Subscribing to fuzzy temporal aggregation of heterogeneous sensor streams in real-time distributed environments

J. Medina*,[†], L. Martínez and M. Espinilla

Department of Computer Science, University of Jaén Campus Las Lagunillas, s/n, 23071. Jaén, Spain

SUMMARY

Because of the deployment of heterogeneous sensors in intelligent environments, the fusion and information processing means an arduous and complex process. The data fusion of sensors and the design of processing information in real time are key aspects in order to generate feasible solutions. In order to shed light on this context, we present an approach for distributing and processing heterogeneous data based on a representation with fuzzy linguistic terms. In this way, the heterogeneous data from sensor streams are computed and summarized based on fuzzy temporal aggregations ubiquitously within mobile and ambient devices. This innovative approach provides an intuitive linguistic representation of mobile and ambient sensors as well as implies a drastic reduction of the communication burden. In order to provide high scalability in network communication, the information from sensor is spread under the publication-subscription paradigm, where subscribers receive asynchronous events when the aggregation degree of the linguistic terms overcomes a threshold (*alpha*-cut). Finally, in order to illustrate the usefulness and effectiveness of our proposal, we present the results of the fuzzy temporal aggregation of sensor streams with *alpha*-cut subscriptions in a case study where an inhabitants performs an daily activities in an intelligent environment. Copyright © 2016 John Wiley & Sons, Ltd.

Received 15 December 2015; Revised 20 September 2016; Accepted 9 October 2016

KEY WORDS: heterogeneous sensors; intelligent environments; information fusion; fuzzy temporal aggregation; linguistic terms

1. INTRODUCTION

Intelligent environments are interactive spaces where technological devices are adapted to solve everyday problems of the people. They are developed under networks of physical objects, so Internet of Things (IoT) [1] has awhile arisen as a new paradigm where ambient intelligence [2] and ubiquitous computing [3] converge [4] to provide connected smart things within intelligent environments.

These recent paradigms properly locate the information processing in the everyday objects, but new challenges are needed in order to develop systems for expediting the communication between humans and devices. Following, the paradigms, which have been developed significant advances in different areas of knowledge in this context, are discussed.

A first drawback, whose intelligent environments deal with, is the *data fusion from heterogeneous sensors* [5]. Currently, a wide range of heterogeneous sensors are deployed in intelligent environments to collect multiple and heterogeneous data from mobile, wearables, and ambient devices [6]. The mobile devices integrate user-interaction data (movement, location, etc.) by means of multiple sensors and mobile communication protocols (bluetooth, NFC, Wi-Fi, etc.). Meanwhile, ambient devices are collecting information from our environment using low-power protocols

^{*}Correspondence to: J. Medina, Department of Computer Science, University of Jaén, Jaén, Spain.

[†]E-mail: jmquero@ujaen.es

(Z-Wave, ZigBee, and bluetooth) in light sensors, or WLAN (6LoWPAN) protocols in vision or audio sensors.

Because of this diversity, a second key aspect is the *models and structures of knowledge representation*. They are required by information processing of sensors to generate richer and higher-level information [7]. Several approaches have been proposed to manage these issues, in which the efforts have been focused on providing general or ad hoc models. In respect of the standard structural models, it is noteworthy SensorML [8], developed by Open Geospatial Consortium that includes geolocation or discovering by means of XML schema. Concerning the semantic annotating for sensors, it is relevant the W3C Semantic Sensor Networks specifications [9, 10]. On data modeling, the development of ontologies have been adapted to (i) requirements of a particular domains [11–15] and (ii) general models to provide scientist interoperability [16] and enterprise interoperability [17].

Furthermore, it is essential to consider a third key aspect that is *the distributed information processing of sensors*. The adequate distribution of services in ambient environments is crucial to provide sensitivity to real time [18] when the information processing is distributed in different central processing units [19, 20]. In this area, the *middleware* highlights as an infrastructure, in which the sensor streams from ambient and mobile devices are connected by remote services in distributed environments. Specifically, the Publish-Subscribe paradigm has been demonstrated to be adequate for IoT [21], where subscribers receive asynchronous messages from publisher, highlighting the implementations in the standards: (i) Data Distribution Service for Real-time Systems by Object Management Group[‡], (ii) MQ Telemetry Transport [22], a lightweight messaging protocol for sensors and devices formalized recently as ISO standard[§], or (iii) Constrained Application Protocol [23], a request/response transfer protocol for constrained sensors and networks over HTTP.

There are several relevant open middlewares that provide sturdy solutions to ubiquitous and ambient contexts. Among them, it is noteworthy ZeroC Ice [24], an object-oriented distributed computing tool with support for several languages and platforms, Global Sensor Network (GSN) that applies sliding window management [25] in changing data stream [26] and, finally, OpenIot, which increases the characteristics of GSN adding top-k subscriptions over sliding windows (top-k/w subscriptions) [27] and semantic open-linked data techniques.

In the context of this state of the art, the aim of this paper is to provide a straightforward approach that processes heterogeneous sensor streams using linguistic terms in real-time intelligent environments. The proposed data modeling is based on the fuzzy linguistic approach and fuzzy logic [28] because it has provided successful results in developing intelligent systems using the data provided by the sensors [11, 29–33] highlighting the development of renewable energy and energy-saving systems [34–37].

The managing of uncertainty and vagueness is key in intelligent systems, such as activity recognition [38], to obtain high performance and results [39]. In this way, sensor-based activity recognition is a relevant research topic that integrates pervasive and mobile computing [40, 41], context-aware computing [42–44], and ambient assisted living [45, 46].

In order to illustrate the usefulness and effectiveness of our proposal, we present the results in terms of reduction of the communication burden of the fuzzy temporal aggregation of sensor streams with alpha-cut subscriptions in a case study where an inhabitants performs an daily activities in an intelligent environment. This case study has been chosen because they generate much interest in recent years within the research community, which aims to recognize the actions and goals of one or more people within the environment based on the changes in the environment that are collected by the sensors [47].

The remainder of the paper is structured as follows: in Section 2, the proposed approach for distributing and processing of data based on a representation with fuzzy linguistic terms is presented; in Section 3, a case study of daily living activities in a intelligent environment is described where the proposal of fuzzy processing and distributing of data is evaluated. Finally, in Section 4, conclusions and future works are pointed out.

[‡]http://www.omg.org/spec/DDS/1.2/.

http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/mqtt-v3.1.1.pdf.

2. APPROACH FOR DISTRIBUTING AND PROCESSING OF HETEROGENOUS DATA BASED ON A REPRESENTATION WITH FUZZY LINGUISTIC TERMS

In this section, a light scalable approach in which the heterogeneous sensors are modeled by linguistic terms to dynamic subscribers in real-time is presented. To do so, the main features and advantages of our proposal are indicated, where the fuzzy linguistic term and fuzzy aggregation operations for processing sensor streams are presented in detail. Finally, the proposed fuzzy model is focused on defining a α -subscription of linguistic terms in distributed environments.

