Object Hallmarks: Identifying Object Users Using Wearable Wrist Sensors

Juhi Ranjan and Kamin Whitehouse
University of Virginia
Charlottesville, Virginia (USA)
{juhi,whitehouse}@virginia.edu

ABSTRACT
In order for objects to perform personalized or contextual functions based on identity, they must solve what we call the *object user identification* problem: understanding who is actually using them. In this paper, we propose a new technique that uses data from wearable wrist sensors to perform object user identification. We hypothesize that objects have unique hallmarks that are imprinted in the hand gestures of its users. By detecting the presence of an object's hallmark in the wrist sensor data, we can identify who used the object. We evaluate this concept with a smart home application: recognizing who is using an object or appliance in a multi-person home by combining smart meter data and wearables. We conduct three different studies with 10 participants: 1) a study with scripted object use 2) a study with high-level tasked activities and unscripted object use, and 3) a 5-day in-situ study. These studies indicate that our approach performs object user identification with an average accuracy of 85-90%.

Author Keywords
Wearable Devices; Energy Apportionment; Object User Identification; Smart watch; Fitness Tracker

ACM Classification Keywords
H.1.2. User/Machine Systems: Human information processing

INTRODUCTION
Context-aware computing is now an integral part of many commercial 'smart' products. Some examples of the contexts used are home occupancy [19], road traffic [24], and global position coordinates [7]. Identity of a person is also an important context for many ubiquitous computing applications. Some bathroom scales can recognize the person stepping on them to provide long-term weight and other body measurement trends [2]. Similarly, automobiles can customize the seat and mirror positions based on the driver's identity. If a home could recognize who was using the lights, appliances, or water fixtures, it could give personalized activity reports or energy feedback to its occupants. In order for objects to perform personalized or contextual functions such as these, they must solve what we call the *object user identification* problem: understanding who is actually using a given object.

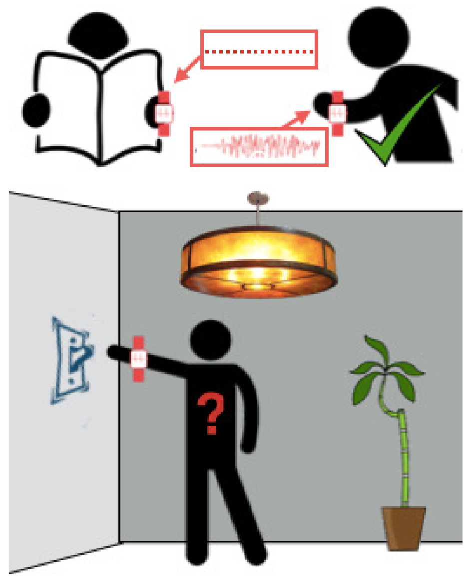

Figure 1. When an object is used, the object user is identified as the person making a hand movement containing the object hallmark. The hand is monitored using wearable devices having IMU sensors, such as fitness trackers (Image ©Juhi Ranjan)

Many techniques have already been designed to solve this problem. Some objects, such as computers or smartphones, identify a user based on a passcode or fingerprint. Other techniques use RFID tags to detect when a person's hand is near an object [4, 20], thermal cameras to detect when a person wearing a unique thermal tag is near an object [11], or room-level tracking systems to detect when a person is in the same room as an object [22, 21]. Other objects that have embedded sensors can recognize the object user based on the unique way in which an object is touched [13] or held [5], which is referred to as 'Object Use Fingerprinting'. Most of the solutions rely on instrumenting objects with additional sensors. Therefore, until every object in the world is fully instrumented, there is not likely to be a "silver bullet" that solves this problem for all object types.

In this paper, we present an object user identification approach that contrasts the approach used in Object Use Fingerprinting. Instead of using an object's sensors to detect

ACM 978-1-4503-3574-4/15/09...$15.00.
http://dx.doi.org/10.1145/2750858.2804263

the unique fingerprint of its user, we use the potential users' wrist motion sensors to detect the unique fingerprint of the object used. Since objects have a unique combination of interface, orientation and location within a space, we hypothesize that object usage hand gestures are unique *object hallmarks*. We explore the idea that the use of an object imprints a unique and identifiable hallmark in the hand movement of its user. By searching the sensed hand gestures of potential object users for the object hallmark, we can identify who used the object. Smart wrist wearable devices such as fitness bands and smart watches, often contain accelerometers, gyroscopes, and magnetometers, referred to as Inertial Measurement Unit, which can sense these object hallmarks. Increasing adoption of commercial wrist wearables, has made sensing a person's hand movements, a more practical approach than instrumenting every object in a space with sensors.

However, recognizing object-use gestures such as the opening of fridge door, is non-trivial because people often perform similar actions throughout the day. To reduce false positives, we only scan the hand movement for a specific object hallmark, at the time of the object's use. The person who makes the hand gesture that most closely resembles the object hallmark is identified as the object user. We leverage existing research in IMU signal processing to implement a simple yet novel technique to identify an object user. Our approach complements existing approaches by recognizing object users under the following conditions: 1) object use can be detected 2) a specific hand gesture is required to use the object, and 3) only a small number of people could possibly have used the object.

To test our hypothesis, we use the concept of apportionment in a home - where the identity of an object user is important from the perspective of attributing the object's energy usage to the user. We instrument 16 objects in a home, including as lights, water fixtures, and major appliances. We perform a series of three feasibility studies with decreasing levels of constraints, ending with a single person in situ study. Some of the research questions we intend to answer with these studies are: whether object hallmarks are unique and identifiable, if object hallmarks are person-independent, and how real world tasks affect the object user identification accuracy?

