

Article

# Assessing Human Activity in Elderly People Using Non-Intrusive Load Monitoring

## José M. Alcalá, Jesús Ureña \*, Álvaro Hernández and David Gualda

Electronics Department, University of Alcalá, Escuela Politécnica, Ctra. Madrid-Barcelona, Km. 33,600, 28871 Alcalá de Henares, Spain; jmanuel.alcala@depeca.uah.es (J.M.A.); alvaro.hernandez@uah.es (A.H.); david.gualda@depeca.uah.es (D.G.)

\* Correspondence: jesus.urena@uah.es; Tel.: +34-918-856-592

Academic Editor: Leonhard M. Reindl Received: 30 November 2016; Accepted: 25 January 2017; Published: 11 February 2017

Abstract: The ageing of the population, and their increasing wish of living independently, are motivating the development of welfare and healthcare models. Existing approaches based on the direct heath-monitoring using body sensor networks (BSN) are precise and accurate. Nonetheless, their intrusiveness causes non-acceptance. New approaches seek the indirect monitoring through monitoring activities of daily living (ADLs), which proves to be a suitable solution. ADL monitoring systems use many heterogeneous sensors, are less intrusive, and are less expensive than BSN, however, the deployment and maintenance of wireless sensor networks (WSN) prevent them from a widespread acceptance. In this work, a novel technique to monitor the human activity, based on non-intrusive load monitoring (NILM), is presented. The proposal uses only smart meter data, which leads to minimum intrusiveness and a potential massive deployment at minimal cost. This could be the key to develop sustainable healthcare models for smart homes, capable of complying with the elderly people' demands. This study also uses the Dempster-Shafer theory to provide a daily score of normality with regard to the regular behavior. This approach has been evaluated using real datasets and, additionally, a benchmarking against a Gaussian mixture model approach is presented.

**Keywords:** non-intrusive load monitoring; activity monitoring; ambient intelligence (AmI); activity recognition (AR); ambient assisted living (AAL)

## 1. Introduction

Life expectancy is becoming higher and higher every year in most developed countries. Although this is clearly a benefit, there are also potential challenges to be faced. For instance, the old-age dependency ratio, the ratio between elderly people and those of working age is rising and expected to double within the next 35 years (from 11.7% to 25.4% of the population). This ratio will exceed 50% in countries such as Japan, Germany, Italy, Spain, and Poland; and the elderly aged 80 years or over will triple in European Union and United States. In fact, this is a challenge on the sustainability of healthcare models that need to care for this community. Furthermore, several studies reveal that the majority of the elderly, at 90%, try to live independently in their own homes, and a shortage of caregivers is expected. Therefore, these two factors, ageing of the population and need of monitoring within the households, are the key issues that have motivated many studies in ambient assisted living (AAL) domain [1].

AAL is based on ambient intelligence (AmI) and it seeks to care for, and monitor, people by adapting their environment and using new technologies. Within this scope, the monitoring of the elderly can be classified into direct and indirect methods and a review can be found in [2,3]. Regarding direct monitoring methods, wearable sensors or body sensor networks (BSN) can be very accurate by registering biomechanical and physiological data, such as blood pressure and heart rate [4].



Due to this accuracy, these sensors are suitable for high-risk subjects that need to be monitored all of the time. For instance, in [5] the author devises flexible graphene electrodes to fully monitor the electrocardiography (ECG) within this community throughout the day. Nevertheless, these sensors are very intrusive, more expensive and, typically, this high level of accuracy is not required for the majority of the population. Consequently, new approaches to indirectly monitor the elderly have been developed, based on the monitoring of activities of daily living (ADLs) by using environmental sensors.

ADL is a term used in healthcare to refer as to the daily activities and routines carried out by people for self-caring. This concept was first proposed by Katz, who stated that age-related diseases have a direct impact on ADLs, and he created ADL indices to measure the dependence level of a person [6]. Furthermore, in [7] instrumental activities of daily living (IADLs) are proposed as more complex activities than ADLs. The performance of these activities is required to determine a successful and independent life.

Concerning the indirect methods and the involved sensors, a good review of previous works is presented in [1,2,8]. Typically, heterogeneous sensor networks are deployed avoiding CCTV cameras and other invasive systems that could threaten the privacy of householders. Hence, there is a mixture of sensors, such as pressure sensors [9], infrared sensors, sound sensors, magnetic switches, optical and ultrasonic sensors, and inertial sensors from smartphones; all of them aiming to capture the human activity within a household. For instance, in [10] inertial sensors are placed in shoes to measure the gait and to foresee the performed ADLs. The procedure requires an activity recognition (AR) algorithm to group event sequences by activities [11]. Then, these activities are assigned to the major ADL activities and they are evaluated to get the ADL indices, also known as circadian rhythms. The composite of major activities based on minor activities, as well as their tracking, have been studied in [12,13]. In [14] a "tape and forget" method to deploy sensors inside the household is presented. In [15] event sequences from heterogeneous sensors are grouped by using support vector machine (SVM). Furthermore, activities can be overlapped over time, which makes the AR clusterization process more difficult. In [16] probabilistic models, such as naïve Bayesian (NB), hidden Markov model (HMM), and hidden semi-Markov model (HSMM), are evaluated to avoid overlapping; whereas in [17] a layered hidden Markov model is used to model the different durations of activities. Further, sensor fusion techniques, such as the Dempster-Shafer theory, are used for AR to determine whether event sequences are overlapped [18,19]. Smartphones are also very compelling to be applied as sensors and in tracking activities [20]. Moreover, activity recognition of ADLs is not sufficient to develop a healthcare monitoring system, it is also necessary to know how activities change over time, which is also a complex task, known as behavior change detection (BCD). In [21] activities are extracted from a heterogeneous sensor network and featured by duration of the activity, sensor density and amount of movement. Then, using two time windows, changes on those features are obtained by means of a permutation-based change detection in activity routine (PCAR) approach and a virtual classifier (VC) approach. Likewise, in [22] daily activity curves are obtained by averaging duration time of those activities over a time window (e.g., a month). Then, using PCAR, a score of the normal daily routine is given.

