Analyzing daily behaviours from wearable trackers using linguistic protoforms and fuzzy clustering

Carmen Martinez-Cruz,
Javier Medina Quero,
Macarena Espinilla Estevez
University of Jan
Departament of Computer Science
Campus Las Lagunillas, 23071 Jan
`cmcruz, jmquero, mestevez@ujaen.es`

Sergio Gramajo
National Technological University
Resistencia,Argentina
`sergio@frre.utn.edu.ar`

Abstract

The proliferation of low-cost wearable trackers are allowing users to collect daily data from human activity in a non-invasive way and outside of laboratory environments. Exploiting these data properly enable the supervision and counseling from experts remotely; however, extracting key indicators from the long datastreams is hard, often based on statistical metrics or clustering from raw data which lack interpretability. To solve it, we propose an interpretable definition of key indicators by means of linguistic protoforms which include fuzzy temporal processing and fuzzy semantic quantification. Moreover, we use the protoforms defined by experts to evaluate the source datastream in order to provide a straightforward description of the daily activity of users. Finally, the degrees of truth of each protoform are analyzed using a fuzzy clustering method to provide an interpretable description of the longterm user activity. This work includes a case study where data from a user activity (heart beats per minute and sleep stages) have been collected by a Fitbit wearable device and evaluated by the proposed methodology.

1 Introduction

The increase of wearable activity trackers has led to a massive growth in their use in the population (Shih et al., 2015). The use of these devices has proved to increase physical activity between young (Heale et al., 2018) and older (Mercer et al., 2016a) adults and promote a health behavior change (Mercer et al., 2016b). Consequently, wearable activity trackers are the key in new interventions to avoid physical inactivity, which contributes to an estimated 3.2 million deaths each year (Lim et al., 2012).

Among the most relevant data recorded by these devices are heartbeats per second, sleep stages duration time and any sleep disturbance, such as the reduction of the overall sleeping hours or the excessive sleep, that would result in an increase in the warning signs to experts. Poor sleep quality is associated with chronic diseases, weight increase and cognitive dysfunction. The National Sleep Foundation emphasizes following the sleep level targets and guideline and study how these technologies can be useful in this sense, such as the smartwatches. In the near future, this kind of devices may tell us the same as a sleep laboratory. Home monitoring trough smartwatch solutions offers the possibility of sleep coaching interventions or performing analysis to detect any other healthy problems designed by experts (Foundation, 2019).

According to this, smart tracking health devices, such as smartwatches and smartphone APPs, have become increasingly popular. These devices claim to monitor several human activities, and one of the most analyzed is the sleep duration of their users. Most of these devices utilize data generated from in-built sensors to determine sleep parameters. There are many studies that evaluate and compare the accuracy of these sleep tracking devices against more conventional methods used to measure sleep duration and quality (Kang et al., 2017; Kolla et al., 2016). In this way, different commercial smartwatches monitor several aspects of human activity and provide statistical information to users (Bai et al., 2018). The usefulness of this kind of devices were evaluated and validated for monitoring participant sleep levels outside the laboratory environment (Dickinson et al., 2016). Thus, the use of low-cost sleep monitoring devices like Fitbit can help to assess sleep trends where clinical accuracy of laboratory is not necessary (Dickinson et al., 2016) nowadays. Notwithstanding the results indicate that a reasonable degree of sleep staging accuracy can be achieved using a wearable device, which may be of utility in longitudinal studies of sleep habits (Beattie et al., 2017).

Besides, the relative efficacy of different approaches to improve physical activity and sleep using technology-based methods have been examined, although their relatively efficacy to improve these behaviors has not been directly compared (Duncan et al., 2016). In addition to this, some tools have been designed to help the understanding of the sleep quality through contextual information obtained by data from wearables devices (Liang et al., 2016).

In this paper, we present a methodology to analyze the daily activity of users which have been collected by a wearable tracker. A linguistic approach allows to describe the resulting health key indicators (HKIs) and a fuzzy clustering process has been applied to detect some user behaviour patterns. The key points of the proposed methodology are the following:

- A reliable wearable device Fitbit is used to obtain activity tracking data through a bluetooth/wireless connection (Diaz et al., 2015). It provides a HTTPS Web API for accessing data from Fitbit, i.e., automatic activity logs and manually entered records. An application to access and analyze the Fitbit user's data, specially, those related with heart-beats per minute rate and the duration of the sleep activity status (wake, restless, light sleep, REM phase, deep sleep) has been developed here.

