International Journal of Pervasive Computing and Communications

Emerald Article: Dynamic similarity-based activity detection and recognition within smart homes

Xin Hong, Chris D. Nugent, Maurice D. Mulvenna, Suzanne Martin, Steven Devlin, Jonathan G. Wallace

Dynamic similarity-based activity detection and recognition within smart homes

Xin Hong
Institute of Electronics, Communications and Information Technology, Queen's University Belfast, Belfast, UK

Chris D. Nugent and Maurice D. Mulvenna
School of Computing and Mathematics, University of Ulster, Belfast, UK

Suzanne Martin
School of Health Sciences, University of Ulster, Belfast, UK, and

Steven Devlin and Jonathan G. Wallace
School of Computing and Mathematics, University of Ulster, Belfast, UK

Abstract

Purpose – Within smart homes, ambient sensors are used to monitor interactions between users and the home environment. The data produced from the sensors are used as the basis for the inference of the users' behaviour information. Partitioning sensor data in response to individual instances of activity is critical for a smart home to be fully functional and to fulfil its roles, such as correctly measuring health status and detecting emergency situations. The purpose of this study is to propose a similarity-based segmentation approach applied on time series sensor data in an effort to detect and recognise activities within a smart home.

Design/methodology/approach – The paper explores methods for analysing time-related sensor activation events in an effort to undercover hidden activity events through the use of generic sensor modelling of activity based upon the general knowledge of the activities. Two similarity measures are proposed to compare a time series based sensor sequence and a generic sensor model of an activity. In addition, a framework is developed for automatically analysing sensor streams.

Findings – The results from evaluation of the proposed methodology on a publicly accessible reference dataset show that the proposed methods can detect and recognise multi-category activities with satisfying accuracy, in addition to the capability of detecting interleaved activities.

Originality/value – The concepts introduced in this paper will improve automatic detection and recognition of daily living activities from timely ordered sensor events based on domain knowledge of the activities.

Keywords Activities of daily living, Ubiquitous sensing technology, Sensor segmentation, Activity detection and recognition, Sensors, Pervasive computing

Paper type Research paper

1. Introduction

Smart homes are emerging as environments where conventional living facilities are augmented with pervasive computing technologies, comprising sensors, actuators and knowledge-based data processing. Activity detection and recognition in such

This work was supported in part by the Nestling Technology Project.

International Journal of Pervasive Computing and Communications
Vol. 8 No. 3, 2012
pp. 264-278

1742-7371
DOI 10.1108/17427371211262653

environments has been recognised as a fundamental research area. Within this one specific application area for the research is in the monitoring of activities of daily living (ADL) for older people at home in support of their independent living. Research developments in gerontology have established standards to assess everyday functional competence of older people through measures of their ADLs (Katz, 1983; Lawton and Brody, 1969). ADL covers a wide range of living functions for example bathing, dressing, toileting, preparing meals, preparing drinks, use of telephone, taking medication, to name but a few.

Capturing knowledge about how people perform ADLs has been a focus of research, centering on the monitoring of people's interactions with domestic objects in association with the completion of the activities (Philipose *et al.*, 2004; Tapia *et al.*, 2004; van Kasteren *et al.*, 2008; Hong *et al.*, 2009). A rudimentary state-change sensor can be used to detect any change of state of an object in the home, which subsequently reflects the interactions of a person with the object (Wilson and Atkeson, 2005). Nevertheless, the situation is often more complex than this, requiring the use of intelligent computing to deduce hidden activities from recorded sensor events.

From a data analysis point of view, the challenge is two-fold. The first is how to automatically partition timely ordered sensor events into segments each corresponding to the process of an activity. The second is to identify the activity, the happening of which has been detected by a segment of sensor events. Solving this challenge can encounter a variety of uncertainties due to characteristics of ADLs and variations of home environments and sensorising facilities. People carry out ADLs in many different ways. For example, some people prefer to add milk and sugar in their tea, but some do not. Two or more activities may be carried out interleaving. For example, during preparing dinner the person may receive a phone call. She/he therefore breaks from the process of preparing dinner and after the telephone call comes back to it.

In this paper we explore methods for analysing time-related sensor activation events in an effort to undercover hidden activity events through the use of generic sensor modelling of activity based upon the general knowledge of the activities. Two similarity measures are proposed to compare a time series based sensor sequence and a generic sensor model of an activity. In addition, a framework for automatically analysing sensor streams is developed.

The remainder of this paper is organised as follows. Section 2 presents related research in the areas of ADL monitoring and time series data similarity measures. Section 3 presents a formal description of object-based activity monitoring with state-change sensors within a smart home environment, from the characteristics of sensor observations to the problem of sensor data analysis. In Section 4 two similarity measures are proposed to compare activities in the form of sensor sequences. In Section 5 we present a framework for detecting and recognising activities from sensor streams based on similarity measures. Section 6 presents results of implementing the methodology on a publicly available reference dataset. The paper is concluded in Section 7.