Overall, AR based on heterogeneous sensor networks presents three major issues that lead to a low implementation in real scenarios [11]: overlapped activities, heterogeneous duration, and the deployment of a complex WSNs, which, in most cases, is intrusive. To increase the acceptance, sensor minimization, low maintenance, and simplicity is pursued. In [23,24] a new approach of monitoring ADLs through electrical signatures of appliances captured by plug-meters is proposed. This monitoring method has been well accepted by subjects, as well as by professionals, during experiments since it is less intrusive. Additionally, daily activities are strongly connected to the usage pattern of appliances and, consequently, they are a good way to infer the human activity. Electrical events are mapped over daily activities, and ADL indices and circadian rhythms are obtained and presented as a decision support tool for experts. The appliance usage patterns are again proved to be correlated with ADLs [25], where a method known as latent dirichlet allocation (LDA) for text analysis is used to map appliance events with ADLs.

In alignment with this, the widespread installation of smart meters is well underway and this has also boosted the development of a new machine learning paradigm: non-intrusive load monitoring (NILM). This is the process of disaggregating the total electrical consumption from the smart meters into individual appliances [26,27]. The smart meter data are distributed through the home area network (HAN) at a low sampling rate (typically about 0.1 samples per second). Thus, probabilistic temporal models have been developed to predict the appliances states (e.g., on or off), such as factorial hidden Markov models (FHMM) [28] or deep learning [29]. These are called eventless NILM algorithms and achieving high performance is still a challenge. Moreover, at higher frequency rates (i.e., higher than 1 Hz), there are event-based NILM algorithms that achieve better performance. These are compounded by an event detector [30–32] and a load identification algorithm [33,34] (i.e., around 90% accuracy).

NILM applied to ADLs monitoring is very compelling since it is a completely non-intrusive system, which is a desired feature for large deployments. Additionally, as was mentioned before, ADLs are strongly related to appliance usage patterns and these can be obtained via NILM. To the best of our knowledge, the first use of disaggregation of smart meter data for health monitoring purposes was carried out in [35]. They performed an iterative time-dependent hidden Markov model to disaggregate appliances based on a priori knowledge of inhabitant's activities. Then, the disaggregation was carried out and every appliance was assigned to a certain activity that could be monitored. In [36] a log Gaussian Cox process is used to monitor the human activity based on the usage pattern for a single appliance. This approach models daily and weekly patterns, as well as uncertainty. The monitored appliance has to be relevant in the daily routine of the householder and has to be daily and manually operated. In this case, a kettle was chosen in an UK dataset. In previous works, the monitoring of the kettle or the TV set has been considered as a relevant variable to detect changes in routines [37] whereas, in [38], the usage pattern of the kettle and fridge during nighttime is used to detect sleep disorders, characteristic of dementia. A multi-appliance monitoring method is proposed in [39], based on a Bayesian approach, which uses Gaussian mixture models for daily and weekly usage patterns. This proposal merges the pattern analysis of all appliances and provides a single score about the normality of the behavior every day.

This work presents a novel algorithm to monitor the human activity through NILM and to evaluate its performance over time. NILM has been already studied widely and it has different resources, such as datasets [40], released open software [41,42] (i.e., in Python and MATLAB, respectively), and a large set of third-party commercial devices [43]. Hence, the focus here is on the activity monitoring based on already disaggregated data. Instead of directly tracking ADLs by mapping event sequences into main activities, as in previous works [36,39], the indices of usage patterns are used to evaluate the performance and to detect whether the subject has deviated from their routine. Thus, the common problems of tracking activities are avoided: overlapped activities, uncertainty and heterogeneity of activity durations, composite of activities, etc. The aim is to simplify the process and directly use the appliance usage patterns as a proxy for health and welfare, detecting anomalies and changes in the routine patterns, which might be a symptom of potential elderly disorders, and reporting them to relatives and caregivers. It is not intended to achieve the accuracy of BSN or WSN approaches, and it is assumed that false alarms may occur with the proposed method. Nevertheless, the cost of a false alarm is low (i.e., it might result in a phone call to check in the house) compared with the benefit of a completely non-intrusive system that does not need maintenance and can be massively deployed. In contrast to [23,24,36,39], the routine evaluation is completely based on NILM, the usage pattern is extracted from multiple appliances and uncertainty is modeled using the Dempster-Shafer Theory, which is commonly used in fusion sensor data. Furthermore, a thorough analysis using real datasets and the Gaussian Mixture model in [39] as a benchmark is carried out to show the benefits of the proposed approach.

The rest of the manuscript is organized as follows: the model description for the algorithm is presented in Section 2; some results and the comparison between models are shown in Section 3; and, finally, conclusions and future work are discussed in Section 4.

#### 2. Model Description

The Dempster-Shafer theory (DST), also known as the evidence theory, is widely used in sensor data fusion. It is a generalization of the Bayesian theory [44], where, instead of probability distribution functions, belief functions or mass functions are handled. The disaggregated consumption data from each appliance can be considered as a reading from an independent sensor, which provides a belief about the normality in the usage pattern for that specific appliance by using a mass function. By means of evidence theory, all beliefs are merged, thus obtaining a general belief about the normality in the use of all appliances. This allows to score daily routines and to detect deviations.

The hypotheses  $2^{\Omega}$  in the DST are increased, compared to the hypotheses  $\Omega$  in the Bayesian theory. For instance, for a presence detector, the hypotheses in the Bayesian domain is the universe  $\Omega = \{h_1(Presence), h_2(Non \ presence)\}$ . Unlike, in the DST domain the hypotheses are the power set  $2^{\Omega} = \{\emptyset, h_1(Presence), h_2(Non \ presence), H_x\}$ , where  $\emptyset$  is the empty set and  $H_x$  is a subset for all the potential combinations, in this case  $H_x = \{h_1(Presence) \cup h_2(Non \ presence)\}$ . Thus, the parameter x models the uncertainty that is hardly represented in the Bayesian Theory by assigning a probability of 0.5.