- Collected data has been used to define the user most relevant HKIs using protoforms. These protoforms has been designed by the expert knowledge of the supervisor of the user activity. Protoforms summarize the collected information and select the time interval of the day which better suits with the expert criteria using linguistic temporal terms and linguistic quantifiers that provides expressiveness and semantic to the result.

- A fuzzy clustering process is applied to the aggregated truth degree of each day protoform, in order to analyze the common activity patterns for a given user. The suitable relation among the proposed clusters and protoforms enables an interpretable representation of daily activity of users which is meaningful to the supervisor of the user activity.

A review of previous researches related to our work have been included in Section 2. The rest of the paper is organized as follows: in Section 3, we present a methodology to analyze data and represent this knowledge through linguistic protoforms. Some experimental results performed on a dataset from a Fitbit data stream are shown in Section 4. Finally, the conclusion and future work is provided in Section 5.

2 Related works

On the first hand, knowledge-driven methods have been proposed to describe daily human activity by means of sensors (Chen et al., 2011; Medina-Quero et al., 2016). The difficulty of these approaches lies in translating the expert knowledge into a computational method in a flexible, interpretable and rigorous way. Among the wide range of approaches, fuzzy logic (Zadeh, 2006) has been described as an high-interpretable knowledge model for reasoning (Zadeh, 2002) and aggregating (Kacprzyk and Yager, 2001) data under uncertainty. Moreover, the use of protoforms and fuzzy logic (Zadeh, 2002) have provided encouraging results integrating expert knowledge to process sensor data streams in multiple areas, such as, weather forecast (Ramos-Soto et al., 2014), prediction of the urgency demand within smart cities (Medina Quero et al., 2018), providing linguistic summaries from heart rate streams (Peláez-Aguilera et al., 2019) or monitoring of patients with preeclampsia in wearable devices(Espinilla et al., 2017).

On the other hand, data-driven methods have proliferated relating features from sensor streams (Okeyo et al., 2014) to human activity by means of Machine Learning approaches (Minor et al., 2015; Choi et al., 2017). A large majority of these works have focused on supervised learning, where an extensive labelled dataset is required to classify targeted human behaviours (De-La-Hoz-Franco et al., 2018). This requirement faces up with the individual learning of key interest indicators since it is not agile collect and label huge datasets for each person and indicator. In this way, egocentric daily activity recognition (Yan et al., 2015) is mainly supported by non supervised methods, where clustering algorithms (Xu and Tian, 2015) provide the discovering of daily patterns and tasks from users. We highlight the integration of fuzzy approaches with clustering methods, specifically the fuzzy C-means algorithm (Bezdek et al., 1984), which have provided suitable methods to extract meaningful patterns from sensors (Moreno-Cano et al., 2015).

Here, we propose an agile linguistic description of HKIs through the use of linguistic protoforms to pool the advantage of knowledge-drive and automatic learning. These protoforms are computed over the sensor data stream collected by a Fitbit wearable tracker. The daily aggregation of protoforms is subsequently evaluated by fuzzy clustering, which provides the most relevant user behaviour patterns.

3 Methodology

In this section we present a methodology to analyze synchronized source streams from wearable trackers under a semi supervised approach. Firstly, in Section 3.1 we describe the source streams and platform tools to collect activity data. Secondly, in Section 3.2, a linguistic approach based on protoforms is presented to define the HKIs from the source streams. Third, in Section 3.3, the degree of truth of each daily instanced protoform is aggregated and evaluated by a fuzzy clustering process. The resulting clusters expose a linguistic and visual framework to identify the behaviour patterns of an user.

3.1 Collecting source streams from wearable trackers

Fitbit IONIC smartwatch has monitorized and collected sleep stages duration and heart rate activities for an user in a period of time for this study. This device has been chosen because of its many advantages over others models related with the features of: multi-day battery life, accuracy, all sleep stages, accessibility to use datasets in cloud to preprocessing of data, compatibility, data storage and integrated GPS antenna (Bai et al., 2018).

In a formal way, each source stream s^l is represented by a 2-tuple value $\bar{s}_i^l = \{s_i^l, t_i\}$, where s_i^l defines a given value collected by the wearable tracker s^l and t_i its time-stamp. Hence, the long-term information from a given user is composed of a data stream $S^l = \{\bar{s}_0^l, \ldots, \bar{s}_i^l, \ldots, \bar{s}_n^l\}$, which is collected by the wearable device.

