Up in the Air: When Homes Meet the Web of Things

Lina Yao[1], Quan Z. Sheng[1], Boualem Benatallah[2], Schahram Dustdar[3], Ali Shemshadi[1], Xianzhi Wang[1], and Anne H.H. Ngu[4]

[1]The University of Adelaide, Adelaide, SA 5005, Australia
{lina.yao, michael.sheng, ali.shemshadi, xianzhi.wang}@adelaide.edu.au
[2]The University of New South Wales, Sydney, NSW 2052, Australia
boualem@cse.unsw.edu.au
[3]Vienna University of Technology, Austria
dustdar@dsg.tuwien.ac.at
[4]Texas State University, USA
angu@txstate.edu

ABSTRACT

The emerging Web of Things (WoT) will comprise billions of Web-enabled objects (or "things") where such objects can sense, communicate, compute and potentially actuate. WoT is essentially the embodiment of the evolution from systems linking digital documents to systems relating digital information to real-world physical items. It is widely understood that significant technical challenges exist in developing applications in the WoT environment. In this paper, we report our practical experience in the design and development of a smart home system in a WoT environment. Our system provides a layered framework for managing and sharing the information produced by physical things as well as the residents. We particularly focus on a research prototype named WITS, that helps the elderly live independently and safely in their own homes, with minimal support from the decreasing number of individuals in the working-age population. WITS enables *device-free*, unobtrusive monitoring of elderly people in a real-world, inhabituated home environment, by leveraging WoT technologies in building context-aware, personalized services.

1. INTRODUCTION

Worldwide, the population in developed (and also developing) countries is aging due to increasing life expectancy and low birth rate. With recent developments in cheap sensor and networking technologies, it has become possible to develop a wide range of valuable applications such as the remote health monitoring and intervention. These applications offer the potential to enhance the quality of life for the elderly, afford them a greater sense of security, and facilitate independent living[3,18,25]. For example, as depicted in the following scenario, by monitoring the daily routines of a person with dementia, an assistant service can track how completely and consistently the daily routines are performed, and determine when the resident needs assistance.

SCENARIO 1. *Thomas is a 77-year-old frail man lives alone in a small apartment. He is making a cup of tea and his kitchen knows it. Tiny sensors monitor his every move and track each tea-making step. If he pauses for too long, a nearby computer reminds him about what to do next. Later that day, Thomas's daughter accesses a secure website and scans a check-list, which was created from the computer in her fatherâĂŹs apartment. She finds that her father took his medicine on schedule, ate normally, and continued to manage his daily activities on his own. This puts her mind at ease.*

There are much significant research efforts on building an intelligent ambient environment, i.e., smart homes, such as House_n in MIT[11], CASAS[21], Adaptive House[19] facilitated the study of smart home technologies in more depth and in contexts that closely resemble real world, domestic spaces. All these efforts focus on people's direct interaction with the technologies. The recently emerging Web of Things (WoT) will bring the opportunity to push such an ambient intelligence to move forward quickly[9]. WoT provides the necessary infrastructure to transparently and seamlessly glue heterogeneous resources together by accessing sensors and actuators using standardized across different protocols, platforms and locations. Achieving ambient intelligent context-recognition system calls for a new opportunistic activity recognition paradigm in which, it can successfully infer what a person is doing or attempting to do by adapting to available sensor data. Most of existing ambient intelligence either heavily on people's involvement such as wearing battery-powered sensors, which might not be practical in real-world situations (e.g., people may forget to wear sensors), or lacking a synthetic method of deploying ubiquitous available sensor data. Recognizing and understanding the activities performed by people is a fundamental research topic in this area. To date, significant research work has been done in different ways. Computer vision related human activity recognition and tracking is one of main directions, but unfortunately, such solutions demand high computational cost for machine interpretation. In addition, the performance of such vision-based approaches depends strongly on the lighting conditions (e.g., hard to monitor sleep postures at night), angles (e.g., there exist some places not be covered by camera), and cameras are generally considered to be intrusive to people's *privacy*.

With the growing maturity of sensor, radio-frequency identification (RFID), and wireless sensor network technologies, activity recognition and tracking based on inertial, unobtrusive sensor readings has become a popular research area in the last few years. Inertial sensors are the most frequently used wearable sensors for

human activity recognition[4,8,11,27]. Although sensor-based activity recognition can better address issues such as privacy than conventional computer vision-based approaches, most work from sensor-based activity recognition require people to wear the inertial sensors and RFID tags. The main drawbacks of such solutions is that they need users' cooperation and maintenance (e.g., battery change). As a result, these approaches are not always practical, particularly for monitoring elderly persons with cognitive disabilities.

In this paper, we present a case study by enabling the smart home monitoring system with the support of the Web of Things platform. Our system monitors the daily human behaviors and object use, helps the elderly residents be aware of their surroundings and make better decisions. In the system, sensing activities are exposed as services and resources for higher level service composition and mashup via a dedicated Web-based interface.

The innovations of our system against the existing works (discussed in Section 2) are as the following:

- We develop a device-free dictionary-based learning approach to uncover structural information between hybrid sensor signals of different activities. The dictionaries are learned by an unsupervised sparse coding algorithm. Compared to the existing activity recognition approaches, our dictionary-based approach achieves more compact representation of the activities while preserve richer information, thereby underpinning an efficient and robust recognition of human activities.

- We propose a bilinear SVM based localization algorithm, in which sensor signal and time stamps are formed as a two dimensional matrix. The localization is developed to detect people presence to assist context-aware activity recognition, such as if a person is detected to fall, we can identify where he falls, e.g., in toilet or bathroom etc. The proposed method shows more robust results compared with sliding window based algorithms

- We embed our proposed activity recognition system into the Web of Things framework[16]. Our system leverages the flexibility of WoT, and provides necessary infrastructure to transparently access sensors, processors, and actuators using standardized protocols regardless of hardware, operating systems, and platforms.

- We further develop a graphical Web-based mashup interface allowing users to set up a series of higher level rules via a Web browser without programming efforts by integrating the inferred contextual information e.g., people's location and postures along with object usages detected. In this way, many enchanting applications of ambient intelligence, e.g., fall detection and alarm, can be easily realized in a user-friendly way.

The remainder of the paper is organized as follows. We first overview the related work in Section 2. The proposed system and technical details are described in Section 3. In Section 4, we report the experimental results. Finally, Section 5 wraps up the entire paper and highlights some future research directions.

2. RELATED WORK

In this section, we review some representative work closely related to our work.

2.1 Web of Things Middleware

With billions of things interconnected and present over the Web, there are significant challenges in developing WoT applications, due to their unique and inherent characteristics. The SENSEI project[1] proposes an architectural framework that focuses on addressing scalability issues in wireless sensor and actuator networks. The SemSorGrid4Env[2] develops a service-oriented architecture and a middleware that assists developers to build large-scale semantic-based sensor network applications. Both projects, however, deal with the connectivity issues WoT: how to connect heterogeneous things to the Web rather than how to describe and model things. The recent research and development activities at CSIRO[24] offer some interesting experience in applying WoT in a number of application domains such as smart farming. An ontology-enabled architecture has been developed where the sensor observations are published as linked data cube for long-term data analysis and sharing at the national scale. The system does not provide sufficiently suitable integrated abstractions for things.