The architecture for distributing and processing of heterogeneous data based on a representation with fuzzy linguistic terms are illustrated in Figure 1. We show how the raw data streams are translated to linguistic terms and the degree is used to filter relevant messages to subscribers based on α -cuts. In this way, each subscriber is able to obtain expressive summaries of sensor stream and to reduce the communication burden dynamically using α -subscription. Following, the proposed architecture is described with its main features.

Our linguistic approach is proposed to facilitate the fusion and description of heterogeneous data from sensors, using linguistic terms that provides an intuitive and interpretable representation without structures or ontologies. Specifically, fuzzy linguistic terms describe the state of sensor stream, such as *movement is high* or *inhabitant is close to kitchen* together with fuzzy linguistic terms, such as, *awhile*, *just now*, or *just now*.

The proposed structure of linguistic terms are related to *protoforms*, which were proposed by Zadeh [48] as useful knowledge model for reasoning [49], summarization [50], and fusion [51] of data under uncertainty. We have started from the classical protoform in the form of *X* is *A* to include a linguistic temporal term generating protoform instances in the shape of: *motion is just now low*. We note the relevance of temporal processing, integrated in this contribution by means of fuzzy logic, which has offered an intuitive and flexible representation of temporal knowledge in several contexts [52–54]. This issue will be further discussed in Section 2.1.

On middleware layer, previous solutions deal with data using heavy computing techniques in desktop computers, delegating light devices to wrappers of raw data. For example, GSN uses MySQL to manage and store the data streams or OpenIot provides subscription services just in server side [55]. Moreover, there are some problems to integrate GSN in light devices within other platforms or languages [56]. Our approach aims to focus the initial information processing in mobile and ambient computers, providing more scalable solutions without persistence of databases.

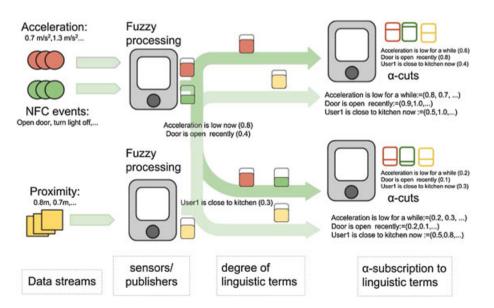


Figure 1. Fuzzy linguistic processing and α -subscriptions.

On subscription services, streaming sensor data under subscription paradigm and temporal windows has become an active research area related to ubiquitous and ambient environments because it enables the real time capabilities in intelligent systems. In these contexts, the processing of data from ambient, mobile, and wearable sensors has been focused on statistical and spectral features [57, 58], which are calculated in short window sizes, such as fixed static window size [59] and dynamic window size [60].

In this work, we propose an approach focused on the flexibility in defining subscriptions to sensors through linguistic terms. Specifically, clients subscribe to sensors defining a *protoform* by means of one linguistic and one temporal a term, for example, *movement is just now high* or *kitchen door has been awhile opened* together with an *alpha*-cut threshold and a minimal time interval to notify subscriptors. Each sensor calculates the fuzzy degrees, which represent a fuzzy aggregate value of the matching of these linguistic terms with the data stream of the sensor, sending a real-time response to subscribers when the degree overcomes the *alpha*-cut. In this way, each sensor processes and distributed information from the data streams in real time using a similar model to top-k/w subscriptions [27] under a fuzzy approach. The α -subscription of linguistic terms is detailed in Section 2.2.

2.1. Fuzzy linguistic terms and fuzzy aggregation of sensor streams

In this section, we describe the use of linguistic terms to describe sensor streams in the proposed approach that uses fuzzy logic to provides the capability for fusion heterogeneous sensor data to achieve high-level inferences about objects and situations [29, 61].

The aim of the fuzzy linguistic processing of this work is relating mobile and ambient sensors to linguistic variables, which are defined by a set of linguistic terms. In the fuzzy logic context, the semantic of the linguistic terms is given by fuzzy sets described by membership functions [28]. Formally, the membership function $\mu_{\tilde{A}}(x)$ describes the membership degree of the elements x of the base set X in the fuzzy set A, $\mu_{\tilde{A}} : X \to [0, 1]$.

A common representation to define a membership function in continuous domains is the trapezoidal function, whose representation is achieved by a 4-tuple TS(a,b,c,d) where b and c indicate the point in which the membership value is 1, with a and d indicating the left and right limits of the definition domain of the membership function outside where its value is 0. A linear function that changes smoothly from 0 to 1 is associated with domain values between (a, b] and (b, d]. The trapezoidal membership functions are defined formally as follows:

$$\mu_{\tilde{A}}(x) = \begin{cases} 0 & if \quad x \leq a \\ \frac{x-a}{b-a} & si \quad x \in (a,b] \\ 1 & if \quad x \in (b,d] \\ \frac{c-x}{c-d} & si \quad x \in (d,c) \\ 0 & if \quad x \geq d \end{cases}$$
(1)

In our approach, each sensor data stream s^j is represented as a set of measures $s^j = \{m_i^j\}$, where each measure is represented by $m_i^j = \{v_i^j, t_i^j\}$; v_i^j represents a value that depends on each sensor, for example, temperature and heart rate; and t_i^j represents the timestamp, the time at which this value, v_i^j , was provided by the device. So, the superscript is related to a given sensor j and the subscript i is related a given term for this sensor s^j .

So, we propose to model data stream from each sensor by means of linguistic terms. To achieve this, we define a self-dependent processing of each component v_i^j and t_i^j and a subsequent method to fuse both linguistic terms.

Firstly, for each sensor data stream s^j , we define a fuzzy linguistic variable, the number of terms, their linguistic terms, and their membership functions, which are all based uniquely on the nature of the sensor. The general process is described as translating the measures to a new fuzzy set $V = \{V_r; r = 1, ..., n\}$, which represents the linguistic terms. V_r is characterized by a membership function $\mu_{\tilde{V}_r}(v_i^j)$, which is interpreted as the degree of membership of an value v_i^j in the fuzzy set V_r for each $v_i^j \in m_i^i \in s^j$. For sake of simplicity, we write V_r instead of $\mu_{\tilde{V}_r}(v_i^j)$.

In Figure 2, an example is illustrated by the linguistic variable *movement* from a mobile accelerometer sensor with two linguistic terms: $V = \{V_1, V_2\} = \{low, high\}$ with two trapezoidal membership functions.