For the aim of this work, the two target source streams collected by the wearable tracker Fitbit are:

- s^{as} defines the activity status (AS) by means 5 discrete values: wake (WK), restless (RS), light sleep (LS), REM phase (RP), deep sleep (DS); so $s_i^{as} \in \{WK, RS, LS, RP, DS\}$.

- s^{hr} defines the heart rate (HR) by a natural number which represents the beats per minute (bpm) in the human range; so $s_i^{hr} \in [40, 220]$.

3.2 Protoform for evaluating source streams from wearable trackers

The proposed methodology aims to define the HKIs from Fibit source streams using a set of a protoform instances, which are straightforwardly defined by the supervisor of user activity who has the expertise knowledge in this context.

First, we introduce an ad-hoc *protoform*, to integrate an interpretable and rich-expressive approach that models the expert knowledge in a linguistic way, P_o in the shape of:

$P_o(\bar{s}_i^l)$: V_r T_j Q_k

Where:

- V_r defines a fuzzy linguistic term to evaluate the data-stream.

- T_j defines a Fuzzy Temporal Window (FTW) where the term V_r is aggregated. The FTWs are described straightforwardly according to the distance from the current time t^* to a given timestamp t_i as $\Delta t_i = t^* - t_i$ using the membership function $\mu_{T_j(\Delta t_i)}$. The aggregation functions of V_r over T_j for a given $\bar{s}_i^l$ are defined by the following t-norm and t-conorm:

$$V_r \cap T_j(\bar{s}_i^l) = V_r(s_i^l) \cap T_j(\Delta t_i) \in [0, 1]$$

$$V_r \cup T_j(\bar{s}_i^l) = \bigcup_{\bar{s}_i^l \in S^l} V_r \cap T_j(\bar{s}_i^l) \in [0, 1]$$

where a fuzzy weighted average (Dong and Wong, 1987) (FWA), which is defined in appendix Abbreviations, is proposed to model these functions (Peláez-Aguilera et al., 2019).

- Q_k defines a fuzzy quantifier to evaluate the *intensity* of the linguistic term V_r within the temporal window T_j (Medina-Quero et al., 2016). The quantifier applies a transformation $\mu_{Q_K} : [0, 1] \rightarrow [0, 1]$ to the aggregated degree of $\mu_{Q_K}(V_r \cup T_k(s_i^l))$.

This shape of protoform has been successfully developed in summarizing the stream data from health devices (Peláez-Aguilera et al., 2019; Medina-Quero et al., 2016) using a linguistic approach.

In Figure 1, an example of the instantiated protoform *Activity status is deep sleep around 2 and 4*
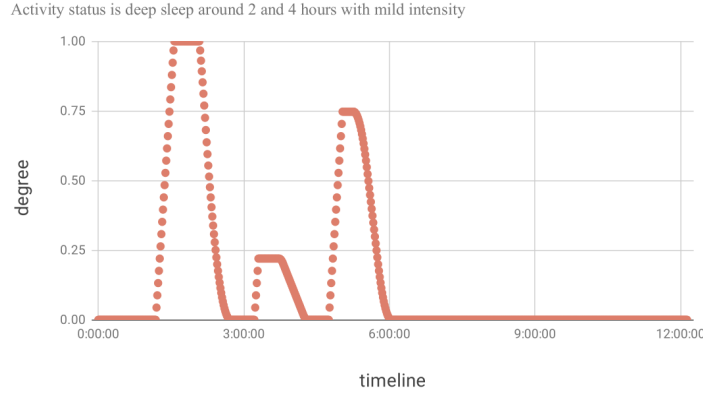
Figure 1: Degree of truth of the protoform *Activity status is deep sleep around 2 and 4 hours WITH mild intensity* that summarizes 12 hours of source stream.

hours with mild intensity is shown, which is computed over a fragment of the source stream from Fitbit.

Moreover, protoform $P_o(\bar{s}_i)$ can be combined using fuzzy logical operators to increase the linguistic expressiveness of the model, i.e, negation, union or intersection operators (Peláez-Aguilera et al., 2019) can be included in the final description of the stream. In addition to this, the set of protoform instances can be replaced by a shorter linguistic expressions, much closer to natural language, e.g. *deep restful sleep* that describes the previously analyzed protoform instance *Activity status is deep sleep around 2 and 4 hours with mild intensity*.

3.3 Fuzzy clustering to detect behaviour patterns

In this section, we aim to identify patterns from the users daily activity collected by the wearable tracker. To do that, the membership degree described by protoform instances is analyzed to provide a linguistic description of the most relevant user activity thought of the discovered patterns.