2. Background

2.1 ADL sensing and recognition

For the purpose of monitoring ADLs, a variety of sensing technologies are available, such as cameras, audio devices and state-change sensors. Among these, state-change sensors are a popular choice due to their characteristics, which include minimum obtrusion, lower cost and simple installation. State-change sensors distributed throughout a smart home are able to monitor the interaction between users and their home environment.

Currently, there are two main themes which are being followed in the development of activity recognition based on object interactions detected by ambient sensors (Philipose *et al.*, 2004). First, there are approaches which emphasise the use of empirical data to model activities in an effort to recognise future instances of activities. Machine learning techniques such as naïve Bayes (Tapia *et al.*, 2004), hidden Markov models (HMM) (Logan *et al.*, 2007; Patterson *et al.*, 2005) and conditional random fields (CRF) (Hu and Yang, 2008; Vail *et al.*, 2007) to name but a few, are widely deployed in this area. In their recent work, Rashidi *et al.* (2011) proposed an unsupervised method to discover unlabelled activity classes and recognise activity instances by learning patterns from the frequency of occurrence of activities. One main disadvantage of this approach was that performance was heavily dependent on the availability of large amounts of observational data. Observational data is, however, extremely expensive for real world applications, especially within the realms of independent living for older people.

Second, there are approaches which involve the development of activity recognition focused on the use of domain knowledge. Advances in web technologies have encouraged research in this direction. Perkowitz *et al.* (2004) built a probabilistic model of an activity by mining the definitions of the activity from the web. In the concepts of evidence theory and ontologies, Hong *et al.* (2009) developed evidential ontology models to represent relations of sensors, objects and activities quantitatively in addition to modelling uncertainty coherent with activities. In Chen *et al.* (2012) semantic web techniques were employed for explicit context and activity modelling while semantic reasoning and classification were used for activity inferencing. In their recent work Gu *et al.* (2010) advanced the concept of emerging patterns into contrast patterns to identify frequent sensor sequences associated with an activity for activity recognition. Different from the aforementioned approaches that address recognising the activity for each segment of sensors manually extracted from a sensor stream, the work is somewhat advanced beyond this. Information related to object relevance is used to segment a sensor sequence and detect the boundary of two adjacent activities.

The current work proposed within this paper falls within the knowledge driven category. It aims to partition a sensor stream into segments and recognise the activity associated with each segment. Generic activity models with regards to related sensors are built from domain knowledge of interactions with household objects mined from many sources, for example online concepts and instruction books and are then subsequently realised into the setting of an environment. Distinct from other approaches, our method only requires loosely coupled information relating to sensor interactions for the purposes of building an activity model.

2.2 Time series data analysis

Initially we consider the scenario of a streaming time series V with a sequence of data $v_1, v_2, \ldots$, where new data are continuously appended as time progresses. Data may be categorical or numerical and range from one to multiple dimensions as determined by the number of variables used for observation (Hong *et al.*, 2009; Last *et al.*, 2004). Formally a streaming time series V is recorded as an ordered set of (data, timestamp) pairs, where the order is governed by the timestamps.

Different kinds of sequence analysis may be used to measure the similarities in and between time series data containing rich contextual information observed over a period of time. The fundamental approach of sequence analysis is to construct metrics and similarity measures pertaining to the attributes of pairs of sequences. A number of approaches to this problem exist for example dynamic time warping (DTW) and neighbourhood counting matrix (NCM).

Distance measures are extensively used in finding the similarity/dissimilarity between two time series. DTW is a popular approach and is based on a dynamic sequence alignment technique (Keogh and Ratanamahatana, 2005). This approach finds all possible paths and selects the one that yields a minimum distance between the two time series using a distance matrix. NCM (Wang and Dubitzky, 2005) is another approach. A contextual probability counts the impact of neighbourhoods by an intuition. If we assume a two-dimensional plane that accommodates several data points, then for a data point p, the set N represents all its neighbourhoods. It is obvious that data points closer to p should be included in more of N, than points that are not close to p. To use NCM on two data sequences, all common neighbourhoods, i.e. neighbourhood combinatorics, are counted. NCM is applicable to both nominal and ordinal attributes and can be implemented in a uniform way. To measure sequence similarity, NCM counts all common subsequences (ACS) as all common sequence-based neighbourhoods (Wang, 2007).

In addition to sequence analysis approaches, some form of detection framework that identifies events arising from sequential data is required. Generally, a time series data stream is the record of direct observations of an event or set of events over time. Data analysis approaches are an important process to extract high-level knowledge about the events from the data stream, which can then be understood and used as the basis for further automated service delivery. It is usually assumed that the raw data sequence is processed to generate a sequence of events, from which patterns of the problem can be mined. Therefore, one of the many interests in the data analysis relates to the derivation of an event sequence from a raw data stream.

