

# Probabilistic Analysis of Temporal and Sequential Aspects of Activities of Daily Living for Abnormal Behaviour Detection

Konios, A., Garcia-Constantino, M., Christopoulos, S., Mustafa, M., Ekerete, I., Shewell, C., Nugent, C. & Morrison, G.

Author post-print (accepted) deposited by Coventry University's Repository

## Original citation & hyperlink:

Konios, A, Garcia-Constantino, M, Christopoulos, S, Mustafa, M, Ekerete, I, Shewell, C, Nugent, C & Morrison, G 2020, Probabilistic Analysis of Temporal and Sequential Aspects of Activities of Daily Living for Abnormal Behaviour Detection. in 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). IEEE, IEEE International Conference on Ubiquitous Intelligence and Computing , Leicester, United Kingdom, 19/08/19.

<https://dx.doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00158>

DOI 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00158

ISBN 978-1-7281-4034-6

Publisher: IEEE

**© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.**

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

# Probabilistic Analysis of Temporal and Sequential Aspects of Activities of Daily Living for Abnormal Behaviour Detection

A. Konios<sup>†</sup>, M. Garcia-Constantino<sup>\*</sup>, S.-R. G. Christopoulos<sup>†</sup>, M. A. Mustafa<sup>‡¶</sup>  
I. Ekerete<sup>\*</sup>, C. Shewell<sup>\*</sup>, C. Nugent<sup>\*</sup> and G. Morrison<sup>§</sup>

<sup>†</sup>Institute for Future Transport and Cities, Coventry University, Coventry, United Kingdom.

<sup>\*</sup>School of Computing, Ulster University, Jordanstown, United Kingdom.

<sup>‡</sup>School of Computer Science, University of Manchester, Manchester, United Kingdom.

<sup>¶</sup>imec-COSIC, KU Leuven, Leuven, Belgium.

<sup>§</sup>The Lava Group, Belfast, United Kingdom.

**Abstract**—This paper presents a probabilistic approach for the identification of abnormal behaviour in Activities of Daily Living (ADLs) from dense sensor data collected from 30 participants. The ADLs considered are related to preparing and drinking (i) tea, and (ii) coffee. Abnormal behaviour identified in the context of these activities can be an indicator of a progressive health problem or the occurrence of a hazardous incident. The approach presented considers the temporal and sequential aspects of the actions that are part of each ADL and that vary between participants. The average and standard deviation for the duration and number of steps of each activity are calculated to define the average time and steps and a range within which a behaviour could be considered as normal for each stage and activity. The Cumulative Distribution Function (CDF) is used to obtain the probabilities of abnormal behaviours related to the early and late completion of activities and stages within an activity in terms of time and steps. Analysis shows that CDF can provide precise and reliable results regarding the presence of abnormal behaviour in stages and activities that last over a minute or consist of many steps. Finally, this approach could be used to train machine learning algorithms for abnormal behaviour detection.

**Index Terms**—Activities of Daily Living, ADLs, Cumulative Distribution Function, CDF, Probabilistic Analysis

## I. INTRODUCTION

Activities of Daily Living (ADLs) play an important role in the design and implementation of reliable sensor systems that can effectively monitor the activities of human population and are also capable of detecting unusual activities that may affect a person’s well-being. In general, the detection of abnormal behaviour in ADLs can be an indicator of a progressive health problem taking place (dementia, osteoporosis, arthritis, etc.) or the occurrence of a hazardous incident (a falls/burns/cut, food/smoke intoxication, etc.). Monitoring ADLs for detecting abnormal behaviour is of particular importance due to the potential life changing consequences that could result from not acting timely. Using sensors that can be placed within the environment and in appliances of interest, as opposed to wearable sensors, has the main advantage of not being intrusive for the users. This work continues the research presented in [4] and [5]; the former uses Petri nets to model and verifies ADLs, and the latter presents a probabilistic approach based on the temporal analysis of ADLs in order to identify abnormal behaviour. The current work extends the probabilistic analysis applied to the sensor data collected from participants for the ADL of “preparing and drinking a hot beverage” with the variants of tea or coffee, which is presented in [5]. Specifically, a probabilistic analysis is conducted with respect to the step sequences of the activities to identify any abnormality that could potentially be associated to health issues or hazards. Thus, as a means to identify normal or abnormal behaviour, the research

presented considers: (i) the number of actions (steps) followed by the participants in each stage of an activity, (ii) the total number of actions (steps) of an activity, (iii) the duration of actions that are part of an activity, and (iv) the total duration of an activity. While “preparing and drinking a hot beverage” with two choices of drink is a relatively simple and common ADL, the sequences of actions and their durations for different users provides an interesting insight and granularity from which abnormal behaviours can be detected.

In this case, it is assumed that a user completes an activity regardless of the number and sequence of steps involved, and of the time taken for the completion. Particularly, it is the combination of the temporal (i.e. duration) and sequential (i.e. number of steps) aspects of the activities that is used to find patterns that could help in the abnormal behaviour detection. If there are patterns found from sensor data outside the norm for an ADL that a person within a certain age range performs, then it could be inferred that this person has a health problem and must be attended by carers. In addition, the sensor system could send prompts to the user suggesting the completion of an activity. Detection of active appliances which could represent a hazard when unattended (e.g. stove or oven) could trigger alert messages sent to the user, relatives and/or carers.

The main contribution of this paper is the probabilistic analysis of the temporal and sequential aspects of the actions performed by 30 volunteers for the identification of patterns that could indicate abnormal behaviours. Although the dataset collected from the experiment conducted includes data from a number of sensors (contact, thermal and accelerometer) for each user, this paper considers only the contact sensor data as it provides more accurate and appropriate data regarding the users’ interactions with the appliances and resources involved in the ADLs for the type of analysis presented. The thermal and accelerometer data will be used in future work.

The remainder of the paper is organised as follows: Section II presents the related work in the areas of activity recognition and probabilistic analysis of intelligent systems. Section III describes the environment setup, participants and scenarios considered for the data collection of ADLs. Section IV discusses the temporal analysis results of the ADLs performed by the participants, while Section V presents the results of the sequential analysis. Section VI discusses the correlation between the temporal and sequential analysis. Finally, Section VII presents the conclusions.