First, we performed a scripted study in which 10 study participants performed specific object usage. Our results show that we can classify object hallmark with 95% accuracy when the objects are used in isolation. Then, we asked 5 groups of two individuals (the same 10 participants) to complete a list of 80 real-world tasks within two hours. This resulted in a large number of fixture usages in quick succession, which could change the nature of the gestures and also creates a challenge for the system since both participants were always moving. Despite the high frequency of object usage, we were able to identify the correct individual in 85% of the total 986 object usage events. To evaluate a more realistic scenario in which people spend a large fraction of the day resting or still, we collected 30 hours of in-situ data from a single participant who lived in the test home for 5 days. We used this data to emulate a multi-person home and the results show that our approach can correctly identify the object user in 90% of the total 3378 object usage events in a 2-person scenario and 84% accuracy for 22587 objects in a 3-person scenario.

RELATED WORK

There are many different approaches for recognizing who is using an object. For example, Object Use Fingerprint research shows that it is possible to recognize an individual based on how they touch objects by equipping them with pressure sensors or other smart surfaces [13]. Chang et. al use accelerometers embedded in a television remote control or mobile device to identify household members, based on the unique way each person uses the remote [5, 26]. They show that accelerometers, touch screens and software keyboards, can be used to differentiate different test subjects based on the unique interaction characteristics of each subject. These approaches work well when an object contains the sensors required to uniquely differentiate object users but does not extend as easily to other objects such as home appliances, light switches, or water fixtures.

Other proposed methods recognize an object user in a house by instrumenting all objects or fixtures with RFID tags[4, 20]. A specialized wrist device consisting of an accelerometer and/or a RFID reader on it is worn by the occupants. The RFID reader tells us which person is touching an object, and the accelerometer characterizes how the object is being used. Cheng et. al use coordinate level tracking in a study and apportion energy usage for users in an office building [6]. To determine which person is using an object or appliance, they use a proximity detection system where users carry a magnetic beaconing system which are detected by special receivers near appliances. The identification policy is simple - assign the energy used to the nearest person detected. While this is expected to be a very accurate method of determining the object user, it is also an expensive technique which requires carrying specialized wearables at home. A similar approach is used by HeatProbe, which uses a thermal camera to detect person presence and appliance usage, but doesn't identify users individually. Our approach mines sensor data from commercially available smartwatches worn by people, without the need to install special hardware on objects, so long as there is an object monitoring system such as NILM based systems that can detect when an object is used [9, 16, 14, 18, 23, 27, 3].

Hay et al. used access logs to infer the occupants of a building working at different shifts in the day [10]. They explored the usage of different apportionment policies to break down electricity consumption among the different occupants of the building. The problem with this approach is that it cannot be used to pinpoint the actual object usage to any occupant, and is therefore not useful for any application other than energy disaggregation among people. Lee et al. use tracking systems to monitor people's movements in the rooms of a house [15]. They identify the occupants' individual bedrooms based on their frequency of movement to the rooms, and assign all object usage in those rooms to the person. In shared spaces, they use tracking to identify people present in the room when an object is used. If multiple people are present in the room, they simply split the energy usage equally between all people present in the room. Therefore, they only perform partial object user identification in a home. In our previous work [21], we explored the idea that we can use a heuristics augmented coarse-grained room level localization system to identify the object user and achieve results comparable to more fine-grained localization system. In general, the use of location tracking systems in homes require special infrastructure just for tracking people and/or a long learning time, which makes the system more expensive for widespread adoption.

OUR APPROACH

Sensing Object Hallmarks

Since objects have a unique combination of interface, orientation and location within a space, the hand gestures for using them in terms of the hand's acceleration signature, tilt and compass direction are also unique hallmarks. In order to sense these three parameters of a hand's motion, we use the nine-axis Inertial Measurement Unit (IMU) sensor of a wrist wearable device. The triaxial accelerometer in an IMU reports values composed of linear acceleration of the hand, as well as acceleration due to earth's gravity. The tilt of a hand can be inferred by measuring the amount of gravity acting upon each of the three axes. In order to separate gravity and linear acceleration, we apply a low pass filter on the accelerometer values. As illustrated in Figure 2, if we have a acceleration vector $acc^{x,y,z}_{(t_1,t_2)}$, and a gravity vector $gra^{x,y,z}_{(t_1,t_2)}$, then linear acceleration vector $linA^{x,y,z}_{(t_1,t_2)}$ can be obtained by $linA^{x,y,z}_{(t_1,t_2)} = acc^{x,y,z}_{(t_1,t_2)} - gra^{x,y,z}_{(t_1,t_2)}$

Data from accelerometer, gyroscope and magnetometer can be merged to infer the 3-D compass direction of the hand. In our work, we used pre-processed direction information available from Android Wear API [1].

Let the sensor data between time t1 and t2 from the wearable, be represented as

$sen_{(t_1,t_2)} = (acc^{x,y,z}_{(t_1,t_2)} , gra^{x,y,z}_{(t_1,t_2)} , orient^{x,y,z}_{(t_1,t_2)})$

where, $acc^{x,y,z}_{(t_1,t_2)}$ is the accelerometer data along x, y and z axes between time t1 and t2, $gra^{x,y,z}_{(t_1,t_2)}$ is the gravity data along x, y and z axes between time t1 and t2, and $orient^{x,y,z}_{(t_1,t_2)}$ is the hand's 3-D compass direction along the planes xy, yz and xz.