Consider a dynamic problem whose behaviour changes enough over time to be considered as a significant change. Each change can therefore be described as an event. An example from the biomedical signal processing domain is the change in a heartbeat as evidenced by the electrocardiogram in the P, QRS and T waves (Last *et al.*, 2004). Identifying the time points at which the changes, i.e. events, occur is referred to as the change point detection problem.

Activity monitoring is an applied problem related to time series data analysis. State-change sensors attached to objects can unobtrusively observe the state change of an object. An example is a contact sensor attached to the door of a fridge. Each time when the fridge door is opened or closed, the sensor's state is recorded as changing from "on" to "off" or vice versa. Object interactions in turn reflect activity occurrences. For example, it is obvious that having lifted the telephone and dialled numbers indicates that a person is using the telephone.

This research inspired from DTW and NCM, proposed two similarity measures for comparing timely ordered sensor sequences and sensor models of activities. The special attention was paid to the dynamic of sensor sequences. Our approach to activity detection and recognition can handle ambiguity caused by multiple ways of undertaking an activity and decomposition of interwoven activities.

3. Activity detection and recognition

The aim of the current study was to develop an approach to detect and recognise activities from sensor activation events with the assistance of domain knowledge in terms of the involvement of object interactions during the undertaking of activities.

3.1 Sensor representation of activity

State-change sensors attached to objects can unobtrusively observe the state change of an object. An activity can be conceptually modelled by a set of sensors installed on objects. Taking medicine involves, for example, picking up a glass, filling the glass with tap water, removing a pill from a medicine box (and swallowing it with water). If three sensors s_1, s_2 and s_3 have been attached to the glass, tap and medicine box, the set $\{s_1, s_2, s_3\}$ is considered a representation of the activity. Uncertainty is present in this way of modelling activities. Consider preparing a cup of tea through taking a cup from the cupboard, filling the kettle with tap water, boiling the water, taking a tea bag and dropping it into the cup, lifting the kettle and pouring hot water into the cup and finally adding milk and sugar. Assume that of all objects have been assigned a sensor. The set of six sensors including cup, tap, kettle, tea, milk and sugar represent the complete model of the activity. Milk and sugar may or not be preferred and may vary from person to person. In addition, the order of sensor activations is not fixed.

3.2 Activity modelling

In smart homes, activity monitoring incorporates a variety of contexts such as locations, sensors, objects and activities. For the purpose of this study we particularly focus on three contextual domains: sensors, objects and activities. In the following paragraphs we formally describe these domains and their relationships with the following definitions.

Definition 1. The *frame* Θ is a set of mutually exclusive and exhaustive context values that the problem domain can hold.

In addition, a frame is the representation of a contextual domain. The frames for the three concerned domains are as follows:

$$\begin{aligned}\text{Sensor frame } \Theta_S &= \{s_i : 1 \le i \le N_S\},\\ \text{Object frame } \Theta_O &= \{o_\mathrm{j} : 1 \le j \le N_O\},\\ \text{Activity frame } \Theta_A &= \{a_p : 1 \le p \le N_A\}\end{aligned}$$

where N_S, N_O and N_A are positive integers, s_i is an installed sensor, o_j is an object onto which a sensor in Θ_S has been attached, and a_p is an activity being monitored. Within a sensing environment frame Θ_A represents a set of activities being monitored. Θ_O represents a frame which contains the objects involved in performing the activities in Θ_A. Θ_S is a frame which represents a set of sensors attached to the objects in Θ_O and which are being observed for state changes.

Interactions with objects are detected through sensor activation events. Activities are carried out through a series of object interactions. The relations of the three domains can be described by multivalued mappings.

Definition 2. Let Θ_E and Θ_H be two frames. A *multivalued mapping* Γ is a mapping function from Θ_E to 2^{Θ}_H, which assigns to each element e_i of Θ_E a subset $\Gamma\,(e_i)$ of Θ_H:

$$\Gamma(e_i) = h_j, \quad e_i \in \Theta_E, \quad i = 1, \ldots, |\Theta_E|, \quad h_j \subseteq \Theta_H, \quad j = 1, \ldots, 2^{|\Theta_H|}.$$

Concerning the relations of the three domains there exist four main multivalued mappings as presented in Table I. For simplicity, Θ_S, Θ_O and Θ_A are written as S, O and A, respectively.

A sensor is attached to one and only one object. The mapping between frame S and O is one to one. An activity may involve interactions with a series of objects and the interaction with an object may be required in several activities. Therefore, the mapping between frame O and A is one to many. The mappings between frame S and A is also one to many.

Based on the aforementioned relations, it can be derived that an activity can be conceptually modelled by a set of sensors attached to the objects being involved in the activity.

Definition 3. The *generic model GM* of an activity is a set of sensors, the activations of which are concerned with the undertaking of the activity:

$$GM = \{s_i : s_i \in S, \ 1 \le i \le N_s\} \ \text{for} \ a_j \in A.$$

It is obvious that a generic model is a subset of S, $GM \subseteq S$.