## II. RELATED WORK

Activity recognition using sensors is typically classified in terms of wearable sensors versus dense sensors. Wearable sensors can be worn by users in parts of their body or in their clothes [3], [7], [18]. Dense

This work was supported by Invest Northern Ireland under the Competence Centre Programs Grant RD0513853 - Connected Health Innovation Centre.

sensors are attached to objects in the environment with which the user interacts (e.g. kitchen, bathroom, kettle) [17].

The use of sensors for activity recognition in ADLs to detect and predict abnormal behaviour has been investigated in the literature. A wireless sensor network to monitor and detect behaviour of elderly people is presented in [16]. The system generates sensor activity patterns which are used to predict abnormal behaviour based on a classification model of regular and irregular sensor activity. In [10], a sensor network to monitor elders' behaviour and to identify and prevent abnormal behaviour using neural networks is introduced. The results presented in [10] suggest that duration of activities is the most effective way to represent and handle a large sensor dataset.

In [11], occupancy data represented as temporal sequences of activities for identifying and predicting behaviours is investigated. The sensors used for detecting occupancy are Passive Infra-red (PIR), door/window entry point sensors and bed/sofa pressure sensors. An automated method based on Markov Chains Model (MCM) to classify abnormal sequences of activities is presented in [19]. The probability distribution of the spatio-temporal activity data of an elderly user over 128 days is analysed. Another probabilistic approach is presented in [1], in which data from motion, door and pressure sensors is collected from 40 households and presented in an abstract layer. A probabilistic spatio-temporal model is then produced to present a summary of daily behaviour where significant changes are defined as anomalies.

The Online Daily Habit Model and Anomaly Detection model [12] is comprised by an Online Activity Recognition model and a Dynamic Daily Habit Modeling component. This model aims to present real-time personalised ADL recognition, habit modeling and anomaly detection for elderly people. In [4], Petri nets are used to model three ADLs for the identification of the completion or non-completion of the activities as indicators of normal or abnormal behaviour, respectively. The description of the sensor system considered and the developed Petri net model and its verification are presented. The work presented in [8] introduces a unification approach to help in the detection of normal and abnormal behaviour based on the performance of ADLs by elders in care homes equipped with different sensors. Unlike the existing solutions, this paper addresses two core aspects of ADLs (i.e. temporal and step sequence) that could be used to indicate the presence or not of an abnormality while they are executed by random populations.

### III. ENVIRONMENT SETUP, VOLUNTEERS AND ADLS

The ADLs considered for the data collection and analysis are: (i) preparing and drinking tea, and (ii) preparing and drinking coffee. These ADLs are usually carried out in the kitchen and are based on the ones used in [2]. The data collection for the kitchen ADLs was conducted at the smart kitchen lab of the Smart Environments Research Group (SERG)<sup>1</sup> at Ulster University (see layout in Fig. 1).

#### A. Setting up the Environment

Three types of sensors (contact, thermal and accelerometers) were installed to capture the activities and collect the respective data. For the experiment, the contact sensors were attached to objects with which the user has interaction in the kitchen in the context of the ADLs considered in this paper: doors, cupboards, refrigerator, cups, containers (tea, coffee, sugar and milk), and location areas (kitchen worktop, table and sink). The contact sensors combine wireless transmitters and magnetic switches. The signals from the contact

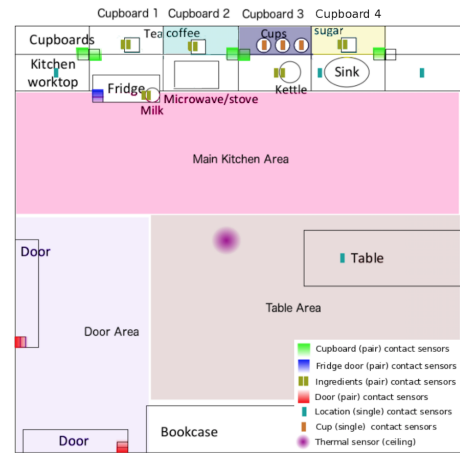


Fig. 1: Smart kitchen layout at SERG.

sensors have two possible states (“on” (or 1) and “off” (or 0)) and are monitored and collected by SensorCentral [13], a sensor data platform, for further processing and data analysis. The contact sensors are represented in Fig. 1 as rectangles divided into two parts that can be separated (‘on’ state) or joint (‘off’ state). The colour codes in the legend indicate the objects to which these sensors are attached. The details of the data collection, preprocessing and analysis of the thermal sensors and accelerometers are beyond the scope of this paper and will be addressed comprehensively in future work.

#### B. Recruiting Volunteers

Thirty volunteers were recruited for this project following specific selection requirements related to their age, gender, and health. All the volunteers were between 18 and 45 years old. This age-group was selected because the majority of the people at these ages are usually healthy and can also exhibit a behaviour that could potentially reveal progressive or temporary health issues [15]. Also, for the experiments conducted, both genders were considered, as the behaviour exhibited usually differs [6], which could disclose useful information about the way in which the ADLs are executed and also the time they last. The aim here was to collect information about these two factors, which could indicate abnormal behavioural patterns that could be linked to medical conditions or hazards. Actually, the purpose of the project is not to examine a specific medical condition, but a sample that contains people who are either healthy or have any medical condition. This could result in useful observations about the behaviour that is displayed in each of these cases.

Following the aforementioned criteria, a random sample of 30 participants was generated consisting of 16 males and 14 females, the youngest and oldest of whom were 22 and 43 years old respectively. Twenty eight of them (~ 93%) stated that they are healthy, i.e. they do not suffer from any chronic medical condition, while the remaining two (one male and one female) are known to have a chronic medical condition. Although it would be interesting to observe and analyse how these two volunteers behave compared to all the others, it is out of the scope of this paper and will be addressed in the future.

#### C. Defining the Activities of Daily Living

To investigate the behaviour of the participants, two ADLs have been considered, the preparation of coffee and tea, respectively. These two activities were chosen due to the fact that they can exhibit multiple or similar occurrences during a typical day of a person or can be met as part of other daily activities, such as breakfast, lunch or dinner.

