IDAct: Towards Unobtrusive Recognition of User Presence and Daily Activities

Hanchuan Li
CSE, University of Washington
Microsoft Corporation
Seattle, USA
hanchuan.li@microsoft.com

Chieh-yih Wan
Intel Corporation
Portland, USA
chieh-yih.wan@intel.com

Rahul C. Shah
Intel Corporation
Santa Clara, USA
rahul.c.shah@intel.com

Alanson P. Sample
CSE, University of Michigan
Ann Arbor, USA
apsample@umich.edu

Shwetak N. Patel
CSE, University of Washington
Seattle, USA
shwetak@cs.washington.edu

***Abstract*—The Internet of Things (IoT) promises to revolutionize the way people interact with their surrounding environment and the objects within it by creating a ubiquitous network of physical devices. However, recent advancements have been focused on creating battery-powered electronics. There remains a huge gap between the collection of smart devices and the massive number of everyday physical objects. In this work, we bridge this gap by enhancing the sensing capabilities of everyday objects using commercial long-range RFID. We apply signal processing and machine learning techniques towards its communication channel parameters to detect the presence of users and to understand their daily activities. Different from prior work, our system can adapt to different environments and objects types. In a naturalistic user study deployed in a home environment, IDAct detected user presence with an F1 score of 96.7% and recognizes 24 different daily activities with an F1 score of 82.8%.**

***Index Terms*—Activity Recognition; Object Usage Sensing; Presence Sensing; RFID**

I. Introduction

Advances in cloud computing, mobile computing and embedded systems have enabled rapid adoption of the Internet of Things. People are now surrounded by a variety of smart devices in their daily lives. These connected devices lay the foundation of the IoT, which revolutionized the way people interact with their surrounding environments and fused applications spaces such as smart home and city, intelligent transportation, and connected health care.

Current advancements of the IoT have been heavily relying on these new and smart electronic devices. It was estimated that 14.2 billion devices will be connected in 2019 [5]. However, there still remain hundreds of billions of everyday objects (e.g. clothes, cookware, furniture etc.) that people use on a daily basis left out of this picture. These objects do not have embedded electronic pieces to support wireless communication and it is challenging to enhance their sensing capabilities in an easy and scalable way. There is an ever-increasing gap between the smart connected electronics and the massive number of everyday non-smart objects.

Fig. 1. Example IDAct Applications, **left**: sensing daily activities, e.g. taking medicine; **right**: sensing user presence using interference with ambient objects

Given the ubiquity of these objects, there are significant opportunities in enhancing their sensing capabilities and creating interactive applications around them. Connecting everyday objects will bring context-awareness into the details of everyday living, creating a truly immersive IoT experience. Imagine a world where your pill bottle keeps track of your medication intake and water glass monitors your hydration level. Even your yoga mat is aware of your exercises and could adjust lighting, temperature and background music accordingly.

In this paper, we propose IDAct, a system that enhances the sensing capabilities of everyday objects in an easy and scalable way to detect the presence of users and understand their daily activities. IDAct utilizes passive long-range RFID, However, different from existing RFID based solutions [12], [13], [19], our approach does not require extensive training for specific objects or the sensing environment. The interaction detection algorithm utilized in IDAct allows it to adapt to different user and RF environments, lowing its deployment barrier. In addition, our system supports sensing human presence without requiring active usage of objects. When compared to existing

978-1-7281-1210-7/19/$31.00 ©2019 IEEE

wearable solutions [14], [15], [17], IDAct does not require any user instrumentation. This significantly reduces user burden, and in turn, helps improve the technology acceptance and compliance [9].

IDAct opens the door for many IoT sensing applications by powering a network of physical objects. For example, in a smart home setting, temperature, humidity, and lighting conditions can be adjusted according to user location and activities, saving energy while allowing for customized personal comfort. IDAct can also enable dietary monitoring or health care applications by detecting what food users are preparing in the kitchen and when they are taking medications. In the field of senior care, automated recognition of daily activities provided by IDAct can be combined with assisting technology to improve the life quality for seniors who live alone [3].

We validate our techniques in a naturalistic study deployed in a home environment. Data from tagged objects were collected by a single RFID reader as participants performed activities that they would normally do in their everyday lives (e.g., cooking, watching TV). We reviewed recorded videos and manually annotated ground truth to evaluate the effectiveness of IDAct. In the 26 hours of data collected with 110 objects across 10 participants, our analysis shows that IDAct can detect user presence with an F1 score of 96.7%, and recognize 24 different daily activities with an F1 score of 82.8%. Our contributions in this paper are as follows:

- A fine-grain long-range RFID-based activity recognition system which could adapt to different RF environments, users, and object types.
- An algorithm that senses user presence from changes in the RF environment surrounding each object without requiring explicit object usage.
- A user study in a natural living environment to quantify the effectiveness of our system at detecting object usage, sensing human presence, and recognizing daily activities.

II. Related Work

Automatic means of activity recognition is one of the key building blocks needed to enable context-aware ubiquitous computing applications. There is a variety of prior work that aims to accomplish this goal by wearable or and distributed sensing leveraging active and passive sensors each with their own unique strengths and weaknesses.