3.3 Formal description of the problem

This research studies the problem of discovering hidden ADL events from temporally ordered sensor activation observations. Therefore, the purpose of analysing sensor observations is to partition a sensor stream into segments each of which corresponds with an ADL having taken place.

Let S be the set of sensors installed in a smart home for ADL monitoring tasks, $S = \{s_i$: i is an integer, $1 \le i \le N_S\}$. A is the set of activities that the system is monitoring, $A = \{a_j$: j is an integer, $1 \le j \le N_A\}$. Over a period of time, a sensor sequence SQ is collected, $SQ = \{sq\,(t)$: $sq\,(t) \in S$, t is a time variable, $1 \le t \le T\}$. If we consider each time at which a sensor recording takes place as a data point, SQ can be represented as $SQ = \{s_l$: l is an integer, $1 \le l \le L$, $s_j \in S\}$.

The problem is two-fold. First it is to partition the sensor events in SQ into segments, so we have a new set SG, $SG = \{sg_k$: k is an integer, $1 \le k \le K\}$ satisfying:

(1) $sg_k \subseteq SQ$

(2) $\cup_{k=1}^{K} sg_k = SQ$

(3) $sg_p \cap sg_q = \phi, \quad \text{for} \ p \ne q, \ 1 \le p, q \le K.$

Second it is to deduce the activity set AQ, $AQ = \{aq_k$: $aq_k \in A$, k is an integer, $1 \le k \le K\}$, from SG (the segmented SQ). SG and AQ maintain a one to one mapping relationship, that is $sg_k \rightarrow aq_k$ for $1 \le k \le K$.

Mapping	Function representation
Γ: $S \rightarrow O$	$\Gamma(s_i) = o_j, \ s_i \in S, \ o_j \in O, \ 1 \le i \le N_s, \ 1 \le j \le N_o$
Γ: $O \rightarrow A$	$\Gamma(o_i) = A_j, \ o_i \in O, \ A_j \subseteq A$ and $A_j \ne \phi, \ 1 \le i \le N_o, \ 1 \le j \le 2^{N_A} - 1$
Γ: $S \rightarrow A$	$\Gamma(s_i) = A_j, \ s_i \in S, \ A_j \subseteq A$ and $A_j \ne \phi, \ 1 \le i \le N_s, \ 1 \le j \le 2^{N_A} - 1$
Γ: $A \rightarrow S$	$\Gamma(a_i) = S_j, \ a_i \in A, \ S_j \subseteq S$ and $S_j \ne \phi, \ 1 \le i \le N_A, \ 1 \le j \le 2^{N_s} - 1$

Table I. Main mappings of frames S, O and A

4. Activity similarity

The first task of automatic sensor analysis is to partition a sensor stream into segments. Each segment is a specific description of an activity in sensor activation events, where the sensors are ordered by the times recorded and some sensors appear several times at different time points.

4.1 Comparing sensor sequences of activity

As previously described, in a sensorised environment an activity is correspondingly reflected by a sensor activation sequence. If the order of sensor appearances is not considered nor the recurrence of any sensor, we can have a generic description of an activity in terms of sensor activations, $GM_j = subS$, $subS \subseteq S$.

To decide if a sensor sequence in a sensor stream can be counted as a segment representing an activity, we compare the sensor sequence with the generic sensor model of the activity. When the order of sensors and repeated sensor events are not considered, it is obvious that a sensor sequence should be the same in terms of sensors involved as the activity model if it does correspond with an activity having taken place.

There are several characteristics with sensor sequences for instances of an activity, including varying sizes of sensor sequences, changeable orders of sensor appearances and irregularly repeated appearances of a sensor.

Two issues, therefore, have to be considered when comparing a sensor segment as an instance of an activity with the generic model of the activity. These are:

(1) the order of sensors in a generic model has no meaning at all, the order of sensors in a specific sequence does; and

(2) a generic model includes a sensor only once, sensors in a specific sequence may appear any number of times.

These can be illustrated by three sensor sequences $sg_1 = \{s_1, s_2, s_3, s_4\}$, $sg_2 = \{s_1, s_3, s_1, s_2\}$, $sg_3 = \{s_1, s_5, s_1, s_2\}$ and a generic activity model $GM_0 = \{s_1, s_2, s_3\}$. The first three elements in sg_1 are exactly the same as the elements in GM_0, however, the last element does not exist in GM_0. sg_2 contains three sensors the same as GM_0, however, they are ordered differently than in GM_0 and sg_2 has one further element given that s_1 appears twice. Finally, sg_3 has two elements the same as GM_0 but one of them appears twice, and it also holds an element not included in GM_0. The question can then be asked, does sg_1 or sg_2 or sg_3 more closely (even equally) match GM_0? To answer the question, we propose a novel similarity measure approach for activities in the format of a sensor sequence.