System Requirements

Our main hypothesis of using object hallmarks to detect object user, depends on the ability to sense a hand's movements in terms of linear acceleration, and two types of orientation - tilt with respect to gravity, and orientation in a 3-D fixed world reference. While linear acceleration and tilt w.r.t. gravity can be obtained by an accelerometer alone, orientation in 3-D space requires more sensors. Android Wear merges accelerometer, gyroscope and magnetometer (together referred to as an Inertial Measurement Unit or IMU) data to establish a 3-D fixed world orientation system, and compute the hand's orientation in it [1]. Our system can be used with any wrist worn device having a nine-axis IMU and a processor that can support sampling three sensors at 33Hz each. We don't expect this sampling rate to be a bottleneck, as most commercial wearables are capable of higher sampling rates. Most commercial wearables such as, fitness trackers and smart watches have Bluetooth or BLE or Wifi that lets them communicate raw data to a more resourced device such as a laptop or a smart phone. To determine if a person is in the same space as the objects, we expect the smart phone to have GPS or cell-tower based location information.

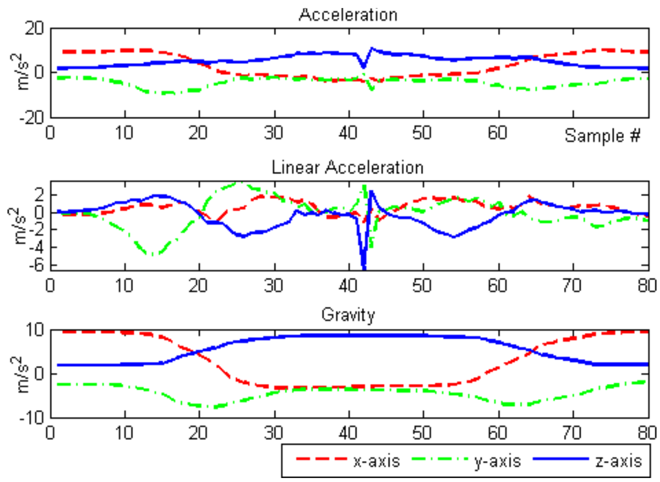

Figure 2. Total acceleration across the x, y and z axis is decomposed into gravity and linear acceleration along the three axes. The gravity component indicates how a hand is oriented, and the linear acceleration defines the movement of the hand

Overview of Object User Identification

To perform object user identification, our approach co-relates the timestamp of the object usage to the hand gestures made by the possible users of the object at the same time. The main insight behind this approach is based on the observation that object usage gestures are unique and repeatable hallmarks. Therefore when an object is used, the person making a hand gesture containing the object hallmark, can be uniquely identified.

In the rest of this section, we discuss how preliminary filtering is performed to determine the subset of possible object users, from the set of people in physical proximity to the object. We also discuss the method to identify the object user from the set of possible object users, by detecting the object hallmark in their hand gestures.

1. *Creating Reference Dataset*: A dataset of object hallmarks of each possible object user is recorded for reference.
2. *Determining Possible Object Users*: A set of possible object users are identified by determining people whose hands were moving at the time of an object usage.
3. *Identifying Object User*: If more than one person is active during an object usage, we process the sensor data further to determine which person made a hand movement that most closely resembles the object gesture in the training dataset.

1. Creating Reference Dataset
Every person has their own 'fingerprint' of operating an object [13]. Conversely, the object hallmarks imprinted in the hand gesture of every person will be different as well. In the current version of the system, we reference object hallmarks of every person for all monitored objects from a pre-recorded dataset to perform object user identification.

The data collection for creating the reference dataset, is a one-time activity undertaken by every possible object user. In this stage, all potential object users are expected to record a reference object hallmark for each monitored object. To simplify the process of collecting and annotating object gesture, we follow a 'no-motion rule' before and after every object usage. In this, the object user is expected to rest the hand for five seconds before and after using the object. This ensures that the data recorded by the wearable sensor around the time-of-use of every object, only contains a single and clean object hallmark.

$training^{O}_{p_x}$ is the reference dataset for person p_x for each monitored object o_i in the set of all monitored objects O.

$$training^{O}_{p_x} = (hallmark^{o_1}_{p_x}, hallmark^{o_2}_{p_x}, .., hallmark^{o_n}_{p_x})$$

2. Determining Possible Object Users
When an object is used, there may be multiple people present in its physical proximity. We use the following criteria to determine potential object users for an object usage: at the time the object is used, a significant hand movement must have been made by the person. This step is based on the observation that people are resting or inactive most of the time [8]. By filtering out the inactive people, we are left with a subset of fewer candidates for the object's user. To determine significant hand movement, we extract sensor data from the smart watch around the time that the object is used. Then we separate the linear acceleration components from the total acceleration using a low pass filter. Finally, we calculate if the linear acceleration is significant enough to label the person as a potential object user.

Let $e_{(o_i,t)}$ be a object usage event, where o_i is a monitored object, and t is the timestamp when the object is used, as reported by the object monitoring system. Next, the system extracts the sensor data from the wearables of the individuals in the house. If there is perfect time synchronization between the wearables and the object monitoring system, then we could simply extract the sensor data captured at time t. However, different systems are often not perfectly time-synched. Also, there may be some delay in detecting object usage due to an inherent signal processing methodology of the object monitoring system. Therefore, the sensor data must be processed for a larger time window around the reported object usage time.

Let the uncertainty in the time difference between the wearable and the object monitoring system be represented by td. Larger the time difference uncertainty between the systems, larger will be the time window of sensor data that will be extracted from the wearable. Therefore when a object usage event $e_{(f,t)}$ takes place, we extract sensor data $sen_{(t-td,t+td)}$ from the wearable.