Applying DST to our case, the power set is  $2^{\Omega} = \{\emptyset, h_1(normal \ pattern), h_2(abnormal \ pattern), h_3 = \{h_1 \cup h_2\}(normal \ or \ abnormal \ pattern)$ . Then, a belief mass function for every element in the power set has to be defined according to Equation (1). A belief mass function is formally called basic belief assignment (BBA), if it meets Equations (2) and (3). The BBA for a certain hypothesis  $h_1$  represents the proportion of belief for the current state to be such.

$$m: 2^{\Omega} \to [0, 1] \tag{1}$$

$$m(\emptyset) = 0 \tag{2}$$

$$\sum_{A \subseteq 2^{\Omega}} m(A) = 1 \tag{3}$$

Belief (or support) and plausibility define the minimum and maximum of the confident interval for the hypothesis *A*, respectively. Hence, the probability for the current state to be *A* is given by an uncertainty interval described in (4), where belief *bel* and plausibility *pl* are defined in Equations (5) and (6). The belief in *A* is evaluated by adding the belief masses from all the subsets of *A*, whereas the plausibility is the sum of all belief masses of subsets where there is an intersection with the set *A*. For instance, the belief of a pattern to be normal or abnormal is  $bel(H_x = h_1 \cup h_2) = m(h_1 \cup h_2) + m(h_1) + m(h_2)$ .

$$bel(a) \le P(A) \le pl(a)$$
 (4)

$$bel(A) = \sum_{B|B \subseteq A} m(B)$$
(5)

$$pl(A) = \sum_{B|B \cap A \neq \emptyset} m(B)$$
(6)

Let us consider two appliances whose hypotheses are to be merged. For that, the Dempster's rules of combination are applied as in Equations (7)–(9) for the hypothesis *A*.

$$m_{1,2}(\emptyset) = 0 \tag{7}$$

$$m_{1,2}(A) = (m_1 \oplus m_2)(A) = \frac{1}{1-K} \sum_{B \cap C = A \neq \emptyset} m_1(B)m_2(C)$$
(8)

$$K = \sum_{B \cap C = \varnothing} m_1(B)m_2(B)$$
(9)

5 of 17

where Equation (7) denotes that the mass function assigned to the empty set must be zero, and Equation (8) defines the new mass function for the hypothesis *A* resulting from merging appliances 1 and 2, where 1 - K is a normalization factor. Finally, Equation (9) describes the amount of conflict between the two mass sets.

For instance, considering two appliances X and Y, a certain observation window  $T_i$ , and the power set  $2^{\Omega} = \{\emptyset, h_1(normal pattern), h_2(abnormal pattern), h_3 = \{h_1 \cup h_2\}(normal or abnormal pattern), their BBAs are defined in Table 1. The Demptster's rule of combination, Equations (7)–(9), is applied and the resulting fusion of BBAs is depicted in Table 2. The combination of <math>h_1 \cap h_2$  and  $h_2 \cap h_1$  are conflicts because the current state cannot be normal and abnormal at the same time. Consequently, they are assigned to the empty set. The proportion of belief corresponding to the conflict states is assigned to *K* as in Equation (9). Lately, the parameter *K* is distributed to the remaining compatible hypotheses according to Equation (8). Table 3 is obtained by adding equal hypotheses in Table 2, normalizing by *K* and rearranging. The belief and plausibility are evaluated in Equations (5) and (6) from the mass functions of the fusion. Therefore, Table 3 shows that, for the beliefs of appliances X and Y in Table 1, the probability of a pattern to be normal can be found in the interval  $0.89 \le P(h_1) \le 0.92$ , defined by Equation (4) within the window  $T_i$ .

Table 1. Basic belief assignments for appliances X and Y (example).

<b>BBAs</b> for X in $T_i$	2 <sup>Ω</sup>	<b>BBAs</b> for Y in $T_i$
$m_x(h_1, T_i) = 0.8$	$h_1$	$m_y(h_1, T_i) = 0.6$
$m_x(h_2, T_i) = 0.1$	$h_2$	$m_y(h_2, T_i) = 0.2$
$m_x(h_3,T_i)=0.1$	$h_1 \cup h_2$	$m_y(h_3, T_i) = 0.2$

Table 2. BBA fusion for appliances X and Y.

$\cap$	$X_{h1}$	$X_{h2}$	$X_{h3}$
$Y_{h1}$	$h_1 \cap h_1 = h_1$	Ø	$h_3 \cap h_1 = h_1$
$Y_{h2}$	Ø	$h_2 \cap h_2 = h_2$	$h_3 \cap h_2 = h_2$
Y <sub>h3</sub>	$h_1 \cap h_3 = h_1$	$h_2 \cup h_3$	$h_3 \cap h_3 = h_3$

Table 3. BBA fusion for appliances X and Y.

2 <sup>Ω</sup>	$m_{T_i}$	$bel_{T_i}$	$pl_{T_i}$
$h_1$	0.89	0.89	0.92
$h_2$	0.08	0.08	0.1
$h_3$	0.02	1	1

The example above can be applied to more appliances by accumulating their evidence iteratively. Therefore, considering N appliances, the masses of evidence from appliances 1 and 2 are fused to obtain a new mass:  $m_{T_i}$ , as shown in Table 3. Then, repeating the process, this new mass is merged with the mass from appliance 3, and so on up to the appliance N. This iterative process results in a general belief and plausibility based on the accumulative evidential from the N appliances in a certain time interval  $T_i$ . This assures that, over that time, the probability of a pattern to be *normal* is higher than the belief and lower than the plausibility.