The University of Washington researchers develop a WoT application, which unfortunately only focuses on managing the collected RFID data[28]. Paraimpu[3] provides a social platform for people to connect, compose, and share things. It is unclear on how this platform is realized. In Hyperpipe project[4], things are represented as Web services and connected using pipes so that users can easily compose. However, things are mostly resource-constrained and the traditional SOA standards like SOAP and BPEL may not be applicable. Many research projects are actively solving these challenges and one notable effort is IoT6 project[5], which focuses on the investigation of IPv6 and related standards (e.g., 6LoWPAN, CoAP) to overcome current fragmentation of the WoT.

Existing WoT middleware falls into two categories in general. The first category allows users to add as many sensors as they desire, and then gives users tools (e.g., simple App) to view the raw data collected. These systems usually have limited functionalities when interfacing with other applications or interpreting the data. The second category limits the user on the type and the number of sensors that they can utilize, but enables the user i) to interpret the collected data since possible use cases can be determined and programmed in a-priori, and ii) to interface with many third party applications, usually through certain cloud storage services. In our framework, data collected from sensors not only can be monitored in real-time, but also can be automatically converted into actionable information by our data interpreter based on trigger values or conditions preset by the user. This gives contextual information about the physical devices and enables WoT applications to be developed by accessing those high-level contexts, independent of low level physical properties of the sensors or devices.

2.2 Human Activity Recognition

The goal of activity recognition is to detect human physical activities from the data collected from various sensors. There are generally two main ways for activity recognition: i) to instrument people, where sensors and RFID tags are attached to people, and ii) to instrument the environment, where sensors are deployed inside the environment.

Wearable sensors such as accelerometers and gyros are commonly used for recognizing activities[2,17]. For example, the authors

[1]http://www.sensei-project.eu/

[2]http://www.semsorgrid4env.eu/

[3]https://www.paraimpu.com/

[4]http://geoweb.crs4.it/doku.php?id=hyperpipes

[5]http://www.iot6.eu/

in [12] design a network of three-axis accelerometers distributed over a user's body. The user's activities can then be inferred by learning the data provided by these accelerometers about the orientation and movement of the corresponding body parts. However, such approaches have obvious disadvantages including discomfort of wires attached to the body as well as the irritability that comes from wearing sensors for a long duration. Recently, researchers are exploring smart phones equipped with accelerometers and gyroscopes to recognize activities and gesture patterns [14,15]. In a very recent work, Krishnan et al. propose an activity inference approach based on motion sensors installed in a home environment [13]. An extensive sensor-based activity recognition survey can be found at [6].

Apart from sensors, RFID has been increasingly explored in the area of human activity recognition. Some research efforts propose to realize human activity recognition by combining passive RFID tags with traditional sensors (e.g., accelerometers). In this way, daily activities are inferred from the traces of object usage via various classification algorithms such as Hidden Markov Model, boosting and Bayesian networks [4,23]. Other efforts dedicate to exploit the potential of using "pure" RFID techniques for activity recognition [30,32]. For example, Wang et al. [26] use RFID radio patterns to extract both spatial and temporal features, which are in turn used to characterize various activities. However, such solutions require people to carry RFID tags or even readers (e.g., wearing a bracelet).

Recently, there have emerged research efforts focusing on exploring *device-free* activity recognition. Such approaches require one or more radio transmitters, but people are free from carrying any receiver or transmitter. Most device-free approaches concentrate on analyzing and learning distribution of radio signal strength or radio links. For instance, Youssef et al. [33] propose to pinpoint people's locations by analyzing the moving average and variance of wireless signal strength. Zhang et al. [34] develop a sensing approach using an RFID tag array. However, most of these efforts focus on localization and tracking. There are not much work on study device-free activity recognition. To our knowledge, the work by Hong et al. [10] is the only such effort, which proposes a solution on device-free activity recognition by using a sensor array.

3. SYSTEM DESIGN

In this section, we first overview the architecture of our system, followed by the descriptions of the key components.

3.1 An Overview

Our proposed system, Web-based Internet of Things Smart (WITS) home [6], consists of a hybrid pervasive sensor networks deployed in the house, and an intelligent in-home monitoring system running over the Web of Things framework. Our WITS system enables the seamless information access, exchange, and manipulation between the digital and physical worlds. As depicted in Figure 1, the system provides a layered framework for managing and sharing the information produced by physical things. It is developed using the Microsoft .NET framework and SQL Server 2012. Physical things and inferred location and activity events are mapped to corresponding virtual resources, which can be aggregated and visualized via a range of software components. We also adopt a rule-based approach to aggregate individual things for building context-aware, personalized new value-added services. We implement the prototype in a real-world, inhabited home environment, where residents can access, control, and compose physical resources in a variety of contexts.

WITS provides a basic online-storage service for end users, with
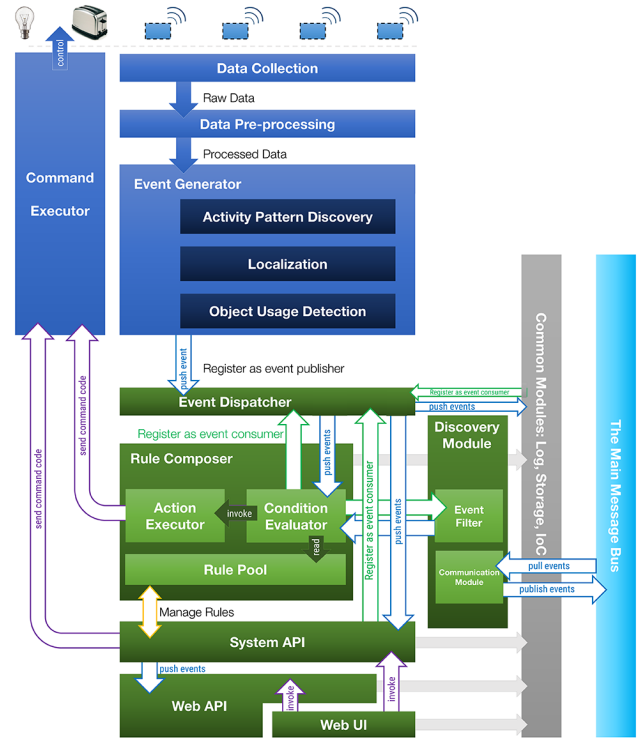
---

[6]www.linayao.com/wits

Figure 1: Overview of the WITS system

which any user can sync their data to the cloud, which includes Event Logs, Rules, and Visual Device Settings. Such information will be the basis of other online services, for example, user can access and manage their home online from a Web browser. WITS system offers a Web interface that allows remote management through cloud services. WITS systems also has an advanced data analytics service, which can perform recommendations to end users such as rules, device settings based on the analytics of user activity logs, device events and rules.