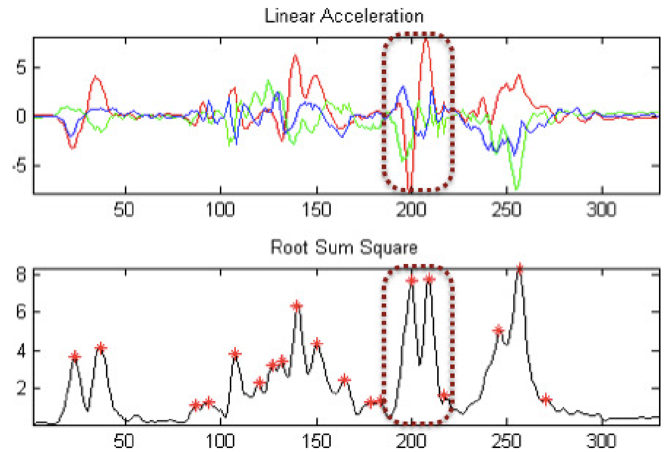

Figure 3. Segmentation: Linear acceleration across x, y and z axis are filtered and combined using Root Sum Square (RSS). Peak detection is performed on RSS to determine bursts of acceleration. Sensor data of a fixed window size around each peak is used for feature extraction

Having extracted the linear acceleration for all the people at home, we now determine the people who are active between the time $t - td$ to $t + td$. To determine if a person is at rest at the time that an object is used, we measure the standard deviation of the linear acceleration, and see if it exceeds a certain threshold value:

$$stdDev(linA^{x,y,z}_{(t-td,t+td)}) > THRESHOLD$$

If there is only one person active in the duration that the object is used, we identify the same person as the object user. However, if multiple people are active, we further process their sensor data to make this determination.

3. Identifying Object User
To identify the object user from the set of possible object users, we search for the object's hallmark in their sensor data. The person performing a gesture that most closely matches the object hallmark is identified as the object user. To do so, we first segment the sensor data using a fixed windows around peaks of linear acceleration. We then perform feature extraction for each data segment, and then compute its distance to the person's reference object hallmark, recorded earlier.

Segmentation of the sensor data is based on the the observation that every time a hand interacts with an interface, there is a small pause (a few milliseconds) before and after the interface is used. This results in a peak in the acceleration. Therefore, to segment the data, we detect the peaks of acceleration, and characterize them by performing feature extraction.

To detect the peaks, we first smooth the linear acceleration along the three axes using Exponential Moving Average filter, and merge them using Root Sum Square (RSS). We then detect local maxima in the RSS vector.

$$RSS_{(t1,t2)} = \sqrt{(linA^{x}_{(t_1,t_2)})^2 + (linA^{y}_{(t_1,t_2)})^2 + (linA^{z}_{(t_1,t_2)})^2}$$

The peak detection on the RSS vector returns the timestamp pk (precision in milliseconds) of each peak, such that

$pk \in peaksRSS = detectPeaks(RSS_{(t-td,t+td)})$, such that $t - td < pk < t + td$

Figure 4. Experimental Setup: (L-R) The fridge, freezer and microwave face the same direction, and have doors that open the same way. The dishwasher door is hinged at the bottom. The bathroom hot and cold faucets pull out from left to right, and right to left respectively. The kitchen sink hot and cold faucets both turn in counter-clockwise direction. Light switches for kitchen and living room are right next to each other. The x, y, and z axis of a LG android smart watch are annotated

Figure 3 show the linear acceleration data of a person opening a microwave door. To open the door, a person places the hand on the door and pauses slightly before pulling the door. When the door opens, he holds his hand firmly to stop the door, resulting in an equal and opposite force. This gesture shows up as adjacent peaks of equal magnitude in the RSS vector.

To perform feature extraction, we process sensor values from a window of fixed size (ws) around each of the peaks pk in $peaksRSS$. Features of each peak of sensor data $features_{pk}$, is characterized by standard parameters - mean, median, standard deviation, 25th percentile and 75th percentile of $(sen_{(pk-ws,pk+ws)})$

For each peak $pk \in peaksRSS$ in the sensor data of a potential object user p_x, we determine the distance $dist_{pk}$ of the peak's features $features_{pk}$, to the object o_i's hallmark $hallmark^{o_i}_{p_x}$ in the reference dataset $training^{O}_{p_x}$.
$dist_{pk} = distance\_function(features_{pk}, hallmark^{o_i}_{p_x})$, where $distance\_function$ is a euclidean distance function. Further, for each potential object user, we select a candidate object usage hand gesture as the peak pk with least $dist_{pk}$ to the object hallmark. From the set of candidate hand gestures from all potential object users, the individual having a hand gesture with the smallest value of $dist_p$ or, the best resemblance to the object hallmark, is identified as the object user.

EVALUATION

Experimental Setup

To test our hypothesis, we used systems that could monitor the use of objects in homes, and smart wrist devices to sense hand gestures. We instrumented 16 major appliances, light and water fixtures in a two-bedroom apartment. The relative locations of all objects are shown in Figure 5.