Likewise, if we are to evaluate a daily pattern, we should proceed by accumulating evidences over the time intervals. Thus, if a daily time *T* is divided into *I* periods of time  $T_i$  in Equation (10), a general mass function  $m_{T_i}$  for each interval *i* can be obtained as explained above. These are accumulated following the same process over time, instead of appliances. This results in the belief and plausibility for the performed daily routine within the household.

$$\boldsymbol{T} = [T_1, \dots, T_i, \dots, T_I] \tag{10}$$

## Basic Belief Assignments and Weighing

Regarding the basic belief functions (BBAs), they are modeled as follows. For every appliance, a BBA is obtained taking into account the day of the week and the time interval  $T_i$  of the day. It has been empirically proved that the appliance usage patterns not only vary during hours of the day, but it also does in a significant way depending on the day of the week. Likewise, during the training process, the occurrences for each appliance are binned by the time interval  $T_i$  and the day of the week. The number of occurrences in each bin for a certain day of the week is divided into the total number of that specific day of the week in the training set. This denotes the probability of the appliance to be used at that time of the day for that day of the week. Following this process, similar BBAs to those in Figure 1 are obtained, where the time interval  $T_i$  was fixed to three hours. Note, for instance, that here the probability  $P(T_i)$ , where  $T_i \in [9, 12)$  is the time interval of the day when using the kettle between 9 h and 12 h, is roughly 0.85. Additionally, there is another time interval with high probability of using the kettle between 18 h and 21 h (0.6 approximately). During the test, BBAs are based on this probability, so they are weighed depending on the presence or absence of events within the considered interval  $T_i$  as in Equations (11)–(13). Thus, the constant  $C_0$  represents a certainty on the probability  $P(T_i)$  for an arrival event, whereas the constant  $C_1$  shows the certainty in absence of event. The value of these two constants ranges from 0 to 1, and they depend on the application. More details about the certainty constants is discussed in Section 3. In order to meet the requirements in Equations (1)–(3), the mass  $m_{T_i}(h_3)$  should be defined as in Equation (12) to hold the uncertainty produced by  $C_0$  and  $C_1$  in Equations (10) and (11).

$$\boldsymbol{m}_{\boldsymbol{T}_i}(h_1) = \begin{cases} P(T_i) \times C_0, \text{ if event} \\ (1 - P(T_i) \times C_1, \text{ if not event} \end{cases} \nabla T_i \in [0, 24)$$
(11)

$$\boldsymbol{m}_{\boldsymbol{T}_{i}}(h_{2}) = \begin{cases} 1 - P(T_{i}) \times C_{0}, \text{ if event} \\ P(T_{i}) \times C_{1}, \text{ if not event} \end{cases} \nabla T_{i} \in [0, 24)$$
(12)

$$m_{T_i}(h_3) = 1 - (m_{T_i}(h_1) - m_{T_i}(h_2))$$
(13)

The generation of mass functions and basic belief assignments functions is subjective and it does not have to follow this method. For the purpose of this study, the usage pattern of appliances is well represented by the probability of using those appliances within an interval time  $T_i$  of the day and, therefore, they model the BBAs. Furthermore, an uncertainty in the model is also desired as the usage pattern could be an indicator of somebody's healthy routine, but its absence does not necessarily imply the opposite. Thus, the constants  $C_0$  and  $C_1$  encode that uncertainty.



**Figure 1.** Kettle's BBA on Mondays with a time interval  $T_i$  of 3 h.

#### 3. Experimental Results

## 3.1. Datasets, Preprocessing, and Selection of the Training Samples

In order to evaluate the performance of the proposed algorithm, two different datasets have been considered: the Household Electricity Survey dataset [45] and the UK Domestic Appliance-Level Electricity (UK-DALE) dataset [46]. Both are real collected data from the aggregated and disaggregated energy consumption from UK households. The former contains a year data from three single pensioner households, which are the targeted community in this study; whereas the latter one is a two-year collection data from a family household (two adults, two children and a dog). The purpose of presenting the performance over the latter dataset (i.e., a family house) is to enhance the outperformance of the proposed algorithm over the first community (the elderly). As mentioned in Section 1, elderly disorders may imply deterioration in performing daily tasks. Therefore, the performance of elderlies is contrasted against a family house that has a known, strict, and regular routine.

The approach presented here is based on the employment of a NILM algorithm and, therefore, only the disaggregated data from datasets are used. Presenting a new NILM algorithm for disaggregation is not the topic of this study and it has been fully studied in previous works already mentioned in Section 1. In this way, the relevant events are extracted from the appliance-level consumption. Relevant events are those whose evaluation can result in the inferring of any human activity inside the household. Consequently, only manually activated appliances with a repetitive pattern over time are considered here, discarding appliances such as the fridge with a continuous and automatic consumption. The relevant events are the recorded timestamp of switching-on for these appliances. Although, the time duration of appliance usages is also a good indication about the routines, this implies a more complex pattern to infer the real human activity, as they could be overlapped. Furthermore, the duration does not really imply any human activity as the appliance could have been left on. Nevertheless, switching-on events do require human activity. For instance, Figure 2 shows the switching-on events in household No. 101017 from the HES dataset over time. Each spot is an event and the *y*-axis represents the appliance code, where non-relevant appliances have been already filtered out. The first 60% of total days are used for training, whereas the remaining 40% is for testing. It is worth noting that some deviation from the routine over time is expected. This is why the model should be trained with earlier days than those used for test, in order to learn the starting normal routine. As can be observed in Figure 2, if the ratio between training and test samples is less than 60%, then some appliances do not have enough samples for training.



Figure 2. Training and test event samples for household No. 101017 in HES dataset.

Regarding the configuration of the DST algorithm, the observation window  $T_i$  and the certainty constants  $C_0$  and  $C_1$  have been empirically fixed. A six-hour interval has been used for the observation window  $T_i$ , so every six hours the presence of each manual appliance (switching-on event) is evaluated and the mass scores in Equations (11)–(13) are obtained. Likewise, the certainty constants are set as follows:  $C_0 = 0.9$  and  $C_1 = 0.1$ . This means that there is a 10% uncertainty ( $(1 - C_0) \times 100$ ) in case that a switching-on event of a certain appliance is presented; and a 90% of uncertainty otherwise. Since the *normality* of a pattern is evaluated by switching-on events, it cannot be stated that, if an appliance is not used during a certain time interval, there is an *abnormal* pattern. Note that the human behavior is non-deterministic and, therefore, a high uncertainty is assigned in absence of events. Indeed, there is lack of information, which the DST algorithm models as uncertainty, whereas a Bayesian approach can seldom model it by using only probabilities. The inverse logic can be also applied: the occurrence of an event does not necessarily mean that the pattern is *normal*, but there is more information and that leads to assign only a 10% of uncertainty. These values are empirically obtained and could be changed regarding the confidence on the usage pattern for appliances.

