Automatic And Non-Invasive Analysis Of Behavioural **Routines Using BLE Technology**

Javier Gaviña Rueda^{1,*,†}, Raúl Montoliu Colás^{1,†}, Emilio Sansano-Sansano^{1,†}, Marina Martínez-García^{$1,\dagger$}, Sergio Lluva Plaza^{$2,\dagger$}, Ana Jiménez Martín^{$2,\dagger$}, José M. Villadangos^{2,†} and Juan Jesús García Domínguez^{2,†}

¹Universitat Jaume I(UJI University), Avinguda de Vicent Sos Baynat, s/n, Castellón de la Plana, 12006, Spain ²Universidad Alcalá de Henares (UAH University) Pza. San Diego, s/n, Alcalá de Henares (Madrid), 28801, Spain

Abstract

Monitoring the activities and routines of the elderly requires continuous work and effort on the part of the care staff. It is vital to follow up on each user since abnormal behaviour could reveal the presence of a medical condition and, in these cases, early intervention is of utmost importance. In this paper, we deploy a room-level symbolic location system based on Bluetooth Low Energy (BLE) technology and present the results obtained from applying several methods capable of determining daily and weekly habits, non-invasive, using association rules and decision trees.

Keywords

Sensor networks, People Monitoring, Indoor localisation, Association rules, Decision trees

Acknowledgments

This work has been partially supported by the PID2021-122642OBproject: C44/AEI/10.13039/501100011033/ FEDER, UE and PID2021-122642OB-C41/AEI/10.13039/501100011033/ FEDER, UE.



Proceedings of the Work-in-Progress Papers at the 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN-WiP 2023), September 25 - 28, 2023, Nuremberg, Germany

^{*}Corresponding author.

[†]These authors contributed equally.

martigar@uji.es (M. Martínez-García); sergio.lluva@uah.es (S. L. Plaza); ana.jimenez@uah.es (A. J. Martín);

jm.villadangos@uah.es (J. M. Villadangos); jjesus.garcia@uah.es (J. J. G. Domínguez)

[🏶] www.linkedin.com/in/javier-gaviÃśa-rueda-a41052204 (J. G. Rueda); http://www.raulmontoliu.com

⁽R. M. Colás); https://www.uah.es/es/estudios/profesor/Sergio-Lluva-Plaza/ (S. L. Plaza);

https://www.uah.es/es/estudios/profesor/Ana-Jimenez-Martin/ (A. J. Martín);

https://www.uah.es/es/estudios/profesor/Jose-Manuel-Villadangos-Carrizo/ (J. M. Villadangos); https://www.uah.es/es/estudios/profesor/Juan-Jesus-Garcia-Dominguez/ (J. J. G. Domínguez)

D 0009-0000-2988-2199 (J. G. Rueda); 0000-0002-8467-391X (R. M. Colás); 0000-0002-7814-0383

⁽E. Sansano-Sansano); 0000-0002-2228-4396 (M. Martínez-García); 0000-0002-0050-2807 (S. L. Plaza); 0000-0003-0713-0054 (A. J. Martín); 0000-0001-5900-0978 (J. M. Villadangos); 0000-0002-7121-8651 (J. J. G. Domínguez)

^{© 02023} Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0). CEUR Workshop Proceedings (CEUR-WS.org)

1. Introduction

In today's age of technology, monitoring and tracking users' daily behaviours and activities are of great importance for medical professionals and care staff. In particular, monitoring the habits of the elderly is crucial to identify any changes in their behaviour and prevent possible diseases [1]. Location-based services (LBS, *Location Based Systems*) have undergone great development in recent decades [2]. Typically, researchers apply methods based on proximity or fingerprinting techniques [3] to estimate the position, either based on a coordinate system or symbolically (room, area, etc.). For this purpose, there are a number of technologies based on the deployment of radio frequency devices (Bluetooth, Wi-Fi, RFID, etc.), either transmitting or receiving beacons, located in known positions.

This work introduces a novel approach focused on this methodology using association rules and decision trees that allow an automatic and non-invasive determination of the daily and weekly habits of each user. In this study, a set of emitting beacons using BLE technology is used, distributed in the different rooms where the presence of people needs to be detected. Subsequently, localisation algorithms are used to obtain the symbolic location of the participants. The information obtained is acquired through the smart watches worn by the participants in this trial.

To evaluate the proposal presented in this work, we use data from a nursing home, where a set of residents wore a smartwatch throughout the study[4]. The results obtained demonstrate that by using the proposed methods, it is possible to determine the daily and weekly habits of the users in an accurate and non-invasive way, which can help prevent diseases and improve the quality of life of the elderly.

The rest of the article is organised as follows: Section II provides an exploration of the prior research and related work; Section III presents the methods used in this work and the proposed methodology; Section IV deals with the data used, the experimental phase and the results obtained; finally, Section V describes the conclusions reached and future research lines.