A. Activity Recognition using Wearable Devices

The objects we utilize in daily living provide rich contextual information about the activities performed. In recent years, deep learning techniques have demonstrated promising results in understanding user's daily activities using first person cameras and sensors like IMU [2], [22]. In particular, Castro et al. [2] presented a deep learning activity recognition system based on egocentric cameras. This work achieved an accuracy of 83% classifying 19 daily activities. However, this approach lacks support for recognizing fine-grained activities (e.g. which dish is being prepared). In addition, privacy concerns of first-person visual recordings limit the usability of computer vision systems in living environments. In general, wearable so+lutions have the advantage of being mobile which extend the physical space for activity recognition. However, the increased user burden due to body-mounted devices limits the technology acceptance and compliance especially for applications in living environments [9].

B. Activity Recognition using Sensor Networks

Researchers also explored distributing sensors in the physical environments for activity recognition. In particular, binary 'state-change' sensors based on piezo-electric switches has been studied to detect object usage [18], [20]. Each object is assigned a wireless sensor that is triggered when the object is moved to infer daily activities. However, the form factor and unit cost is still not ideal for sensing user interactions with everyday objects. In general, techniques based on active sensor networks demonstrate promising results in activity recognition, but at the same time limited by their high per unit cost and the requirement for battery replacement. To mitigate this challenge, researchers explored integrating a large number of sensing modalities [10] (infrared, IMU, microphone, magnetometer, etc.) into a single sensor and demonstrated recognition of activity at a category level. However, fine-grain recognition of daily activities requires knowing the instance and specific identity of the object being manipulated. For example, sensing medicine intake requires knowing which pill bottle was picked up among the many similar pill bottles. Knowing the precise identity of the pill requires higher granular of sensing capabilities.

C. Activity Recognition using RFID

The small and low-cost nature of passive RFID tags makes it easy for instrumenting everyday objects. Philipose et al. developed a glove-based near-field RFID reader that could identify the read/no-read states of tags attached to objects in close proximity [16]. Objects held in hands are recognized by the wrist-worn reader, over time, the usage of these objects can be utilized to infer higher level activities.

The recent advances in commercially available passive long-range RFID systems give rise to their popularity in distributed sensing applications. These systems operate at 900Mhz Ultra-High Frequency (UHF) ISM band. The RF channel parameters reported in the Gen 2 UHF RFID readers [6] provides insights into the state of the tag and its surrounding environments. This system is previously studied for health and wellness sensing applications [7], [13], [19], [21]. However, the object usage sensing methods described in their systems are based solely on received signal strength (RSSI). In addition to requiring dense antenna deployment, these prior work require excessive training for each object, making it a challenge to scale to a large number of objects. Prior work also explored levering RF channel parameters to sense user interactions with objects [12] in forms of motion and touch. This system only requires a single antenna and a single tag on each object. However, it was still challenging to adapt to different object types and new physical environments. Improving upon the prior work, IDAct

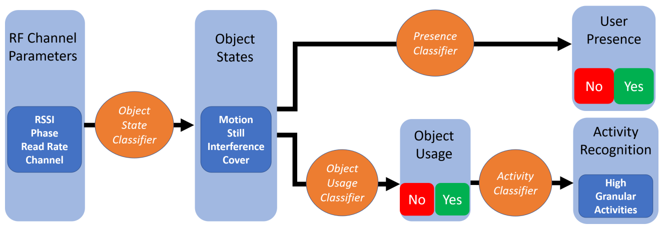

Fig. 2. An overview of our activity recognition pipeline: from RF channel parameters to sensing object usage, detecting user presence and recognizing daily activities.

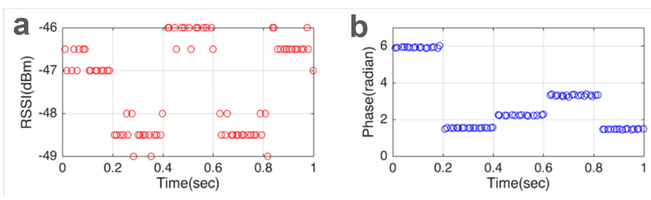

Fig. 3. Discontinuity in RSSI and phase samples due to carrier frequency sweep. **(a)**: RSSI vs. time, **(b)**: Phase vs. time

tackles variations in ambient RF environments using signal processing and machine learning techniques, which allows it to adapt to new object types and physical environments after a easy calibration process. We conducted a study to quantify our system in presence sensing and activity recognition in the evaluation section.

III. System Overview

Figure 2 outlines the data flow of our system. We first describe the back-scattered RF channel parameters utilized in this work. Then we extract features from those RF channel parameters to create an object state classifier. We further explore the correlation between object states and user interactions to determine object usage. Afterwards, we describe our heuristic approach for sensing user presence leveraging states of the object in the surrounding environments. Finally, we discuss how different granular of user activities can be inferred based on object usage.