Secondly, we associate fuzzy linguistic terms $T = \{T_k; k = 1, ..., m\}$ with the temporal component of the data stream. The degree of this temporal linguistic term is obtained by the temporal membership function and the temporal period that is defined by the distance $\Delta t_i^j = t_0 - t_i^j$, $t_0 > t_i^j$ from a reference point of time t_0 to the timestamp of the measurement t_i . Each temporal linguistic term T_k relates the timestamp of the measurement t_i^j to a fuzzy set T_k , which is characterized by a membership function $\mu_{\tilde{T}_k}(t_0 - t_i^j)$. For a given temporal linguistic term, we can write T_k instead of $\mu_{\tilde{T}_k}(\Delta t_i^j)$, $\Delta t_i^j = t_0 - t_i^j$.

In Figure 3, an example of temporal linguistic terms, where the universe of discourse is measured as seconds since the current time, is illustrated with two linguistic temporal terms: $T = \{T_1, T_2\}$ ={*just now, awhile*}, which represent a fuzzy time interval up to 5 and 20 s respectively between current time and the timestamp of the measure.

Once, the linguistic information for each sensor has been defined, it is necessary aggregating the sensor streams in a time period. To do so, we propose to compute the relevance of a linguistic value term $V_k(v)$ in a fuzzy temporal term $T_k(t)$ in two steps:

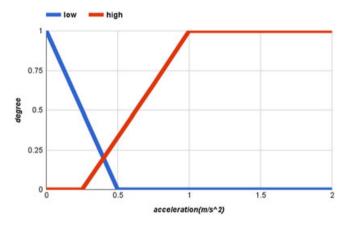


Figure 2. Example of membership functions for the linguistic variable *movement* and the linguistic terms $V = \{V_1, V_2\} = \{low, high\}$ by means of accelerometer data from a mobile sensor measure in m/s².

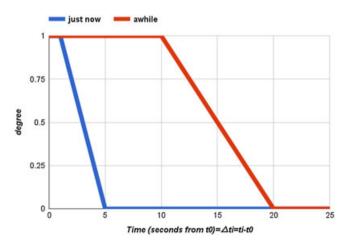


Figure 3. Example of temporal linguistic terms $T = \{T_1, T_2\} = \{just now, awhile\}$.

• Firstly, for each sensor measurement, we define an *t*-norm operation [62] to fuse both fuzzy sets, jointly relating the value and temporal components:

$$V_r \cap T_k(m_i^j) = V_r\left(v_i^j\right) \cap T_k\left(\Delta t_i^j\right) \in [0, 1], m_i^j = \left\{v_i^j, \Delta t_i^j\right\}.$$
(2)

In Eq. (2), the independent terms V_r and T_k are applied to each measure m_i^j in the sensor data stream providing transformation from $\{v_i^j, \Delta t_i^j\}$ to the membership degree of linguistic term $V_r \cap T_k$, for example, $m = \{0.2 \text{ m/s}, 5 \text{ s}\} \rightarrow is$ awhile low (0.8).

• Secondly, the degree of membership of all measures in the sensor data stream are aggregated using the *t*-conorm operator in order to obtain a single degree of fuzzy sets $V_r \cap T_k$ for the given sensor data stream over a period of time:

$$V_r \cup T_k(s^j) = \bigcup_{m_i^j \in s^j} V_r \cap T_k\left(m_i^j\right) \in [0, 1]$$
(3)

The linguistic terms T_k and V_k , which are instanced in the sensor stream s^j , represent a protoform in the shape of: s^j is $T_k V_r$. For example, the protoform instance: *inhabitant motion is awhile low*. Each protoform instances is measured with the degree of fuzzy aggregation from sensor data in accordance with Eq. (3).

In this way, we highlight that protoforms require an accurate assessment to express properly their semantic and granularity. Basically, in sensor streams, we propose aggregate measures depending on the nature of the sensor stream. So, we propose two models of solving Eq. (3) in order to compute the aggregation degree of data streams based on the variable and terms included in the protoform instances:

• *Min-max*. In this first case, the aggregation degree is evaluated in Eq. (3) using the maximal and minimal operators, being $\bigcup = max$, $\cap = min$, which calculated the higher degree with *co-operator=max* of each measured evaluated with the *t-operator=min*. If we replace Eq. (3) with operators $\bigcup = max$, $\cap = min$, we obtain:

It represents the best value of the measures in the time term:

$$V_r \cup T_k(s^j), \bigcup = max, \cap = min \Rightarrow \tag{4}$$

$$V_r \cup T_k(s^j) = \max_{m_i^j \in s^j} \left(\min\left(V_r\left(v_i^j\right), T_k\left(\Delta t_i^j\right) \right) \right) \in [0, 1]$$
(5)

The semantic of this aggregation is useful to aggregate binary open sensors or the best references to other objects, for example, the next protoform instances: *the door is awhile open*, *inhabitant location is awhile close to microwave*, where if a door is opened, or the user has been closed at least once, the degree of linguistic term is raised.

• Fuzzy weighted average. In this second case, the aggregation degree is evaluated in Eq. (3) using $\bigcup = \sum_{n \in \mathbb{N}} (1 - 1) = \sum_{n \in \mathbb$

$$\frac{1}{\sum T_k(\Delta t_i^j)} \tag{6}$$

If we replace Eq. (3) with operators $\bigcup = \frac{1}{\sum T_k(\Delta t_i^j)} \sum_{j \in I_k} (\Delta t_i^j)$.

$$V_r \cup T_k(s^j), \bigcup = \frac{1}{\sum T_k(\Delta t_i^j)} \sum_{j \in \mathcal{N}} (\beta - \lambda) = 0$$
(7)

$$V_r \cup T_k(s^j) = \frac{1}{\sum T_k\left(\Delta t_i^j\right)} \sum_{m_i^j \in s^j} V_r\left(v_i^j\right) \times T_k\left(\Delta t_i^j\right) \in [0, 1]$$
(8)

Copyright © 2016 John Wiley & Sons, Ltd.

This semantic is related to fuzzy weighted average [63], whose range $\in [0, 1]$ represents the frequency degree of the linguist term V_r weighted by the term T_k in the sensor stream.

This fuzzy aggregation operations are suitable to aggregate measures of high sample rates, for example, we arable or mobile sensors, where the degree is related to the presence of relevant measure in the temporal term, as the next protoform instances: the movement is just now high or the heart rate is awhile low.

In this section, we have provide two aggregation operators: *min-max* or *fuzzy weighted average* but other operators could be considered according to the needs.

2.2. α -subscription of linguistic terms

In this section, we detail the proposed approach based on the Real-time Publish-Subscribe model with M2M [64] communication, in which the degrees of aggregation of protoforms from sensors are distributed under subscriptions.

This paradigm focuses the information processing of linguistic terms on mobile or ambient devices where the sensors are integrated. To do so, we propose that each sensor will be represented as a publisher to which clients subscribe to a linguistic terms which configure a protoform and an $\alpha - cut$ threshold to be notified asynchronously.