2. Related Work

Numerous works and approaches have utilized BLE (Bluetooth Low Energy) for various applications, such as signal intensity monitoring [5], trajectory analysis in cultural sites [6], digital contact tracing [7], and detection of routine changes [8]. The authors in [5] introduce a costeffective BLE-based sensor platform for efficient Received Signal Strength (RSS) monitoring. During performance testing, the platform exhibited high precision, with an average localization error of 0.8 meters. Moreover, the battery-powered devices demonstrated remarkably low energy consumption, totaling less than 0.2 W, making it significantly more efficient than comparable Wi-Fi CSI-based setups. Notably, the platform's design allows implementation across a wide range of BLE SoCs available in the market, rendering it an attractive cost-effective option.

Regarding the work presented in [6], it consists on an intensive study on visitor trajectory analysis in cultural sites. To achieve this, the authors employ low-cost BLE beacons for indoor positioning, enabling precise user localization in a 2D Cartesian space. Leveraging a data-driven modeling approach, a set of Fuzzy Rule Classifiers (FRC) is generated for indoor trajectories in

cultural sites. However, although the study focuses on analyzing various types of trajectories in these venues, it does not delve into daily routines or room-level routines.

On the other hand, the authors in [7] propose an interdisciplinary approach to comprehend the epidemiological, social, and technical aspects of digital contact tracing solutions for combating the COVID-19 pandemic. Utilizing BLE as the technology for digital contact tracing, the work capitalizes on its prevalence in mobile phones and its ability to provide proximity detection signals both indoors and outdoors, with relatively consistent distance estimation. Digital contact tracing is employed to identify and notify individuals who have had close contact with someone who tested positive for COVID-19, enabling them to take preventive measures.

Finally, in the work presented in [8], a BLE-based approach is proposed for detecting changes in daily routines in a controlled environment. It utilizes the 105 iBKS model and iBKS PLUS as transmitter beacons and the BQ Aquarius Plus as the receiving node. The system estimates the symbolic user's location and detects routine changes based on the average time spent in each room. Furthermore, they define a routine day when the reproducibility coefficient is below 9, corresponding to a 15% deviation of total time spent in each room. Although this approach shows promise as an initial version for detecting routine changes, its analysis is simplified to a single parameter and does not provide an easily interpretable descriptive method for routines, nor does it determine routines based on specific days of the week.

Our proposed approach leverages BLE technology to estimate room-level positions of elderly individuals residing in a nursing home. The primary goal is to discern and illustrate weeklylevel routines in a more intelligible and visually descriptive manner, employing logical rules or first-order axioms. To achieve this, we utilize association rule algorithms such as Apriori and decision trees to generate these descriptive rules

3. Methods

3.1. Technology

This project was initiated in the year 2020 within the following publication [9]. To facilitate the monitoring of the elderly, Bluetooth technology was selected as the preferred option. To achieve this objective, the beacon model 105 of the company iSBK [10] was employed. As depicted in the Figure 1, it has a small size, so it allows to be displayed in various locations in the environment. The elderly were equipped with a Sony smartwatch 3 model, running Android Wear 6.0.1. The firmware of the smartwatch has been customized to support an application that consistently scans at the highest permissible sampling rate dictated by the operating system. The collected information is stored on a microSD card. The estimated battery life ranges from 10 to 12 hours.

The foremost advantages of utilizing BLE technology over WiFi are as follows:

- 1. **Enhanced battery life**: BLE technology offers superior battery longevity compared to WiFi, with an approximate duration of 2-3 months before requiring a recharge or replacement.
- 2. Flexible beacon placement: Unlike WiFi routers, BLE beacons can be positioned freely without restrictions, allowing for more versatile and adaptable deployment options

In our approach, we employ a passive beacon methodology wherein the smartwatch autonomously captures the Received Signal Strength Indication (RSSI) of the beacons



Figure 1: [9] On the right-hand side, we have the iSBK 105 model, positioned alongside a coin for size comparison purposes. On the left-hand side, we have the Sony smartwatch 3

3.2. Indoor Positioning

Some indoor positioning algorithms aim at obtaining the symbolic location of a subject in an indoor space. In this work, a localisation method based on fingerprinting techniques is used, which is explained below. First of all, it is necessary to create the radio map $\mathcal{D} = \{\mathcal{F}, \mathcal{L}\}$ which consists of capturing, at different indoor positions, the received signal strength (RSSI) by each beacon available in the environment. In our case, the beacons are BLE. The set \mathcal{F} is defined as follows:

$$\mathcal{F} = \{\lambda_1, \lambda_2, \dots, \lambda_n\} \tag{1}$$

This set is composed of n vectors or fingerprints, stored as vectors of RSSI measurements $(\lambda_i = \{\rho_{1,i}, \rho_{2,i}, \ldots, \rho_{q,i}\}, i \in [1, n])$, where q represents the number of beacons used in the experiment, and $\rho_{q,i}$ represents the RSSI measurement of beacon q associated with sample i. On the other hand, there exists the set of \mathcal{L} annotations, which consists of an n-dimensional vector associating each fingerprint to a location. We define it as:

$$\mathcal{L} = \{\tau_1, \tau_2, \dots, \tau_n\} \tag{2}$$

The main objective of the positioning is to estimate the location of a user using the \mathcal{D} radio map as a training database. For this purpose, it is very common to use the *k*-NN algorithm which, in its most elementary version, given a test fingerprint, obtains the *k* most similar fingerprints, based on a distance function, among those included in the \mathcal{D} radio map and estimates the user's location (τ_i) as the centroid of the locations associated to the closest fingerprints.