The system provides two ways to percept the physical objects and human activities.

- Object use detection. For identifying the status of physical objects (e.g., in use/not in use), the first is to use the radio frequency identification (RFID) technology, where physical objects are attached with RFID tags and interrogated by RFID readers. The second is to attach sensors with objects [31].

- Human activity detection. To realize the device-free activity detection, we propose that the environment itself should be augmented with a variety of sensors, such as RFID readers, passive RFID tags and a set of motion sensors mounted in different rooms. The installed locations are arbitrarily as long as the mounted sensors can create a signal field covering the house area. It is noted that the optimal configuration of mixed sensors are not our target in this work. Figure 2 shows our sensor deployment in part of the kitchen area.

The raw data captured by RFID readers and sensors are continuously transferred to a local server to be processed further. In the following, we will focus on introducing the key system modules of WITS and their technical details.
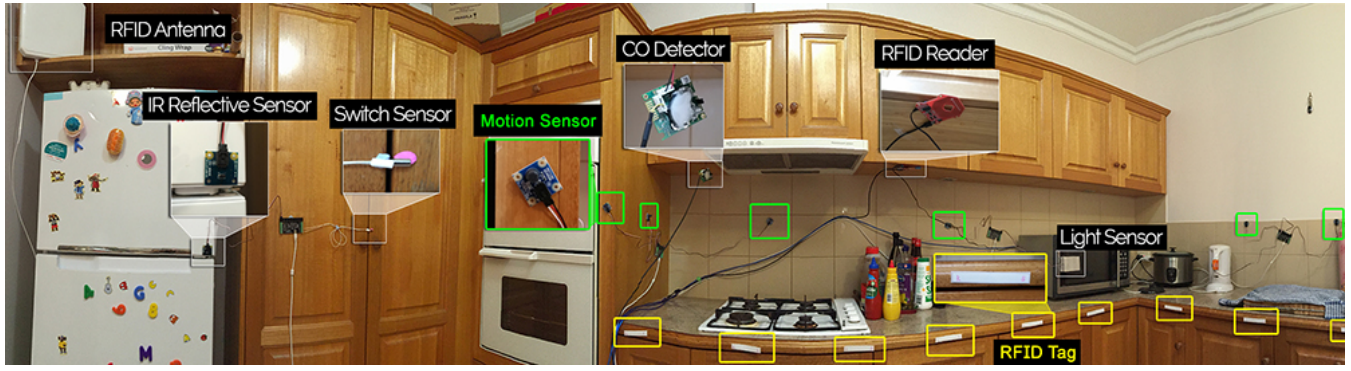
3.2 Data Collection and Pre-Processing

Figure 2: Deployment of sensing units inside the kitchen area of a house

The system provides two ways to identify physical objects and connect them to the Web. The first one is to use the RFID technology, where the physical objects are attached with RFID tags and interrogated by RFID readers. The second one is to combine sensors with objects to transfer the raw data. The raw data captured by readers and sensors will be further processed. In particular, the component of Data Collection takes care of i) managing RFID tags and sensors associated with physical things, ii) collecting raw RFID and sensor data, while the component of Data Pre-Processing processes the collected raw data and provides a universal API for higher level programs to retrieve the status of things. Due to inherent characteristics of RFID and sensor data (e.g., volatile)[22], this layer contains several software components for filtering and cleaning the collected raw data, and adapting such data for high-level applications. This task greatly benefits from our extensive experience in a large RFID research project[29].

The event generater layer is responsible for generating events based on the collected information (e.g., refrigerator door opens), which can be directly used by high-level applications or further processed by other modules (e.g., event dispatcher). This layer focuses on event processing that automatically extracts and aggregates localization, activity and object use events based on the data feeds from the data pre-processing layer in a pipelined fashion. The technical details on localization, activity recognition, and object usage detection will be discussed later respectively.

3.3 Object Usage Detection

The object usage detection layer focuses on the event processing that automatically extracts and aggregates things usage events based on the data feeds from the pre-processing layer in a pipelined fashion. The pipeline consists of three main phases: the event detection, the contextual information retrieval, and the event aggregation.

The event detector captures and decides whether a physical thing is in-use. In our design, there are two ways to detect usage events of things: *sensor-based approach* for detecting state changes and *RFID-based approach* for detecting mobility. In the sensor-based detection, the usage of an object is reflected by changes of the object's status, e.g., the status of a microwave oven moves from ideal to in-use. In the RFID-based detection, the movement of an object indicates that the object is being used. For example, if a coffee mug is moving, it is likely that the mug is being used.

The event aggregator indexes and stores all the events and services, together with their related information in a database, which can be mined for various purposes (e.g., finding hidden correlations among things, recommendation). A list of elements is constructed,

storing the identifiers of objects, their types and values, as well as the calculated contextual information. In this way, applications can focus on the functionalities without worrying about operations such as connecting to the database, opening connections, querying with specified languages and handling the results (normally they are raw data and inconvenient to access). End users also can access and query data through the provided user interface.
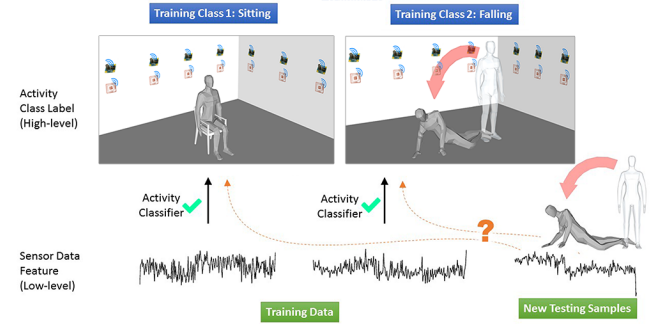

Figure 3: The activities can be recognized by analyzing the corresponding signal strength fluctuation of hybrid sensor network

3.4 Activity Monitoring

Our system for human activity recognition (see Figure 3) consists of three main stages:

- Processing the raw signal streaming data from various sensor inputs into individual segments, and then extracting low-level statistical features from each segment,

- Learning the overcomplete dictionary for each activity using the extracted features, and

- Given a new signal streaming data, finding the dictionary from the learned activity dictionaries that best approximates the testing sample.