The proposed approach provides the following three advantages:

- The information processing is distributed in several central processing units, such as mobile and ambient computers. In this way, each device (i) describes the linguistic terms related to their sensors, (ii) computes the aggregation of protoform defined by subscribers, and (iii) publishes degrees to their subscribers using a M2M connection. For example, the mobile devices calculate the degrees of aggregation of protoforms related to each accelerometer sensor notifying in real time to several mobile, desktop, or ambient subscribers. In the same way, each ambient sensor is evaluated by means of remote services that can be deployed in distributed ambient computers.
- The subscription parameters are intuitive, because subscribers merely choose a relevant linguistic terms. In addition, subscribers can defined a fuzzy threshold called $\alpha - cut$, in order to receive just the degree of protoforms that overcomes the $\alpha - cut$. It enables subscribers optimizing the asynchronous events received and reducing sending irrelevant degrees of protoforms in an individual way.
- The subscribers obtain an understanding output; it is a fuzzy degree of aggregation of the linguistic terms relevant to subscribers, facilitating the fusion of interpretable information from heterogeneous sensors.

Following, we present the parameters of the remote services defined by the publisher side and the subscriber side.

Publisher side

- s^{j} that is a label to represent the sensor stream, such as *movement* or *temperature*. Based on the nature of the sensor stream s^j , an $\cap - \cup$ aggregation operator is related, *min-max* or fuzzy weighted average, in order to calculate the degree of the Eq. (3).
- $V^j = \{V_1^j, \dots, V_n^j\}$ that is a set of linguist terms, such as *is high* or *is normal*
- $T^{j} = \{T_{1}^{j}, \dots, T_{n}^{j}\}$ that is a set of temporal linguist terms, such as, *awhile* or *just now*.
- $\cap \cup$ aggregation operator. Subscribers specify the operators to aggregate the date: (i) fuzzy weighted average or (ii) min-max.

Subscriber side The parameters of the service to be subscribed to sensor publisher are

- s^{j} that is the label of the sensor to subscribe.
- $V_r^j \in V$ that is the linguist term from the set V^j .
- T^j_k A temporal linguist terms from the set T^j.
 α cut that is a threshold which have to overcome the degree of aggregation of the linguist terms V^j_r and T^j_k related to Eq. (3).

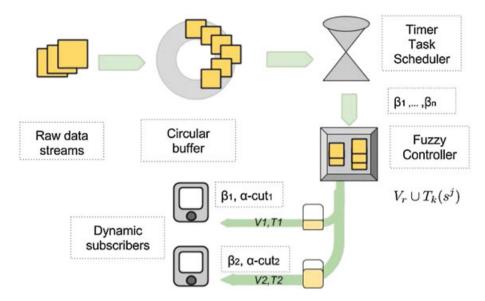


Figure 4. Flow of data in sensors, from raw stream data to publishing to subscribers degrees of linguistic terms in real time.

• β_t that is the minimal time interval in milliseconds that the publisher wait to process and to notify again to the subscriber.

In this way, the information processing of the sensor streams is developed in each publisher by means of four modules responsible, which are illustrated in Figure 4, of collecting and synchronizing the data stream and publishing the fuzzy linguistic processing to subscribers.

- Dynamic collection module. A dynamic collection of data from the sensor stream using circular buffer [65], being the size of its buffer enough to represents the fuzzy temporal terms, allocating a minimal time of data $\Delta t_{max} = max(\Delta t_i^J), \forall T_k \in T, T_k(\Delta t_i^J) \neq 0.$
- Timer task scheduler module. In order to handle the notification time for their subscribers, notifying to them by means of asynchronous events when the interval time between the last sent event and the current time overcomes β_t .
- Dynamic list module. This module can be dynamically subscribed and unsubscribed to using a M2M service.
- Fuzzy controller module. In this module, the aggregating membership degree of the fuzzy linguistic and temporal terms is calculated according to Eq. (3) and the $\cap \cup$ aggregation operator of the subscriber.

3. CASE STUDY

Intelligent systems for smart environments require a proper uncertainty treatment to obtain a high performance and accurate results [39] in relevant topics, such as activity recognition [38, 66], which provides a key successful solution to the aging population [67] helping elderly people to stay with the best quality of life as long as possible in their homes. So, the fuzzy linguistic processing of this work has been focused on describing daily activities from heterogeneous sensor of a complex scene to provide an interpretable, intuitive, and flexible representation that can be analyzed by humans and intelligent computer systems.

In this section, a description of a daily activity in an intelligent environment is presented, which is modeled under our proposed approach for fuzzy linguistic processing and the subscription model.

Firstly, the smart lab and the set of deployed sensors are presented, then the features of the middleware that has been developed to this case study is described. Following, the fuzzy linguistic terms of the set of sensor are indicated and, finally, the results in terms of reduction of the communication burden of the fuzzy temporal aggregation of sensor streams are provided and discussed.



Figure 5. Sensors deployed in the smart lab: accelerometer (BQ Aquarius E5), BLE proximity (iBeacon), vision (PTZ D-link), open/close (Tynetec), lights (Philips Hue), and marked shirts.

The presented scene shows two inhabitant in a smart lab provided by mobile and ambient sensors.

3.1. Smart lab and deployed sensors

The case study was carried out in the smart lab of the Center for Advanced Studies in Information Technology and Communication of University of Jaén[¶].

At the scene, an inhabitant A is resting on the living room while another inhabitant B enters into the lab. Inhabitant B turns the hall lights on and then, he cooks in the dinner using the kitchenware and microwave. Inhabitants A and B meet in the kitchen when inhabitant B goes to drink water and afterwards inhabitant B backs into the room. Finally, the inhabitant B eats his dinner in the living room beside the inhabitant A.

To carry out the aforementioned scene, we have included several ambient and mobile sensors whose objective is describing the activities of users from different perspectives. The aim is receiving a homogeneous linguistic descriptions from a wide heterogeneous range of sensor. To do this, we have included the following sensors from ambient and mobile devices, which are illustrated in Figure 5:

- Vision sensors. Two vision sensors in order to capture the user localization by means of ambient vision devices. To do so, we have integrated a proposal where inhabitants dress shirts with four markers that allow vision sensors to recognize who and where the inhabitants are located, using the open-source project: Minimal library for Augmented Reality applications (Aruco) [68] to detect markers in the shirts of inhabitants.
- Mobile sensors. Inhabitants A and B hand a mobile device that collects several data streams from the next sensor:
 - Accelerometer sensor, which describes the inhabitant motions in the scene.
 - Bluetooth location sensor, combined with the iBeacons, provides closeness to rooms and objects using the Bluetooth Low Energy (BLE) protocol [69].
- Open/close sensors, which are located at the doors of some home appliances (microwave, fridge, etc.) as well as at the doors of kitchen furniture.
- Z-Wave lights. The smart lab integrate remote control of lights by means of Z-Wave protocol [70], which can be turned on/off from the mobile devices of inhabitants.

The plane of the smart lab and the locations of set of sensors are illustrated in Figure 6.