<sup>1</sup><https://www.ulster.ac.uk/research/institutes/computer-science/groups/smart-environments>

For the execution of these two activities, the volunteers had the initiative with respect to the way that they can prepare their drink and the time that they will spend to drink it. Consequently, each participant was able to repeat any actions he/she wanted, but it is assumed that he/she could only use the ingredients and equipment that were available for the activity (i.e. coffee, tea, sugar, milk, cups and kettle). The order in which all these items would be used was exclusively dependent on the participant's will. The only restrictions imposed on the participants regarding the completion of the activity were: i) each participant can prepare only one drink, ii) coffee/tea should be drunk at the table and iii) the cup is placed in the sink after finishing the drink.

For the experiments, the volunteers had to follow a general scenario. In this scenario, each participant had to first enter the kitchen using one of the doors, prepare the drink of his/her preference, sit at the table to drink it, leave the cup in the sink when the drink is finished and finally exit the kitchen. This scenario was introduced because the activity steps which the volunteers follow to perform the ADLs can be traced more accurately, thus supporting the data analysis process. Note that taking into account the initiative of the volunteers and the number of items used in each of the activities, the construction of all the paths (i.e. different sequences of actions) that can be potentially followed for the successful completion of each activity gives more than hundred different ways for the preparation of each drink. Thus, if the repetition of some actions (steps) occur, then this number may increase exponentially. For instance, an initial calculation of the total number of unique sequences of steps/actions (i.e. paths) showed that there exist around 120 different ways to perform the coffee or tea activity respectively (including no repetition of actions). Examples of such paths are presented below for each activity:

Cof: *Door1* → *Fridge\_Door* → *Milk* → *Cupboard2(Coffee)* → *Coffee* → *Cupboard3(Cup)* → *Cup* → *Cupboard4(Sugar)* → *Sugar* → *kettle* → *Hot\_water*

Tea: *Door1* → *Cupboard3(Cup)* → *Cup* → *Fridge\_door* → *Milk* → *Cupboard1(Tea)* → *Tea* → *Cupboard4(Sugar)* → *Sugar* → *kettle* → *Hot\_water*

Identifying all the sequences of action for both activities will help with their temporal and step analysis as they could define both the time and number of actions required for the completion of an activity. Finally, having set all the conditions and requirements for the experiment, data were collected from the contact sensors with respect to the ADLs and then were preprocessed (see [5] for the details) in order to convert their raw form into the appropriate format for the data analysis. Thus, 30 CSV files were created containing only the information that is required for the recreation of the scene for the activity performed by each user, such as dateAndTime, name (i.e. object that user interacts with) and eventCode. Actually, the resulting files show a sequence of interactions of the participants with the objects, and when and how they were used based on the timestamp and event code (i.e. "on" and "off" states) respectively. For the temporal and step analysis of those sequences, each activity (i.e. coffee or tea) was broken into four main stages: i) entering the room, ii) drink preparation, iii) drinking and iv) exiting the room. This approach was used to explicitly examine how each participant behaves in each stage in terms of time and steps. The purpose of the temporal and step analyses is to associate the duration and number of steps of the subtasks or entire activity with potential abnormal behaviour respectively. In this work, the abnormal behaviour is defined as a value (i.e. duration or number of steps) that deviates considerably from the average sample values and also lies outside a

predetermined value range that defines the spectrum of all the other behaviours.

#### IV. TEMPORAL ANALYSIS OF ADLs

For the temporal aspects of the ADLs, the duration of each stage needs to be explicitly defined for this analysis. Thus, the duration of the entering process is defined as the time past from the activation of the door sensor until the first reading recorded from any of the object sensors involved in the preparation of the chosen drink. Similarly, the duration of the drink preparation stage can be counted by getting as starting point the first sensor reading related to the preparation process (i.e. ingredient or equipment reading) and as ending point a cup sensor reading that is followed by the prolonged use of that cup. The duration of the drinking process is calculated by considering the timestamps of the cup sensor readings (i.e. event codes), which indicate a prolonged use of the cup. This prolonged use is defined by the continuous changes to the value of the event code of the cup sensor from 0 to 1 and vice versa. Finally, the duration of the exiting process is computed by taking the timestamp of the last zero value of the cup sensor event code as the starting point of this stage and the timestamp of the last door sensor reading as the ending one. Note all the repeated actions have been included in the calculation of the duration of all the different activity stages. Hence, following the rationale described above, the duration of the activities and stages is calculated for each volunteer, as shown in Table I.

TABLE I: Activities and Stages Duration (in sec.)

Vol. ID	Entering	Preparation	Drinking	Exiting	Tot. Time
1	5	144	670	21	840
2	44	110	481	21	656
...	...	...	...	...	...
29	56	84	356	20	516
30	8	170	759	4	941
Aver.	35.07	150.27	577.47	16.27	779.08
St. Dev.	19.26	43.93	240.28	8.09	231.99

Next, the average and standard deviation of the durations are calculated to define the average time of execution per stage and the range under which a behaviour can be considered as "normal" for each stage and activity respectively. In more details, the range given by the average value and the standard deviation (i.e.  $\bar{x} \pm \sigma$ ) denotes the durations that correspond to normal behaviour. Thus, any duration that falls outside this range could potentially represent an abnormal behaviour for (a stage of) the entire activity [8].

To ensure that a normal behaviour will not be accidentally identified as abnormal due to its boundary value (i.e. the duration value is close to the minimum or maximum range values but outside that range), the  $1.5\sigma$  is chosen for the creation of the normal behaviour range. For this sample size, the selection of the  $1.5\sigma$  gives a representative range of values for the duration of normally executed activities. Also,  $1.5\sigma$  enables the identification of potential abnormality in a more realistic way compared to the choice of  $\sigma$  or  $2\sigma$ , which would give too many or zero and one abnormality results respectively. This range (i.e.  $\bar{x} \pm 1.5\sigma$ ) in the normal distribution covers the 86.6% of the sample. The reference to the normal distribution is made because of the fact that it is initially expected, for this sample, the duration of the activities (and/or their stages) to follow that distribution.

To further analyse the activities considered, the Cumulative Distribution Function (CDF) is used [14]. The CDF gives the probability that an activity/stage duration takes a value that is less or equal to a certain value [9]. Thus, to determine whether a behaviour is abnormal or not, an upper and lower bound duration is defined for each activity and stage (given by the  $1.5\sigma$ , which is defined in the CDF graphs as

the range from 0.067 to 0.933). CDF shows the probability of getting abnormal behaviours related to the early and late completion of an activity or stage respectively.