In the absence of any commercially available Non Intrusive Load Monitoring systems at the time this study was conducted, we placed direct sensors on all objects to monitor their usage. We used magnetic reed sensors on appliances with hinged doors, such as Fridge (F), Microwave (M), Freezer(Fr) and Dishwasher(D) (Figure 4). We also

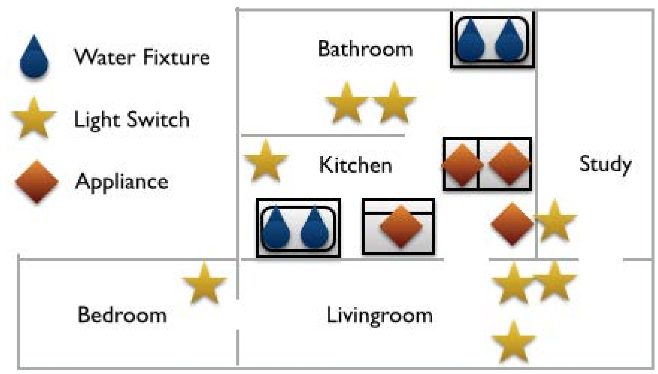

Figure 5. Object Layout: All the major appliances, light and water fixtures of a two-bedroom apartment were instrumented for the study. Quite often, two objects with similar interfaces are co-located, such as the kitchen sink hot and cold faucets

used magnetic reed sensors to instrument the bathroom hot (BH) and cold (BC) faucets, as well as the kitchen hot (KH) and cold (KC) faucets. The reed sensors were plugged into HOBO UX90-001 data loggers [25]. We used light on/off sensing HOBO UX90-002M data loggers [12] to sense the Lights(L) in the house. These loggers are installed right next to the light bulbs, and have a programmable threshold for light intensity to detect when a light is on.

To sense the hand gestures, we used LG G watch [17], which runs on Android Wear platform. We wrote an app for the watch which collected IMU based sensor data (accelerometer, gravity and orientation vector) at 33 Hz and transmitted it to its paired smartphone. The smartphone had a listener app which logged the received data on the phone. For each sensor data sample, the app recorded the timestamp, the sensor type, and the set of raw sensor values along x, y and z axes. The app assigned a sequence number to each received sample as well to detect packet loss. The sequence number helped us detect a 30 second ($\tilde{3}000$ data points) data loss during one of our trials.

We invited 10 participants (7 F, 3 M) to use the objects of the instrumented home. They participated in two different studies - a 40 minute individual scripted study, and a 2 hour real world task based study performed in five groups of two people each. We also collected in-situ data for 30 hours (6 hours every day, for five days) from a single participant who lived in this house.

Study Design

To evaluate the feasibility and inform the parameters of our approach, we structured the experiments into three different studies:

- Scripted Study
- Real World Task based Study
- In-situ Study

Scripted Study

The main premise of our approach relies on the ability to detect and differentiate between hallmarks of different objects. Therefore, to ensure feasibility of our approach, we first evaluated our approach in a scripted study. In this study the participants were instructed exactly which objects to use in a predetermined, scripted order. The aim of this study was to answer the following questions:
1. Do people use objects using gestures that are distinguishable and learnable?
2. Is it possible to learn the object hallmarks in a person-independent manner?
3. Can we differentiate between hallmarks of objects with the same interface? For e.g. the fridge and the freezer doors have similar handles.

The procedure of this study required each participant to follow a script that made them operate each of the 16 objects in the house in a fixed order. The participants were free to operate each object in a manner they liked. All the objects in the home were labeled for easy reference. To discourage participants from mechanically using the objects in the same way, we designed the script to have consecutive object usages in different rooms. They were asked to perform the entire object usage script ten times, therefore logging 160 object usages each.

In order to obtain clean gesture data for this study, the participants were instructed to pause for 5 seconds before and after using each object. This made annotation of gesture data extremely easy, and we could automate the process. Every time that an object was used, a +/- 5 seconds segment of sensor data from the object usage timestamp was extracted. The only gesture in this time window was the one used for operating the object. Root sum square based peak detection was used to automatically detect the gesture and annotate it. The processed sensor data was presented to the researcher who now simply had to validate the annotated gesture. This saved us a lot of time as opposed to using a camera based annotation system, which requires more manual effort for annotation. For ground truth on object usage timestamps, an observer manually noted the start and end time for every repetition of the script.

Real World Task based Study

To determine the accuracy of our object user identification approach, we conducted a two hour study in a home environment, with pairs of individuals operating objects for performing real world tasks.

Real life object usage is very noisy, making the sensing of hand movements challenging. In a real life scenario, we can imagine that people do not rest their hands before and after using an object. In fact, object usages are preceded and succeeded by different hand actions. People also vary the force with which they use objects at different times. A object usage could comprise of a complex combination of actions, such as open door, close door and push buttons in case of a microwave. Also, the object monitoring system and the wearable can have non synced time stamps.

To encapsulate aspects of real life situations and interaction between multiple people when they share the same set of objects between them, we wanted to ensure that the following conditions were present in the study:

- Possible simultaneous usage of similar objects (such as fridge and microwave)
- No simultaneous usage of same object
- Performing actions before and/or after an object usage (such as opening the kitchen faucet, washing hands, and closing the kitchen faucet)

Therefore, to evaluate our approach in the presence of noise, we designed this study to mimic the real world object usage. We organized the 10 participants into 5 groups of two individuals each. Pairs of participants were each given an identical list of tasks that involved using objects similar to real life scenarios. Some examples of tasks given to them are, take out food from the fridge, wash face in the bathroom sink, take a cup of water from the kitchen sink and heat it in the microwave, wash dishes in the kitchen sink. The participants were free to decide how, in which order and when they wanted to perform each of the tasks in the list, within the two hour limit. The list contained sufficient number of tasks to ensure that each participant interacted with objects at least 80-90 times, within the two hours.