3.3. Association Rules

Association rules is an artificial intelligence technique widely used in Data Mining, which consists of detecting and extracting intrinsic structures or patterns in the data, in an unsupervised

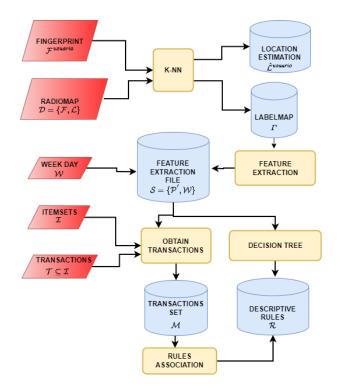


Figure 2: Outline of the proposed methodology.

way, by generating first-order logic. The rules are of the form $\{A_1, A_2, A_3, \ldots, A_m\} \Rightarrow B$, where each rule has an antecedent and a consequent. The antecedent consists of a series of predicates that are disjunctively interpreted: $A_1 \land A_2 \land A_3 \ldots \land A_m$. On the other hand, the consequent is the predicate that is concluded when the antecedent is true.

3.4. Decision Trees

A decision tree is a predictive model, quite well known within the field of machine learning, which consists in performing a series of divisions of the data, based on a measure of impurity [11]. It aims to generate a tree-like structure, where the nodes are partitions of the data, and the branches are the rules that are applied to split the data. The potential of the decision tree lies in the interpretation of the results. This allows, within the field of data mining, to detect intrinsic data patterns.

3.5. Proposed Methodology

Figure 2 shows a schematic of our proposal to obtain a set of rules to define the behaviour of users during their daily life in the nursing home. The different elements are explained below.

3.5.1. Localization

Given a user, and from the RSSI measurements obtained during the data collection campaign, we obtain one fingerprint per minute. The set of all test fingerprints is denoted by \mathcal{F}^{user} . The objective is to obtain, for each fingerprint, the location of the user (symbolic in our case), i.e. the set $\hat{\mathcal{L}}^{user}$. For this purpose, the *k*-NN algorithm with k = 1 has been used. By means of the above process, we obtain the user's location in each of the 1440 minutes of the day.

We call the above data *Labelmap* Γ and it can be expressed as shown below:

$$\Gamma = \{\varphi_1, \varphi_2, \dots, \varphi_d\} \tag{3}$$

Where *d* is the total number of days on which data has been collected from a user and φ_k is a vector of 1440 location annotations estimated ($\tau_{1,k}^{user}$) on a given day such that:

$$\varphi_k = \{\hat{\tau}_{1,k}^{user}, \dots, \hat{\tau}_{L,k}^{user}\}, k \in [1,d], L = 1440$$
(4)

3.5.2. Features extraction

Once the Labelmap set (Γ) is defined, feature extraction is applied by partitioning temporary windows according to the chosen requirements. For each partition, the number of occurrences of each symbolic location for each φ_k is counted.

 $C_{(t_0,t_1)}$ is an occurrence function such that:

$$\mathcal{C}_{(t_0,t_1)} : \varphi_k \to \{ f_{(t_0,t_1)}^{k,1}, f_{(t_0,t_1)}^{k,2}, \dots, f_{(t_0,t_1)}^{k,c}, \dots, f_{(t_0,t_1)}^{k,R} \}
\forall_{t_0,t_1} \in [1, 1440], \ t_1 > t_0, \ k \in [1,d], \ c \in [1,R]$$
(5)

where:

- (t_0, t_1) is the time interval (in minutes of the day) at which we calculate the occurrence of a symbolic location.
- $f_{(t_0,t_1)}^{k,c}$ the number of minutes the user is in room c, on day k, for the time window (t_0,t_1) .
- *R* is the number of different locations in this study.