First, we compute the truth degree (P_o^T) of each protoform for each day. To compute it, the aggregation function $\bigcup$ (implemented by the max function) is applied to get the degree of truth of the protoform $P_o(\bar{s}_i)$ for each fragment of day T:

$$P_o^T : \bigcup P_o(\bar{s}_i), \bar{s}_i^l \in T$$

Second, the degrees of truth of the protoform instances, that represents the value of an user HKIs in a given day, are evaluated to extract the common patterns in a long term evaluation. To do

Table 1: Degree to which daily protoform instances $P_{1...5}$ belong to each cluster.

Cluster	P_1	P_2	P_3	P_4	P_5
Short description	high HR	low-intensity HR	deep sleep	light sleep	rest nap
C1	0.05	0.02	0.02	0.03	0.15
C2	0.86	0.50	0.54	0.85	0.91
C3	0.63	0.53	0.56	0.90	0.03

that, we applied fuzzy clustering using Fuzzy C-means algorithm (Bezdek et al., 1984) over the maximal daily degree of the protoform instances. It is important to notice that the aim of combining these representation of HKIs with fuzzy clustering relies in obtaining a pattern which represent the degree of relevance for each protoform in each cluster. For example, the values of three fuzzy clusters ($K = 3$ in the clustering algorithm) computed from a source stream computed for five protoform instances $P_{1...5}$ defined by the expert criteria, are illustrated in Table 1. The values of these fuzzy clusters correspond to the maximal aggregated degree of each protoform in a day. A bar chart of these results is shown in Figure 2).

4 Case study

In this section we analyze with real data how to describe linguistically the daily behaviour of an user monitored by the wearable tracker Fitbit. The architecture of the system and the process of identifying the patterns are defined as well.
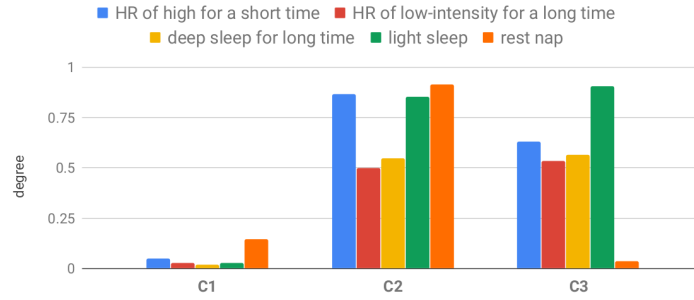
Figure 2: Visual representation of the three clusters obtained from five protoform instances that shows the daily behaviour pattern of an user.

4.1 Data acquisition and processing

This case study has been tested with 200 days data of real activity from an user that wear a Fitbit Ionic device all day. This device has generated a source stream with 700.526 samples with AS, sleep stages and heart rate from the user and sent all these collected activity records to the Cloud.

Thus, to obtain this information, we have developed a web application called *Monwatch* that allows to synchronize and view the information collected by the wearable (and sent by Fitbit Smartphone APP) from the Fitbit Cloud. Monwatch has been developed using the Django Framework version 2 and Python 3.6. All data is synchronized via HTTPS requests according to Fitbit API requirements and the returned data is stored locally in a MySQL Database. To do that Fitbit API requires the use of OAuth 2.0 authorization framework that enables a third-party application to obtain limited access to an HTTP service, either on behalf of a resource owner by orchestrating an approval interaction between the resource owner and the HTTP service, or by allowing the third-party application (our application) to obtain access. OAuth 2.0 is a authentication standard defined in RFC 6749 (D. Hardt, 2019).

4.2 System architecture

The complete scheme of the proposal from the smartwatch real-time monitoring to the gathering and processing of that data is shown in Figure 3. First, the smartwatch collects the raw data and sends it via Bluetooth to the Smartphone using the Fitbit APP which shows several activities in statistical way. Next, data are sent to the Fitbit cloud from where, through a API Service, we are able to extract raw data to Monwatch. So, Monwatch : i) processes data, ii) stores the records in

an internal database, iii) generates the instances of protoforms and their truth degree, iv) classifies the protoform instances according with the resulting clusters and v) exports and displays the results by linguistic expressions.

4.3 Protoform definition and their linguistic representation

A supervisor has defined the five HKIs that accurately describes the source streams of Fitbit using the protoform $V_r\ T_j\ Q_k$ described in Section 3.2 :

1. HR is high around 15 and 30 minutes with normal intensity

2. HR is low around 2 and 4 hours with normal intensity

3. AS is deep around 2 and 4 hours with moderate intensity

4. AS is (rem AND light) around 2 and 4 hours with moderate intensity

5. AS is (restless AND asleep) around 30 minutes and 2 hours with moderate intensity

These HKIs or protoform instances contains several terms, FTW and quantifiers which have been defined by fuzzy sets. The trapezoidal membership function for each fuzzy set defined in this proposal is described in Table 2 [1].