Although signal strength variation of hybrid sensors reflects the uncertainty and non-linear distributed patterns, we still can observe some interesting characteristics. More specifically, we discover that the variations of signal strength reflect different patterns, which can be exploited to distinguish different activities in different locations. Figure 4 shows the distinctive fluctuation patterns of signal strength collected from different activities, and different rooms, respectively.
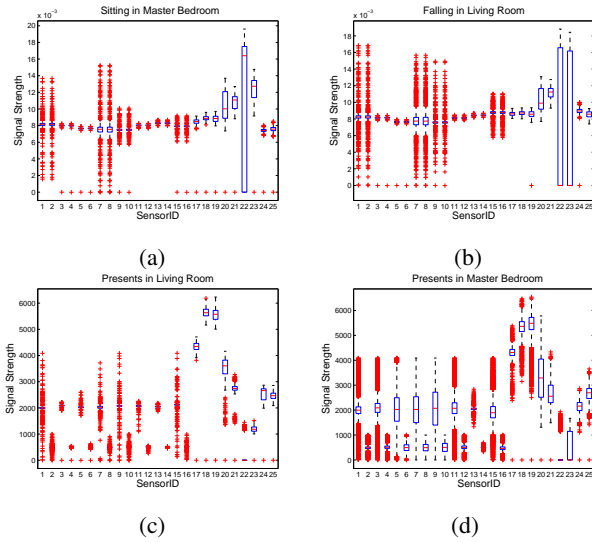
Figure 4: Different signal strength distribution patterns of (a) *sitting*, **(b)** *falling* **in the living room, and presenting in (c) the living room, and (d) the master room**

As shown in Figure 4, the signal fluctuations induced from resident activities (e.g., walking around), collected from two testing rooms, are significantly distinguishable. We propose a sparse representation based approach to recognize human activity by investigating signal fluctuations. We learn one single dictionary for each activity, which is formed by a set of basis vectors learned by solving a sparse optimization problem. Each basis vector can effectively capture part of key structural information of given training data from each activity.

There are several advantages in learning activity dictionaries. Firstly, the dictionary for each activity is learned from a collection of training samples via solving ℓ_1 optimization problem[5], which represents structural information of signal strength data in a more compact and informative way. Secondly, the dictionary learning and training process of each activity is independent from other activities, which makes an activity recognition system flexible and scalable, as no change is needed on the existing activity dictionaries when a new activity is added. Finally, each dictionary can be trained and learned by using only very small training samples, which can effectively relax the heavy workload on labeling and annotating training data in the activity recognition, as required by the most existing approaches.

Assuming we have K types of activities, and we construct K dictionaries (one dictionary for each activity). After that, a new signal strength vector is measure by using the coefficients of K dictionaries. We propose to compare the largest absolute value of coefficients of different dictionary for given new signal, larger of which indicates that the new testing signal sample fits better to the specific corresponding dictionary than others.

Let $\mathbf{O}^k = \{\mathbf{o}_1^k, \mathbf{o}_2^k, ..., \mathbf{o}_i^k\}$ be the training sample from activity class $\mathcal{C}^k$, to learn and encode the information of the testing samples belonging to a particular activity class, we first construct an overcomplete dictionary $\mathcal{D}^k$ for each class $\mathcal{C}^k$. Recall the set of training samples from k^{th} activity as $\mathbf{O}^k = \{\mathbf{o}_1^k, \mathbf{o}_2^k, ..., \mathbf{o}_N^k\}$, where $\mathbf{o}_i^k \in \mathbb{R}^m$, m is the feature dimensions. We intend to find a dictionary matrix $\mathbf{D}^k \in \mathbb{R}^{m \times K}$ having $K(K > m)$ vectors $\{\mathbf{d}_1^k, ..., \mathbf{d}_K^k\}$, over which $\mathbf{O}^k$ has a sparse representation $\mathbf{X}^k = \{\mathbf{x}_1^k, ..., \mathbf{x}_N^k\}$, where $\mathbf{x}_i^k \in \mathbb{R}^K$. In this case, the original training matrix $\mathbf{O}^k$ can be represented as a linear combination of no more

than $\tau_0^k (\tau_0^k << K)$ dictionary vectors. The optimization problem can be formalized as:

$$\min_{\mathbf{D}^k, \mathbf{X}^k} ||\mathbf{O}^k - \mathbf{D}^k\mathbf{X}^k||_2^2, \quad \text{s.t.} \quad ||\mathbf{x}_i^k||_0 \leq \tau_o^k \qquad (1)$$

We adopt the K-SVD algorithm[1] to solve this problem, which performs two steps iteratively until converged. The first stage is the sparse coding stage, $\mathbf{D}^k$ is kept fixed and the coefficient matrix $\mathbf{X}^k$ is computed by orthogonal matching pursuit algorithm, and then the dictionary $\mathbf{D}^k$ is updated sequentially allowing the relevant coefficients to be unique to K-SVD and resulting in a faster convergence.

As mentioned above, one advantage of having class-specific dictionaries is that each class is modeled independently from the others and hence the painful repetition of the training process can be avoided when a new type of activity is added to the system. After profiling the dictionary for each activity, for a given query feature vector of signal samples $\mathbf{o}^*$, The activity label is associated with the training samples have the largest absolute value of coefficients of $\mathbf{X}^k$:

$$l_{\mathbf{o}^*} = l(\max_i |\mathbf{X}|_i^k) \qquad (2)$$

Our proposed activity classification is summarized in Algorithm 1.

Algorithm 1: Activity Classification

Input: Sensor samples $\mathcal{S} = \mathbf{S}_1, ..., \mathbf{S}_K$, where K is the number of activity classes;
Querying signal samples $\mathbf{S}^* = \{\mathbf{s}_1^*, ..., \mathbf{s}_I^*\}$
**Output**: Activity label $\mathbf{l}^* = \{l_1^*, ..., l_I^*\}$ of $\mathbf{S}^*$

1 Extracting N^k feature vectors of signal samples from each activity class C^k using the proposed feature representation;
2 Constructing K activity-specific dictionaries $\mathbf{D} = \{\mathbf{D}^1, ..., \mathbf{D}^K\}$;
3 **for** *i = 1: I* **do**
4   Transform $\mathbf{S}^*$ to features $\mathbf{O}^*$;
5   Computing sparse representation x_i^* of $\mathbf{s}_i^*$ using K dictionaries $\mathbf{D}$;
6   Output activity label by $l_{\mathbf{o}^*} = l(\max_k |\mathbf{X}|_k^i)$;
7 **end**

3.5 Localization

Localization is critical to track and locate people in indoor environments. Monitoring the movements of an elderly person inside the home is specially important to spot abnormal behaviors, e.g., staying in the toilet over 30 minutes. Such contextual information is also important for the system to perform commonsense reasoning. For example, when an old person is detected lying down, an alarm should be produced if his current location is a place other than the bedroom (e.g., the kitchen). Our system can provide coarse-grain location support, e.g., positioning which room a person presents. Different from traditional work using one dimensional signal values as feature vector and then classifying the signal streams using sliding windows, we propose a two dimensional feature matrix by coupling time dimension with signal feature vector, which shows more robust results.