*Evaluating the Temporal Analysis Outcomes:* Using the duration values of Table I for the CDF, different graphs were produced for the analysis of tea and coffee activities. Specifically, these graphs show the probability over the duration for each activity and stage.

To examine the probability of getting abnormal behaviours that can be used as an indicator for progressive health issues or hazard, the sample of the 30 volunteers was organised into groups according to their gender, age and activity. Thus, the different groups for each of these activities are defined as: i) male and female, ii) below and over 30 years old and iii) coffee and tea. The actual numbers of participants for these groups are: i) 16 males and 14 females, ii) 18 and 12 participants below and over the 30 years old respectively and iii) 12 participants execute the coffee activity while the remaining 18 participants perform the tea one.

Now, to find the projection values that correspond to the abnormal behaviour expected for each of these groups, the number of participants needs to be multiplied by the probability of abnormal behaviour as is defined by the  $1.5\sigma$ . Since the  $1.5\sigma$  covers the range of normal behaviour (i.e. 86.6%), this implies that the remaining 13.4% indicates the potential abnormal behaviour. Therefore, the obtained projection values of the abnormal behaviour expected for each group are as follows: i) 2.14 for males and 1.88 for females, ii) 2.41 and 1.61 for participants below and over 30 years old respectively and iii) 1.61 for participants executing the coffee activity and 2.41 for participants performing the tea activity. To conduct the temporal analysis of the ADLs and their stages, these projection values are compared with the CDF results in order to identify the abnormal behaviour exhibited by the participants in each case.

Starting with the overall duration of the activities (see graphs in Fig. 2), it is observed, for example, in the coffee activity that only one participant exhibits behaviour that can resemble an abnormal behaviour as the total time required for the completion of his/her activity is considerably high. Specifically, this value lies outside the  $1.5\sigma$  as it is greater than the upper bound of 93.3%. Comparing the CDF result with the projection value for the coffee activity, it is noticed that they are quite close. This indicates that either one or two abnormal behaviours were expected according to the probability, which is also observed in the CDF result that shows that one participant was found to be out of the spectrum of normal behaviour (see Fig. 2(a)). Similarly, for the tea activity, three participants are outside the  $1.5\sigma$  range as shown in the graph of Fig. 2(a), which number is very close to that of the projection, indicating the abnormal behaviour with quite a good accuracy.

Examining the volunteers' behaviour with respect to their gender, the graph in Fig. 2(b) shows that three male (resp. one female) participants (for both activities) are not in the normal behaviour spectrum defined. Comparing this value with the probability projection for the males (resp. females), it can be concluded that the real behaviour can be precisely depicted using the CDF for these subgroups of the main sample. Similarly, for the under and over 30 age-groups, the projection values are close to the number of participants whose behaviour is considered as abnormal, which is three and one respectively (as shown in Fig. 2(b)).

Apart from the probabilistic analysis, CDF can also provide very useful information about potential group of values by examining their density and distance from each other. This information could be used as an additional factor that could result in a more accurate

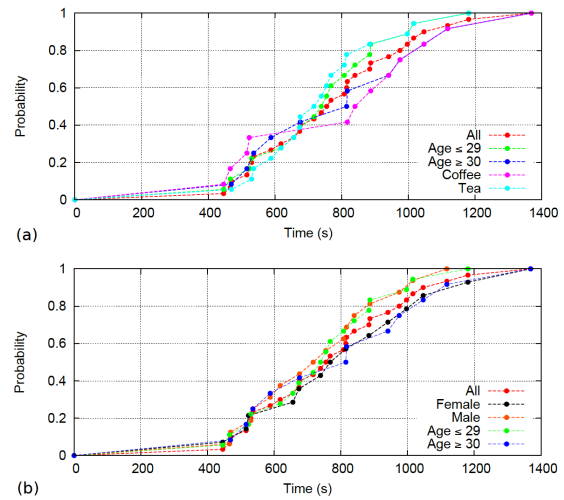


Fig. 2: CDF of Duration for (a) activity performed and age-group and (b) each gender and age-group respectively.

identification of the abnormal behaviour. Furthermore, it can show potential behavioural patterns of the participants by examining the comparison of the groups that were defined earlier for the participants. Note that the density of the values for each participant groups (e.g. tea, over 30 y.o., etc.) is quite high (i.e. very close to each other) for the short lasting cases, i.e. the activities last less than 480 seconds. This indicates that although the probability gives a certain number of possible abnormal behaviours, the adjacent values of these behaviours should be examined for abnormality as they are marginally close to them. For example, in Fig. 2(b), CDF shows that three males exhibit abnormal behaviour (i.e. two and one participants for early and late completion respectively). However, since the density is high for the group of values below 480 seconds, then the adjacent value of the two values that are considered as abnormal behaviour should also be examined, which means that the number of abnormal behaviours for the males can potentially be increased to four.

Examining the distance of the sample values, it is noteworthy that there exist no long distances between them in terms of time regarding the different groups of participants. The only case where the distance between the values is significantly big, is in the coffee activity, where it seems that two behavioural groups can be considered, one that executes the activity in less than 500 seconds and another that finishes it after the 800 seconds showing that a gap (i.e. distance) of almost 300 seconds existing between the two behavioural patterns. Now, comparing these two groups with the values of the tea activity, it is observed that the first group is consistently faster than the respective values of the tea activity, while the latter is slower than the remaining tea participants. Finally, it is noticed that generally the males execute the activities faster than the females (see Fig. 2(b)) revealing two different behavioural patterns for these groups.