In addition to this, each protoform instance is represented by a set of linguistic expressions closer to natural language. These expressions that improve the semantic expressiveness of the protoform instances and shorten its length in most of the cases,

---

[1]TS, TR and TL are described in the appendix Abbreviations A
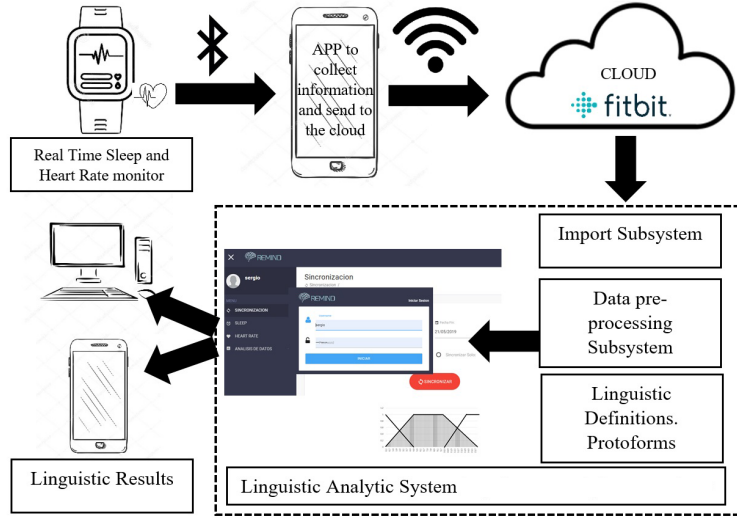
Figure 3: System architecture.

Table 2: Trapezoidal membership functions for terms, FTWs and quantifiers defined for the protoform $V_r\ T_j\ Q_k$

.

Textual description in natural language	Type	μ_T
hr is low	V_r	$TL(s_i^l)[60bpm, 70bpm]$
hr is high	V_r	$TR(s_i^l)[80bpm, 90bpm]$
around 15 and 30 minutes	T_j	$TL(\Delta t_i)[15m, 30m]$
around 2 and 4 hours	T_j	$TL(\Delta t_i)[240m, 480m]$
around 30 minutes and 2 hours	T_j	$TL(\Delta t_i)[30m, 240m]$
normal intensity	Q_k	$TR(x)[0.25, 0.75]$
moderate intensity	Q_k	$TR(x)[0, 0.5]$

Table 3: Short linguistic description of each HKIs.

Linguistic description	Protoform instance ($V_r\ T_j\ Q_k$)
HR of high-intensity for a short time	HR is high around 15 and 30 minutes with normal intensity
HR of low-intensity for a long time	HR is low around 2 and 4 hours with normal intensity
deep sleep for long time	AS is deep around 2 and 4 hours with moderate intensity
light sleep	AS is (rem AND light) around 2 and 4 hours with moderate intensity
rest nap	AS is (restless AND asleep) around 30 minutes and 2 hours with moderate intensity

are defined in Table 3. According to this, our system stores the data knowledge using a set of protoform instances in a first place and, in the second one, a linguistic summary, more suitable to the final user, is provided by the system (Marín and Sánchez, 2016).

4.4 Identification of patterns

The degree ofB truth of each protoform instance is aggregated for a day and results have been evaluated by the fuzzy C-means clustering algorithm to extract some pattern behaviours of HKIs. This algorithm has been executed with different number of clusters ($N = 4, 5, 6, 7$) to analyze these protoform instances behaviour. The clusters obtained are shown in Figure 4.

In Figure 4, we can observe the different patterns

of HKIs that can be found for an user in a day in each cluster. There are four scenarios according with the analyzed number of clusters. For example, in the first scenario of the figure A) (with $N = 4$) the first cluster represents the days where the user mainly developed a high-hr session and a long rest nap at once, this is characterized as a pattern and its corresponding linguistic description is: *Day with HR of high-intensity for a short time and a rest nap*. Following the same example, there are also another pattern that shows those days where the user has not a deep sleep for long time but the remaining HKIs has been activated.

Regarding the analysis of the number of clusters in this real example, a solution with $N = 7$ disintegrates the values too much, giving us too specific patterns with isolated HKIs. Solutions with $N = 5$ or $N = 6$ are more inclusive with all the HKIs and both give us a good solution.