IV. Sensing Object States

This section describes our approach to classify four different states of RFID tags using RF channel parameters. These states are closely related to object usage events. We pay special attention to detecting the unintended interference created by the human body as they move in the ambient environments. Understanding its difference with intended object motion is crucial towards detecting user presence and lowering false positives in activity recognition.

A. Tag States

(a) Moving: Tags under motion is a strong indication of object usage.
(b) Interfered: RF Interference generated by human motion is a strong indicator of human presence.

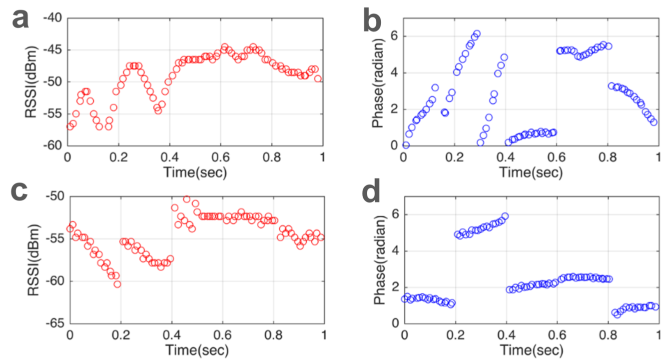

Fig. 4. Raw channel parameters of tag moving and interfered states. **(a)** RSSI-motion **(b)** phase-motion **(c)** RSSI-interference **(d)** phase-interference

(c) Still: Having no motion or external interference indicates that the corresponding object is not being used.
(d) Blocked: RFID reader unable to read tag is likely a result of occlusion from the human body or objects with high dielectric.

B. RF Signatures

In this section, we describe the set of features we use to classify different tag states. We utilize Impinj R420 reader in this work, which give us access to low-level channel parameters including RSSI and RF phase. Figure 4 is a visualization of RSSI and phase for a one second period. FCC regulations require RFID readers in the 915 MHz ISM band to pseudo-randomly change their transmit frequency in order to minimize interference with other devices. To satisfy this requirement, RFID readers frequency sweep across 50 channels from 902 MHz to 928 MHz (in the USA). The changes in carrier frequency happen at roughly 0.2-second intervals, causing the discontinuities in RSSI and phase. In the first row (Figure 4a, 4b), the tag is being moved. In the second row (Figure 4c, 4d), the tag is still but impacted by user moving nearby. When we compare RSSI of these 2 states to when tags are still (Figure 3a), variations in each channel is dramatically increased. We use the standard deviation (SD) of the RSSI as one of the features to differentiate between still and motion/interfered states.

Feature 1: RSSI SD for each channel frequency (every 0.2 seconds)

$$RSSI\,SD = SD(RSSI(channel == channel(X))) \tag{1}$$

When we compare the phase of moving and interfered states (Figure 4b, 4d) to motionless tags (Figure 3b), the phase difference within each channel is dramatically increased (in each 0.2 second intervals). This observation is described in feature 2

Feature 2: Phase difference for any channel frequency

$$\begin{aligned} Phase\,Dif = |&Phase(channel == channel(X))(end) \\ &- Phase(channel == channel(X))(1)| \end{aligned}$$

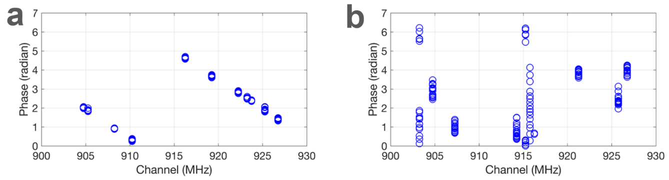

Fig. 5. Phase-channel relation **(a)** linear relation when still **(b)** non-linear under motion

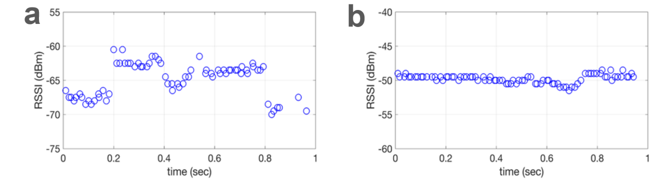

Fig. 6. RSSI variations of 2 tags attached to two objects of different materials at 2 different locations

For a passive RFID system, the primary signal path where most power is transmitted is the line-of-sight path. Under most conditions, there will be at least one, if not more, multipath reflections. The motion of the object will directly change the length of all signal paths, which result in variations of RF channel parameters such as RSSI and phase. Interference created by human proximity will either attenuate the signal, detune the tag, or create additional multipath trajectories between the tag and the reader. These changes will also contribute to the variations in RF channel parameters. When we compare tag interference (Figure 4c, 4d) with tag motion (Figure 4a, 4b), we can see that the RF phase is much more sensitive to motion when compared to interference. Our hypothesis is that when the human body is interfering the signal path, the presence of the body will attenuate the signal or detune the tag, which leads to higher variability in the RSSI of the back-scattered signal. However, the phase is independent of signal attenuation and primarily determined by the length of the line-of-sight path, so it remains relatively stable. We calculate the ratio between the variation of the two signals above to characterize this observation.