#### 3.3. The Benchmark's Model

In [39] a Bayesian approach is presented to model the usage patterns of appliances by means of Gaussian mixture models and the union probability to score the routine. This model has been improved by dividing the training data into the days of the week, instead of dividing it into working days and weekends. As was proved in [36], some usage patterns have also a weekly repetition, so, for instance, Mondays can be different from Wednesdays. Thus, the accuracy of the GMM model has been improved in comparison with previous results in [39]. This model is used as benchmark in this study.

#### 3.4. Analysis of DST and GMM Scores

Figures 3–5 depict the analysis of the human activity for the three single household pensioners over the test data (145 days roughly in each case) and Figure 6 presents the results for the same house of Figure 5 zoomed to a particular week. Similarly, Figure 7 shows the family household activity during the test (315 days), where it is possible to observe the sensitivity of the proposed algorithm to pattern deviations. Apart from some exceptions, the family house keeps a strict routine and, consequently, the proposed algorithm generates fewer alarms, as shown in the plots in Figures 3a, 4a, 5a, and 6a.

Figures 7a and 8a are similar to the one in Figure 2, where the *x*-axis represents time and the *y*-axis means the different manual appliances, whereas the switching-on events of appliances are drawn as red dots. Due to the fact that, the higher the number of appliances is, the more difficult the labelling of the *y*-axis becomes. Table 4 shows the correspondence between the *y*-axis values and the appliance labels for clarity's sake. A long-term observation of the score (several months) can help to detect deteriorations in someone's activity performance. As described in Section 1, these might be a symptom about the apparition of degenerative diseases as dementia, whose early detection could be very beneficial. On the other hand, short-term deviations (several hours or days) are more difficult to interpret, although they should be also watched as they might denote an emergency. Following, some potential cases of deterioration detected by the DST algorithm are presented. Its performance is evaluated in comparison with the GMM algorithm.



**Figure 3.** Score for test events in household No. 101017 from HES dataset: (**a**) Test events; (**b**) DST score every 6 h; (**c**) DST score by day; (**d**) union probability score by day; and (**e**) union probability score by week.



**Figure 4.** Score for test events in household No. 103034 from HES dataset: (**a**) test events; (**b**) DST score every 6 h; (**c**) DST score by day; (**d**) union probability score by day; and (**e**) union probability score by week.



**Figure 5.** Score for test events in household No. 102003 from HES dataset: (**a**) test events; (**b**) DST score every 6 h; (**c**) DST score by day; (**d**) union probability score by day; (**e**) union probability score by week.



**Figure 6.** Score for test events in household No. 102003 from HES dataset during a week: (**a**) test events; (**b**) DST score every 6 h; (**c**) DST score by day; (**d**) union probability score by day; and (**e**) union probability score by week.



**Figure 7.** Score for test events in household no. 1 from UKDALE dataset: (**a**) test events; (**b**) DST score every 6 h; (**c**) DST score by day: (**d**) union probability score by day; and (**e**) union probability score by week.



**Figure 8.** Score for test events in household No. 1 from UKDALE dataset: (**a**) test events; (**b**) DST score every 6 h; (**c**) DST score by day; (**d**) union probability score by day; and (**e**) union probability score by week.

	Appliance Label				
Appliance ID	Household no. 101017	Household no. 103034	Household no. 102003	Family House	
1	Cooker	Kettle	Cooker	Boiler	
2	Kettle	Microwave	Kettle	Solar thermal station	
3	Microwave	Toaster	Microwave	Washer drier	
4	Toaster	Cooker extractor fan	Toaster	Dish washer	
5		Washing machine	Cooker extractor fan	Television	
6	Tumble dryer	-	Washing machine	Light 1	
7	Washing Machine	DVD player	TV-LCD	HTPC	
8	Set Top Box	TV-CRT	TV-LCD 2	Kettle	
9	TV-LCD	VCR	Computer desktop	Toaster	
10	Laptop	Lamp 1	Computer monitor	Fridge freezer	
11	Lamp 1	Lamp 2	Printer inkjet	Microwave	
12	Lamp 2	Lamp 3	Lamp 1	Computer monitor	
13	-	Lamp 4	Lamp 2	Breadmaker	
14	Vacuum cleaner	-	Lamp 3	Audio amplifier	
15	TV-LCD	Vacuum cleaner	Lamp 4	Light 2	
16			Lamp 5	Soldering iron	
17			Ĩ	Ethernet switch	
18			Vaccuum cleaner	Vaccum cleaner	
19				Light 3	
20				Light 4	
21				Light 5	
22				Light 6	
23				Active subwoofer	
24				Light 7	
25				Radio	
26				Light 8	
27				Phone charger	
28				Light 9	
29				Phone charger 2	
30				Light 10	
31				Coffee maker	
32				Radio 2	
33				Phone charger 3	
34				Hair dryer	
35				Hair straighteners	
36				Clothes iron	
37				Oven	
38				Light 11	
39				Baby monitor	
40				Light 12	
41				Light 13	
42				Computer desktop	
43				Fan	
44				Printer	

Table 4. Appliance labelling in the *y*-axis by house.

3.4.1. Single Pensioner Household No. 101017 in the HES Dataset

Figure 3a shows the test event samples of manually operated appliances that show a repetitive pattern: cooker, kettle, microwave, toaster, lamps, and others. Figure 3b represents the score provided by the DST algorithm with a 6-h observation window  $T_i$ , whereas Figure 3c depicts the date score by accumulating evidences over time as was explained in Section 2. In both plots, there is a green filled area limited by the belief and plausibility curves. This is the uncertainty area and it becomes larger as the number of events decreases. The uncertainty is clear in Figure 3b, but the accumulative evidence over the day makes the uncertainty area thinner in Figure 3c, thus obtaining a more accurate score. Figure 3d,e show the results after applying the improved GMM algorithm to score a day, and the average score over the week, respectively. The red line in Figure 3b–e are the empirical thresholds for the DST and the GMM algorithms, where any score bellow those thresholds is considered an anomaly in the behavior. For this specific household, a deviation from the regular routine over months is described with a lower score in both algorithms. These types of long-term deviations are better

observed in the DST day score and in the GMM week score. Nevertheless, the former seems to be more sensitive to this kind of pattern, thus rapidly dropping to zero what allows an earlier detection of pattern deviations than the one achieved by the GMM algorithm. The DST algorithm is also able to detect three isolated days of *abnormal* activity. Furthermore, it is worth noting that the oscillation in the daily score for the GMM (see Figure 3c) is high and it makes it difficult to configure the threshold to avoid false alarms. From a manual inspection of data, it has been verified that this deviation is due to those appliances related to cooking, such as the kettle, microwave, and toaster, whose occurrences decrease during that period. Consequently, the ADL of feeding might be deteriorated.