The graphs (a) and (b) in Fig. 3 represent the duration of the entering and exiting stages of the activities, respectively. The CDF results reveal that one and three participants exhibit abnormal behaviour for the coffee and tea activities of both stages, respectively. This shows that the number of abnormal behaviour exhibited in reality is very close to the projection values of these samples. Moreover, for the age-groups of those stages, it is observed that three persons from the under 30 group present an abnormal behaviour, while for the other group, only one person lies outside the  $1.5\sigma$  range. Note that

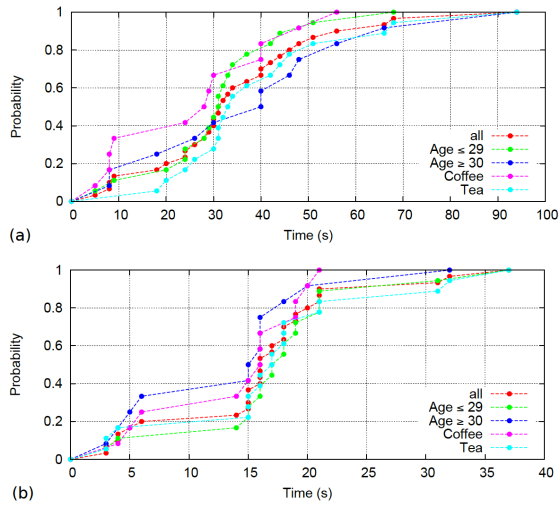


Fig. 3: CDF of (a) entering and (b) exiting kitchen for age-groups and activities performed.

although the results show that the projection values are very close to the CDF ones, this does not necessarily mean that an “abnormal” behaviour is related to a health issue as the participants may hurry. Consequently, from the analysis of the entering and exiting stages, it could be concluded that CDF cannot give a very clear outcome regarding the abnormal behaviour related to the early completion of these phases, but it could give useful information about their late completions, which are highly-likely to be linked to a health issue or hazard.

Now, regarding the density and distance of the values for the entering stage, it is noticed that although the values around the lower and upper bounds are not very dense (in terms of time) to need further examination of the adjacent values for potential abnormality, there exist notable distances between three different groups of values. These groups consist of the values that are in the ranges from 0 to 9, 18 to 56 and 66 to 70 seconds, respectively. Specifically, the distance between the first and second group, and the second and third group is 9 and 10 seconds, respectively. These distances could be used as reference/indication points for the early and late completion of this stage. Regarding the exiting phase, it becomes apparent that the high densities and long distances observed in terms of time can indicate that the adjacent values of the lower and upper bounds should be examined further for potential abnormalities and that may exist different behavioural patterns that are defined by the group values that belong to the ranges from 0 to 6, 14 to 21 and 31 to 33 seconds, respectively (see Fig. 3(b)). At this point, it should be mentioned that in both stages, there exists a participant who is extremely late (exhibiting abnormal behaviour) compared to all the other group of values. This is denoted by the distances shown in Fig. 3(a) and (b). Comparing the CDF results for the different groups of participants, it is observed that these stages show clear behavioural differences between certain groups. For instance, the participants of the coffee activity seem to execute these stages faster than those of the tea activity. Finally, the participants that are over 30 year old finish their exiting earlier than the participants of the other age-group, which does not clearly happen during the entering process.

In the last part of this section, the probability of having abnormal behaviours during the preparation and drinking stages of the activities is discussed. In the preparation of the drinks (resp. drinking process),

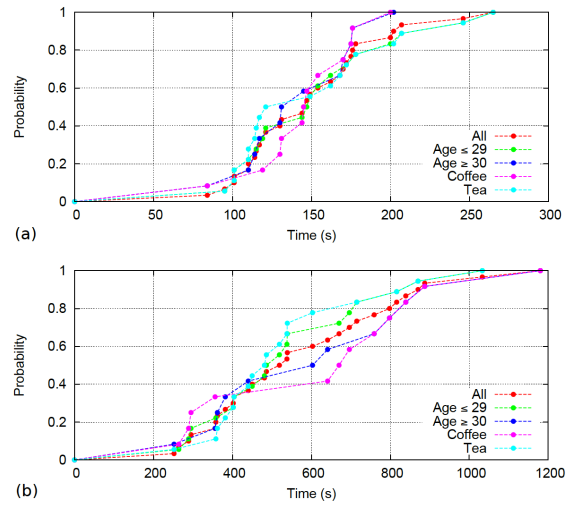


Fig. 4: CDF of (a) preparing and (b) drinking coffee or tea for age-groups.

the number of behaviours that are outside the  $1.5\sigma$  is equal to one and three (resp. 1 and 3) for the coffee and tea respectively. These values are close to the values given to the projection values calculated earlier. Similarly for the two age-groups, where the graph shows that one and three (resp. 1 and 3) participants exhibit abnormal behaviour for the over and below 30 year old group respectively. This again indicates that the abnormal behaviour in reality “matches” the one given by the projection in the drink preparation (resp. drinking) stage. Examining the density for these two stages, it is noticed that the values around the lower and upper bounds are not very dense (in terms of time) for the preparation stage in order to indicate that further examination of the adjacent values should be considered. On the contrary, in the drinking process, there exist high density regarding the lower bound values for the participants that are below 30 years old or drink a coffee, which implies that the adjacent values should be further examined in order to avoid the case of false positives or negatives as regards the abnormalities. Now, for both stages, it is observed that in general there are no significantly large distances between the values of the different groups of participants, except for two cases of the drinking stage. In the first case, a distance of over 200 seconds is noticed for the coffee activity, which creates two different groups of values, which could correspond to different behaviours. One group where all the participants drink their coffee in less than 400 seconds and another group in which all the participants need more than 600 seconds for the drinking stage (see Fig. 4(b)). Similarly, for the participants that are over 30 years old, there exist a distance of over 150 seconds that splits this group into two subgroups that drink their beverage in less than 420 seconds and more than 600 seconds respectively. This could be an indication of having two different behaviours within the same group. Finally, it should be noted that no conclusions can be drawn with respect to the clear identification of behavioural patterns between the different groups, as from the comparison of their values results that in most cases the behaviours of these groups overlap or are quite similar.