3.2. Middleware deployment

Because each sensor depends on each manufacturer platform, a metalanguage middleware is required in order to deploy the subscription services.

In this case study, we have implemented the model of fuzzy linguistic terms to describe sensor streams, detailed in Section 2.1, in Java code language as well as the α -subscription of linguistic terms described in Section 2.2 using ZeroC Ice. This implementation has been integrated into: (i) Android devices, which monitorize the mobile sensors and (ii) an ambient computer, which monitorizes the ambient sensors.

[¶]http://ceatic.ujaen.es/.

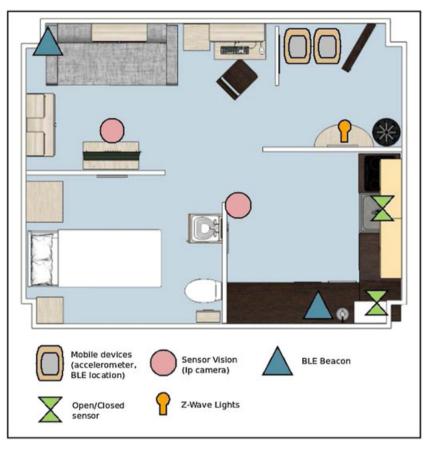


Figure 6. Location of the sensors in the smart lab.

The distribution of real-time data has been implemented from a previous work by the authors [71]. This proposal uses an object-oriented middleware, ZeroC Ice [24], providing the spread of data in real time under the paradigm Data Distribution Service, using subscriptions and sending data through channels of real-time events. It also supports transparent communication for several languages (C++, .NET, Java, Python, Objective-C, Ruby, PHP, and JavaScript) and protocols (TCP, UPD, and SSL).

3.3. Fuzzy linguistic terms of sensors

In this section, we detail the fuzzy linguistic terms related to temporal terms and the mobile and ambient sensor of the smart lab.

In this case study, we propose the use of *protoforms* that are specified by means of variables and linguistic terms. As we have detailed previously, we have started from the classical protoform in the form of *X* is *A*, which represents possibilities constraints. Based on this prototypical form, other extensions result useful, for example, in quantification [72] or in time series [73].

In this work, we propose the use of protoform s^j is $T_k^j V_r$ to represent the linguistic terms V_k^j and the temporal terms T_k^j that describe the sensor stream s^j , for example, the protoform instance: *inhabitant motion is awhile low.* The degree of the protoform instances are computed as the degree of fuzzy aggregation of sensor data related in Eq. (3).

To instance the protoforms, it is necessary to define the variables and linguistic terms related to each sensor stream, following this information is provided for each sensor.

• Accelerometer mobile sensor that measures linear acceleration in m/s^2 and represents the (*A/B*) *inhabitant motion*, with two linguistic terms: *low* and *high*, whose membership functions are illustrated in Figure 7.

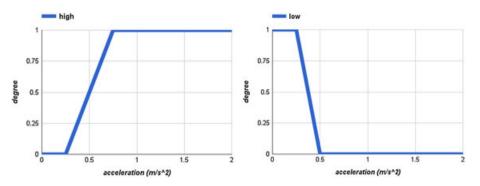


Figure 7. Linguistic terms and membership functions of the accelerometer mobile sensor.

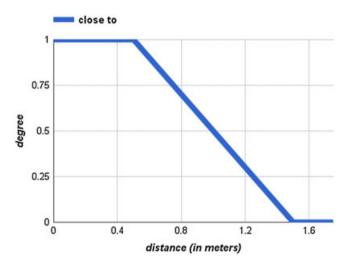


Figure 8. Linguistic terms and membership function of the proximity mobile sensor.

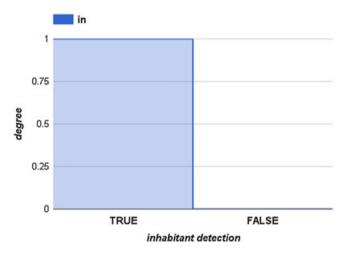


Figure 9. Linguistic terms and membership function of the vision sensor.

• BLE proximity sensor that measures the distance from mobile to a beacon sensor in meters, *m*, and represents *close to (microwave/sofa)* with the linguistic term *close*, whose membership function is illustrated in Figure 8. We note the same linguistic term can be used for the rooms kitchen and living room because of their similar sizes.

- Vision sensor to detect the t-shirt mark of (A/B) inhabitant *true* that represent that (A/B) inhabitant location in (kitchen/living room), whose membership function is illustrated in Figure 9.
- Switch sensor to detect of open/close of the (*microwave/ cupboard*) status with the linguistic term *open*, whose membership function is illustrated in Figure 10.

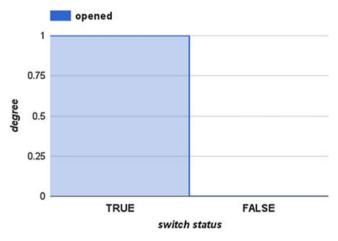


Figure 10. Linguistic terms and membership function of switch sensor.

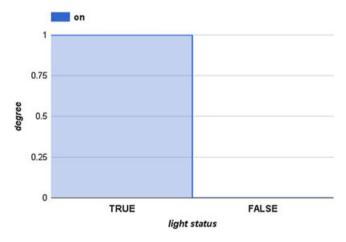


Figure 11. Linguistic terms and membership function of light sensor.

Variable <i>s^j</i>	Term V _r	Terms T_k	$\cap - \cup$
A inhabitant motion	high	awhile	fuzzy weighted average
B inhabitant motion	low	awhile	fuzzy weighted average
A inhabitant location	close to microwave	awhile	fuzzy weighted average
B inhabitant location	close to sofa	awhile	fuzzy weighted average
A inhabitant location	in kitchen	awhile	min-max
B inhabitant location	in kitchen	awhile	min-max
A inhabitant location	in livingroom	awhile	min-max
B inhabitant location	in livingroom	awhile	min-max
hall light	on	just now w	min-max
cupboard	open	just now	min-max
microwave	open	just now	min-max

Table I. Variables, terms, and aggregation operator.

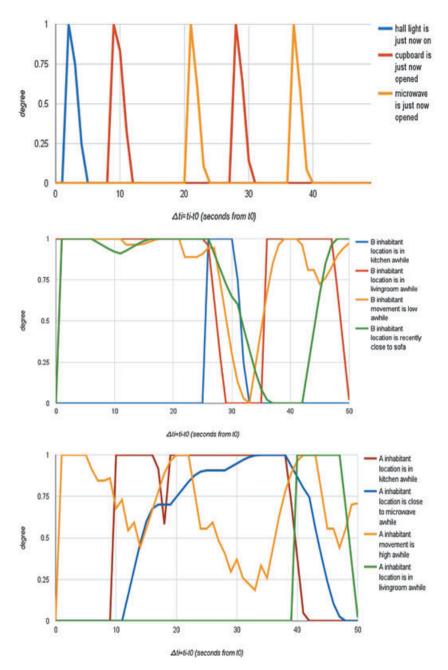


Figure 12. Degree of linguistic terms in protoforms provided by fuzzy processing of sensor streams in real time. (a) Protoforms related to ambient sensors, (b) protoforms related to inhabitant A, and (c) protoforms related to inhabitant B.