The dataset resulting from feature extraction, \mathcal{P} , can be obtained as follows:

$$\mathcal{P}_{k} = \bigcup_{j=1}^{z} \mathcal{P}_{k}^{j} = \bigcup_{j=1}^{z} \mathcal{C}_{(t_{0}^{j}, t_{1}^{j})}(\varphi_{k}) = \begin{cases} f_{(t_{0}^{1}, t_{1}^{1})}^{k,1}, f_{(t_{0}^{1}, t_{1}^{1})}^{k,R}, f_{(t_{0}^{1}, t_{1}^{1})}^{k,1}, f_{(t_{0}^{2}, t_{1}^{2})}^{k,1}, f_{(t_{0}^{2}, t_{1}^{2})}^{k,2}, \dots, f_{(t_{0}^{2}, t_{1}^{2})}^{k,R}, \\ \dots, f_{(t_{0}^{2}, t_{1}^{2})}^{k,1}, f_{(t_{0}^{2}, t_{1}^{2})}^{k,2}, \dots, f_{(t_{0}^{2}, t_{1}^{2})}^{k,R}, \\ \{\forall_{t_{0}, t_{1}} \in [1, 1440], \ \forall_{j} \in [1, z] \ | \ t_{1}^{j} > t_{0}^{j}, \ t_{0}^{j+1} > t_{1}^{j} \} \\ k \in [1, d], \ \varphi_{k} \subseteq \Gamma \end{cases}$$

$$(6)$$

Where:

- $j \in [1, z]$ corresponds to temporary partition j over z total partitions.
- (t_0^j, t_1^j) is the time interval in which we calculate the occurrence of a symbolic location in the time partition j.
- $C_{(t_n^j, t_n^j)}$ (see Eq. 5) is the occurrence function in time partition j.
- $f_{(t_0^j,t_1^j)}^{k,c}$ is the total number of minutes in which the user is in the room.

Finally, we discretise the set \mathcal{P} by applying the following:

$$f_{(t_0^j, t_1^j)}^{'k, c} = \begin{cases} 1 & if \quad f_{(t_0^j, t_1^j)}^{k, c} \ge t_{\varepsilon} \\ 0 & if \quad another \ case \end{cases} ; t_{\varepsilon} \in [1, 1440]$$
(7)

With t_{ε} as the threshold time, in minutes, for which we consider whether a subject is at a location. We apply this discretisation to the set \mathcal{P} and obtain a data set \mathcal{P}' . This set contains d rows, where each row \mathcal{P}'_k is a vector of size zR as described by the equation (8).

$$\mathcal{P}'_{k} = \{ f'^{k,1}_{(t^{1}_{0},t^{1}_{1})}, f'^{k,2}_{(t^{1}_{0},t^{1}_{1})}, \dots, f'^{k,R}_{(t^{1}_{0},t^{1}_{1})}, \dots, f'^{k,1}_{(t^{2}_{0},t^{2}_{1})}, \dots, f'^{k,R}_{(t^{2}_{0},t^{2}_{1})} \}$$

$$k \in [1,d], \quad c \in [1,R], \quad j \in [1,z], \quad f'^{k,c}_{(t^{j}_{0},t^{j}_{1})} \in \{0,1\}$$

$$(8)$$

Where:

- $\mathcal{P}' \in \mathbb{R}^{d \times zR}$ is the dataset resulting from Γ feature extraction (see Eq. 6) and discretisation (see Eq. 7), with d rows and zR columns.
- *z* number of partitions.
- $f_{(t_0^j,t_1^j)}^{'k,c}$ is a dichotomous variable indicating whether on day k, in the time interval (t_0^j,t_1^j) , the user is in room c, $f_{(t_0^j,t_1^j)}^{'k,c} = 1$, or he/she is not, $f_{(t_0^j,t_1^j)}^{'k,c} = 0$.

Finally, we obtain the annotated feature extraction set S, where for each day k we have annotated the corresponding day of the week $w_k \in W$:

$$S = \{\mathcal{P}', \mathcal{W}\}\tag{9}$$

Being $W \in \{Monday, Tuesday, \dots, Sunday\}$, a vector of annotations of dimension $\mathbb{R}^{d \times 1}$.

The dataset obtained with the extraction of annotated features S is the one we will use to extract the association rules later on.

3.5.3. Set of Transactions

In order to use the association rule algorithm, *a priori*, we have to transform S (see Eq. 9) into a set of \mathcal{M} transactions. For this, we define \mathcal{I} as the set of m distinct attributes (or items) and \mathcal{T} as the transactions containing a set of \mathcal{I} items, such that $\mathcal{T} \subseteq \mathcal{I}$ [12]. We define \mathcal{M} to be the set of all transactions \mathcal{T} such that:

$$\mathcal{M} = \{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_d\}$$
(10)

We apply a transformation to obtain the set of itemsets \mathcal{I} . Each row of \mathcal{S} is represented by a vector of size zR + 1, where each attribute is a dichotomous variable $f_{(t_0^j, t_1^j)}^{'c} \in \{0, 1\}$ indicating presence in a room, and \mathcal{W} (day of the week) is a categorical variable of seven possible values. For each dichotomous variable, two items (dichotomous complementary attributes) $f_{(t_0^j, t_1^j)}^{'c} \rightarrow \{I_1, I_2\}$ are generated. The transformation generates the vector of attributes that compose the itemset $\mathcal{I} \in \mathbb{R}^{2zR+7}$, a set of 2zR + 7 items that represent all the possible combinations that a user can make in a transaction.