## V. SEQUENTIAL ANALYSIS OF ADLS

As with the temporal analysis, the number of steps performed in each stage needs to be explicitly defined in order to conduct the step analysis of the ADLs. Consequently, the number of steps for the entering process is defined as the number of door sensor activations,

counting from the first activation of the door sensor until the first reading recorded from any of the object sensors involved in the preparation of the chosen drink. Usually, this involves the “on” and “off” readings recorded for the door sensor, which are considered as one single step. Similarly, the number of steps for the drink preparation stage is counted by getting as starting point the first sensor reading related to the preparation process (i.e. ingredient or equipment reading) and as ending point a cup sensor reading that is followed by the prolonged use of that cup. Once again, the pair of the “on” and “off” values of the event code of each activated sensor are considered as a single use of that sensor (i.e. a single step). Now, the number of steps for the drinking process is calculated by considering the total number of the “on” and “off” values of the cup sensor readings (i.e. event codes), which indicate a prolonged use of the cup. This prolonged use is defined by the continuous changes to the value of the event code of the cup sensor, where the pair of the “on” and “off” values constitutes a single use of the cup. It should be mentioned that the drinking process is considered as one single step that consists of a number of sips defined by the pair values of the event code of the cup sensor. Finally, the number of steps for the exiting process is computed by taking the last “off” (or zero) value of the cup sensor event code as the starting point of this stage and counting the pairs of values of the event code for door sensor reading until the last value of the door event code is set to “off”. Note that for the calculation of the number of steps of all the different activity stages, all the repeated actions have been included. Hence, following this logic, the number of steps of both the activities and their stages were calculated for each volunteer, as shown in Table II.

TABLE II: Activities and Stages Steps

Vol. ID	Entering	Preparation	Drinking	Exiting	Tot. No. of Steps
1	1	11	1 (17 sips)	1	14
2	1	8	1 (13 sips)	1	11
...	...	...	...	...	...
29	1	7	1 (10 sips)	1	10
30	1	21	1 (14 sips)	1	24
Aver.	1.03	8.87	14.5	1.13	12.2
St. Dev.	0.18	3.58	7.87	0.73	3.85

For the visualisation of the total number of steps performed by the participants in each activity, a sunburst diagram was generated (see Fig. 5). This graph depicts all the different paths/sequences followed during the activities execution and also indicates cases where participants have identical behaviours. Note that the last step of each path in the graph corresponds to the participant id and is not considered as an activity step.



Fig. 5: Graphical representation of step sequences of ADLs.

According to the sunburst diagram, there exist only two pairs of participants which exhibit the same behaviour, as they follow identical

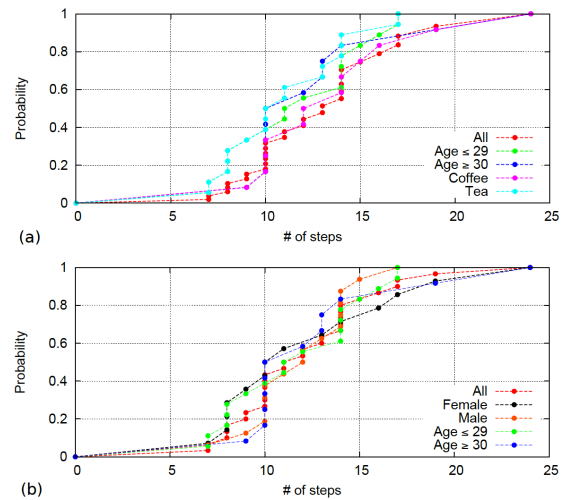


Fig. 6: CDF of total no. of steps for (a) activity performed and age-group and (b) each gender and age-group respectively.

sequences of actions with respect to the entire activities (see black frames in Fig. 5). On the contrary, more patterns can be identified regarding the steps of the main stages (e.g. entering, preparing, etc.). Note that this diagram includes all the repetitions of steps/actions and can be used to validate the results of the CDF performed on the steps, as they can be compared to the produced paths of the sunburst diagram to identify abnormal behaviour.

Applying CDF to the sequences of steps, the abnormal and normal behaviour is determined based on the defined lower and upper bounds for each activity and stage given by the  $1.5\sigma$  range. Specifically, in this case, the CDF shows the probability of getting abnormal behaviours related to the short and long paths (i.e. number of steps) of an activity or stage, respectively.

*Evaluating the Step Analysis Outcomes:* For the creation of the CDF diagrams for the step analysis of the tea and coffee activities, the number of steps presented in Table II is used. Actually, these graphs show the probability over the number of steps for each activity and stage. Note that the step analysis considers the groups of participants and projection values defined in the temporal analysis.

Analysing the CDF results for the coffee (resp. tea) activity in terms of the total number of steps, it is shown that one participant (resp. three participants) falls out of the  $1.5\sigma$  range (see Fig. 6(a)). These values are very close to the projection values for the coffee (resp. tea) activity confirming the accuracy of the CDF. Now, analysing the volunteers’ behaviour according to their gender, the graph in Fig. 6(b) shows that three male (resp. one female) participants (for both activities) lie outside the range of normal behaviour. Comparing the CDF outcome with the projection values for both gender, it can be concluded that the real behaviour can be described in a quite accurate way when using the CDF for these samples. Similarly, the projection values for the participants of the two age-groups are very close to the number of participants whose behaviour is considered as abnormal, which is three and one for the below and over 30 years old group respectively.

Examining the density of the values for the graphs in Fig. 6, it is noticed that the values around the lower bound of the range are quite dense (in terms of number of steps) for all the groups considered (i.e. age-group, gender and activity) as in some cases there exist only one

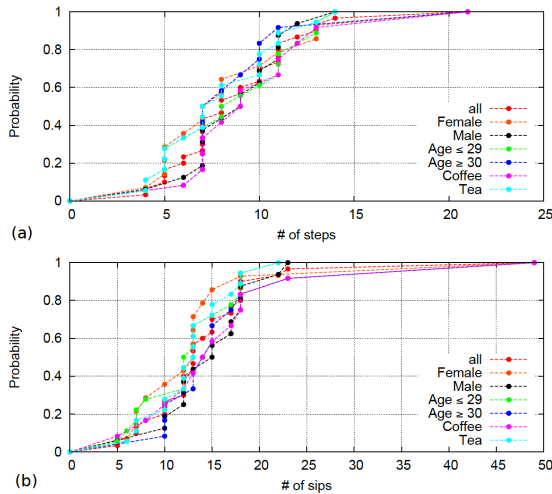


Fig. 7: CDF of (a) Preparing and (b) Drinking coffee or tea for age-groups, genders and activities.

step difference between the abnormal and normal values. This means that further examination should be conducted in order to avoid the case of a statistical error. Now, the analysis of the distance between the sample values cannot give any useful information with respect to the identification of abnormal behaviour as the distances in both graphs seem to be equal or very similar. On the contrary, comparing the different samples (i.e. groups), it is shown that all the participants that drink tea execute considerably less steps than those who drink coffee, which can represent potential behavioural patterns for these two activities.