We decompose the continuous signal stream collected from each room into every 30 seconds interval. Since the common sampling rate is 0.5 seconds in this work, it turns out to be 60 time frames in each time interval. Therefore, each interval is formed as signal-vs-time $\mathbf{O}^k \in \mathbb{R}^{m^k \times t}$, where m^k is the dimension of features extracted from hybrid sensors in each room k, and t is the number
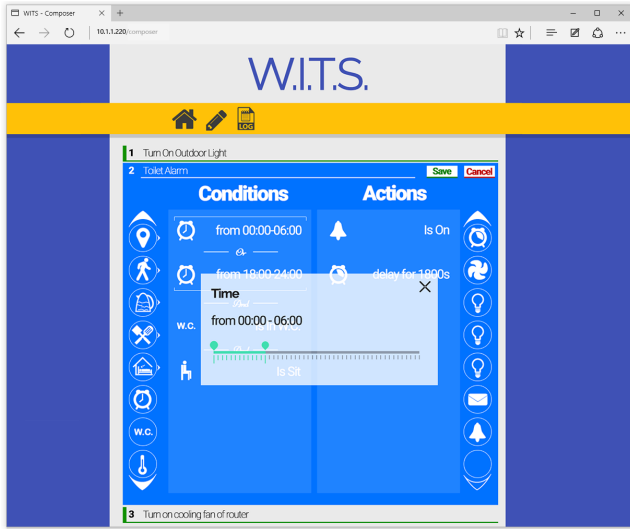
Figure 5: Rule composer interface. For example, for editing a rule like "send an alarm when a person stays in the toilet for over 30 minutes". A user needs only to drag the person, toilet and clock icons to the Conditions subpanel, the alarm icon to the Actions subpanel and performs some simple adjustments (e.g., adjust the clock slider to set the time period).

of samples during each 30 seconds time window ($t = 60$). As a result, each feature matrix contains the signal values in every 30 seconds. The coupling feature matrix is more informative and robust than one-dimensional feature vector. Since our objective is to identify whether a person presents in the room, our room-level localization problem can be formulated as: given N sample features $\{\mathbf{X}_i, y_i\}_{i=1}^{N}$ where $y_i \in \{-1, +1\}$ of the ith sample, indicating whether the person is in a room or not.

Let $\mathbf{O} \in \mathbb{R}^{m \times n}$ be the coupling feature matrices every 30 seconds. To identify whether a person presents, we adopt the bilinear classifier, which can be formulated as:

$$
\begin{aligned}
\hat{y} &= tr(\mathbf{W}^T \mathbf{O}) + b \\
&= tr(\mathbf{W}_n^T \mathbf{W}_m \mathbf{O}) + b \\
&= tr(\mathbf{W}_m^T \mathbf{O} \mathbf{W}_n) + b
\end{aligned}
\tag{3}
$$

where $\mathbf{W} \in \mathbb{R}^{m \times n} = \mathbf{W}_m \mathbf{W}_n^T$, $\mathbf{W}_m \in \mathbb{R}^{m \times d}$, $\mathbf{W}_n \in \mathbb{R}^{n \times d}$, and $d \leq \min m, n$. We can solve this problem under the maximum margin framework by measuring the margin of the bilinear classifier in Equation 3 by the matrix trace norm, e.g., sum of singular values for minimizing the matrix rank, which results in the following optimization problem:

$$
\min_{\mathbf{W}, b} \frac{1}{2} tr(\mathbf{W}^T \mathbf{W}) + C \sum_{i=1}^{N} \max \left(0, 1 - y_i \{ tr(\mathbf{W}^T \mathbf{O}_i) + b \} \right)
\tag{4}
$$

Equation 4 is convex and can be solved using coordinate descent along with SVM solver[20].

3.6 High-Level Service/Resource Composition

The WITS system offers a rule-based reasoning for high level service (or resources) composition, by developing a Web-based rule editing interface. The interface is used to set up a set of rules. A rule consists of two parts: a *condition* and an *action*. A condition is a composition of a set of simple boolean expressions. An action is simply a set of settings of devices. For example,

Figure 6: 3D scene of the real-world on the Web browser: the microwave oven will be in a highlighted status (yellow) while it is being used; a steam is shown on the kettle icon when the real kettle in the kitchen is boiling water.

`Toilet.Occupying=true & Duration>= 30mins`, an action, `SendAlert`, will be triggered, to send an alert to corresponding agent (e.g. a caregiver). By combining simple boolean expressions together, the application can setup a complex rule to make devices "smarter".

The rule engine can aggregate the inferred contextual information (location information, object usage events and inferred person's activity etc) for building context-aware, personalized new value-added services without any programming efforts. It consists of three main components: the *rule composer*, the *condition evaluator*, and the *action executor*. The rule composer is a Web-based application implementing a user-friendly GUI for rule creation and action setup, in a drag-and-drop fashion. The condition evaluator is a software that receives the string expressions of rules from the rule composer. It analyzes and annotates the string statement based on a state machine. The string expression is then translated to a list of annotated objects. The action executor is implemented based on the shunting-yard algorithm. It first compiles each part of the input sequence into a .NET Expression object. Then, it combines all such objects together into a complex Expression Tree, which will be compiled into a Lambda Expression. This Lambda expression object will be stored in memory when the system is running. It can be invoked when a device status changes or time elapses. If the Lambda expression returns true, a corresponding action will be called.

3.7 Real-time 3D Web Presence

This layer sits on the top of the framework and provides the access to the management of things of interest (for example, connection, monitoring, control, mashup, and visualization). The Web-based interface (Web UI) offers a 3D scene in a Web browser. We particularly adopt the Web Graphics Library (WebGL) in HTML5 to enable 3D scene recreation. The 3D models are stored as Digital Asset Exchange (DAE) files, and imported and rendered by using three.js[7] with plugins. Things are visualized and managed by device plugins. Each visualized thing is considered as a device plugin, which contains one or more 3D model or animation settings. For instance, the kettle will show steam when it is boiling water. We use the ShaderParticleEngine plugin[8] for three.js to create the steam effect for the kettle (see Figure 6). Each device plug-in also provides a serial of APIs (i.e., Web APIs), to communicate with the service

---

[7]http://threejs.org

[8]https://github.com/squarefeet/ShaderParticleEngine

Table 1: WoT Middleware Comparison

Categories	Prototypes	Functionality							Interface		Implementation		
		S	P	RM	RE	CAR	SN	MSS	M	W	RDL	API	PL
Smart Energy	Webnergy	–	–	–	–	–	–	–	●	●	JSON/XML	REST	Java
Tagging Objects	EPCIS Web Adapter	–	–	–	–	–	–	–	–	●	JSON/XML	REST	–
IoT/WoT Middleware	Zetta	–	–	●	–	–	–	●	–	●	–	REST	–
	Kaa IoT Platform	–	–	●	–	–	–	●	–	–	–	–	–
	w3c/web of things	●	●	–	●	–	–	–	●	●	JSON/TDL/LD	REST	–
	ThngSpeak	–	–	●	–	●	–	–	–	●	–	REST	–
	EVRYTHNG	●	●	●	–	–	●	●	–	●	JSON	REST	–
	OpenIoT	●	●	●	–	–	–	●	–	●	SSN Specification	–	–
SmartHomes	openHAB	–	–	●	●	–	–	–	–	–	–	–	–
	Eclipse Smart Home	–	–	●	–	–	–	●	●	●	–	–	Java
	OpenDomo	●	●	●	●	–	–	●	–	●	–	–	–
	FreeDomotic	●	●	●	●	–	–	–	–	●	–	–	–
	Calaos	–	–	●	–	–	–	–	●	●	–	–	–
	MisterHouse	–	–	–	●	–	–	–	–	●	–	–	Perl
	Wosh	–	●	●	–	–	–	●	–	●	XML	SOA	C++
	IoTivity	●	●	–	–	–	–	●	●	●	JSON	REST	–
	RaZberry	–	–	–	–	–	–	–	●	●	JSON	–	JavaScript
	The Things System	●	●	●	●	–	–	●	–	●	JSON	–	node.js
	PrivateEyePi	●	●	●	●	–	–	●	●	●	–	–	Python
	IoTSyS	–	–	●	–	–	–	●	–	●	–	–	Java
	House_n	–	–	–	●	●	–	●	–	–	–	–	–
	CASAS	–	–	–	●	●	–	●	–	●	–	–	–
Hybrid project	**WITS**	●	●	●	●	●	●	●	●	●	JSON	REST	C#/Java