Feature 3: Ratio between phase difference and RSSI variation

$$Ratio = Phase\,Dif/RSSI\,SD \tag{2}$$

When the tag is blocked, the reader cannot communicate with the tag. For any given time t_0 to t_1, blocked states can be determined by whether the reader receives any readings from the tag as described by equation 3.

Feature 4: Detection of tag presence (covered state)

$$presence = isempty(Channel\,Parameters(t_0 - t_1)) \tag{3}$$

In addition to time domain features, we also explore the relationship between channel parameters and their corresponding transmit frequencies. We visualize the relationship between the channel and phase signals. In Figure 5a, the tag is motionless and generates phase samples which linearly correlate with the channel frequency. This observation is consistent with the relationship expressed in Equation 2. Motion (Figure 5b) creates variation in the tag - reader distance that breaks this linear relation. The difference between the 2 states can be represented by the root mean square error of linear regression between the channel and phase, where the error of motion states will be much higher when compared to errors of still states. We capture this observation in Feature 5.

Feature 5: RMS error of phase-channel linear regression.

C. *Object State Classifier*

One challenge in detecting object state is that channel parameters may exhibit different patterns when attached to objects with different materials. For example, an object with a high dielectric constant may detune the tag, increasing the noise level in the reflected signal. In addition, tagged objects at different locations may also have different RF environments. So, the challenge lies in creating a unified object state classifier that can adapt to different RF environments and object types. For example, Figure 6 compares the RSSI variations of the RFID tag when attached to 2 different objects at 2 different locations. Both objects in Figure 6a and 6b are still. The variations of the RSSI within each channel (every 0.2 seconds) are bigger for object a when compared to object b. Object materials, tag-reader distance, and multipath reflections can all contribute to differences like this, making it a challenge to create a unified classifier. On the other hand, training specific machine learning models for each individual object is computationally expensive and does not scale well.

To mitigate the challenges of variable object types and RF environments, we combine signal processing and machine learning to account for different RF environments around each tag while making it possible to scale to a large number of objects. More specifically, we combine a unified SVM classifier with a moving window that could adjust to the RF channel parameters noise level of each tag. Our approach accounts for the variations in RF channel parameters associated with objects and environments. In this section, we discuss the implementation of this classifier. We later present a study to evaluate its performance.

The incoming RF channel parameters are segmented using a 0.5-second sliding window with overlapping of 0.25 seconds. First, we determine if the tag is visible to the reader (Equation 3). If it is, we create a 10-second buffer to capture the variations of raw RSSI and phase within each channel. We calculate the SD of RSSI and phase. For still objects, we assume that the measurements error of RSSI and phase are random. And when the tag is still, the measurements of these two channel parameters will follow Gaussian distributions centered around the real value. The probability of each parameter deviating from the mean by more than 3 times the standard deviation (SD) is 0.2%. In other words, if either parameter contains samples that are out of the 3 times of SD on either side of the Gaussian distribution, it is very likely that this data point is generated by non-still object states (e.g. motion or interference). Given that the RF baseline of different tags is dependent on object type, location and time, we continuously update the 10-second buffer according to current RF channel parameters. Details of this method are described in Algorithm 1.

Algorithm 1: Determine if the tag is still

```
if present == 1 then
    /*Initialize 10 second buffer for RSSI and Phase
    which include 50 channels (0.2s/channel)*/
    for each of the 50 channels
    Buffer_Phase_SD = SD(Phases in each channel);
    Buffer_RSSI_SD = SD(RSSIs in each channel);
Phase_SD = mean(Buffer_Phase_SD);
RSSI_SD = mean(Buffer_RSSI_SD);
if (|phase(now) - mean(phases in he current channel))| <= 3 *
  Phase_SD&&|(RSSI(now) - mean(RSSI in the current channel)| <= 3 * RSSI_SD)
  then
    still = true;
    /*Update buffer for RSSI and Phase */
    remove(Buffer_Phase_SD(1));
    Buffer_Phase_SD(1 : end - 1) = Buffer_Phase_SD(2 : end);
    Buffer_Phase_SD(end) = SD(phases in the current channel);
    remove(Buffer_RSSI_SD(1));
    Buffer_RSSI_SD(1 : end - 1) = Buffer_RSSI_SD(2 : end);
    Buffer_RSSI_SD(end) = SD(RSSIs in the current channel);
else
    still = false;
return still ;
```

Note that we classify the tag state as still only when both the phase and RSSI stay within the $[mean-3SD, mean+3SD]$ range. This was an explicit decision to improve sensitivity for non-still states, which in turn improves sensitivity for object usage detection. In addition, all objects should not be moved during the 10-second initialization so that our classifier could function accurately.

Under the condition that the tag is present and not still, we utilize the five features in the previous section to implement an SVM classifier with RBF kernel to differentiate between the motion and the interfered states. The parameters of the RBF kernel are optimized using training data in the evaluation study.