Next, the abnormal behaviour regarding the preparation and drinking stages of the activities is examined. Starting with the preparation of the drinks, as is noticed from the graph in Fig. 7(a), the number of behaviours that lie outside the  $1.5\sigma$  range is one and three for the coffee and tea, respectively, which are once again close to the projection values of these two samples. Also, examining the two age-groups, it is shown that one and three participants exhibit abnormal behaviour for age-group over and under 30 years old, respectively. Comparing those values with the projection ones, it is observed that the abnormal behaviour in reality “matches” the one given by the probabilistic analysis of the preparation stage. Finally, regarding the genders, it is noticed that the graph of Fig. 7(a) indicates that the number of abnormal behaviours for the males and females is equal to three and one, respectively, which is again close to the values given by the probabilities.

Now, analysing the density of the values for the preparation stage, it is observed that the values around the lower bound present a high density for all the groups considered (i.e. age-group, gender and activity) as for these groups there exist one or two steps difference between the abnormal and normal behaviour. This means that further examination should be useful for the more accurate identification of the abnormal cases. Now, checking the distance between the sample values, it does not give any useful information regarding the identification of abnormal behaviour as all the distances in the graph of Fig. 7(a) are equal or very close. Finally, analysing the samples of the preparation stage, it is observed once again that the participants that prepare tea consistently execute less steps than those who prepare coffee exhibiting two potential behavioural patterns.

Continuing with the drinking process, it is worth noting that since the

drinking is considered as a single step process in the step sequences (see Fig. 6), the CDF for the graph of Fig. 7(b) is conducted on the number of sips performed by the participants in each activity<sup>2</sup>. The CDF for the below and over 30 years old age-groups reveals that volunteers that exhibit an abnormal behaviour while drinking their beverage is 3 and 1 respectively. Comparing this value to those given by the probability, it is shown that the expected abnormal behaviour is efficiently identified in reality for these samples. Now, for the drinking of coffee and tea, 1 and 3 abnormal behaviours are noticed in each activity respectively (see Fig. 7(b)), which is also confirmed by the probability values. Finally, regarding the two genders, CDF shows that 3 male and 1 female participants exhibit abnormal behaviour, which are quite similar to the indication of the projection values.

Examining the density of the values for the drinking stage, it is observed that the values around the lower and upper bound show a high density for age-group and activity samples as in some cases there exist one or two sips difference between the abnormal and normal behaviour (e.g. female and tea). Thus, further examination would help towards a more accurate identification of the potential abnormalities. Regarding the distance, the graph of Fig. 7(b) provides no useful information regarding the identification of abnormal behaviour as the distances do not differ significantly. Finally, analysing the samples of the drinking stage, it is noticed that the female and tea participants drink their beverage using less sips than the male and coffee participants respectively, which shows that there exist different behavioural patterns regarding the drinking process.

It is important to mention that step analysis of the entering and exiting stages using CDF has not been conducted due to the fact that all but one participants execute these stages in only one step. Therefore, CDF could not provide any accurate and useful information about the potential abnormal behaviour of these two stages. Another reason is because these stages cannot be defined as a collection of distinct sub-steps that could be further analysed, as is with the drinking stage.

## VI. COMPARING TEMPORAL AND STEP ANALYSIS

Having presented the temporal and step analysis for the identification of abnormalities during the ADLs and their stages, it is examined if there exists any correlation between the time and steps that could make the identification of abnormal behaviour more accurate. For that reason, the graphs of Fig. 8 were created considering only the stages for which the CDF was applied.

Examining the graph of Fig. 8(a), it is noticed that there is no clear correlation between the time and number of steps for the identification of abnormal behaviour when it comes to the preparation of the drinks. This is concluded from the fact that a behaviour that is identified as abnormal due to its long (resp. short) duration does not necessarily mean that this behaviour will consist of a big (resp. small) number of steps. The same applies to a behaviour that is identified as abnormal due to its big (resp. small) number of steps, which does not necessarily imply that it will last too much (resp. too little). Representative examples of such cases are the coffee preparation that lasts 170 seconds using 21 steps and the tea preparation that lasts 265 seconds using 10 steps in Fig 8(a), which imply that a big number of steps can be executed in a relatively small time and also that a small number of steps can take significantly long time to be performed. Thus, it is concluded that there is no clear pattern that can efficiently associate the time with the number of steps for the detection of abnormal behaviour using CDF.

<sup>2</sup>This occurs because the number of sips represents actual sub-steps of the one-step drinking stage.



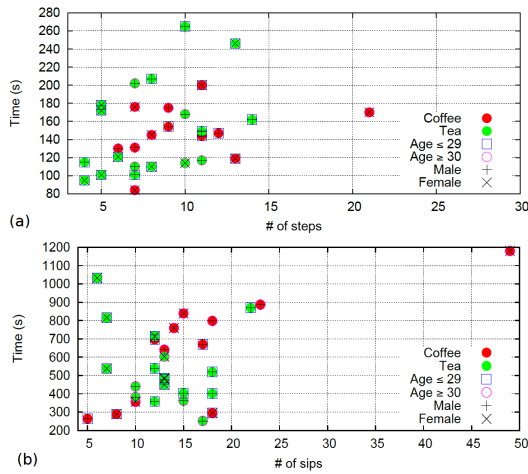


Fig. 8: Associating time and steps for (a) preparation and (b) drinking processes.

Now, for the drinking stage, it seems that there exists a correlation between time and steps (i.e. sips), but only for the coffee activity, in which the time required is interrelated to some extent to the number of sips made. This happens due to the fact that in this case the duration of the sips does not significantly vary among the participants (see Fig. 8(b)). On the contrary, this does not apply to the case of the tea activity, where the time and number of sips are not clearly linked. This occurs because of the fact that in this activity the duration of each sip differs for most of the participants.