4.2 Activity similarity measures

Consider distinct characteristics of a generic sensor model of activity and a specific sensor sequence, we define two similarity measures with different views of neighbourhoods in comparing a specific sensor sequence with the generic sensor model of an activity.

Let GM be the generic sensor model of activity a, and sg be a specific sensor sequence, $sg = \{s_j: s_j \in S, j \text{ is an integer}, 1 \leq j \leq L\}$. When measuring how similar the two sets of sensors are, it is a natural approach to consider how many elements are in both sets. The more elements that are in both sets, the more similar the two sets are. Consider that an element in a sensor sequence may appear more than once, any element in the generic model sensor set only exists once. We can, however, obtain two values for the amount of elements that are in a sensor sequence *sg* and the generic model *GM* of an activity *a*. The first value, denoted δ, is the amount of the elements in *sg* that are included in *GM*:

$$\delta = \sum_{j=1}^{|sg|} \phi_j$$

$$\text{where} \quad \phi_j = \begin{cases} 1 & \text{if} \quad s_j \in GM \\ 0 & \text{if} \quad s_j \notin GM \end{cases} \quad \text{for} \quad s_j \in sg.$$

The second value σ is the number of the elements in *GM* that are also in *sg*:

$$\sigma = \sum_{r=1}^{|GM|} \varphi_r$$

$$\text{where} \quad \varphi_r = \begin{cases} 1 & \text{if} \quad s_r \in sg \\ 0 & \text{if} \quad s_r \notin sg \end{cases} \quad \text{for} \quad s_r \in GM.$$

Equivalently $\sigma = |GM \cap sg|$. δ and σ have the relationships:

- $\delta > \sigma$, if any elements from the intersection of *GM* and *sg* exist more than once in *sg*; and
- $\delta = \sigma$, otherwise.

The degree of similarity of *sg* to *GM* are measured in two different ways as follows.

Definition 4. The first similarity measure, denoted *sim*, is the rate at which elements in *GM* are included in *sg* without considering the recurrences of elements in *sg*:

$$sim = \frac{\sigma}{|GM|}.$$

Definition 5. The second similarity measure, denoted $\hat{sim}$ is the rate at which elements in *sg* are also the member of *GM* when taking into account repeated elements in *sg*:

$$\hat{sim} = \frac{\delta}{|sg|}.$$

sim measures the degree of similarity of *sg* to *GM* by looking at the number of the elements of *GM* that appear in *sg* not taking into account the recurrences of elements in *sg*. $\hat{sim}$ measures the degree of similarity of *sg* to *GM* by looking at the amount of the elements of *sg* that are included in *GM* taking into account the recurrences of elements in *sg*.

When the value of *sim* is equal to 1, the sensor sequence is considered to be an exact replica of the activity model. When the value of *sim* is less than 1, however, not 0 and the value of $\hat{sim}$ reaches 1, the sensor sequence may not be a complete replica, however, a good proportion of the activity model.

5. Our approach

The development of our approach relies on two hypotheses:

H1. The more elements of the generic sensor model of an activity that are contained in a sequence of sensor events, then the more likely the sensor sequence indicates the occurrence of the activity.

H2. The more frequently the elements of the generic sensor model of an activity appear in a sequence of sensor events, then the more likely the sensor sequence indicates the occurrence of the activity.

Let the set {A B C D} be the generic sensor model of an activity. With a timely ordered sequence of sensor events $\{A\ A\ B\ E\ C\ F\ A\ D\}$, three segments can be produced each of which, to some extent, reflects the model, or one large segment consisting of the above three smaller segments may be formed. It is obvious that each of the three smaller segments is less likely than the segment combining three smaller ones to address the occurrence of the activity. Figure 1 shows this statement.

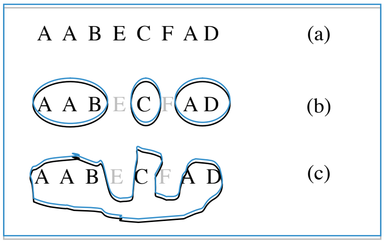
Our solution is presented in three stages. Prior to beginning the process the generic sensor model is prepared for each activity to be monitored with domain knowledge of object interactions and the realisation of sensor installations in the environment. The mapping for a sensor in relation to activities involved can therefore be attained.

Pre-stage. Represent sensor recordings in two sets: the sensor sequence set $SQ = \{s_l: s_l \in S, l \text{ is an integer}, 1 \leq l \leq L\}$, and the sensor time set $SQT = \{t_l: t_l \text{ holds a time value}, l \text{ is an integer}, 1 \leq l \leq L\}$.