V. Activity Recognition in a Home Environment

To validate the effectiveness of our approach, we conducted a naturalistic study across 4 different spaces in a living environment including: a kitchen (Figure 8a), a bathroom (Figure 8b), a dining room (Figure 8c), and a living room (Figure 8d). We instrumented 110 commonly used objects with RFID stickers. RFID reader with a single antenna is placed in each of these environments, the positions of the antennas are highlighted with the red boxes in Figure 8. 10 participants were recruited, who are undergraduate and graduate students (7 males, 3 females), aged from 21 to 30. The study happened in 7 different days during a period of 1 month. On each day, there are either 1 participant individually or 2 participants working together to finish all or a subset of the 24 activities outlined in Figure 7.

A. Data Collection and Processing

26 hours of RFID data in total was collected as participants went through 24 different activities in these environments. These 24 different activities belong to a few categories of ADLs and instrumental ADLs, referring to people's daily self-care activities proposed by Katz et al. [8]. These activities cover categories including preparing meals, eating, taking medication, housework, reading, personal hygiene, mobility, and entertainment. Figure 7 provides more details on these activities. Examples activities are shown in Figure 8. A ceiling-mounted camera is deployed to collect ground truth information about object usage and activities. We created a video analysis tool to manually annotated object usage and activities happened in the study. The data was annotated using 1-second sliding windows. 2 annotators were employed to label these videos independently and the sections in which their labeling are consistent was utilized as ground truth (95%) and the inconsistent sections are ignored (5%).

B. Sensing Object Usage

To improve performance in object usage sensing, daily objects are grouped into 2 categories and different tagging strategies were applied. The 2 categories include objects that are mobile while being used and objects that are static while being used. For example, floor mops are under motion when performing cleaning tasks, while the sitting bench is static when seated.

1) Tagging strategy: Mobile objects are instrumented with a single tag. Most of the objects used in the evaluation study come from this category. For objects that are static under usage, a tag is attached every 30 cm on the object; a single tag is applied to objects that are smaller than 30 cm.

2) Tag states and object usage correlation: Object usage was detected leveraging its correlation with object states. In Figure 9, we visualize the recorded time series states of a floor mop (Figure 9a) and a sitting bench (Figure 9b) for 100 seconds. For each object, the data is segmented into 2 sections. In the first 50 seconds, the object is not in use while in the second 50 seconds it is being used. For mobile objects, the motion is a direct indication of object usage. However, blocked, still and interfered states were also detected. This is a typical situation for mobile object since the pauses in between motions will generate still states while interference and blocking from the human body are also common. The distinctive signature for usage is the increased variations in object states when compared to sections where the object is still. We apply a moving buffer with a duration of X seconds to capture this variation and determine usage when the SD of object states (0 to 3) in the buffer reaches threshold P. The buffer size X and the threshold P will be optimized using training data. For tags that are attached to static objects, such as the ones on the bench, usage can be determined by the absence of tag reflected signals when seats are occupied.

3) Evaluation: In order not to overfit our model towards any specific environment, we first train our classifier in 3 different environments including an office, a hardware lab and a living environment using a small dataset collected from a staged study including 300 instances of object state data collected from 3 different environments. Note that this data collection process is separated from the 7-day evaluation study.

Preparing meals	1. Blend smoothie, 2. Make cereal, 3. Make a sandwich, 4. Cook rice, 5. Garlic shrimp, 6. Fried egg with tomato, 7. Roasted chicken and potato, 8. Sweet and sour ribs, 9. Spinach almond salad; 10. Pan seared salmon; 11. Caesar salad; 12. Potato salad; 13. Pork chops; 14. Make tea.
Eating	15. Eat dinner
Mobility	16. Climb stairs
Medication	17. Take vitamin pills
Housework	18. Mop floor
Entertainment	19. Watch TV; 20. Read a book
Personal hygiene	21. Wash hands; 22. Brush teeth; 23. Wash face; 24. Dry hair

Fig. 7. Activities conducted in the user evaluation study

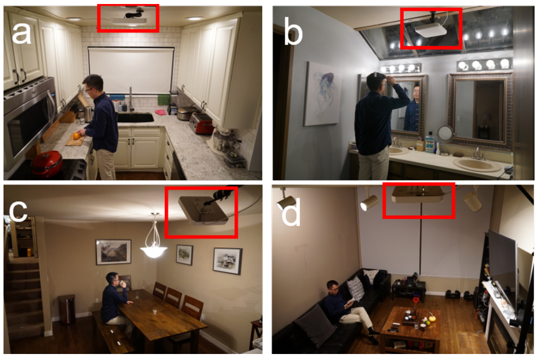

Fig. 8. In home study environments and example activities **(a)** kitchen: prepare smoothie; **(b)** bathroom: comb hair; **(c)** dining room: take medicine; **(d)** living room: read book