*General Remarks:* From the evaluation of the results presented in the temporal and step analysis, it could be concluded that the CDF can provide quite accurate results regarding the presence of abnormal behaviour in stages/activities that are not short in terms of duration and number of steps. In case of having short stages or activities, the abnormal behaviour given from the CDF should be further examined for more accurate results, analysing the density and distance of the sample values. Apart from the probabilistic analysis, CDF can be used to compare the tendencies of the examined samples providing useful information about potential behavioural patterns.

Moreover, examining whether a combination of time and steps could be used for a more precise detection of potential abnormal behaviours using CDF, it is shown that this approach does not constitute an efficient way as there is no clear correlation between these two factors. However, the CDF can be used separately for the temporal and step analysis of ADLs, providing very accurate and useful outcomes. Finally, it could be claimed that these analysis models could be used for the training of machine learning algorithms to improve their capability of predicting abnormal behaviour with respect to the execution time and number of steps involved in the ADLs.

## VII. CONCLUSION

This paper proposed a probabilistic approach that is based on the Cumulative Distribution Function (CDF) for the temporal and step analysis of ADLs conducted by 30 volunteers in a lab environment. The methodology used introduces an effective way of indicating the probability of an executed activity to be considered as a normal or abnormal behaviour with respect to its duration and total number of steps. This analysis framework for the ADLs could help to the training of monitoring system to effectively detect abnormal

behaviours that are related to temporal and sequential aspects of the executed ADLs for a given population. For the further improvement of the proposed probabilistic analysis of the ADLs, a profile analysis of the participants will be conducted in the future trying to associate it with their temporal and step analysis in order to identify with more precision the presence of abnormalities related to health issues.

## REFERENCES

- [1] O. Aran, D. Sanchez-Cortez, M. Do, and D. Gatica-Perez, "Anomaly detection in elderly daily behavior in ambient sensing environments", Inter. Workshop on Human Behavior Understanding, pp. 51-67, 2016.
- [2] L. Chen, C.D. Nugent, and H. Want, "A Knowledge-Driven Approach to Activity Recognition in Smart Homes", IEEE Transactions on Knowledge and Data Engineering, IEEE, Vol. 24, No. 6, pp. 961-974, 2012.
- [3] M. Ermes, J. Pärkkä, J. Mäntyjärvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions", IEEE Transactions on Information Technology in Biomedicine, Vol. 12, No. 1, pp. 20-26, 2008.
- [4] M. Garcia-Constantino, A. Konios, and C. Nugent. "Modelling Activities of Daily Living with Petri nets", Advanced Technologies for Smarter Assisted Living solutions: Towards an open Smart Home infrastructure (SmarterAAL). 16th IEEE Inter. Conf. on Pervasive Computing and Communications, pp. 866-871, 2018.
- [5] M. Garcia-Constantino et al. "Probabilistic Analysis of Abnormal Behaviour Detection in Activities of Daily Living", 17th IEEE Inter. Conf. on Pervasive Computing and Communications, pp. 461-466, 2019.
- [6] R. C., Gur and R. E., Gur. "Complementarity of Sex Differences in Brain and Behavior: From Laterality to Multi-Modal Neuroimaging". Journal of Neuroscience Research, Vol. 95, No. 1-2, pp. 189-199, 2017.
- [7] Y. Hong, I. Kim, S. Ahn, and H. Kim. "Activity recognition using wearable sensors for elder care", Second International Conference on Future Generation Communication and Networking, IEEE, Vol. 2, pp. 302-305, 2008.
- [8] A. Konios, Y. Jing, M. Eastwookd, and B. Tan. "Unifying and Analysing Activities of Daily Living in Extra Care Homes", IEEE 16th International Conference on Pervasive Intelligence and Computing, IEEE, 2018.
- [9] E. L. Lehmann and G. Casella, "Theory of Point Estimation", 1998.
- [10] A. Lotfi, C. Langensiepen, S.M. Mahmoud, and M.J. Akhlaghinia. "Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour", Journal of Ambient Intelligence and Humanized Computing, Springer, Vol. 3, No. 3, pp. 205-218, 2012.
- [11] S. Mahmoud, A. Lotfi, and C. Langensiepen. "Behavioural Pattern Identification and Prediction in Intelligent Environments", Applied Soft Computing, Elsevier, Vol. 13, No. 4, pp. 1813-1822, 2013.
- [12] L. Meng, C. Miao, and C. Leung. "Towards Online and Personalized Daily Activity Recognition, Habit Modeling, and Anomaly Detection for the Solitary Elderly Through Unobtrusive Sensing", Multimedia Tools and Applications, Vol. 76, No. 8, 2017.
- [13] J. Rafferty, J. Synnott, A. Ennis, C. Nugent, I. McChesney, and I. Cleland. "SensorCentral: A Research Oriented, Device Agnostic, Sensor Data Platform", International Conference on Ubiquitous Computing and Ambient Intelligence, pp. 97-108, 2017.
- [14] N. V., Sarlis and S.-R. G., Christopoulos. "Natural time analysis of the Centennial Earthquake Catalog", Chaos: An Interdisciplinary Journal of Nonlinear Science, Vol. 22, No. 2, pp. 23-123, 2012.
- [15] C. D., Sherbourne, E., Keeler, J. Unützer, L., Lenert and K. B. Wells. "Relationship between age and patients' current health state preferences". Gerontologist, Vol. 39, No. 2, pp. 271-278, 1999.
- [16] N.K. Suryadevara, A. Gaddam, R.K. Rayudu, and S.C. Mukhopadhyay. "Wireless Sensors Network Based Safe Home to Care Elderly People: Behaviour Detection", Procedia Engineering, Vol. 25, pp. 96-99, 2011.
- [17] E. Tapia, S. Intille, and K. Larson, "Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors", Pervasive Computing, Springer, Vol. 4, pp. 158-175, 2004.
- [18] L. Wang, T. Gu, X. Tao, H. Chen, and J. Lu, "Recognizing multi-user activities using wearable sensors in a smart home", Pervasive and Mobile Computing, Elsevier, Vol. 7, No. 3, pp. 287-298, 2011.
- [19] T. Zhao, H. Ni, X. Zhou, L. Qiang, D. Zhang, and Z. Yu, "Detecting Abnormal Patterns of Daily Activities for the Elderly Living Alone", International Conference on Health Information Science, Springer, pp. 95-108, 2014.