S: Security (authentication, https etc)   **P:** Privacy (access control etc.)   **RM:** Real-time Monitoring   **RE:** Rule Engine   **CAR:** Complex Activity Reasoning   **MSS:** Multi-source Support   **SN:** Social Network (Twitter etc.)   **M:** Mobile interface   **W:** Web interface   **RDL:** Resource Description Language   **PL:** Programming Language   **API:** application programming interface   ●: Function enabled   –: Not Available/Not Applicable

layer for status changes of the corresponding things, and to reflect such changes on the Web browser. All the control and data flow can be manipulated through this lightweight Web interface, which also provides an administrative portal for things management and activity reasoning (e.g., connecting and disconnecting things, and viewing event logs).

4. EXPERIMENTS

This section is devoted to the validation study of our proposed system. We will report two main studies: i) a solid comparison and evaluation of our system with the existing WoT middlewares; and ii) an extensive performance study of several main modules of the WIITS prototype (i.e., activity recognition, localization).

The WITS system has been successfully deployed in the first author's home. In particular, we attached RFID tags and sensors to 127 physical things (e.g., microwave oven, fridge, coffee mug, laptop, couch) in several different places in the house (e.g., bedroom, bathroom, garage, kitchen) to capture the status of real-world physical objects as well as residents. In our implementation, things were exposed on the Web using RESTful Web services, which can be accessed and discovered from a Web-based interface.

4.1 WoT Middleware Evaluation

Existing IoT middleware falls into two categories in general. The first category allows users to add on as many sensors as they desire, and then gives users some tools (simple App or Web browser) to view the raw data that the sensors are collecting, but usually has limited functionalities when it comes to interfacing with other applications or interpreting the data. The second category limits the user on the type and the number of sensors that they can utilize, but enables the user to interpret the collected data - since possible use cases can be determined and programmed in a-priori - and to interface with many third party applications, usually through some cloud storage services. In order to leverage the advantages of scalability and usability, in our proposed WITS, data collected from sensors not only can be monitored in real-time, but also automatically converted into actionable information by our running intel-

ligent event generator or conditions/rules pre-defined by the user. This gives contextual information about the physical devices and enables IoT applications to be developed by accessing those high-level contexts independent of low level physical properties of the sensors or devices.

In addition, we propose a benchmarking methodology to compare the proposed system with the existing systems. To build the foundation of the benchmarking methodology in terms of IoT/WoT middleware and Smart Home projects, we analyze some of the related works in this area. The criteria from literature can fall within different categories as follows:

- Compliance with standards includes HTTP, scripts or using standard platforms (RFID EPCIS etc.)

- Versatility of the features supports for different features such as what kind of user interfaces, whether using Rule engine and scripting, or whether supports mobile access.

- Technology stack support enables versatility of the supported technologies, interfaces and compatibility with other packages, such as what supported platforms, using REST or SOA, what kind of developing language etc.

- Desirability of the outcome. This depends on the flexibility of the application in smart home scenario. The following criteria can be addressed in this regard[7]:

  - Allow for the organic evolution of routines and plans
  - Easily construct new behaviors and modify existing behaviors
  - Understand periodic changes, exceptions and improvisation
  - Design for breakdowns
  - Account for multiple, overlapping and occasionally conflicting goals
  - The home is more than a location
  - Participate in the construction of family identity

The detailed comparison is shown in Table 1.
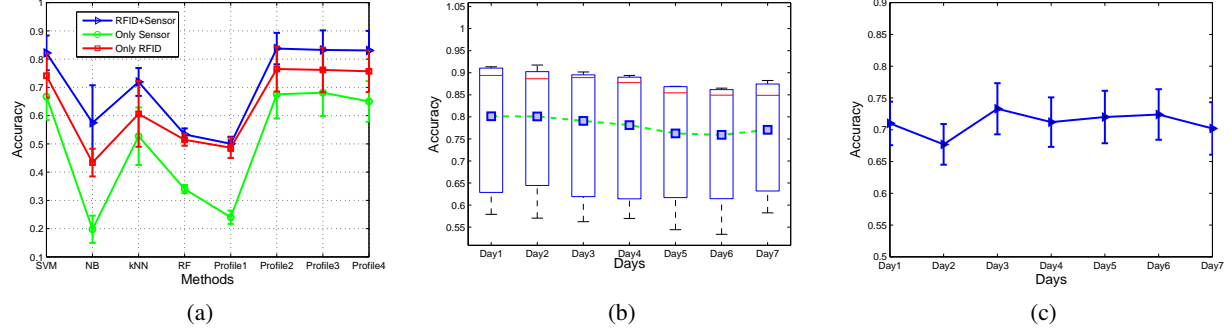
(a)   (b)   (c)

Figure 7: (a) Activity recognition performance comparison on hybrid sensor networks (RFID+Sensor), compared with unitary sensor network and unitary RFID network; (b) Low-level activity recognition performance consistency over 7 days evaluation; (c) High-level activity reason accuracy over 7 days.

4.2 Activity Recognition Evaluation

To evaluate the performance of activity recognition, we carried out the following two experiments. In the first experiment, we compared the performance of set of proposed dictionary-based approaches with a few state-of-the-art methods. We conducted extensive empirical study to determine the parameters of activity recognition method, the optimal settings are: the top k features $k = 30$, the dictionary size $d = 6$, and the training/testing ratio of data $= 20\%$. All the experiments were conducted under person-independent strategy, where we used the data from the subjects as training examples to train our algorithm and models, and the data from the left-out subject was used for testing.

Except comparing with other general methods of activity recognition in smart homes, except predicting the activity by utilizing largest absolute value of coefficients (Profile 2) in Section 3.4, we also developed a set of other strategies to infer the activity by fully exploiting the coefficient $\mathbf{X}^k$.