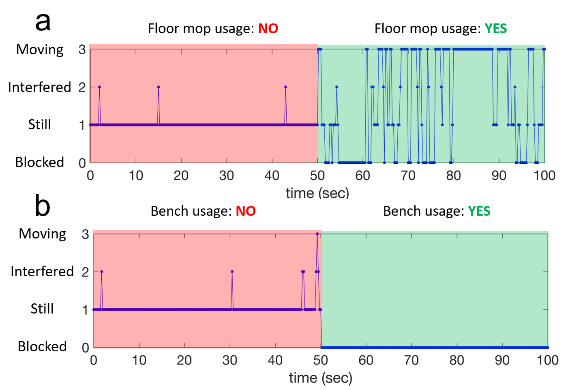

Fig. 9. Correlation between tag states and object usage: **(a)** example mobile object: floor mop, **(b)** example static object: sitting bench

We then use this classifier to process these 7 days of raw RFID data collected in the living environment into 4 different object states. 1 day of annotated object usage data was used as training data to optimize parameters in the object usage classifier by maximizing object usage detection F1 score. We then evaluate this classifier ($X = 21s, P = 0.94$) on other 6 days of annotated object usage data. The classifier achieved a precision of 92.4% and a recall of 90.9% for detecting object usages. It is worth noting that the false positive is low (0.03%). During the study, each object will generate on average about 1 false positive every hour.

C. User Presence Sensing

In this section, we discuss how we determine the presence of users by leveraging the states of tagged objects in the ambient environment. We first determine The room level location by the corresponding RFID antenna. While object usage is a strong indication of user presence, it is not required for user presence sensing. People's motion in the ambient environment will trigger interference in the tags nearby which can be utilized to infer user presence. Details of this method are described in Algorithm 2. Note that there is one parameter, *ratio*, that needs to be optimized in this algorithm to provide a balance between precision and recall in our presence detection. *ratio* is the threshold of the percentage of tags that are in motion or interfered required to trigger a positive presence classification. Having a low *ratio* will improve the sensitivity of presence detection, however, it will trigger more false positives in the user presence classifier. Having a large ratio will improve the precision in detecting user presence, while at the same time increase the possibility to generate false negatives. This trade-off is optimized through training data collected in the evaluation study.

Algorithm 2: Determine user presence

```
n = number of objects present in the ambient environment
   TagStateBuffer = TagStateNow(1 : n);
if Count(TagStateBuffer == (moving || interfered)) > n * ratio then
   | UserPresence = 1;
else
   | UserPresence = 0;
return UserPresence;
```

1) Evaluation: We utilized data collected in our activity recognition study to evaluate the user presence sensing approach. In addition to 26 hours of activities when users are present in these environments, we also collected 8 hours of data, 2 hours from each room when they are empty to evaluate the false positives of our presence sensing approach. Our sensing parameters *ratio* in Algorithm 2 are optimized using 1 hour of labeled user presence and 1 hour labeled user, not presence data and evaluated on the rest of the data. Our result yielded an F1 score of 96.7% with a precision of 96.2% and a false positive rate of 2.8%. At this point, $ratio = 15\%$. Object usage contributed to 65% of presence detection. During the other 35% of the time when no object usage is observed, user presence was captured by the RF interference as they move in the physical space.

D. Activity Recognition Approach

Previous literature has demonstrated how to recognize daily activities from object usage [1], [15]. In this work, We employ one commonly used methods: Hidden Markov models (HMM). The contribution here is not meant to be of machine learning, but it serves as an important piece to demonstrate the feasibility of activity recognition based on long-range RFID.

1) HMM Emission & Transition Probabilities: The HMM classifier considers the activity sequence to be the hidden states and the object used to be the emission states. There are three variables that need to be optimized from the training data, the initial probability of all activities, the transition probabilities between different activities and the emission probability from the activates to the objects. These concepts are explained using the following example. Consider a list of two activities, each with a set of utilized objects

- Blend smoothies: Orange, Apple, Cutting Board, Ice, Blender, Milk, Water Cup
- Make a sandwich: Cutting Board, Plate, Bread, Cheese, Ham, Plate

Here, the activities are hidden states, and their corresponding objects are the emission states. The emission probabilities for activity "blending smoothies" refers to the conditional probability of each object observed given the activity. (E.g. the probability of using "orange" given activity "blending smoothies"). The transition probability controls how the activity is chosen given the state of the previous activity. For example, what is the chance that a user is making a sandwich given that he was previously blending smoothies? The initial probability is the prior likelihood distribution of the two activities. These parameters are learned using training data in the evaluation section. The system infers the most likely sequence of activities by maximizing the joint probability of the activity sequence and the observations of objects. The most likely sequence of activities for any given sequence of objects can be effectively determined using the Viterbi algorithm [4]. For more details on the HMM-based activity recognition, please refer to the evaluation section.