Finally, in addition to the extraction and visualization of daily pattern behaviours from clusters, the trend of the membership degree for each day and clusters of a given time period provides a very rich and representative information for the supervisor of the activities. For example, in Figure 5, we show the membership degree between clusters and 30 consecutive days which provide a visual relation of the trend of behaviour.

5 Conclusions and ongoing works

In this work a combination of semi supervised analysis of user behaviour is proposed. For that, a wearable tracker provides the source streams about sleep stages duration and heart rate from users in a non invasive way. The aim of the methodology has been focused on integrating expert criteria by means of protoforms, which evaluate the source streams computing the degree of the protoform instances in full time line. The second contributions lies in extract behaviours patterns for each day using fuzzy clustering.

The results from the case study, which was developed for more than 200 days, shows a promising capability to aggregate data, extract patterns and provide a linguistic and visual representation due to interpretability of the protoforms.

In on going works, we will focus on evaluating a wide range of users, which could provide several behaviours patterns. The analysis between the clusters from several users will provide a suitable comparative between user profiles.

References

Yang Bai, Paul Hibbing, Constantine Mantis, and Gregory J. Welk. 2018. Comparative evaluation of heart rate-based monitors: Apple watch vs fitbit charge hr. *Journal of Sports Sciences*, 36(15):1734–1741. PMID: 29210326.

Z Beattie, Y Oyang, A Statan, A Ghoreyshi, A Pantelopoulos, A Russell, and C Heneghan. 2017. Estimation of sleep stages in a healthy adult population from optical plethysmography and accelerometer signals. *Physiological Measurement*, 38(11):1968–1979.

James C Bezdek, Robert Ehrlich, and William Full. 1984. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2-3):191–203.

Liming Chen, Chris D Nugent, and Hui Wang. 2011. A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24(6):961–974.

R. Choi, W. Kang, and C. Son. 2017. Explainable sleep quality evaluation model using machine learning approach. In *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 542–546.

Ed. D. Hardt. 2019. The oauth 2.0 authorization framework. RFC 6749.

Emiro De-La-Hoz-Franco, Paola Ariza-Colpas, Javier Medina Quero, and Macarena Espinilla. 2018. Sensor-based datasets for human activity recognition–a systematic review of literature. *IEEE Access*, 6:59192–59210.

Keith M Diaz, David J Krupka, Melinda J Chang, James Peacock, Yao Ma, Jeff Goldsmith, Joseph E Schwartz, and Karina W Davidson. 2015. Fitbit®:
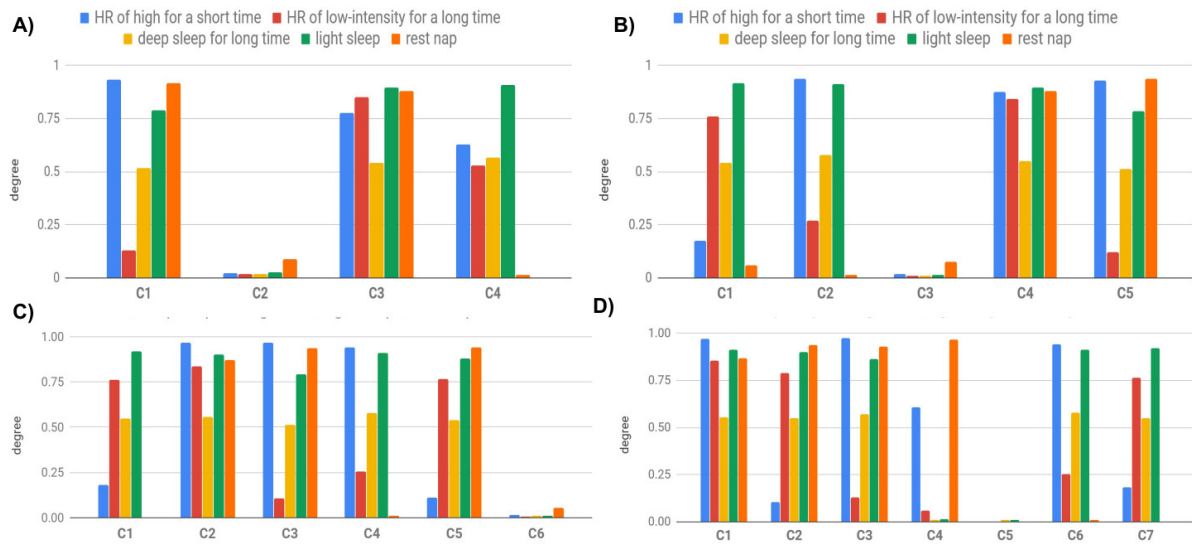
Figure 4: Clusters obtained by fuzzy C-means algorithm of the main HKIs for different values of N: A) $N = 4$, B) $N = 5$, C) $N = 6$ and D) $N = 7$
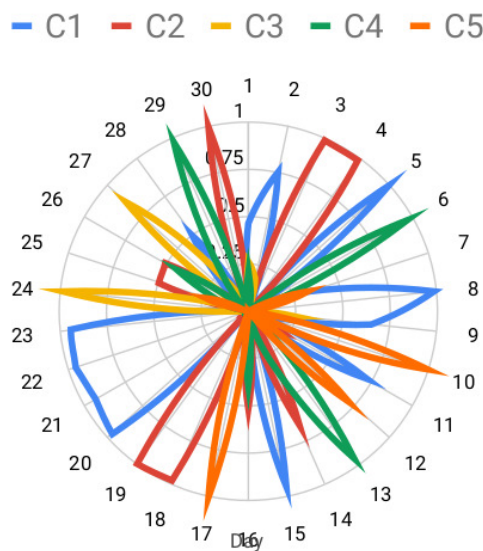