## 3.4.2. Single Pensioner Household No. 103034 in the HES Dataset

Similar deviations can be found in Figure 4c where the deterioration is due to those appliances related to a range of activities: some days caused by cooking appliances as seen before, others due to the use of the laptop, TV set, or the lighting system. This implies more noise for the GMM approach (see Figure 4d) and the decrease of the mean per weeks becomes very slight (see Figure 4e). Again, the threshold in the daily score for the GMM (Figure 4d) is difficult to be set a priori. Furthermore, looking into month intervals such as the one from 29 January to 19 February of 2011 in Figure 3, short-term anomalies can be detected in the DST algorithm, whereas the GMM algorithm does not show these deviations.

#### 3.4.3. Single Pensioner Household No. 102003 in the HES Dataset

Additionally, there is a case in Figure 5c where a single pensioner has recovered his routine after a deviation. The pattern routine carried by the person involves days of inactivity, which cannot be properly modeled by the Bayesian theory resulting in oscillations from *normality* to an anomalous behavior score provided in Figure 5. This can be closely observed in Figure 6, a zooming zone of a week from Figure 5. Starting on Monday 4 April 2011, this inhabitant spends most of their time out during Fridays, Saturdays, and Sundays. This has been previously learned by the DST algorithm from the training data and it is modeled as an increasing of the uncertainty in Figure 6b, which decreases in Figure 6c due to the accumulated evidence. For instance, on 9 April 2011 in Figure 6b, the uncertainty is such that most time the threshold lies into the area between the belief and the plausibility and, hence, one cannot state whether the pattern is *normal* or *abnormal*. Only from 12 h to 18 h it is feasible to conclude that the period is *normal*, since the threshold lies below both belief and plausibility. The GMM algorithm, which is a classical Bayesian network, cannot handle uncertainties, so it assigns a probability that decreases with the lack of activity. This leads to a higher number of false alarms.

#### 3.4.4. Family House in UKDALE Dataset

Finally, a case of a family with two adults, two children, and a dog is analyzed in Figure 7. Firstly, it is worth noting the increasing number of events coming from different (manually operated) appliances. Furthermore, the scores provided by Figure 7c–e reflect a very strict routine. This routine does not deteriorate over time and only shows *abnormal* intervals that may correspond with holidays or absences in the house. However, these deviations remain undetected by the GMM algorithm. Regarding the threshold, this can be easily set for the DST algorithm, whereas it becomes more difficult for the GMM one. Anyway, it should be close to 1 in order to detect some anomalies.

Zooming into a week period as in Figure 8, it is shown that the DST algorithm is sensitive to the anomalies, whereas the GMM cannot detect them. Since the score in GMM is a union probability, a single appliance which keeps a strict routine can saturate the score and, consequently, masks other anomalies in the remaining appliances. This does not happen in DST where the accumulative evidence of the hypotheses with *abnormal* pattern is higher than the evidence of *normal* patterns. Therefore, it properly models these effects. Furthermore, it is worth noting that the uncertainty is smaller in Figure 6b than its corresponding in Figure 3b, Figure 4b, and Figure 5b due to the increasing number of appliances per 6-h observation window.

Therefore, experimental results show that the DST algorithm behaves better than the GMM algorithm in detecting not only short-term but also long-term deviations. Additionally, the lack of activity is muffled by the uncertainty as in Figure 6, whereas the high activity does not mask other anomalies in the pattern (see Figure 8).

## 4. Conclusions

This paper presents a novel approach for activity monitoring of the elderly to detect deviations from their daily routines. Many health problems and welfare issues are directly related to these deviations, thus, it is a powerful tool for experts, caregivers and relatives. Thanks to the use of NILM, the proposed approach is non-invasive and, therefore, it enables the possibility for larger deployments, which have always been an issue in ADL monitoring systems. The obtained experimental results are promising and they point at the Dempster-Shafer Theory as a very suitable method to score the non-deterministic human behavior thanks to the encoding of uncertainty. Further, modelled uncertainty in the DST algorithm is suitable to be used along with low-performance NILM algorithms (i.e., as those discussed in Section 1) because, in the case of missed events, the uncertainty increases, but the normality score does not decrease. Thereby, its performance has been evaluated with real data and benchmarked with a Bayesian approach, concluding that the DST algorithm outperforms in the detection of short-term as well as long-term routine pattern deviations. The proposed DST approach shows to be more sensitive to pattern deviations and less susceptible to false alarms due to long periods of inactivity. Furthermore, the case of a family house, where a strong routine, is followed is presented to enhance the applicability in scenarios where there are deteriorations and deviations, mostly caused by elderly disorders but not exclusively by them.

Comparing with the previously mentioned BCD algorithms, the approach presented here has the novelty and the advantage of avoiding the use of many sensors and, consequently, the use of AR approaches, whose weaknesses have been pointed out in Section 1, thus the potential scalability is higher. Nevertheless, this also brings drawbacks: there are less information than in other BCD approaches where the duration of the activity and the amount of movement are also analyzed. In conclusion, the proposal is suitable to carry out a coarse monitoring over most elderly persons at large scales, although other BCD methods should be used in order to improve accuracy. In future studies, it is planned to compare the classical BCD methods with the one proposed here; for that, it is necessary to use a dataset similar to [47], merging not only health events but also energy consumption information.

Acknowledgments: This work has been supported by the Spanish Ministry of Economy and Competitiveness (TARSIUS project, ref. TIN2015-71564-C4-1-R, and SOC-PLC project, ref. TEC2015-64835-C3-2-R MINECO/FEDER).