E. Activity Recognition Evaluation

In this section, We leverage the object usage data generated in the previous section to classify 24 different activities outlined in Figure 7. Given that there is a dedicated antenna installed in each room, we can refer to the corresponding antenna number to narrow down the search space for activity recognition. In particular, activity #1 to #13 is conducted in the kitchen, activity #14 to #17 is conducted in the dining room, activity #18 to #20 is conducted in the living room and finally, activity #21 to #24 is conducted in the bathroom. Next, time series data of object usage are segmented into "episodes". For example, if the object is continuously used, for 30 seconds, that will be considered a single episode of object usage. Each episode will be considered one emission state from an activity (a hidden state) in the HMM model which are ordered by their starting time. The activity recognition results will be compared with manually labeled ground truth for evaluation. For example, given a continuous usage of a "floor mop" from 0 to 30 seconds, the HMM inferred the most likely activity for this 30-second episode is "cleaning the floor". However, ground truth indicate that from 20 to 30

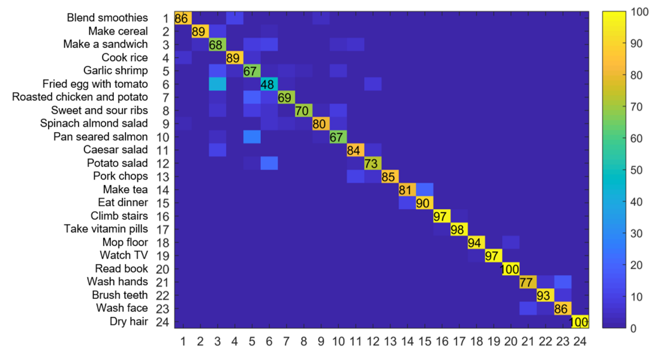

Fig. 10. Activity classification result confusion matrix

seconds, the participant was "watching TV". So that period will be considered as false positives for "cleaning floor" and false negatives for "watching TV".

We applied leave-one-day-out cross-validation where the HMM is trained on data from 6 days and then evaluated on the data collected on a remaining day. Given that different activities have different time durations. We normalize activity recognition results by time and evaluate the average F1 score accuracy across all 24 daily activities. A random guess, in this case, would yield an F1 score of 4.2%, provide an upper bound for the activity classification, we first evaluate the HMM classifier using human labeled object usage ground truth as input. In this case, the HMM classifier achieved an average F1 score of 91.1%. When we use the object usage automatically generated by our RFID based approach. Activity recognition accuracy range from 46% to 100% with an average F1 score of 82.8% across 24 different daily activities. We further break down this result in Figure 10 which shows the classification F1 score for each of the 24 activities described in Figure 7. In general, activities that share physical space and objects have lower recognition accuracy. For example, the food preparation section is low when compared to all other activities since they take place in the same environment and share a number of objects, making it a challenge to differentiate between these activities. Activities that take place in other environments have a comparatively distinctive set of objects, making it easier to accurately recognize and classify these activities.

VI. Conclusion and Discussion

In this work, we presented a method for detecting user presence and understanding daily activities leveraging unmodified long-range RFID systems. Our approach requires no user instrumentation and single antenna coverage for operation. We evaluated our system performance in a study deployed in a living environment and achieved an F1 score of 82.8% recognizing 24 different daily activities. In addition, we demonstrated the user presence sensing capability of our system with an F1 score of 96.7%. We achieved similar results when compared with previous work where multiple antennas and extensive training data are required.

Prior RFID based solutions require excessive training on the objects or the ambient RF environment. In this work, we combine signal processing with machine learning to monitor the noise level of the ambient environment of each object, which helps IDAct to adapt to different object types and environments. In addition, our algorithm allows differentiation of intentional object usage and accidental user interference, allowing IDAct to detect user presence without requiring object interactions. Given the limited sensing area of each antenna (around 200 sqft), we instrument RFID antennas at a room level which reduced the number of potential activities. As a result, we achieved promising recognition results given the relatively small set of activity data we collected in our study. The training data and testing data are generated by different participants for our evaluation study, which demonstrate the potential of IDAct to adapt to new users without the requirement of retraining.

Our current system has the following limitations. Passive tags utilized in this work will suffer from a decreased sensing range when attached to conductive materials (e.g. metal objects). Sufficient read rate of 10 reads/second or more per tag is required to guarantee high sensing accuracy. In our evaluation study, each antenna is turned on sequentially in each room to guarantee a high reading rate. When the reader is simultaneously powering multiple antennas, the individual read rate for each tag will drop, which may impact the sensing accuracy for object usage detection. In addition, the current framework of our activity recognition approach does not support detecting multiple users conducting activities simultaneously.

In the future, We are interested in further exploring reducing the training required for activity recognition. Instead of using a list of predefined activates, We want to explore the possibilities to automatically cluster activities into different categories leveraging the object state and human presence data provided by our system. In addition, We would like to investigate how to detect and differentiate multiple activities conducted by multiple users simultaneously. We consider these as important next steps towards the goal of creating an intelligent system that provides ubiquitous and fine-granular activity recognition as a service to support smart environment, personal wellness and assisted living applications.

ACKNOWLEDGEMENT

We thank Intel Labs for their generous support of this work. We are thankful to Ying Wang who provided expertise that greatly assisted the research. Part of the materials contained in this paper were previously included in the first author's PhD thesis [11].

REFERENCES

[1] Michael Buettner, Richa Prasad, Matthai Philipose, and David Wetherall. 2009. Recognizing Daily Activities with RFID-Based Sensors. *Ubicomp* (2009), 51–60.