• Light (Z-Wave) to detect the status of lights, with the linguistic term *on*, whose membership function is illustrated in Figure 11.

In this case study, two temporal linguistic terms ={*just now, awhile*} has been defined to obtain a flexible representation of fuzzy temporal intervals where the sensor streams are relevant (Figure 3).

3.4. Results and discussion

In this section, the results on fuzzy linguistic processing of protoforms and the events generated by sensor streams in the case scene are exposed and discussed.

Firstly, we have subscribed to several protoforms structured by fuzzy linguistic terms using a subscriptor of one-second time interval $\beta_t = 1000$ and $\alpha - cut = 0$. The linguist terms and the $\cap - \bigcup$ aggregation operators related to the fuzzy linguistic processing are shown in Table I.

Based on the value of sensor streams, the subscriptors received the in real-time evolution of membership degrees related to fuzzy aggregation operation of the linguistic terms. The degrees are represented in a baseline in Figure 12.

On the sensor data, we highlight the sturdiness of integrating two location systems: BLE and vision tracking. Vision tracking provides a prompt response but with losing tracking in some inter frames because of non-frontal captures of shirt markers. Meanwhile, BLE provides a consistent stream locations even thought it presents a delay in adapting the closeness to beacons. Both are complementary to obtain a robust accuracy of the location of inhabitants in real time. In Figure 12C, due to the inhabitant B is opening the door of the smart lab in the hall and he is going to the kitchen between the time interval of [0,10] seconds, we show a delay of presence detection in kitchen, but we receie high movement because he is walking before entering.

The accelerometer sensor by means of linear acceleration of mobile devices present fluctuating data, which are need to be integrated with other sensor data to be useful in human activity descriptions. Ultimately, the ambient binary sensors (Z-Wave light and open/door) provide no-latency reliable data of smart objects. We note how the linguistic transformation binary sensor, where each V_r represents a crisp state such as *hall light is on*, shapes a fuzzy evolution when they are aggregated with temporal terms *hall light is just now on*.

On the fuzzy information processing, we note the relevance of the fuzzy aggregation to soft the degrees of the linguist terms in the time scene. In this way, the fuzzy processing provides (i) the temporal persistence of the crisp binary sensor by means of a temporal linguistic term, such as ambient sensor, (ii) the smoothing of noise or sawtooth values related to high-frequency sample of data, such as motion or BLE closeness, and (iii) the decreasing in the communication burden using time interval subscriptions without losing relevant information.

In addition, the decreasing in the communication burden can be tuned depending on context by adjusting the value of the minimal time interval β_t or the threshold α -cut. In Table II, we compare the number of raw data developed by each sensor with regard to the generated events by computing linguistic protoforms when their degree overcomes the α -cut. In this way, these data represent the ratio reduction over the data/event stream, which we obtain when distributing raw data of sensor versus events of degree of linguistic proforms. In addition, we have included different values of α -cut of subscribers ($\alpha = 0, 0.5, \text{ and } 0.9$) in order to show its influence in the ratio reduction.

We highlight the reduction in sensor with high sample rate, where the fuzzy aggregation process with α -subscriptions is very significant. However, in binary sensors, the temporal persistence of fuzzy linguistic terms increase the numbers of received events, but they provides a membership degree aggregation of linguist terms richer than a single sample.

Sensor	Raw data	$\alpha = 0.0$	$\alpha = 0.5$	$\alpha = 0.9$
Vision sensor for inhabitant (A/B) in (kitchen/	Frames = 860	Events = 200 , ratio reduction = 4.3	Events = 101 , ratio reduction = 8.51	Events = 93, ratio reduction = 9.2
livingroom)				1440 1044040 J.2
Motion sensor for inhabitant (A/B)	Samples = 2013	Events = 100, ratio reduction = 20.13	Events = 60, ratio reduction = 33.55	Events = 40, ratio reduction = 50.33
BLT sensor for inhabitant (A/B)	Samples = 100	Events = 100, ratio reduction = 1	Events = 65, ratio reduction = 1.53	Events = 45, ratio reduction = 2.22
Binary sensors (microwave/ cupboard/light)	Samples = 5	Events = 50, ratio reduction = 0.1	Events = 10, ratio reduction= 0.5	Events = 5, ratio reduction= 1

Table II. Sensor, raw data from sensor, and ratio reduction of α -subscribers.

Copyright © 2016 John Wiley & Sons, Ltd.

4. CONCLUSIONS AND FUTURE WORK

In this work, we propose the use of variables and linguistic terms to represent and process the stream data from heterogeneous sensors. We relate the linguistic terms to protoforms where temporal representation is highlighted. These linguistic expression provide the advantage of interpretability and understanding close to human knowledge over data or semantic structures focused on processing of computers. At the same time, we propose two models of computing the aggregation degree of data stream based on the variable and terms included in the protoform instances.

In addition, the proposed model has been deployed under Real-time Publish-Subscribe model enabling subscriptors define the protoforms, aggregation operators, the *alpha*-cut, and minimal time interval β_t to be notified. In the case study, it provides an encouraged stream reduction from raw data to high rate asynchronous events of $\beta_t=1$ s. The data reduction overcomes up to 1/50 ratio in function of the *alpha*-cut.

The fuzzy linguistic approach presented in this work develop descriptions closed to the language used by human user, but in other contexts, they could need to be complemented with structured data or statical summaries. For example, a heart rate monitoring system could include as input the aggregation degree of the protoform instance *heart rate is awhile high*, but it is naive that the average, maximal or standard deviation of heart rate must not be required in higher layer of information processing.

A useful future approach of this work is developing fusion subscribers to integrate heterogeneous linguistic terms from different sensor that match the assessment of fuzzy IF-THEN rules. This kind of subscribers would generate new high-level knowledge or activate alert processes focusing on the inference of expert knowledge, for example, *IF heart rate is awhile low AND oxygen level is a while scarce THEN send supervisor sms*.

ACKNOWLEDGEMENTS

This contribution has been supported by research projects: UJA2014/06/14 and TIN-2012-31263.