Finally, each transaction $\mathcal{T}_k \subseteq \mathcal{I}$ is a set of items observed on day k. If we put all the transactions together we obtain the set \mathcal{M} that we can see represented in Table 1.

Table 1

Example of a set of transactions \mathcal{M}

Transactions	Items list
\mathcal{T}_1	$\{Item_1,Item_2,Item_5,Item_{Monday}\}$
$ $ \mathcal{T}_2	$\{$ Item ₁ , Item ₂ , Item ₅ , Item _{Monday} $\}$ $\{$ Item ₃ , Item ₂ , Item ₄ , Item _{Tuesday} $\}$
\mathcal{T}_3	$\{Item_2, Item_5, Item_{Wednesday}\}$
\mathcal{T}_d	$\{Item_6, Item_7, Item_{Monday}\}$

3.5.4. Rule extraction file

For automatic pattern determination, a rule-based descriptive method is used. To describe a subject, we have to obtain the rule extraction file $\mathcal{R} = \{r_1, r_2, \ldots, r_p\}$, in which each rule r_i has the form shown in equation (11).

$$r_1: \{A_1, A_2, \dots, A_n\} \mapsto B \tag{11}$$

There are three metrics used in data mining to evaluate the performance of a rule extracted from data:

- Support. Proportion of transactions containing $A \cup B$ versus the total number of transactions: $\frac{|A \cup B|}{|M|}$
- Confidence. Proportion of transactions containing $A \cup B$ against the support of the antecedent: $\frac{|A \cup B|}{|A|}$
- Lift [13]. It measures the distance between the observed and expected support under the assumption of independence between the elements: $\frac{support(A \cup B)}{support(A) \cdot support(B)}$

For the automatic extraction of rules we have considered two alternatives. The first one consists of using *a priori* association rule algorithms. The second one consists in training a decision tree and, with the partitions made by the tree, converting them into rules and calculating the three metrics mentioned above.

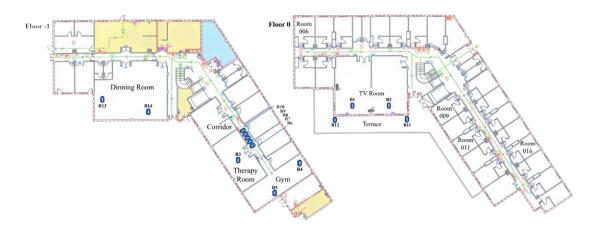


Figure 3: Illustration of a residential map depicting the spatial distribution of beacons across rooms

4. Experiments and Results

4.1. Data

In order to evaluate the proposal presented in this paper, the sixth campaign of the database published in [14] has been used. During this campaign, six volunteers were provided with a smartwatch. These volunteers were required to wear the watch at all times, except when the watch was removed by the care staff to be recharged (usually before going to bed). During the day the watch is continuously acquiring the RSSI of the different beacons deployed throughout the nursing home as shown in Figure 3. This campaign lasted 8 weeks.

4.2. Experiments

To adjust the hyperparameter k of the k-NN algorithm, we used a validation set included in the database [14] in which a set of fingerprints was obtained in each of the locations included in the nursing home (room, dining room, gym,...). By means of cross-validation, the performance of the k-NN algorithm was evaluated for different values of k, being k = 1 the value for which the best result was obtained, where the accuracy of the classifier, for the validation set, is found to be very high (100%).

It is with this trained model that we will obtain all the symbolic annotations of the users of the residence in order to obtain the *labelmap* Γ shown in Figure 4. As this figure shows, it can be seen that practically all users go to the dining room (orange colour) at the same time. However, the activities before or after the meals vary greatly between users.

After obtaining Γ , we divide the time windows into two partitions: morning (9:00-13:30) and afternoon (13:31-16:50). To do so, we use the values shown in Table 2.

We apply these values to the occurrence function $C_{(t_0^j, t_1^j)}$ (see Eq. 5) and obtain the feature extraction file \mathcal{P} . Finally, we discretise it with $t_{\varepsilon} = 5$ minutes, join it to the weekly annotations \mathcal{W} and obtain the discretised and annotated feature extraction file \mathcal{S} .

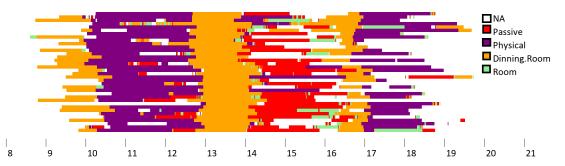


Figure 4: Visual representation of the *labelmap* of one of the users of campaign 6 where the rows are the days and the columns are the minutes of the day. Each colour corresponds to a room. The minutes where no data collection takes place are represented as white colour.

Table 2

Partitions t_0^j y t_1^j to obtain \mathcal{P} .

	Partition (j)	t_0^j (minutes)	t_1^j (minutes)
Morning	1	540	810
Afternoon	2	811	1010

The next stage is the clustering of locations with similar activities and at a close distance, to define three groups:

- Physical activity: Gym, corridor, therapy room.
- Passive activity: TV room, terrace.
- Room: Room.