- Reconstruction error (Profile 1): its reconstruction error for the kth activity ($k \in [1, K]$) can be calculated as:

$$e_k = ||\mathbf{o}^* - \mathbf{D}^k \mathbf{X}^k||_2 \qquad (5)$$

Then the activity label of $\mathbf{o}^*$ can be assigned using:

$$l_{\mathbf{o}^*} = l(\min_k e_k) \qquad (6)$$

- Maximal sum of coefficients (Profile 3): the activity label is the top label with the maximal sum of coefficients of $\mathbf{X}^k$:

$$l_{\mathbf{o}^*} = l(\max \sum_i \mathbf{X}_i^k) \qquad (7)$$

- Concatate coefficients (Profile 4): we stack the learned coefficients with original features together to obtain new feature vectors, and then feed the enhanced features into SVM for classification.

We evaluate our approach by examining the recognition accuracy in terms of low-level activities (inferred from signal fluctuations), and high-level activities (inferred from low-level activities, along with object usage and location information (Section 3.6)). The detailed activities evaluated in this work is summarized in the Table 2.

Figure 7 shows the results on the activity recognition accuracy. From the figure, we can draw the following four key observations:

Table 2: Activities Used in Our Experiments

Index.	Recognizable Activities	
	Low-level Activities	**High-level Activities**
1	Sitting	Cooking
2	Standing	Eating
3	Sitting to Standing	Watching TV
4	Walking	Reading magazine
5	Arm Movement	Cleaning table
6	Kicking	Vaccuming
7	Bending	Bathing
8	Crouching	Toileting
9	Falling	Sleeping

- Our dictionary-based approach outperforms a set of state-of-the-arts. Specially the one with making use of the largest absolute value of coefficients performs the best result across all the compared methods using RFID only, sensor only and hybrid RFID+sensor networks, as shown in Figure 7, the average accuracy reaches over 83% and shows good performance of standard deviation as well.

- To test the consistency of the proposed approach, we also continuously evaluated it over 7 days, we can observe the recognition accuracy has some slight ups and downs, but the overall performance is quite stable and consistent with days (Figure 7 (b) and (c));

- We examined the positiveness of using hybrid RFID sensor network, other than only one of them in single. The performance using hybrid sensor sources is better in a full spectrum of recognition approaches (Figure 7 (a));

- The average recognition accuracy of high level activities is generally stable around 70% by adopting the generic Bayesian inference (Figure 7 (c)).

Figure 8 shows a detailed example of a sequence of activities performed by subjects. Our proposed method can identify these activities successfully, only with some minor misclassification during the activity transitions.

4.3 Localization Evaluation

In this experiment, we study the performance of localization of our system of detecting person presents in room-level. We collect
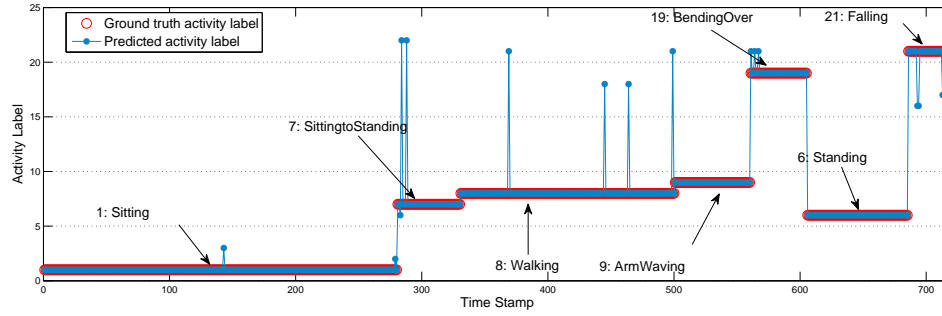
Figure 8: An illustrative example of detected activity sequence

the signal data in terms of empty status (no one is in a room) and person shows up in a room, respectively. Then, followed by person-independent strategy to validate our presence detection approach.

From the results shown in Figure 9, we can see that bilinear classifier works better than the other methods in identifying whether a person presents in a certain room, as the streaming signal data are modeled as the confluence of a pair of signals and time interval, signal-time matrix, which better retain and capture the dependency of data factors compared with using one dimensional signal vector with other linear classifiers. Figure **??** visualizes a subject traces of moving between four testing rooms, which shows that our bilinear method can perform good discriminative results with handling switching rooms as well.

4.4 System Latency Analysis

We conclude this section with some brief discussions about latency handling in our system. Fast detecting of posture changes is critical, particularly for aged care applications. For example, for the fall detection, we could send an alert to notify the care givers as quickly as possible to offer medical assistance for the elderly people after a fall happens. Our system has $4 \sim 4.5$ seconds recognition latency, which results from three main reasons:

- Our system evaluates subject's postures every 0.5 seconds using the latest 2 seconds of signal stream. In other words, if the current system time is at timestamp t, our system will produce the predicted postures in the $[t-2, t-1]$ seconds, and $[t-1, t]$ seconds is used to backtrack check if the predicted label complies with predefined rules. For instance, assume that the label is estimated as: *lie in bed* at $[t-2, t-1]$ interval, if the predicted label in interval $[t-1, t]$ is *nobody*, our system will determine the predicted posture is still *lie in bed*.

- Signal collector is programmed with a timer to poll the signal variations with a predefined order of transmission, which takes around 1 second to complete a new measurement with no workarounds.

- It should be noted that we integrated our system into a Web-based interface, which sends AJAX requests to services for the latest results and then looks up the database to retrieve data for sending back to the Web interface with updating DOM (document object model) element. Completing such a querying process normally takes 300ms to 500ms.

5. CONCLUSION

In this paper, we have proposed the design and development of a smart home system that leverages the emerging Web of Things

(WoT) for providing personalized, context-aware services to residents. Via seamless integration of digital world and the physical world, our WoT-based system can efficiently manage things of interest and access their corresponding services. In particular, the system implements the device-free monitoring of elderly people who live alone, in which both a person's location and activities can be monitored by learning signal strength fluctuations collected by pure passive RFID tags in the WoT environment. We have conducted extensive experiments to validate our proposed system. The practical experience gained from this system is useful for building future WoT applications.

Many challenges still exist in effective development of WoT applications. One such challenge is transaction handling. In WoT, the digital and physical worlds co-exist and interact simultaneously. Most things are resource-constrained, which are typically connected to the Web using lightweight, stateless protocols such as CoAP. In the future, we will focus on investigating novel ways for efficient transaction processing in WoT applications.

References

[1] M. Aharon, M. Elad, and A. Bruckstein. K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*, 54(11):4311–4322, 2006.

[2] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *Proceedings of 2nd International Conference on Pervasive Computing (PERVASIVE)*, pages 1–17. Springer, 2004.

[3] M. Blount, M. R. Ebling, J. M. Eklund, A. G. James, C. McGregor, N. Percival, K. P. Smith, and D. Sow. Real-time analysis for intensive care: development and deployment of the artemis analytic system. *Engineering in Medicine and Biology Magazine, IEEE*, 29(2):110–118, 2010.