**Author Contributions:** José M. Alcalá, Jesús Ureña and Álvaro Hernández devised the algorithm proposal; José M. Alcalá and David Gualda conducted the experiments; and the paper was written by all the authors.

Conflicts of Interest: The authors declare no conflict of interest.

## References

- Debes, C.; Merentitis, A.; Sukhanov, S.; Niessen, M.; Frangiadakis, N.; Bauer, A. Monitoring activities of daily living in smart homes. *IEEE Signal Process. Mag.* 2016, 33, 81–94. [CrossRef]
- Scanaill, C.N.; Carew, S.; Barralon, P.; Noury, N.; Lyons, D.; Lyons, G.M. A review of approaches to mobility telemonitoring of the elderly in their living environment. *Ann. Biomed. Eng.* 2006, 34, 547–563. [CrossRef] [PubMed]
- Ghamari, M.; Janko, B.; Sherratt, R.S.; Harwin, W.; Piechockic, R.; Soltanpur, C. A survey on wireless body area networks for ehealthcare systems in residential environments. *Sensors* 2016, 16, 831. [CrossRef] [PubMed]
- 4. Patel, S.; Park, H.; Bonato, P.; Chan, L.; Rodgers, M. A review of wearable sensors and systems with application in rehabilitation. *J. Neuroeng. Rehabil.* **2012**, *9*, 21. [CrossRef] [PubMed]

- 5. Lou, C.; Li, R.; Li, Z.; Liang, T.; Wei, Z.; Run, M.; Yan, X.; Liu, X. Flexible graphene electrodes for prolonged dynamic ECG monitoring. *Sensors* **2016**, *16*, 1833. [CrossRef] [PubMed]
- Katz, S.; Ford, A.B.; Moskowitz, R.W.; Jackson, B.A.; Jaffe, M.W. Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. *J. Am. Med. Assoc.* 1963, 185, 914–919. [CrossRef]
- 7. LaPlante, M.P. The classic measure of disability in activities of daily living is biased by age but an expanded IADL/ADL measure is not. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* **2010**, *65*, 720–732. [CrossRef] [PubMed]
- 8. Kautz, H.; Etzioni, O.; Fox, D.; Weld, D. *Foundations of Assisted Cognition Systems*; Technical Report; Department of Computer Science & Engineering, University of Washington: Seattle, WA, USA, 2003; pp. 1–25.
- Lim, J.; Jang, H.; Jang, J.; Park, S. Daily activity recognition system for the elderly using pressure sensors. In Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–24 August 2008; pp. 5188–5191.
- Moufawad El Achkar, C.; Lenoble-Hoskovec, C.; Paraschiv-Ionescu, A.; Major, K.; Büla, C.; Aminian, K. Physical behavior in older persons during daily life: Insights from instrumented shoes. *Sensors* 2016, 16, 1225. [CrossRef] [PubMed]
- Massot, B.; Noury, N.; Gehin, C.; McAdams, E. On designing an ubiquitous sensor network for health monitoring. In Proceedings of the 2013 IEEE 15th International Conference on e-Health Networking, Applications & Services (Healthcom), Lisbon, Portugal, 9–12 October 2013; pp. 310–314.
- 12. Rashidi, P.; Cook, D.J.; Holder, L.B.; Schmitter-Edgecombe, M. Discovering activities to recognize and track in a smart environment. *IEEE Trans. Knowl. Data Eng.* **2011**, *23*, 527–539. [CrossRef] [PubMed]
- Kakde, A.; Gulhane, V. Real time composite user activity modelling using hybrid approach for recognition. In Proceedings of the 2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 5–7 March 2015; Volume 32248, pp. 1–6.
- 14. Tapia, E.M.; Intille, S.S.; Larson, K. Activity recognition in the home using simple and ubiquitous sensors. *Pervasive Comput.* **2004**, *3001*, 158–175.
- Chen, B.; Fan, Z.; Cao, F. Activity recognition based on streaming sensor data for assisted living in smart homes. In Proceedings of the 2015 International Conference on Intelligent Environments (IE), Prague, Czech Republic, 15–17 July 2015; pp. 124–127.
- Hoque, E.; Stankovic, J. AALO: Activity recognition in smart homes using Active Learning in the presence of Overlapped activities. In Proceedings of the 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth), San Diego, CA, USA, 21–24 May 2012; pp. 139–146.
- 17. Oliver, E.H.N.; Garg, A. Layered representations for human activity recognition. In Proceedings of the 4th IEEE International Conference on Multimodal Interfaces, Pittsburgh, PA, USA, 14–16 October 2002; pp. 3–8.
- Hong, X.; Nugent, C.; Mulvenna, M.; McClean, S.; Scotney, B.; Devlin, S. Evidential fusion of sensor data for activity recognition in smart homes. *Pervasive Mob. Comput.* 2009, *5*, 236–252. [CrossRef]
- Nugent, C.D.; Hong, X.; Hallberg, J.; Finlay, D.; Synnes, K. Assessing the impact of individual sensor reliability within smart living environments. In Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE 2008), Washington, DC, USA, 23–26 August 2008; pp. 685–690.
- 20. Noury, N.; Quach, K.A.; Berenguer, M.; Bouzi, M.J.; Teyssier, H. A feasibility study of using a smartphone to monitor mobility in elderly. In Proceedings of the 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom), Beijing, China, 10–13 October 2012; pp. 423–426.
- 21. Sprint, G.; Cook, D.; Fritz, R.; Schmitter-Edgecombe, M. Detecting health and behavior change by analyzing smart home sensor data. In Proceedings of the 2nd IEEE International Conference on Smart Computing (SMARTCOMP 2016), St. Louis, MO, USA, 18–20 May 2016; pp. 1–3.
- 22. Dawadi, P.N.; Cook, D.J.; Schmitter-Edgecombe, M. Modeling patterns of activities using activity curves. *Pervasive Mob. Comput.* **2016**, *28*, 51–68. [CrossRef] [PubMed]
- 23. Noury, N.; Berenguer, M.; Teyssier, H.; Bouzid, M.J.; Giordani, M. Building an index of activity of inhabitants from their activity on the residential electrical power line. *IEEE Trans. Inf. Technol. Biomed.* **2011**, *15*, 758–766. [CrossRef] [PubMed]
- Noury, N.; Quach, K.A.; Berenguer, M.; Teyssier, H.; Bouzid, M.J.; Goldstein, L.; Giordani, M. Remote follow up of health through the monitoring of electrical activities on the residential power line—Preliminary results of an experimentation. In Proceedings of the 11th International Conference on e-Health Networking, Applications and Services (Healthcom 2009), Sydney, Australia, 16–18 December 2009; pp. 9–13.