For ground truth, we had an observer making real time entries about the object usage of each individual, using a logging tool that was time-synced with the sensor loggers. To perform object user identification for the tasked study, we used the annotated gestures from the scripted study for the reference dataset. Since, the script was repeated ten times, we obtained ten reference object hallmarks for each object in the home.

We aimed to answer the following questions with this study:

1. How accurately can we identify the object user in objects used for real world tasks?
2. How much reference data is required to be able to identify individuals accurately?
3. How does the time uncertainty between the object usage reporting system and the watch affect the accuracy of recognizing the correct individual?

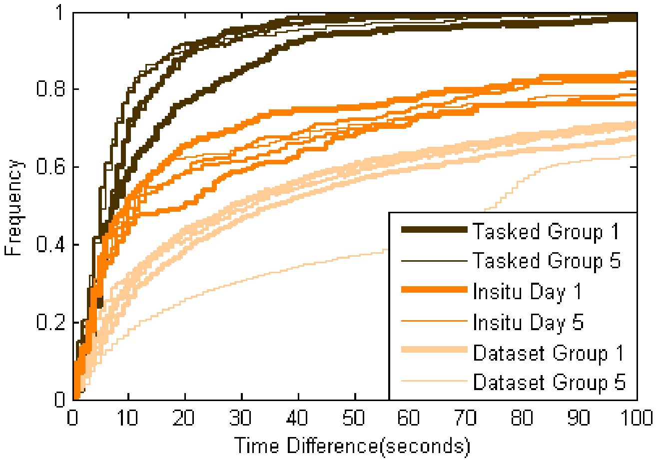

Figure 6. CDF of time differences between object usage: 50% of the object usage in the tasked study happened within 5 seconds of each other

Additionally, we wanted to use this study to perform sensitivity analysis for the technical aspects in our approach:

4. How large should the window size be around each peak of the root sum square of linear acceleration?
5. How does an increase in the time difference uncertainty between object monitoring system and the wearable affect the identification accuracy?

On an average, ~200 object events were performed within 2 hours by each group. To put this into perspective, we compare this frequency to the object usage frequency from an in situ dataset from one of our previous studies [21]. The dataset contained object usage timestamps of 5 groups of 2 individuals, who lived in an instrumented home for 10-12 days. As can be seen in Figure 6, this actually resulted in a high frequency of object usage in our study, with about 50% of the object usages happening within 5 seconds of each other. Whereas, according to the previous dataset, only 15-20% of the object usages happened within 5 seconds of each other.

The actual number of object usages varied with each participant, because the interpretation of generic tasks was left at the discretion of the participants. Across all the five studies, the total number of object usages was 986. Per participant, the minimum number of object usages made was 82, maximum was 108, and median was 98.

In-Situ Study

Although the task based study encapsulated a lot of noise from real life scenarios, we wanted to understand how well our approach worked in presence of noise introduced in an uninstructed, in-situ object usage in a home. Although we wanted to collect data from a home with multiple residents, we were limited by the fact that in a real home setting, it is extremely difficult to collect ground truth on who is using an object without the use of privacy invasive technology, such as cameras.

Therefore, we collected in-situ data from a single-resident home and used this to simulate multiple person (2, 3, 4 and 5) home settings. We asked a participant to live in the instrumented house for five days, and wear the smartwatch for 6 hours every day - three hours in the morning, before leaving for work, and three hours in the evening, after coming from work. The participant was given no specific instructions on what objects to use and what activities to perform in the home. We collected 10 datasets of 3 hours each, consisting of 515 in-situ object usages in total. The main research question that we aimed to answer using this study was:

1. How does the system perform in the presence of more than two individuals in a home?

To simulate multi-person scenarios, we treated each dataset's sensor data as the data of another person. Since there was only one person living in the house, we knew for every object usage in the dataset, who the object user was.

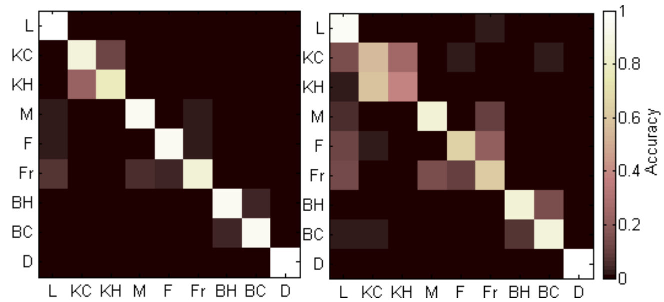

Figure 7. Gesture classification: Confusion Matrix of 10-fold cross validation in scripted study show (a) 95% accuracy for a person dependent gesture model, (b) 88% accuracy for person independent gesture model

RESULTS

We first discuss the results of the feasibility study for determining if object gestures are differentiable and learnable hallmarks. Then we analyze the object user identification accuracy using our rule based approach in the real world task based study and in the in-situ study. Finally, we perform sensitivity analysis to determine the technical parameters of our approach.

Feasibility Study

We used the annotated gesture data from the scripted study to validate the feasibility of using object gestures as hallmarks for object user identification. The metric used in this section is *Classification Accuracy*, which is defined as the sum of correct classifications divided by the total number of classifications. The annotated gesture dataset has an object id for every object gesture, and the mean, median, standard deviation, 25^{th} percentile data point and 75^{th} percentile data point of the three axes of linear acceleration, gravity and orientation as the feature set. For each of the 10 participants' data, we performed 10-fold cross validation on the object gestures, using Nearest Neighbor method. As shown in Figure 7(a), the per-person based classifier achieved 95% classification accuracy using the 15 features calculated on the annotated sensor data. This suggests that gestures used to operate objects can be differentiated. This also suggests that for a given object, the gestures required to operate the object are consistent for the same person.

