terms of MAE, the features performed better than the raw data: 28 cm (range 0–5 m), 8 cm (range 0–3 m), and 8 cm (range 0–5 m). Furthermore, the most essential characteristics of these regression models regard the reduced memory and computation requirements, making them suitable for running on mobile and hardware-constraint devices.

The second category of occupancy detection and counting systems behave more intelligently. In fact, users do not need to carry any device. By using SHPIA, we can detect their presence and number based on the variations of the RSSI measurements associated with the BLE signals received by the data aggregator from Thingy 52 located in the environment. The idea is that RSSI measurement fluctuations are generated by people's presence and movements inside the environment. Figure 6 shows very clearly the difference between nocturnal (red plot [8:00 PM–8:00 AM]) and diurnal (blue plot [8:00 AM–8:00 PM]) RSSI observations at the office shown in Fig. 5. The same concept is applied as regards the occupancy counting scenario (aka., identification of the number of people present inside or in the environment).
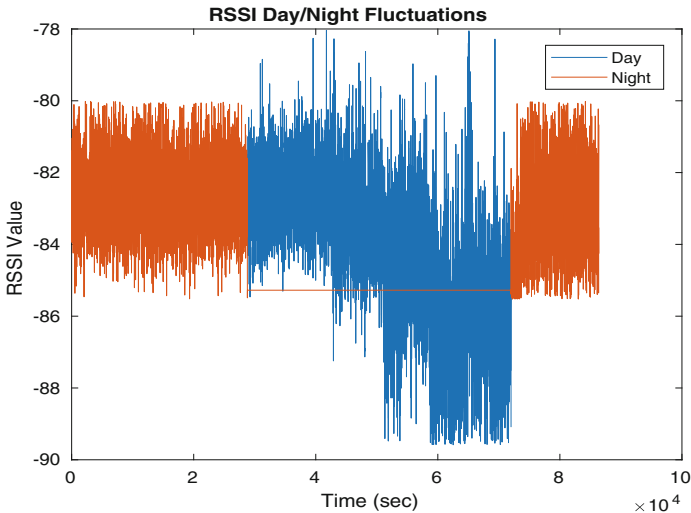
Fig. 6. Office 1.71. RSSI fluctuation between night and day for occupancy detection. (Color figure online)

We carried out tests in a university classroom (8.8 m × 8.6 m) with 15 study stations (chairs + tables) involving six different subjects. One female (29 years, 1.58 m height) and five males (25–29 years, 1.75–1.95 m height) were involved in the experiment. Subjects entered and left the environment in an undefined order with the only constraint that they must stay in the environment at least for one minute. Besides, the following environmental situations were recreated: i) all standing still, ii) all standing in motion, iii) all seated, and iv) some standing in motion and some sitting.

Table 5 presents the achieved results by using five different BLE nodes connected to SHPIA. Tests were performed over five different well-known classification models[3]. Columns two to five show results in terms of specificity, sensitivity, precision, and comprehensive accuracy. Overall, the SVM model with a linear kernel achieved the most noticeable results. Among all the other models, such a model requires higher computational capabilities; however, the Keras library provides a Quasi-SVM model implementation for Android-based mobile devices, thus enabling the SHPIA data collector recognition capabilities. By verifying the classification errors in detail, we observed that the incorrectly classified samples are related to the situation in which people inside the environment are all seated, independently by their number.

Table 5. Occupancy detection results.

Model	Specificity	Sensitivity	Precision	Accuracy
kNN	98.72%	99.10%	99.10%	99.10%
WkNN	98.29%	99.02%	99.03%	99.10%
LDA	99.83%	99.70%	99.70%	99.70%
QLDA	99.78%	99.77%	99.77%	99.77%
SVM	99.82%	99.86%	99.81%	99.82%

Table 6. Occupancy estimation results.

Regression model	Raw data		Features	
	RMSE	MAE	RMSE	MAE
Gradient Boosting	0.9	0.6	0.5	0.3
Random Forest	0.7	0.4	0.5	0.3
Linear	1.4	1.0	1.3	1.0
Ridge	1.4	1.0	2.5	4.2
RANSAC	1.8	1.3	3.3	3.3
Bayesian	1.4	1.0	2.1	2.1
TheilSen	1.9	1.2	2.0	1.8

Table 6 presents the results obtained on raw and features data concerning the occupancy counting scenario. The outcome is an estimation of the number of persons in the environment. The lower the RMSE, the higher the estimation accuracy. In particular, the proposed occupancy counting system, given a set of features identifying a one-second time window of RSSI measurements, estimates the number of people in the environment with an RMSE of 0.5 and an MAE of

---

[3] k-Nearest Neighbor (kNN), Weighted kNN (WkNN), Linear Discriminant Analysis (LDA), Quadratic LDA (QLDA), Support Vector Machine (SVM).