REFERENCES

- Kopetz H. Internet of things. In *Real-time systems*. Springer US: Langford Ln, Kidlington OX5, United Kingdom, 2011; 307–323.
- Zelkha E, Epstein B, Birrell S, Dodsworth C. From devices to ambient intelligence. *Digital Living Room Conference*, Vol. 6, California, USA, 1998.
- 3. Weiser M. The computer for the 21st century. Scientific American 1991; 265(3):94-104.
- 4. Marie P, Desprats T, Chabridon S, Sibilla M. Extending ambient intelligence to the internet of things: new challenges for QoC management. In *Ubiquitous Computing and Ambient Intelligence. Personalisation and User Adapted Services*. Springer International Publishing: Belfast, United Kingdom, 2014; 224–231.
- 5. Bellavista P, Corradi A, Fanelli M, Foschini L. A survey of context data distribution for mobile ubiquitous systems. *ACM Computing Surveys (CSUR)* 2012; **44**(4):24.
- Pei Z, Deng Z, Yang B, Cheng X. Application-oriented wireless sensor network communication protocols and hardware platforms: a survey. In IEEE International Conference on Industrial Technology (ICIT 2008), IEEE, Chengdu, China, 2008; 1–6.
- 7. Haefner K. Evolution of Information Processing Systems: An Interdisciplinary Approach for a New Understanding of Nature and Society. Springer Publishing Company: Incorporated, 2011.
- Botts M, Robin A. OpenGIS sensor model language (SensorML) implementation specification. OpenGIS Implementation Specification OGC 2007; 07–000.
- Compton M, Barnaghi P, Bermudez L, Garcia-Castro R, Corcho O, Cox S, Taylor K. The SSN ontology of the W3C semantic sensor network incubator group. Web Semantics: Science, Services and Agents on the World Wide Web 2012; 17:25–32.
- Neuhaus H, Compton M. The semantic sensor network ontology. In AGILE Workshop on Challenges in Geospatial Data Harmonisation. Hannover: Germany, 2009; 1–33.
- 11. Castro JL, Delgado M, Medina J, Ruiz-Lozano MD. Intelligent surveillance system with integration of heterogeneous information for intrusion detection. *Expert Systems with Applications* 2011; **38**(9):11182–11192.
- Chen L, Nugent C, Okeyo G. An ontology-based hybrid approach to activity modeling for smart homes. *IEEE Transactions on Human-Machine Systems*, 2014; 44(1):92–105.
- Kim J, Kim J, Lee D, Chung KY. Ontology driven interactive healthcare with wearable sensors. *Multimedia Tools and Applications* 2014; 71(2):827–841.

- Pinheiro P, McGuinness DL, Santos H. Human-aware sensor network ontology: semantic support for empirical data collection. *Proceedings of the 5th Workshop on Linked Science*, Bethlehem, PA, USA, 2015.
- Mitchell D, Morrow PJ, Nugent CD. A sensor and video based ontology for activity recognition in smart environments. In Engineering in Medicine and Biology Society (EMBC). 2014 36th Annual International Conference of the IEEE, IEEE, 2014; 5932–5935.
- Nugent CD, Finlay DD, Davies RJ, Wang HY, Zheng H, Hallberg J, Mulvenna MD. home-ML—an open standard for the exchange of data within smart environments. In *Pervasive Computing for Quality of Life Enhancement*. Springer Berlin Heidelberg: Nara, Japan, 2007; 121–129.
- 17. Weichhart G, Molina A, Chen D, Whitman LE, Vernadat F. Challenges and Current Developments for Sensing, Smart and Sustainable Enterprise Systems. Computers in Industry: Groningen, Netherlands, 2015.
- Balan RK, Satyanarayanan M, Park SY, Okoshi T. Tactics-based remote execution for mobile computing. In Proceedings of the 1st International Conference on Mobile Systems, Applications and Services, ACM, San Francisco, USA, 2003; 273–286.
- 19. Verissimo P, Rodrigues L. *Distributed Systems for System Architects*, Vol. 1. Springer Science & Business Media: New York, United States, 2012.
- 20. Liu G. Distributing Network Services and Resources in a Mobile Communications Network. U.S. Patent No. 5,825,759. U.S. Patent and Trademark Office: Washington, DC, 1998.
- Villanueva FJ, Villa D, Moya F, Barba J, Rincn F, Lopez JC. Leightweight middleware for seamless HW-SW interoperability, with application to wireless sensor networks. In *Design, Automation & Test in Europe Conference & Exhibition 2007.* DATE'07. IEEE: Nice, France, 2007; 1–6.
- 22. Locke D. *Mq telemetry transport (mqtt) v3. 1 protocol specification. IBM developerWorks Technical Library*, 2010. Available from: http://www.ibm.com/developerworks/webservices/library/ws-mqtt/index.html (accessed October 23, 2016).
- 23. Shelby Z, Hartke K, Bormann C. The constrained application protocol (CoAP) (No. RFC 7252). 2014.
- 24. Henning M. A new approach to object-oriented middleware. Internet Computing IEEE 2004; 8(1):66-75.
- 25. Salehi A. Design and implementation of an efficient data stream processing system (Doctoral dissertation, cole Polytechnique Fdrale de Lausanne) 2010.
- Kifer D, Ben-David S, Gehrke J. Detecting change in data streams. In Proceedings of the Thirtieth International Conference on Very Large Data Bases, Vol. 30, VLDB Endowment, 2004; 180–191.
- Pripui K, Arko IP, Aberer K. Top-k/w publish/subscribe: A publish/subscribe model for continuous top-k processing over data streams. Information systems 2014; 39:256–276.
- 28. Zadeh LA. Fuzzy sets. Information and control 1965; 8(3):338-353.
- Stover JA, Hall DL, Gibson RE. A fuzzy-logic architecture for autonomous multisensor data fusion. *IEEE Transactions on Industrial Electronics* 1996; 43(3):403–410.
- 30. Zhang Z, Xu L, Li HH, Shi A, Han H, Wang H. An information fusion method for sensor data rectification. *Journal of Electronics (China)* 2012; **29**(1-2):148–157.
- Russo F, Ramponi G. Fuzzy methods for multisensor data fusion. Instrumentation and measurement. *IEEE Transactions on* 1994; 43(2):288–294.
- Fracczak L, Podsedkowski L, Kobierska A. Data fusion using fuzzy logic techniques supported by modified weighting factors (FLMW). *International Journal of Fuzzy Systems* 2016; 18(1):72–80.
- 33. Bouchon-Meunier B (Ed.), 2013. Aggregation and fusion of imperfect information. Physica, Vol. 12.
- Suganthi L, Iniyan S, Samuel AA. Applications of fuzzy logic in renewable energy systems a review. *Renewable and Sustainable Energy Reviews* 2015; 48:585–607.
- Juan YK, Gao P, Wang J. A hybrid decision support system for sustainable office building renovation and energy performance improvement. *Energy and buildings* 2010; 42(3):290–297.
- Ahmed SS, Majid MS, Novia H, Rahman HA. Fuzzy logic based energy saving technique for a central air conditioning system. *Energy* 2007; 32(7):1222–1234.
- 37. Villar JR, de la Cal E, Sedano J. A Fuzzy Logic Based Efficient Energy Saving Approach for Domestic Heating Systems, 2009.
- Krishnan NC, Cook DJ. Activity recognition on streaming sensor data. *Pervasive and Mobile Computing* 2014; 10:138–154.
- Kim E, Helal S, Nugent C, Beattie M. Analyzing activity recognition uncertainties in smart home environments. ACM Transactions on Intelligent Systems and Technology (TIST) 2015; 6(4):52.
- 40. Satyanarayanan M. Pervasive computing: vision and challenges. IEEE Personal communications 2001; 8(4):10-17.
- 41. Varshney U. Pervasive healthcare and wireless health monitoring. *Mobile Networks and Applications* 2007; **12**(2-3):113–127.
- Emmanouilidis C, Koutsiamanis RA, Tasidou A. Mobile guides: taxonomy of architectures, context awareness, technologies and applications. *Journal of Network and Computer Applications* 2013; 36(1):103–125.
- Makris P, Skoutas DN, Skianis C. A survey on context-aware mobile and wireless networking: on networking and computing environments' integration. *IEEE Communications Surveys & Tutorials* 2013; 15(1):362–386.
- 44. Perera C, Zaslavsky A, Christen P, Georgakopoulos D. Context aware computing for the internet of things: a survey. *IEEE Communications Surveys & Tutorials* 2014; **16**(1):414–454.
- Alam MM, Hamida EB. Surveying wearable human assistive technology for life and safety critical applications: Standards, challenges and opportunities. *Sensors* 2014; 14(5):9153–9209.