First stage. A new set SQA is formed, $SQA = \{A_l: A_l = \Gamma\ (s_l)\}$, $A_l \subseteq A$, l is an integer, $1 \leq l \leq L\}$. The sensor sequence SQ then goes through a round of a grouping process following Rule 1:

Rule 1. Group adjacent sensors s_i *and* s_{i+1} if the corresponding activity sets A_i and A_{i+1} are equal and $t_{i+1} - t_i$ is not greater than the greatest threshold among those of the activities in A_i.

Two sets are output from this stage, $SG = \{sg_m: sg_m \subseteq SQ, m \text{ is an integer}, 1 \leq m \leq L\}$, and $SGA = \{A_m: A_m \in SQA, m \text{ is an integer}, 1 \leq m \leq L\}$. The grouped sensor sequence SG then progresses to the next stage of the process.



Notes: (a) a sensor sequence; (b) direct connection based composition of segmentation; (c) multi-component based composition of segmentation

Figure 1.
Illustration of possible composition of sensors

Second stage. Following Rules 2 and 3, two sensor groups are connected together to form a larger multi-component group:

Rule 2. Group adjacent groups sg_i *and* sg_{i+1} if sim or $s\hat{i}m$ for $a_k \in A_i \cap A_{i+1}$ increases and the activation period is within the threshold of a_k.

Rule 3. Group sg_i *and* sg_j if the corresponding activity sets A_i and A_j are equal and the groups are not distant in terms of activation times between the first sensor in the earlier group and the last sensor in the later group being shorter than the threshold of the activity.

Two sets are output from this stage, $SG' = \{sg_n : sg_n \subseteq, n \, \text{is an integer}, 1 \leq n \leq |SG|\}$, and $SGA' = \{A_n : A_n \subseteq SGA, n \, \text{is an interger}, 1 \leq n \leq |SGA|\}$. The newly grouped sensor sequence SG' then progresses to the third and final stage of the process.

Third stage. The activity is identified for each group in SG' by the following rules.

Rule 4. Identify the group sg_i as the activity if the corresponding activity set A_i contains only one element and either sim or $s\hat{i}m$ is greater than 0.5, otherwise as unknown.

Rule 5. Identify the group sg_i as the activity with the largest $sim > 0.5$ if A_i contains more than one element, otherwise as unknown.

The final outputs are two sets, $SG' = \{sg_n : sg_n \subseteq SQ, n \, \text{is an interger}, 1 \leq n \leq K\}, AQ = \{a_n : a_n \in A, n \, \text{is an integer}, 1 \leq n \leq K\}$.

6. Experimental evaluation

We evaluated the performance of our approach for activity detection and recognition on a real data set collected by van Kasteren *et al.* (2008) that is publicly available. In this Section, we present the results of these evaluations.

6.1 Dataset

The data used in our experiments was collected from sensors installed in a three-room apartment occupied by a single 26-year-old male. Contact switch sensors were used to monitor the open/close status of doors within the apartment, including the doors of the apartment, bedroom, toilet, bathroom, fridge, freezer, microwave, dishwasher, washing machine and cupboards of cups, plates, pans and groceries. A tilt sensor was attached to the toilet flush handle to observe the use of the toilet flush. The seven activities under observation, indicated in Table II, were monitored through 14 sensors detailed in Table III. Any period of time at which no activity took place was labelled "Idle". Over a period of 28 days, 1,230 sensor events in the format as depicted in Table IV were recorded. The annotation was provided to declare 245 activities along with 59 idle periods.

1	Leave house
2	Use toilet
3	Take shower
4	Go to bed
5	Prepare breakfast
6	Prepare dinner
7	Prepare drink

Table II. Activities monitored in the dataset

Sensor alias	Description	Activities monitored
A	Detects if the microwave is opened	5, 6
B	Detects if the bathroom door is opened	3
C	Detects if the toilet door is opened	2
D	Detects if a cup is removed from the cup cupboard	7
E	Detects if the fridge door is opened	5, 6, 7
F	Detects if a plate is removed from the plate cupboard	5, 6
G	Detects if the front door is opened	1
H	Detects if the dishwasher door is opened	–
I	Detects if the toilet is flushed	2
J	Detects if the freezer is opened	5, 6, 7
K	Detects if a pan is removed from the pan cupboard	6
L	Detects if the washing machine door is opened	–
M	Detects if dry food is taken from the grocery cupboard	5, 6
N	Detects if the bedroom door is opened	4

Notes: Sensor alias: A – microwave door sensor; B – hall-bathroom door sensor; C – hall-toilet door sensor; D – cup cupboard door sensor; E – fridge door sensor; F – plate cupboard door sensor; G – front door sensor; H – dishwasher door sensor; I – Toilet flush sensor J – freezer door sensor; K – pan cupboard door sensor; L – washing machine door sensor; M – grocery cupboard door sensor; N – hall-bedroom door sensor

Table III.
Summary of the 14 sensors used within the environment

25 February 2008 10:10:15	25 February 2008 10:10:16	B
25 February 2008 10:10:20	25 February 2008 17:00:28	B
25 February 2008 10:19:28	25 February 2008 10:19:32	G
25 February 2008 16:54:37	25 February 2008 16:54:41	G
25 February 2008 16:59:04	25 February 2008 16:59:11	E
25 February 2008 16:59:14	25 February 2008 16:59:42	M
25 February 2008 16:59:48	25 February 2008 16:59:51	E
25 February 2008 17:00:33	25 February 2008 17:00:37	C
25 February 2008 17:01:22	25 February 2008 17:01:23	I

Table IV.
Example of extraction of the sensor dataset

Analysis of the data set unveiled that the activity of “Go to bed” often interleaved with “Use toilet”. In addition, it was found that the activity “Prepare dinner” may interleave with the activity “Use toilet”.