In this work, we have excluded the locations *living room* and *dining room* as we want to focus on the activities they carry out in the morning (before lunch) and in the afternoon.

One way to visualise volunteers' routines is by using frequency tables. If we group by day of the week, we can observe for a specific volunteer, the percentage of times that on a Monday, for example, they are in the room, doing some kind of physical activity or doing a passive activity. Similarly, we can determine for the entire campaign, the percentage of times that each volunteer has been found doing each of the activities. In this way, we can determine which volunteers are more similar or if any volunteer has unusual behaviour.

We have applied the *a priori* algorithm with the package *arules* in *R* [15]. For this problem, we start with transaction data. The first step is to extract daily rules and remove them and then extract weekly rules. To do the former, we use a function *itemFrequency*, which returns the individual support of each item belonging to \mathcal{I} , and we eliminate the items whose support is higher than a certain threshold which, in our case, we have set at 90%.

For the rule generation with a priori we have defined the following parameters:

- **maxlen** = 4 \rightarrow The number of items making up the antecedent does not exceed four
- **support** = $1e-5 \rightarrow$ Minimum support to generate the rule
- **confidence** = $1e-5 \rightarrow$ Minimum confidence to generate the rule

For the decision trees, we have used the *rpart* and *tidyrules* libraries to train the tree and extract association rules from the discretised and annotated dataset. First, we extract the daily rules with *itemFrequency* and then we fit a decision tree for each subject using the CART algorithm [16] with a maximum depth of 3 and a minimum of 4 samples per node. Finally, we extract rules from the tree and obtain the metrics with *tidyRules*. In our case, we have limited the depth of the tree to maintain interpretation without having an excessive number of antecedents in a rule.

4.3. Results

To determine the accuracy of the algorithms used for automatic rule extraction, we have an annotation file in which the care staff detailed the daily and weekly behaviours of each of the participants in the study. In this way, we can extract rules and check whether there is a contradiction with those noted in the file. In this case, we will perform this analysis with user 9FE9 from campaign 6. The rules provided by the care staff that we can extract from the annotation file are the following:

• Every day

- In the morning: the volunteer does physical activity.
- In the afternoon: the volunteer usually does physical or passive activities and goes to his/her room when finished.
- Tuesdays: the volunteer does physical or passive activity in the morning.
- Twice a week: the volunteer spends the whole morning doing physical activity.

To find patterns, we first use a daily frequency table (see Table 3). From user 52EA we noticed that he/she does not do physical activity and he/she is not in leisure areas, spending most of the time in his/her room. In contrast, user 9FE9 is physically active and is usually found engaging in passive activities during the afternoon, such as spending time on the terrace or watching TV. We also use a weekly frequency table (see Table 4) to see which activities are performed on each day of the week. From user 02A8, we can see that he/she engages in less physical activity on Thursdays and spends most of his/her time doing passive activities on Tuesday afternoons. On the other hand, user 9FE9 is not usually in his/her room on Thursdays and Fridays. Instead, he/she is frequently found doing some physical activity or any other type of activity.

Whether we use an association rules algorithm or decision trees, we have first obtained the absolute rules with the *itemFrequency*. In the case of volunteer 9FE9, we extract as a daily rule $Physical_{Morning} = 1$, as we can see in Figure 5. This rule agrees with the annotations provided by the care staff.

The next step is to apply the *a priori* algorithm and a decision tree to volunteer 9FE9. First, we remove the *Physical*_{Morning} attribute, the daily rule obtained with *itemFrequency*, and we obtain a weekly rule set of 315 rules with *a priori*. We select the seven best rules according to the gain and obtain the seven strongest rules that we see in Table 5. Finally, we train a decision tree, as we see in Figure 6 and extract its weekly rules with *tidyRules* and obtain a set of 5 weekly rules. Table 6 shows these rules.

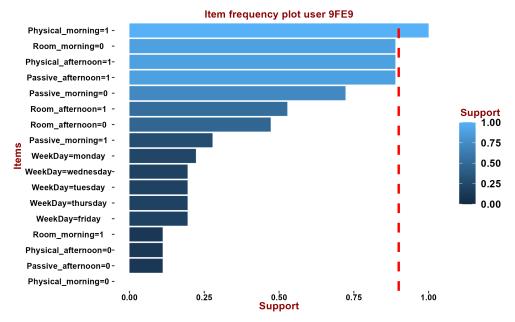


Figure 5: Bar chart showing the proportion of times each item appears in the dataset for volunteer 9FE9. The red line at 0.9 delimits the threshold that we have established for considering the rule

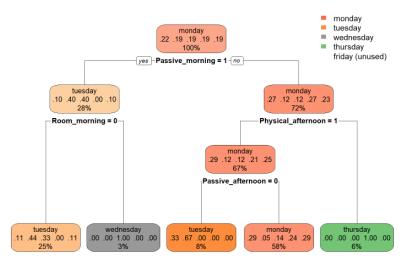


Figure 6: Decision tree trained with the volunteer 9FE9.