- 46. Van Hoof J, Wouters EJM, Marston HR, Vanrumste B, Overdiep RA. Ambient assisted living and care in The Netherlands: the voice of the user. *Pervasive and Ubiquitous Technology Innovations for Ambient Intelligence Environments*, 205 2012.
- Chen L, Hoey J, Nugent CD, Cook DJ, Yu Z. Sensor-based activity recognition. systems, man, and cybernetics, Part C: Applications and reviews. *IEEE Transactions on* 2012; 42(6):790–808.
- Zadeh LA. Generalized theory of uncertainty (GTU) principal concepts and ideas. Computational Statistics & Data Analysis 2006; 51(1):15–46.
- Zadeh LA. A prototype-centered approach to adding deduction capability to search engines-the concept of protoform. In *Fuzzy Information Processing Society*, 2002. *Proceedings. NAFIPS. 2002 Annual Meeting of the North American*, IEEE, Nueva Orleans, USA, 2002; 523–525.
- Kacprzyk J, Zadrony S. Linguistic database summaries and their protoforms: towards natural language based knowledge discovery tools. *Information Sciences* 2005; 173(4):281–304.
- Yager RR. On linguistic summaries of data. In *Knowledge Discovery in Databases*, Piatetsky-Shapiro G, Frawley B (eds). MIT Press: Cambridge, USA, 1991; 347–363.
- 52. Raha S, Ray KS. Approximate reasoning with time. Fuzzy Sets and Systems 1999; 107(1):59-79.
- Ohlbach HJ. Relations between fuzzy time intervals. In 11th International Symposium on Temporal Representation and Reasoning (TIME 2004), IEEE, 2004; 44–51.
- Castillo-Ortega R, Marn N, Snchez D. A fuzzy approach to the linguistic summarization of time series. *Multiple-Valued Logic and Soft Computing* 2011; 17(2-3):157–182.
- 55. Aberer K, Hauswirth M, Salehi A. The global sensor networks middleware for efficient and flexible deployment and interconnection of sensor networks. *Technical Report LSIR-REPORT-2006-006*, 2006.
- 56. Perera C, Zaslavsky A, Christen P, Salehi A, Georgakopoulos D. Connecting mobile things to global sensor network middleware using system-generated wrappers. *In Proceedings of the Eleventh ACM International Workshop on Data Engineering for Wireless and Mobile Access*, ACM, Scottsdale, USA, 2012; 23–30.
- Maurer U, Smailagic A, Siewiorek DP, Deisher M. Activity recognition and monitoring using multiple sensors on different body positions. *In International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006)*, IEEE, 2006; 4-pp.
- Kwapisz JR, Weiss GM, Moore SA. Activity recognition using cell phone accelerometers. ACM SigKDD Explorations Newsletter 2011; 12(2):74–82.
- Foerster F, Smeja M, Fahrenberg J. Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. *Computers in Human Behavior* 1999; 15(5):571–583.
- Alon J, Athitsos V, Yuan Q, Sclaroff S. A unified framework for gesture recognition and spatiotemporal gesture segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2009; 31(9):1685–1699.
- 61. Gibson RE, Hall DL, Stover JA. An autonomous fuzzy logic architecture for multisensor data fusion. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI '94)*, IEEE, 1994; 143–150.
- 62. Gupta MM, Qi J. Theory of T-norms and fuzzy inference methods. Fuzzy Sets and Systems 1991; 40(3):431-450.
- 63. Dong WM, Wong FS. Fuzzy weighted averages and implementation of the extension principle. *Fuzzy Sets and Systems* 1987; **21**(2):183–199.
- Wu G, Talwar S, Johnsson K, Himayat N, Johnson KD. M2M: from mobile to embedded internet. *Communications Magazine*, IEEE 2011; 49(4):36–43.
- Delgado M, Quero JM, Lozano MD, Miranda AV. Integrating multimedia streaming from heterogeneous sources to JavaME mobile devices. *International Journal of Communication Systems* 2012; 25(6):763–778.
- 66. Cook DJ, Krishnan N. Mining the home environment. *Journal of Intelligent Information Systems* 2014; **43**(3): 503–519.
- 67. United Nations. Department of economic and social affairs, population division (2013). World Population Ageing 2013. ST/ESA/SER.A/348.
- Garrido-Jurado S, Muoz-Salinas R, Madrid-Cuevas FJ, Marn-Jimnez MJ. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 2014; 47(6):2280–2292.
- 69. Brandt A, Buron J. Transmission of IPv6 packets over ITU-T G. 9959 Networks 2015.
- Martin P, Ho BJ, Grupen N, Muoz S, Srivastava M. An iBeacon primer for indoor localization: demo abstract. In Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, ACM, Memphis, USA, 2014; 190–191.
- 71. Medina J, Quesada FJ, Espinilla M. A lightweight distributed architecture to integrate fuzzy relevant objects in real-time environments, 2015. UCAmI15. Chile. In press.
- 72. Zadeh LA. A computational approach to fuzzy quantifiers in natural languages. *Computers & Mathematics with applications* 1983; **9**(1):149–184.
- Sanchez ABNMD, Trivino G. Fuzzy knowledge representation for linguistic description of time series.16th World Congress of the International Fuzzy Systems Association. 9th Conference of the European Society for Fuzzy Logic and Technology, Atlantis Press, Gijn (Spain), 2015; 1346–1353.