[4] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall. Recognizing daily activities with rfid-based sensors. In *Proc. of 11th ACM Intl. Conference on Ubiquitous Computing (UbiComp)*, pages 51–60, 2009.

[5] E. J. Candès, J. Romberg, and T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *Information Theory, IEEE Transactions on*, 52(2):489–509, 2006.

[6] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu. Sensor-based activity recognition. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 42(6):790–808, 2012.
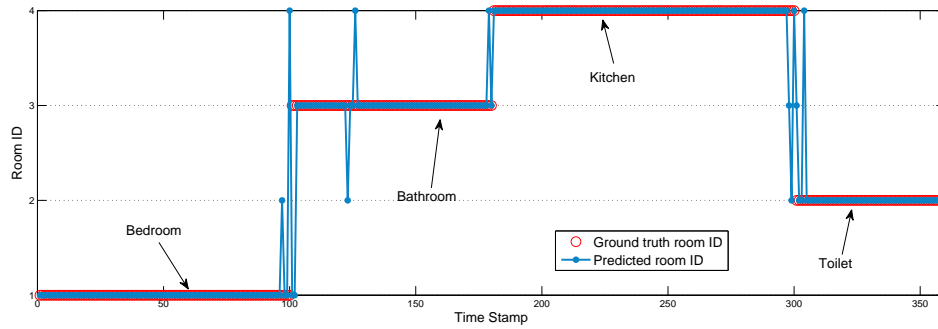
Figure 9: Accuracy of Room-level person present detection

[7] S. Davidoff, M. K. Lee, C. Yiu, J. Zimmerman, and A. K. Dey. Principles of smart home control. In *UbiComp 2006: Ubiquitous Computing*, pages 19–34. Springer, 2006.

[8] K. P. Fishkin, B. Jiang, M. Philipose, and S. Roy. I sense a disturbance in the force: Unobtrusive detection of interactions with rfid-tagged objects. In *UbiComp 2004: Ubiquitous Computing*, pages 268–282. Springer, 2004.

[9] J. Heuer, J. Hund, and O. Pfaff. Toward the Web of Things: Applying Web Technologies to the Physical World. *IEEE Computer*, 48(5):34–42, 2015.

[10] J. Hong and T. Ohtsuki. Ambient intelligence sensing using array sensor: device-free radio based approach. In *Proc. of ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, 2013.

[11] S. Intille et al. Using a live-in laboratory for ubiquitous computing research. In *Proc. of Intl. Conf. on Pervasive Computing (PERVASIVE)*. 2006.

[12] N. Kern, B. Schiele, H. Junker, P. Lukowicz, and G. Tröster. Wearable sensing to annotate meeting recordings. *Personal and Ubiquitous Computing*, 7(5):263–274, 2003.

[13] N. C. Krishnan and D. J. Cook. Activity recognition on streaming sensor data. *Pervasive and Mobile Computing*, 10:138–154, 2014.

[14] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 12(2):74–82, 2011.

[15] N. D. Lane et al. Bewell: A smartphone application to monitor, model and promote wellbeing. In *Proc. of 5th Intl. ICST Conference on Pervasive Computing Technologies for Healthcare*, pages 23–26, 2011.

[16] Q. Z. S. Lina Yao and S. Dustdar. Web-based management of the internet of things. *Internet Computing*, 19(4):60–67, 2015.

[17] P. Lukowicz, H. Junker, M. Stäger, T. von Bueren, and G. Tröster. Wearnet: A distributed multi-sensor system for context aware wearables. In *Proceedings of ACM International Conference on Pervasive and Ubiquitous Computing (UbiComp)*. Springer, 2002.

[18] S. Mennicken, J. Vermeulen, and E. M. Huang. From today's augmented houses to tomorrow's smart homes: new directions for home automation research. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 105–115. ACM, 2014.

[19] M. Mozer. *Lessons from an adaptive house*. PhD thesis, University of Colorado, 2004.

[20] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Bilinear classifiers for visual recognition. In *Advances in neural information processing systems*, pages 1482–1490, 2009.

[21] P. Rashidi and D. J. Cook. Activity knowledge transfer in smart environments. *Pervasive and Mobile Computing*, 7(3):331–343, 2011.

[22] Q. Z. Sheng, X. Li, and S. Zeadally. Enabling Next-Generation RFID Applications: Solutions and Challenges. *IEEE Computer*, 41(9):21–28, 2008.

[23] M. Stikic et al. Adl recognition based on the combination of rfid and accelerometer sensing. In *Proc. of Intl. Conference Pervasive Computing Technologies for Healthcare*, 2008.

[24] K. Taylor et al. Farming the web of things. *IEEE Intelligent Systems*, 28(6):12–19, 2013.

[25] J. Tung, H. Snyder, J. Hoey, A. Mihailidis, M. Carrillo, and J. Favela. Everyday patient-care technologies for alzheimer's disease. *IEEE Pervasive Computing*, 12(4):80–83, 2013.

[26] L. Wang, T. Gu, H. Xie, X. Tao, J. Lu, and Y. Huang. A wearable rfid system for real-time activity recognition using radio patterns. In *Proc. of the 10th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous)*, 2013.

[27] J. A. Ward et al. Activity recognition of assembly tasks using body-worn microphones and accelerometers. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 28(10):1553–1567, 2006.

[28] E. Welbourne, L. Battle, G. Cole, K. Gould, K. Rector, S. Raymer, M. Balazinska, and G. Borriello. Building the internet of things using rfid: the rfid ecosystem experience. *Internet Computing, IEEE*, 13(3):48–55, 2009.

[29] Y. Wu, Q. Z. Sheng, H. Shen, and S. Zeadally. Modeling Object Flows from Distributed and Federated RFID Data Streams for Efficient Tracking and Tracing. *IEEE Transactions on Parallel and Distributed Systems*, 24(10):2036–2045, 2013.

[30] L. Yao, Q. Z. Sheng, X. Li, S. Wang, T. Gu, W. Ruan, and W. Zou. Freedom: Online activity recognition via dictionary-based sparse representation of rfid sensing data. In *Proceedings of the 2015 IEEE International Conference on Data Mining (ICDM 2015)*. IEEE, 2015.

[31] L. Yao, Q. Z. Sheng, A. H. Ngu, and X. Li. Things of Interest Recommendation by Leveraging Heterogeneous Relations in the Internet of Things. *ACM Transactions on Internet Technology*, 2015, to appear.

[32] L. Yao et al. Exploring tag-free rfid-based passive localization and tracking via learning-based probabilistic approaches. In *Proc. of 23rd ACM Intl. Conference on Information and Knowledge Management (CIKM)*, 2014.

[33] M. Youssef, M. Mah, and A. Agrawala. Challenges: device-free passive localization for wireless environments. In *Proc. of 13th ACM Intl. Conference on Mobile Computing and Networking (MobiCom)*, 2007.

[34] D. Zhang, J. Zhou, M. Guo, J. Cao, and T. Li. Tasa: tag-free activity sensing using rfid tag arrays. *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 22(4):558–570, 2011.