0.3. Using the raw dataset, we achieved an RMSE of 0.7 and an MAE of 0.4. As for the occupancy detection scenarios, the estimation error is amplified when all people inside the environment are sitting down.

Automatic Annotation of ADLs: As already mentioned in Sect. 1, one of the most significant limitations in the HAR research area concerns the creation of the learning dataset through a data annotation process. This process usually requires extensive manual work, during which at least two annotators associate data samples (e.g., perceived through inertial sensors) with labels that identify the activity (e.g., sleeping, eating, drinking, cooking, and many others) based on a video recording of the context. SHPIA can automatically annotate these activities by assigning a Thingy 52 device to specific objects or locations in the environment (e.g., by associating the Thingy 52 to the eating table, to the working desk, the bottle of water, the bed, etc.) and by estimating the distance between the data collector (i.e., the smartphone that the user is carrying) and the Thingy 52. Thus, when the user is eating, SHPIA assigns the label "eating" to the data collected from the smartphone, based on the estimated distance between the nearest Thingy 52 (i.e., the one on the table) and the user's smartphone. Preliminary results in this direction [8] showed that the approach works properly for activities requiring more than 30 s to be performed (e.g., intensively washing hands, or cooking).

Virtual Coaching: Virtual coaching capabilities can be easily supported by the SHPIA platform. A virtual coaching system (VCS) is an ubiquitous system that supports people with cognitive or physical impairments in learning new behaviors and avoiding unwanted ones. By exploiting SHPIA, we set up a VCS comprising a set of smart objects used to identify the user needs and to react accordingly. For example, let us imagine a person requiring a new medical treatment based on pill's assumption that initially forgets to respect the therapy. By attaching a BLE tag to the pills container SHPIA can monitor pills assumption. The user carries the smartphone (e.g., into the pocket), and when he/she approaches the pills container, SHPIA estimates the distance between the user and the container. It can also understand when the user opens and closes the cap based on the received motion information emitted by the BLE tag attached to the container. Thus, it is possible to understand, with greater accuracy, whether the user has taken medicines or not. If the person does not take medicines, the system warns him. Otherwise, the system remains silent. A prototype of such a system has been proposed in [10].

5   Conclusions

This paper presented SHPIA, a platform exploiting low-cost BLE devices and an Android mobile application that transforms ODLs into smart objects. It allows effective and efficient data collection for implementing various solutions in smart

home and HAR scenarios. SHPIA works in a ubiquitous and non-invasive way, using only privacy-preserving devices such as inertial and environmental sensors. Its versatility has been evaluated by discussing four monitoring scenarios concerning the automatic data annotation of ADLs, occupancy detection and counting, coaching systems, and environmental monitoring. Moreover, despite the mentioned scenarios, SHPIA can be easily used in other scenarios such as industrial, smart buildings, smart cities, or human activity recognition. Nevertheless, although we have already implemented a computation balancing system that overcomes SHPIAs scalability issue, further work is required to overcome such weakness. Therefore, besides making open-source SHPIA, we intend to integrate BLE broadcasting communication technology in future developments, thus increasing the number of supported BLE devices per smartphone to hypothetically infinite.

References

1. SHPIA (2022). https://github.com/IoT4CareLab/SHPIA/
2. Abdallah, Z.S., Gaber, M.M., Srinivasan, B., Krishnaswamy, S.: Activity recognition with evolving data streams: a review. ACM Comput. Surv. (CSUR) **51**(4), 71 (2018)
3. Acampora, G., Cook, D.J., Rashidi, P., Vasilakos, A.V.: A survey on ambient intelligence in healthcare. Proc. IEEE **101**(12), 2470–2494 (2013)
4. Balakrishnan, M.: ASSISTECH: an accidental journey into assistive technology. In: Chen, J.-J. (ed.) A Journey of Embedded and Cyber-Physical Systems, pp. 57–77. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-47487-4_5
5. Bouchabou, D., Nguyen, S.M., Lohr, C., LeDuc, B., Kanellos, I., et al.: A survey of human activity recognition in smart homes based on IoT sensors algorithms: taxonomies, challenges, and opportunities with deep learning. Sensors **21**(18), 6037 (2021)
6. Cook, D.J., Crandall, A.S., Thomas, B.L., Krishnan, N.C.: Casas: a smart home in a box. Computer **46**(7), 62–69 (2013)
7. Cook, D.J., et al.: Mavhome: an agent-based smart home. In: Proceedings of the First IEEE International Conference on Pervasive Computing and Communications (PerCom 2003), pp. 521–524. IEEE (2003)
8. Demrozi, F., Jereghi, M., Pravadelli, G.: Towards the automatic data annotation for human activity recognition based on wearables and BLE beacons. In: 2021 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), pp. 1–4. IEEE (2021)
9. Demrozi, F., Pravadelli, G., Bihorac, A., Rashidi, P.: Human activity recognition using inertial, physiological and environmental sensors: a comprehensive survey. IEEE Access **8**, 210816–210836 (2020)
10. Demrozi, F., Serlonghi, N., Turetta, C., Pravadelli, C., Pravadelli, G.: Exploiting bluetooth low energy smart tags for virtual coaching. In: 2021 IEEE 7th World Forum on Internet of Things (WF-IoT), pp. 470–475. IEEE (2021)
11. Ehlers, G.A., Beaudet, J.: System and method of controlling an HVAC system, US Patent 7,130,719, 31 October 2006
12. Fernández-Caramés, T.M., Fraga-Lamas, P.: A review on human-centered IoT-connected smart labels for the industry 4.0. IEEE Access **6**, 25939–25957 (2018). https://doi.org/10.1109/ACCESS.2018.2833501