[2] Daniel Castro, Steven Hickson, Vinay Bettadapura, Edison Thomaz, Gregory Abowd, Henrik Christensen, and Irfan Essa. 2015. Predicting daily activities from egocentric images using deep learning. In *proceedings of the 2015 ACM International symposium on Wearable Computers*. ACM, 75–82.

[3] Saisakul Chernbumroong, Shuang Cang, Anthony Atkins, and Hongnian Yu. 2013. Elderly activities recognition and classification for applications in assisted living. *Expert Systems with Applications* 40, 5 (2013), 1662–1674.

[4] G David Forney. 1973. The viterbi algorithm. *Proc. IEEE* 61, 3 (1973), 268–278.

[5] Gartner. 2019. Gartner Identifies Top 10 Strategic IoT Technologies and Trends. https://www.gartner.com/en/newsroom/press-releases/2018-11-07-gartner-identifies-top-10-strategic-iot-technologies-and-trends. (2019). [Online; accessed 10-Feb-2019].

[6] Impinj. 2015. Low Level User Data Support Impinj Support Portal. https://support.impinj.com/hc/en-us/articles/202755318-Application-Note-Low-Level-User-Data-Support/. (2015). [Online; accessed 10-Feb-2018].

[7] Asangi Jayatilaka and Damith C Ranasinghe. 2016. Towards unobtrusive real-time fluid intake monitoring using passive UHF RFID. In *RFID (IEEE RFID), 2016 IEEE International Conference on*. IEEE, 1–4.

[8] Sidney Kaiz. 1983. Assessing Self-maintenance: Activities of Daily Living, Mobility, and Instrumental Activities of Daily Living. *Journal of the American Geriatrics Society* 31, 12 (dec 1983), 721–727. DOI: http://dx.doi.org/10.1111/j.1532-5415.1983.tb03391.x

[9] Marije Kanis, Sean Alizadeh, Jesse Groen, Milad Khalili, Saskia Robben, Sander Bakkes, and Ben Kröse. 2011. Ambient Monitoring from an Elderly-Centred Design Perspective : What , Who and How. (2011), 330–334.

[10] Gierad Laput, Yang Zhang, and Chris Harrison. 2017. Synthetic Sensors: Towards General-Purpose Sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 3986–3999.

[11] Hanchuan Li. 2018. *Enabling Novel Sensing and Interaction with Everyday Objects using Commercial RFID Systems*. Ph.D. Dissertation.

[12] Hanchuan Li, Can Ye, and Alanson P Sample. 2015. IDSense: A Human Object Interaction Detection System Based on Passive UHF RFID. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2555–2564. DOI: http://dx.doi.org/10.1145/2702123.2702178

[13] Xinyu Li, Dongyang Yao, Xuechao Pan, Jonathan Johannaman, JaeWon Yang, Rachel Webman, Aleksandra Sarcevic, Ivan Marsic, and Randall S Burd. 2016. Activity recognition for medical teamwork based on passive RFID. In *RFID (IEEE RFID), 2016 IEEE International Conference on*. IEEE, 1–9.

[14] Omar Oreifej and Zicheng Liu. 2013. HON4D: Histogram of oriented 4D normals for activity recognition from depth sequences. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2013), 716–723. DOI: http://dx.doi.org/10.1109/CVPR.2013.98

[15] Donald J Patterson, Dieter Fox, Henry Kautz, and Matthai Philipose. 2005. Fine-grained activity recognition by aggregating abstract object usage. In *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*. IEEE, 44–51.

[16] Matthai Philipose. 2005. Large-scale human activity recognition using ultra-dense sensing. *The Bridge, National Academy of Engineering* 35, 4 (2005).

[17] Ronald Poppe. 2010. A survey on vision-based human action recognition. *Image and vision computing* 28, 6 (2010), 976–990.

[18] Emmanuel Munguia Tapia, Stephen S Intille, and Kent Larson. 2004. Activity recognition in the home using simple and ubiquitous sensors. In *International Conference on Pervasive Computing*. Springer, 158–175.

[19] Roberto L Shinmoto Torres, Damith C Ranasinghe, Qinfeng Shi, and Alanson P Sample. 2013. Sensor enabled wearable RFID technology for mitigating the risk of falls near beds. In *RFID (RFID), 2013 IEEE International Conference on*. IEEE, 191–198.

[20] Tim Van Kasteren, Athanasios Noulas, Gwenn Englebienne, and Ben Kröse. 2008. Accurate activity recognition in a home setting. In *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 1–9.

[21] Shrenik A Vora, William M Mongan, Endla K Anday, Kapil R Dandekar, Genevieve Dion, Adam K Fontecchio, and Timothy P Kurzweg. 2017. On implementing an unconventional infant vital signs monitor with passive RFID tags. In *RFID (IEEE RFID), 2017 IEEE International Conference on*. IEEE, 47–53.

[22] Shugang Zhang, Zhiqiang Wei, Jie Nie, Lei Huang, Shuang Wang, and Zhen Li. 2017. A Review on Human Activity Recognition Using Vision-Based Method. *Journal of healthcare engineering* 2017 (2017).