Table 3

Percentages of total activities.

Volunteer	Physical morning	Physical afternoon	Passivemorning	Passiveafternoon	Room morning	Room _{afternoon}
02A8	19.35%, n=31	61.29%, n=31	0%, n=31	64.52%, n=31	96.77%, n=31	100%, n=31
402E	94.44%, n=36	97.22%, n=36	30.56%, n=36	36.11%, n=36	19.44%, n=36	38.89%, n=36
52EA	13.33%, n=30	3.33%, n=30	0%, n=30	0%, n=30	96.67%, n=30	93.33%, n=30
682A	41.67%, n=36	2.78%, n=36	50%, n=36	5.56%, n=36	100%, n=36	91.67%, n=36
9FE9	100%, n=36	88.89%, n=36	27.78%, n=36	88.89%, n=36	11.11%, n=36	52.78%, n=36
F176	94.59%, n=37	89.19%, n=37	40.54%, n=37	83.78%, n=37	56.76%, n=37	94.59%, n=37

Table 4

Total weekly percentages.

Volunteer	Day	Physical morning	Physical _{afternoon}	Passivemorning	Passiveafternoon	Room _{morning}	Room _{afternoon}
02A8	Monday	25%, n=8	87.5%, n=8	0%, n=8	50%, n=8	100%, n=8	100%, n=8
02A8	Tuesday	50%, n=4	0%, n=4	0%, n=4	100%, n=4	100%, n=4	100%, n=4
02A8	Wednesday	0%, n=7	85.71%, n=7	0%, n=7	57.14%, n=7	100%, n=7	100%, n=7
02A8	Thursday	16.67%, n=6	16.67%, n=6	0%, n=6	83.33%, n=6	83.33%, n=6	100%, n=6
02A8	Friday	16.67%, n=6	83.33%, n=6	0%, n=6	50%, n=6	100%, n=6	100%, n=6
9FE9	Monday	100%, n=8	100%, n=8	12.5%, n=8	87.5%, n=8	25%, n=8	75%, n=8
9FE9	Tuesday	100%, n=7	85.71%, n=7	57.14%, n=7	71.43%, n=7	14.29%, n=7	42.86%, n=7
9FE9	Wednesday	100%, n=7	85.71%, n=7	57.14%, n=7	100%, n=7	14.29%, n=7	57.14%, n=7
9FE9	Thursday	100%, n=7	71.43%, n=7	0%, n=7	85.71%, n=7	0%, n=7	42.86%, n=7
9FE9	Friday	100%, n=7	100%, n=7	14.29%, n=7	100%, n=7	0%, n=7	42.86%, n=7

Table 5

Rules obtained (sorted by Lift) using **APRIORI** for volunteer 9FE9.

	Antecedent		Consequent	Support	Confidence	Lift	Frequency
			W ID T I	0.020		5.1.4	
r_1 :	$\{Room_{morning} = 1, Passive_{afternoon} = 0\}$	\Rightarrow	WeekDay=Tuesday	0.028	I	5.14	I
r_2 :	$\{Physical_{afternoon}=0, Passive_{afternoon}=0\}$	\Rightarrow	WeekDay = Thursday	0.028	1	5.14	1
r_3 :	$\{Room_{morning}=1, Passive_{morning}=1\}$	\Rightarrow	WeekDay = Wednesday	0.028	1	5.14	1
r_4 :	{Passive _{morning} =0, Physical _{afternoon} =0}	\Rightarrow	WeekDay = Thursday	0.05	1	5.14	2
r_5 :	$\{Room_{morning}=1, Room_{afternoon}=0, Passive_{afternoon}=0\}$	\Rightarrow	WeekDay = Tuesday	0.028	1	5.14	1
r_6 :	{Room _{morning} =1, Passive _{morning} =0, Passive _{afternoon} =0}	\Rightarrow	WeekDay = Tuesday	0.028	1	5.14	1
r_7 :	$\{Room_{morning}=1, Physical_{afternoon}=1, Passive_{afternoon}=0\}$	\Rightarrow	WeekDay = Tuesday	0.028	1	5.14	1

Table 6

Rules obtained (sorted by Lift) using **decision tree** for volunteer 9FE9

	Antecedent		Consequent	Support	Confidence	Lift	Frequency
r_1 :	{Passive _{morning} =0, Physical _{afternoon} =0}	\Rightarrow	WeekDay=Thursday	0.06	0.75	3.86	2
r_2 :	$\{Passive_{morning}=1, Room_{morning}=1\}$	\Rightarrow	WeekDay = Wednesday	0.03	0.67	3.43	1
r_3 :	$\{Passive_{morning}=0, Physical_{afternoon}=1, Passive_{afternoon}=0\}$	\Rightarrow	WeekDay = Tuesday	0.08	0.6	3.09	3
r_4 :	$\{Passive_{morning}=1, Room_{morning}=0\}$	\Rightarrow	WeekDay = Tuesday	0.25	0.45	2.34	9
r_5 :	$\{Passive_{morning}=0, Physical_{afternoon}=1, Passive_{afternoon}=1\}$	\Rightarrow	WeekDay = monday	0.58	0.3	1.37	21