13. Gram-Hanssen, K., Darby, S.J.: Home is where the smart is? Evaluating smart home research and approaches against the concept of home. Energy Res. Soc. Sci. **37**, 94–101 (2018)
14. Granja, C., Janssen, W., Johansen, M.A.: Factors determining the success and failure of ehealth interventions: systematic review of the literature. J. Med. Internet Res. **20**(5), e10235 (2018)
15. Jeong, S., Kim, T., Eskicioglu, R.: Human activity recognition using motion sensors. In: Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems, pp. 392–393 (2018)
16. Keller, M., Olney, B., Karam, R.: A secure and efficient cloud-connected body sensor network platform. In: Camarinha-Matos, L.M., Heijenk, G., Katkoori, S., Strous, L. (eds.) IFIP International Internet of Things Conference, pp. 197–214. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-96466-5_13
17. Kindt, P.H., Yunge, D., Diemer, R., Chakraborty, S.: Energy modeling for the bluetooth low energy protocol. ACM Trans. Embed. Comput. Syst. (TECS) **19**(2), 1–32 (2020)
18. Kirimtat, A., Krejcar, O., Kertesz, A., Tasgetiren, M.F.: Future trends and current state of smart city concepts: a survey. IEEE Access **8**, 86448–86467 (2020)
19. Krishnan, N.C., Cook, D.J.: Activity recognition on streaming sensor data. Pervasive Mob. Comput. **10**, 138–154 (2014)
20. Pourbehzadi, M., Niknam, T., Kavousi-Fard, A., Yilmaz, Y.: IoT in smart grid: energy management opportunities and security challenges. In: Casaca, A., Katkoori, S., Ray, S., Strous, L. (eds.) IFIPIoT 2019. IAICT, vol. 574, pp. 319–327. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-43605-6_19
21. Rachakonda, L., Mohanty, S.P., Kougianos, E.: cStick: a calm stick for fall prediction, detection and control in the IoMT framework. In: Camarinha-Matos, L.M., Heijenk, G., Katkoori, S., Strous, L. (eds.) IFIP International Internet of Things Conference, pp. 129–145. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-96466-5_9
22. Rashidi, P., Mihailidis, A.: A survey on ambient-assisted living tools for older adults. IEEE J. Biomed. Health Inform. **17**(3), 579–590 (2013)
23. Salimi, S., Hammad, A.: Critical review and research roadmap of office building energy management based on occupancy monitoring. Energy Build. **182**, 214–241 (2018)
24. Skubic, M., Alexander, G., Popescu, M., Rantz, M., Keller, J.: A smart home application to eldercare: current status and lessons learned. Technol. Health Care **17**(3), 183–201 (2009)
25. VandeWeerd, C., et al.: Homesense: design of an ambient home health and wellness monitoring platform for older adults. Heal. Technol. **10**(5), 1291–1309 (2020)
26. Wang, J., Spicher, N., Warnecke, J.M., Haghi, M., Schwartze, J., Deserno, T.M.: Unobtrusive health monitoring in private spaces: the smart home. Sensors **21**(3), 864 (2021)
27. Wang, Y., Yalcin, A., VandeWeerd, C.: Health and wellness monitoring using ambient sensor networks. J. Ambient Intell. Smart Environ. **12**(2), 139–151 (2020)
28. Wu, J., Feng, Y., Sun, P.: Sensor fusion for recognition of activities of daily living. Sensors **18**(11), 4029 (2018)
29. Yang, J., et al.: Comparison of different occupancy counting methods for single system-single zone applications. Energy Build. **172**, 221–234 (2018)
30. Zou, H., Zhou, Y., Yang, J., Spanos, C.J.: Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT. Energy Build. **174**, 309–322 (2018)