- Zhang, X.; Kato, T.; Matsuyama, T. Learning a context-aware personal model of appliance usage patterns in smart home. In Proceedings of the 2014 IEEE Innovative Smart Grid Technologies—Asia (ISGT Asia), Kuala Lumpur, Malaysia, 20–23 May 2014; pp. 73–78.
- 26. Bonfigli, R.; Squartini, S.; Fagiani, M.; Piazza, F. Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview. In Proceedings of the 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC), Rome, Italy, 10–13 June 2015; pp. 1175–1180.
- 27. Zoha, A.; Gluhak, A.; Imran, M.A.; Rajasegarar, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors* 2012, *12*, 16838–16866. [CrossRef] [PubMed]
- 28. Parson, O.; Ghosh, S.; Weal, M.; Rogers, A. An unsupervised training method for non-intrusive appliance load monitoring. *Artif. Intell.* **2014**, *217*, 1–19. [CrossRef]
- 29. Kelly, J.; Knottenbelt, W. Neural NILM. In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, Seoul, Korea, 4–5 November 2015; pp. 55–64.
- Anderson, K.D.; Berges, M.E.; Ocneanu, A.; Benitez, D.; Moura, J.M.F. Event detection for non intrusive load monitoring. In Proceedings of the 38th Annual Conference on IEEE Industrial Electronics Society, Montreal, QC, Canada, 25–28 October 2012; pp. 3312–3317.
- De Baets, L.; Ruyssinck, J.; Deschrijver, D.; Dhaene, T. Event detection in NILM using cepstrum smoothing. In Proceedings of the 3rd International Workshop on Non-Intrusive Load Monitoring, Vancouver, BC, Canada, 14–15 May 2016.
- Wild, B.; Barsim, K.S.; Yang, B. A new unsupervised event detector for non-intrusive load monitoring. In Proceedings of the 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Orlando, FL, USA, 14–16 December 2015; pp. 73–77.
- Gao, J.; Kara, E.C.; Giri, S.; Berges, M. A feasibility study of automated plug-load identification from high-frequency measurements. In Proceedings of the 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP), Orlando, FL, USA, 14–16 December 2015; pp. 220–224.
- 34. Hassan, T.; Javed, F.; Arshad, N. An empirical investigation of V-I trajectory based load signatures for non-intrusive load monitoring. *IEEE Trans. Smart Grid* 2014, *5*, 870–878. [CrossRef]
- 35. Song, H.; Kalogridis, G.; Fan, Z. Short paper: Time-dependent power load disaggregation with applications to daily activity monitoring. In Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-IoT), Seoul, Korea, 6–8 March 2014; pp. 183–184.
- 36. Alcalá, J.; Parson, O.; Rogers, A. Detecting anomalies in activities of daily living of elderly residents via energy disaggregation and cox processes. In Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments, Seoul, Korea, 4–5 November 2015.
- Fleury, A.; Vacher, M.; Noury, N. SVM-based multimodal classification of activities of daily living in health smart homes: Sensors, algorithms, and first experimental results. *IEEE Trans. Inf. Technol. Biomed.* 2010, 14, 274–283. [CrossRef] [PubMed]
- 38. Nikamalfard, H.; Zheng, H.; Wang, H.; Jeffers, P.; Mulvenna, M.; McCullagh, P.; Martin, S.; Wallace, J.; Augusto, J.; Carswell, W.; et al. Knowledge discovery from activity monitoring to support independent living of people with early dementia. In Proceedings of the 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), Hong Kong, China, 5–7 January 2012; Volume 25, pp. 910–913.
- Alcalá, J.; Ureña, J.; Hernández, Á. Activity supervision tool using non-intrusive load monitoring systems. In Proceedings of the IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA), Luxembourg, 8–11 September 2015.
- 40. NILM-wiki. NILM datasets. Available online: http://wiki.nilm.eu/index.php?title=NILM\_datasets (accesed on 1 February 2017).
- Batra, N.; Kelly, J.; Parson, O.; Dutta, H.; Knottenbelt, W.; Rogers, A.; Singh, A.; Srivastava, M. NILMTK: An open source toolkit for non-intrusive load monitoring categories and subject descriptors. In Proceedings of the 5th International Conference on Future Energy Systems (ACM e-Energy), Cambridge, UK, 11–13 June 2014; pp. 1–4.
- Beckel, C.; Kleiminger, W.; Cicchetti, R.; Staake, T.; Santini, S. The ECO data set and the performance of non-intrusive load monitoring algorithms. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, Memphis, TN, USA, 3–6 November 2014; pp. 80–89.
- 43. NILM-wiki. Companies Offering NILM Products and Services. Available online: http://wiki.nilm.eu/index.php?title=Companies\_offering\_NILM\_products\_and\_services (accesed on 1 February 2017).

- 44. Rakowsky, U.K. Fundamentals of the Dempster-Shafer theory and its applications to reliability modeling. *Int. J. Reliab. Qual. Saf. Eng.* **2007**, *14*, 579–601. [CrossRef]
- 45. Zimmermann, J.-P.; Evans, M.; Lineham, T.; Griggs, J.; Surveys, G.; Harding, L.; King, N.; Roberts, P. *Household Electricity Survey: A Study of Domestic Electrical Product Usage*; Intertek: London, UK, 2012; p. 600.
- 46. Kelly, J.; Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci. Data* **2015**, *2*, 150007. [CrossRef] [PubMed]
- 47. Cook, D.J.; Crandall, A.S.; Thomas, B.L.; Krishnan, N.C. CASAS: A smart home in a box. *Computer* **2013**, 46, 62–69. [CrossRef] [PubMed]



© 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).