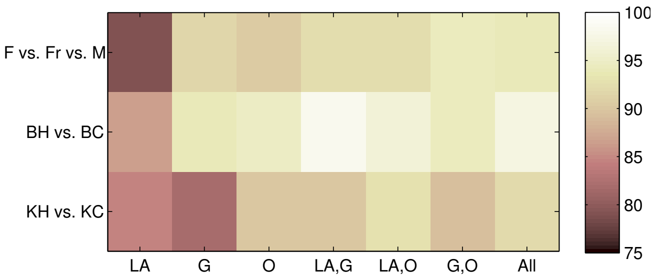

Figure 8. Classifying between objects with similar interfaces: Different features are evaluated to determine how gestures are differentiated between sets of co-located and similar objects

The main reason why we care about being able to identify the gestures used for different objects, is so that when multiple objects are used simultaneously by different individuals in a home, we can determine which individual used which object. For example, if a person A uses a fridge while another person uses the bathroom sink, the ability to tell apart the different objects hallmarks is crucial to the identification of the correct individuals using each object.

To determine if object hallmarks are person dependent or not, we created test cases, where we trained using annotated gesture data from 9 people and tested on the 10^{th} person. As shown in Figure 7(b), on an average, this method was able to differentiate between gestures with 88% accuracy. In Figure 7(b), the precision and recall for kitchen hot and cold faucets (KC and KH) was quite low, with one getting classified as the other, quite often. This was also true for the fridge and freezer, and the bathroom hot and cold faucets. The dishwasher, microwave and lights performed the best in the person-independent classification. Given that certain object hallmarks performed well in the person independent analysis, while others performed poorly, it is hard to conclude whether object hallmarks can be learnt in a person independent manner.

Figure 7(a) showed that a person's gestures for co-located and similar objects are differentiable as hallmarks. We further analyze how each of the three IMU sensor's features: Linear Acceleration (LA), Gravity (G) and Orientation (O), and their combinations contribute to the classification between the following sets of similar objects:

1. Fridge vs. Freezer vs. Microwave (F/Fr/M): All three of these objects have doors that open the same way, and the appliances are co-located and placed against the same wall in the kitchen.
2. Bathroom Hot Faucet vs. Cold Faucet (BH/BC): Both the faucets, are co-located on the same sink, and have handles for operating them. However, the hot faucet is turned right to left, and the cold faucet is turned left to right, to open them.
3. Kitchen Hot Faucet vs. Cold Faucet (KH/KC): Both the hot and cold faucets in the kitchen, have knobs that are turned counter clockwise for the faucets to be opened.

The accuracy of using different features in classifying between sets of similar objects are shown in Figure 8. Results show that orientation (O) features are the most accurate and linear acceleration features are the worst for differentiating between F/Fr/M. This is because all three objects have similar interface - handle, that needs to be pulled outward in the same way. Thereby resulting in similar linear acceleration signatures.

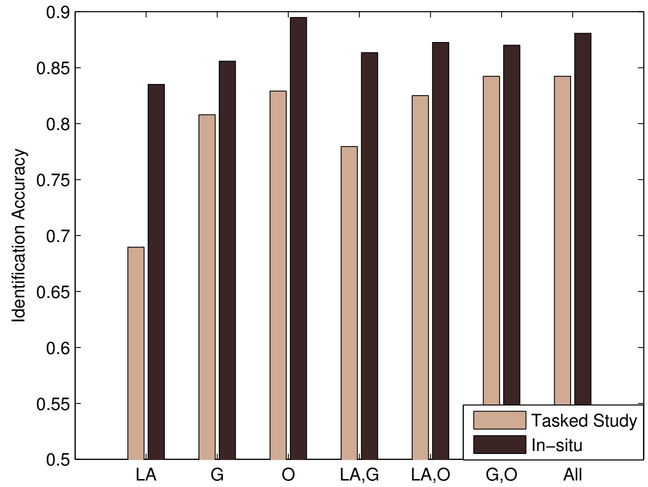

Figure 9. Different sensing features are evaluated to determine person identification accuracy. Linear Acceleration (L) based features do not perform as well as the orientation features (Gravity (G) and 3-D orientation (O))

In differentiating between bathroom hot and cold faucets, the *Linear Acceleration,Gravity (LA,G)* feature set performs the best. Since the hot faucet handle moves from left to right, while the cold faucet handle moves from right to left, we believe that linear acceleration signatures are sufficiently different between these two objects. In classifying kitchen hot and cold faucets, the gravity features creates the most confusion between the two objects. The kitchen hot and cold faucets are both knobs that turn counter clockwise to open. Therefore the hand is tilted the same way to open/close both the objects. In this case, This is further confirmed by the observation that the feature set *Linear Acceleration,Orientation (LA,O)* performs the best in classifying these objects.

Identifying the Fixture User

Given that the feasibility studies show that it is possible to have distinct and repeatable hallmarks for objects in a home, we now want to evaluate our main hypothesis, that is - if we have an object usage event, and a set of people living in a home, how accurately can we identify the object user. For evaluating our hypothesis, we use data from the real world task based study and the in-situ study. The metric used in this section is *Identification Accuracy* which is defined as the number of correct object user identified divided by the total number of object usages.