Figure 5: Evolution of the membership degree between clusters and 30 consecutive days for fuzzy C-means with $N = 5$.

An accurate and reliable device for wireless physical activity tracking. *International journal of cardiology*, 185:138–140.

David Dickinson, Joseph Cazier, and Thomas Cech. 2016. A practical validation study of a commercial accelerometer using good and poor sleepers. *Health Psychology Open*, 3(2):2055102916679012.

WM Dong and FS Wong. 1987. Fuzzy weighted averages and implementation of the extension principle. *Fuzzy sets and systems*, 21(2):183–199.

M. J. Duncan, C. Vandelanotte, S. G. Trost, A. L. Rebar, N. Rogers, N. W. Burton, and W. J. Brown. 2016. Balanced: a randomised trial examining the efficacy of two self-monitoring methods for an app-based multi-behaviour intervention to improve physical activity, sitting and sleep in adults. *BMC public health*, 16(670).

Macarena Espinilla, Javier Medina, Ángel-Luis García-Fernández, Sixto Campaña, and Jorge Londoño. 2017. Fuzzy intelligent system for patients with preeclampsia in wearable devices. *Mobile Information Systems*, 2017.

National Sleep Foundation. 2019. Which sleep tracker is best for you?

Liane D Heale, Saunya Dover, Y Ingrid Goh, Victoria A Maksymiuk, Greg D Wells, and Brian M Feldman. 2018. A wearable activity tracker intervention for promoting physical activity in adolescents with juvenile idiopathic arthritis: a pilot study. *Pediatric Rheumatology*, 16(1):66.

Janusz Kacprzyk and Ronald R Yager. 2001. Linguistic summaries of data using fuzzy logic. *International Journal of General System*, 30(2):133–154.

Seung-Gul Kang, Jae Myeong Kang, Kwang-Pil Ko, Seon-Cheol Park, Sara Mariani, and Jia Weng. 2017. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. *Journal of Psychosomatic Research*, 97:38 – 44.

Bhanu Prakash Kolla, Subir Mansukhani, and Meghna P. Mansukhani. 2016. Consumer sleep tracking devices: a review of mechanisms, validity and utility. *Expert Review of Medical Devices*, 13(5):497–506. PMID: 27043070.

Zilu Liang, Bernd Ploderer, Wanyu Liu, Yukiko Nagata, James Bailey, Lars Kulik, and Yuxuan Li. 2016.

Sleepexplorer: a visualization tool to make sense of correlations between personal sleep data and contextual factors. *Personal and Ubiquitous Computing*, 20(6):985–1000.

Stephen S Lim, Theo Vos, Abraham D Flaxman, Goodarz Danaei, Kenji Shibuya, Heather Adair-Rohani, Mohammad A AlMazroa, Markus Amann, H Ross Anderson, Kathryn G Andrews, et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the global burden of disease study 2010. *The lancet*, 380(9859):2224–2260.

Nicolás Marín and Daniel Sánchez. 2016. On generating linguistic descriptions of time series. *Fuzzy Sets Syst.*, 285(C):6–30.

Javier Medina-Quero, Macarena Espinilla, and Christopher Nugent. 2016. Real-time fuzzy linguistic analysis of anomalies from medical monitoring devices on data streams. In *Proceedings of the 10th EAI international conference on pervasive computing technologies for healthcare*, pages 300–303. ICST (Institute for Computer Sciences, Social-Informatics and .