5. Conclusions and Future Work

This paper discusses the technology used to obtain data, the steps involved in the estimation of symbolic location and the methods applied for the descriptive analysis of people's rule-based habits. The data used in the paper are limited due to time mismatches, excessive missing data and small sample sizes. Rule association algorithms have potential, but require a large amount of sample information and may generate redundant rules. On the other hand, decision trees provide robust rules and a clear visualisation and interpretation of them. As possible lines of future work, we propose to improve the sample quantity by generating synthetic data and the possibility of using Bayesian networks. Alternatively, we are working on a new longer data campaign in order to better verify this proposal.

References

- R. Liu, A. A. Ramli, H. Zhang, E. Henricson, X. Liu, An overview of human activity recognition using wearable sensors: Healthcare and artificial intelligence, in: International Conference on Internet of Things, Springer, 2021, pp. 1–14.
- [2] M. Biswas, A. Rahman, M. S. Kaiser, S. Al Mamun, K. S. Ebne Mizan, M. S. Islam, M. Mahmud, Indoor navigation support system for patients with neurodegenerative diseases, in: Brain Informatics: 14th International Conference, BI 2021, Virtual Event, September 17–19, 2021, Proceedings 14, Springer, 2021, pp. 411–422.
- [3] R. Montoliu, A. Pérez-Navarro, J. Torres-Sospedra, Efficient tuning of *k* nn hyperparameters for indoor positioning with n-tbea, in: 2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), IEEE, 2022, pp. 39–44.
- [4] S. L. Plaza, R. M. Colás, A. J. Martín, J. M. V. Carrizo, E. Sansano-Sansano, J. J. G. Domínguez, Ble rssi database for analyzing routines of community-dwelling older adults, in: 2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), IEEE, 2022, pp. 50–55.
- [5] G. Maus, N. Kuxdorf-Alkirata, D. Brückmann, M. Gemci, A low-cost ble based sensor platform for efficient received signal strength monitoring, in: 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, 2019, pp. 1199–1202.
- [6] M. Orenes-Vera, F. Terroso-Saenz, M. Valdes-Vela, Recite: A framework for user trajectory analysis in cultural sites, Journal of Ambient Intelligence and Smart Environments 13 (2021) 389–409.
- [7] V. Shubina, A. Ometov, A. Basiri, E. S. Lohan, Effectiveness modelling of digital contacttracing solutions for tackling the covid-19 pandemic, The Journal of Navigation 74 (2021) 853–886.
- [8] A. J. Martín, I. M. Gordo, D. G. Gómez, S. G. de Villa, S. L. Plaza, J. J. G. Domínguez, Ble-based approach for detecting daily routine changes, in: 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), IEEE, 2021, pp. 1–6.
- [9] S. L. Plaza, R. M. Colás, A. J. Martín, J. M. V. Carrizo, E. Sansano-Sansano, J. J. G. Domínguez, Ble rssi database for analyzing routines of community-dwelling older adults, in: 2022 14th

International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2022, pp. 50–55. doi:10.1109/ICUMT57764.2022.9943410.

- [10] model 105 iBKS, Accent systems, 2018. Available online: https://accent-systems.com/ product/ibks-105. Last accesed: 15 June 2023.
- [11] R. Buaton, H. Mawengkang, M. Zarlis, S. Effendi, A. M. H. Pardede, Y. Maulita, A. Fauzi, N. Novriyenni, A. Sihombing, K. Lumbanbatu, Decision tree optimization in data mining with support and confidence, in: Journal of Physics: Conference Series, volume 1255, IOP Publishing, 2019, p. 012056.
- [12] S. Kotsiantis, D. Kanellopoulos, P. Pintelas, et al., Handling imbalanced datasets: A review, GESTS international transactions on computer science and engineering 30 (2006) 25–36.
- [13] P. D. McNicholas, T. B. Murphy, M. O'Regan, Standardising the lift of an association rule, Computational Statistics & Data Analysis 52 (2008) 4712–4721.
- [14] S. L. Plaza, R. M. Colás, A. J. Martín, J. M. V. Carrizo, E. S. Sansano, J. J. G. Domínguez, BLE RSSI Database for Analyzing Routines of Community-Dwelling Older Adults, 2022. URL: https://doi.org/10.5281/zenodo.6551614. doi:10.5281/zenodo.6551614.
- [15] The r project for statistical computing, https://www.r-project.org/, 2023. Last accessed: 28 March 2023.
- [16] L. Breiman, J. Friedman, R. Olshen, C. Stone, Classification and regression trees. wadsworth int, Group 37 (1984) 237–251.