6.2 Generic activity models

With the knowledge of how sensors were installed and distributed in the house, the activities which were monitored, and in which room an activity was performed, six sensor sets were generated representing generic models of seven activities. Let a sensor be represented by its alias. {B} is the model of activity “Take shower”, {C, I} “Use toilet”, {N} “Go to bed”, {G} “Leave house”, {D, E} “Prepare drink”, {A, E, F, J, K, M} “Prepare meal”. Here “Prepare breakfast” and “Prepare dinner” share the same model as they involve the same set of object interactions. These two activities are distinguished by time of taking place, i.e. "Prepare breakfast" takes place in the morning hours and "Prepare dinner" takes place in the afternoon or evening hours of the day.

6.3 Results

A total of 1,230 sensor event tuples were presented to the activity detection process and partitioned into segments and subsequently identified as the activities.

The purpose of sensor segmentation is to detect and identify the activity when it happens. An annotated activity contains a set of sensors. There are three scenarios concerning activity detection and identification on sensor streams:

(1) An activity is detected and recognised correctly for the whole set of the sensors.

(2) An activity is detected and recognised correctly for a partial set of the sensors.

(3) An activity is not detected at all for the set of the sensors.

First we measured the accuracy of activity detection with the metrics: precision (*PR*) and recall (*RE*) defined as follows:

$$PR = \frac{TP}{TP + FP} \quad \text{and} \quad RE = \frac{TP}{TP + FN}$$

where *TP* (true positive) is the total number of instances that the approach has detected correctly as of the correct activity class *C*; *FP* (false positive) is the total number of instances that the approach has detected incorrectly as belonging to an activity class *C*; and *FN* (false negative) is the total number of instances that the approach has failed to detect as belonging to an activity class *C*.

We then distinguished what percentage of an activity in terms of sensors has been correctly detected and recognised. We refer to this measurement as Percentage Accuracy (*PA*), formulated as follows:

$$PA = \frac{1}{N_C} \sum_{i=1}^{N_C} \frac{\sum_{j=1}^{N_i} [inferred_C(j) = true_C(j)]}{N_i}$$

where $[inferred_C(j) = true_C(j)]$ is a binary indicator: 1 when it is true and 0 when it is false; N_i is the total number of sensors involved in activity instance *i* of class *C* and N_C is the total number of activity instances of class *C*.

It is worth to note that we define that an activity instance is detected and recognised correctly when only a partial set of the sensors involved has been identified. Otherwise, it is considered as not having been detected. The results are depicted in Table V.

One specific issue that our approach addresses is detecting interleaved activities by separating mixed sensor events. We measured the performance of our approach with 17 interleaved activities within the dataset, including 15 of 24 "Go to bed" and 2 of 10 "Prepare dinner" both of which were interleaved with "Use toilet". Table VI details the results of *PR*, *RE* and *PA* for the group of the interleaved activities of a class and the group of the non-interleaved activities of the same class.

6.4 Discussion

The dataset was collected from a real living environment. There were some minor interruptions to activity performance when the subject had to record information for annotation. Otherwise the subject performed the daily activities in a normal everyday manner. In the sense of evaluating the approach's robustness and scalability, the dataset is better than datasets collected under experimental conditions, such as subjects performing a list of instructed activities within a simulated environment. The seven activities monitored are considerably downscaled compared with all the activities that can normally be performed during daily living. It is considered that 14 sensors are not sufficient to cover the entire environment. Therefore, the dataset may be considered to be less complex than in normal instances of everyday life.

7. Conclusions

State-change sensors provide an unobtrusive means to monitor object-based ADLs in home environments. Accurate recognition of activities is important, however, largely depends on deriving the correct activity sequences from raw sensor data. We proposed novel similarity measures on sensor sequences and algorithms to detect an activity occurrence induced from a segment of sensor data, by employing concepts such as ACS and DTW within a range of similarity measures.