Javier Medina Quero, Miguel Ángel López Medina, Alberto Salguero Hidalgo, and Macarena Espinilla. 2018. Predicting the urgency demand of copd patients from environmental sensors within smart cities with high-environmental sensitivity. *IEEE Access*, 6:25081–25089.

Kathryn Mercer, Lora Giangregorio, Eric Schneider, Parmit Chilana, Melissa Li, and Kelly Grindrod. 2016a. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. *JMIR mHealth and uHealth*, 4(1):e7.

Kathryn Mercer, Melissa Li, Lora Giangregorio, Catherine Burns, and Kelly Grindrod. 2016b. Behavior change techniques present in wearable activity trackers: a critical analysis. *JMIR mHealth and uHealth*, 4(2):e40.

Bryan Minor, Janardhan Rao Doppa, and Diane J Cook. 2015. Data-driven activity prediction: Algorithms, evaluation methodology, and applications. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 805–814. ACM.

Victoria Moreno-Cano, Fernando Terroso-Saenz, and Antonio F Skarmeta-Gomez. 2015. Big data for iot services in smart cities. In *2015 IEEE 2nd World Forum on Internet of Things (WF-IoT)*, pages 418–423. IEEE.

George Okeyo, Liming Chen, Hui Wang, and Roy Sterritt. 2014. Dynamic sensor data segmentation for real-time knowledge-driven activity recognition. *Pervasive and Mobile Computing*, 10:155–172.

María Dolores Peláez-Aguilera, Macarena Espinilla, María Rosa Fernández Olmo, and Javier Medina. 2019. Fuzzy linguistic protoforms to summarize heart rate streams of patients with ischemic heart disease. *Complexity*, 2019.

Alejandro Ramos-Soto, Alberto Jose Bugarin, Senén Barro, and Juan Taboada. 2014. Linguistic descriptions for automatic generation of textual short-term weather forecasts on real prediction data. *IEEE Transactions on Fuzzy Systems*, 23(1):44–57.

Patrick C Shih, Kyungsik Han, Erika Shehan Poole, Mary Beth Rosson, and John M Carroll. 2015. Use and adoption challenges of wearable activity trackers. *IConference 2015 Proceedings*.

Dongkuan Xu and Yingjie Tian. 2015. A comprehensive survey of clustering algorithms. *Annals of Data Science*, 2(2):165–193.

Yan Yan, Elisa Ricci, Gaowen Liu, and Nicu Sebe. 2015. Egocentric daily activity recognition via multitask clustering. *IEEE Transactions on Image Processing*, 24(10):2984–2995.

Lotfi A Zadeh. 2002. A prototype-centered approach to adding deduction capability to search engines-the concept of protoform. In *2002 Annual Meeting of the North American Fuzzy Information Processing Society Proceedings. NAFIPS-FLINT 2002 (Cat. No. 02TH8622)*, pages 523–525. IEEE.

Lotfi A. Zadeh. 2006. *Generalized Theory of Uncertainty (GTU) – Principal Concepts and Ideas*, pages 3–4. Springer Berlin Heidelberg, Berlin, Heidelberg.

A   Appendices

HR	Heart Rate
HKI	Health Key Indicator
FTW	Fuzzy Temporal Window
AS	Activity Status

FWA
$$V_r \cup T_k(s^j) = \frac{1}{\sum T_k(\Delta t_i^j)} \sum_{m_i^j \in s^j} V_r(v_i^j) \times T_k(\Delta t_i^j)$$

TS
$$TS(x)[l_1,l_2,l_3,l_4] = \begin{cases} 0 & x \leq 0 \\ (x-l_1)/(l_2-l_1) & l_1 \leq x \leq l_2 \\ 1 & l_2 \leq x \leq l_3 \\ (l_4-x)/(l_4-l_3) & l_3 \leq x \leq l_4 \\ 0 & l_4 \leq x \end{cases}$$

TR
$$TR(x)[l_1,l_2] = \begin{cases} 1 & x \leq l_1 \\ (l_2-x)/(l_2-l_1) & l_1 \leq x \leq l_2 \\ 0 & l_2 \leq x \end{cases}$$

TL
$$TL(x)[l_1,l_2] = \begin{cases} 0 & x \leq l_1 \\ (x-l_1)/(l_2-l_1) & l_1 \leq x \leq l_2 \\ 1 & l_2 \leq x \end{cases}$$