The first set of evaluations has been conducted on a dataset collected from a real living environment. Initial results are encouraging and achieved above 85 per cent accuracy in detecting activities. The approach has achieved equally good performance in detecting non-interleaved and interleaved activities. In the future we would like to test the scalability of our approach by considering datasets containing increased classes of activities in addition to having multiple types and various amounts of sensors. A comparison with other existing activity detection approaches is also planned for future work. At present the approach targets activity monitoring in single occupancy environments. We are going to investigate how the approach may be extended to support multi-occupancy environments. This type of technology

Table V.
Results of activity accuracy following evaluation with dataset

Activity	*PR* (%)	*RE* (%)	*PA* (%)
1	100	97.06	97.06
2	94.95	82.46	70.72
3	84.00	91.30	91.30
4	95.83	95.83	85.63
5	95.24	100	99.00
6	100	100	100
7	100	60	59.17
Average	95.72	89.52	86.13

Table VI.
Results for interleaved and non-interleaved activities

	Interleaved group			Non-interleaved group		
Activity alias	*PR* (%)	*RE* (%)	*PA* (%)	*PR* (%)	*RE* (%)	*PA* (%)
4	94.74	100	86.39	100	83.33	83.33
6	100	100	100	100	100	100

deployment has the potential to support community based healthcare and support the elderly at home at the very least delaying admission to institutional care if not avoiding it totally.

References

Chen, L., Nugent, C.D. and Wang, H. (2012), "A knowledge-driven approach to activity recognition in smart homes", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 24 No. 6.

Gu, T., Chen, S., Tao, X. and Lu, J. (2010), "An unsupervised approach to activity recognition and segmentation based on object-use fingerprints", *Data & Knowledge Engineering*, Vol. 69, pp. 533-44.

Hong, X., Nugent, C., Mulvenna, M., McClean, S., Scotney, B. and Devlin, S. (2009), "Evidential fusion of sensor data for activity recognition in smart homes", *Pervasive and Mobile Computing*, Vol. 5 No. 3, pp. 236-52.

Hu, D.H. and Yang, Q. (2008), "CIGAR: concurrent and interleaving goal and activity recognition", *Proc. 23rd AAAI*.

Katz, S. (1983), "Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living", *Journal of the American Geriatrics Society*, Vol. 31 No. 12, pp. 721-26.

Keogh, E. and Ratanamahatana, C.A. (2005), "Exact indexing of dynamic time warping", *Knowledge and Information Systems*, Vol. 7 No. 3, pp. 358-86.

Last, T., Nugent, C.D. and Owens, F.J. (2004), "Multi-component based cross correlation beat detection in electrocardiogram analysis", *BioMedical Engineering Online*, Vol. 3 No. 26, pp. 14.

Lawton, M.P. and Brody, E.M. (1969), "Assessment of older people: self-maintaining and instrumental activities of daily living", *Gerontologist*, Vol. 9, pp. 179-86.

Logan, B., Healey, J., Philipose, M., Tapia, E.M. and Intille, S. (2007), "A long-term evaluation of sensing modalities for activity recognition", *Proc. Ubicomp 2007*, pp. 483-500.

Patterson, D., Fox, D., Kautz, H. and Philipose, M. (2005), "Fine-grained activity recognition by aggregating abstract object usage", *Proc. ISWC 2005*, pp. 44-51.

Perkowitz, M., Philipose, M., Fishkin, K. and Patterson, D.J. (2004), "Mining models of human activities from the web", *Proc. WWW 2004*, pp. 573-82.

Philipose, M., Fishkin, K., Perkowitz, M., Patterson, D., Fox, D., Kautz, H. and Hahnel, D. (2004), "Inferring activities from interactions with objects", *IEEE Pervasive Computing*, Vol. 3, pp. 50-57.

Rashidi, P., Cook, D.J., Holder, L.B. and Schmitter-Edgecombe, M. (2011), "Discovering activities to recognize and track in a smart environment", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 23 No. 4, pp. 527-39.

Tapia, E.M., Intille, S.S. and Larson, K. (2004), "Activity recognition in the home using simple and ubiquitous sensors", *Proc. 2nd Int. Conf. Pervasive Computing, IEEE 2004, LNCS 3001*, pp. 158-75.

Vail, D.L., Weloso, M.M. and Lafferty, J.D. (2007), "Conditional random fields for activity recognition", *Proc. AAMAS 2007*.

van Kasteren, T., Noulas, A., Englebienne, G. and Kröse, B. (2008), "Accurate activity recognition in home setting", *Proc. 10th Int. Conf. Ubiquitous Computing, ACM 2008*, pp. 1-9.

Wang, H. (2007), "All common subsequences", *Proc. IJCAI-07*, pp. 635-40.

Wang, H. and Dubitzky, W. (2005), "A flexible and robust similarity measure based on contextual probability", *Proc. IJCAI'05*, pp. 27-32.

Wilson, D. and Atkeson, C. (2005), "Simultaneous tracking and activity recognition (star) using many anonymous binary sensors", *Proc. 3rd Int. Con. Pervasive Computing, IEEE 2005*, pp. 62-79.

Corresponding author
Xin Hong can be contacted at: x.hong@qub.ac.uk