The tasked study had pairs of participants using objects in a home. We simulated two person scenarios from the in-situ data, by treating each of the ten 3-hour datasets as a separate person's dataset and performing object user identification on every combination of two datasets. This resulted in 90 test cases of 3 hours each, which was then evaluated using our approach.

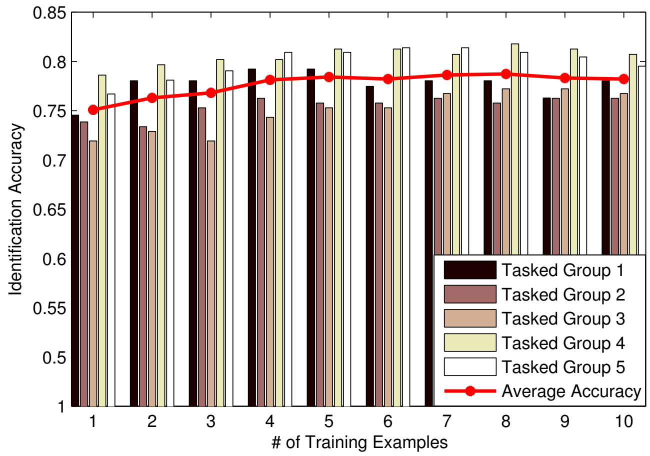

Figure 10. An average accuracy of 80% is achieved with just a single training example for each object gesture. More training examples bring incremental improvements in the accuracy

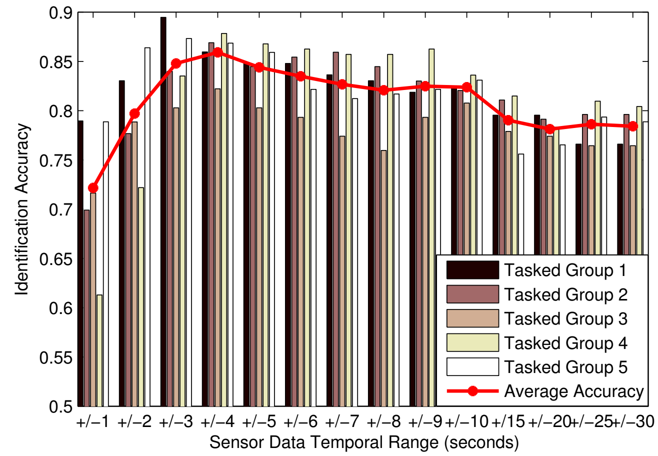

Figure 11. Highest accuracy is obtained when a sensor data size of +/- 4 seconds around the object usage timestamp, is selected. The likelihood of an overlapping event from another person increases with a larger temporal range of sensor data. Accuracy is low for smaller sensor data size because the object monitoring hardware had a 1-3 seconds difference with the time on the watch

Since our approach uses features from the three sensors in an IMU (accelerometer, gravity and orientation), we varied the sensor features in the distance function to study which features performed the best in terms of identification accuracy. In Figure 9, we can see that our approach achieves 90% accuracy in the in-situ study based simulations using only orientation parameters in the distance function. We achieve 85% accuracy for the tasked study, using gravity and orientation parameters in the distance function. The linear acceleration features in the distance function, perform the worst for both the in-situ and the tasked study. This may be explained in part that there might be variations in the force with which a person uses the same object at different times. In general, the orientation features achieve higher accuracy. This is possibly because the hand's orientation, in terms of the gravity and the compass direction in space, remains relatively constant, as the hand is constrained to face a certain way to use an object.

An important factor in our approach is the amount of training data required for the system to identify object users accurately, when an object is used. An ideal gesture recognition system should be able to learn the most frequently occurring distinct gesture as hallmark for each object in an unsupervised manner. However, this is a hard problem, because object usage in real world is noisy and it is hard to separate the signal from the noise. Therefore, we assume that every individual in the home needs to use the objects in the home at least once in a training mode, before being able to use the system.

Figure 10 shows our system's identification accuracy on varying the number of training examples used. With just one set of training data for all the objects in the home, our system can identify the correct person 80% of the time in the tasked study. With additional training data, the identification accuracy improves incrementally. This is possibly because the training data was used from the scripted study which probably did not include all possible variations of people's object usage gestures.

One of the limitations of the tasked study was that it was conducted with sets of only two participants at a time in the test home. This was partly because of the smaller size of the home, and the fact that having three or more people using objects simultaneously around the home would be harder to track for ground truth. We were interested to know how the system would perform in the presence of more individuals in a home. Therefore, we used simulations from the data collected from the in-situ study to create scenarios with 2, 3, 4 and 5 people using objects at the same time. The number of object events for 2, 3, 4 and 5 people simulations were 3378, 22587, 167196, and 673857 respectively. The average person identification accuracy was comparable for 2 and 3 people scenarios, and measured at 90% and 84%. For the 4 person scenario, the average accuracy came down to 65% and was similar for the five person scenario at 63.8%.

Sensitivity Analysis

In this section, we discuss how varying the technical parameters such as temporal range of sensor data, and the window size around each peak in the segmentation section affects the accuracy of the system. The metric used in this section is *Identification Accuracy* which is defined as the number of correct object user identification divided by the total number of object usages.

In our approach, we assume that there exists an system which records and reports every time that an object is used in a home. To know how robust our system is to the differences in time stamps between object monitoring and the wearable, it is important to how this time difference affects identification accuracy. Researchers have explored many systems for performing Non Intrusive Load Monitoring of appliances in a home. These systems are often not very accurate in terms of the reported time of object event. In such a scenario, we evaluate how well our system can perform with increasing uncertainties about the actual time of event. If the uncertainty is

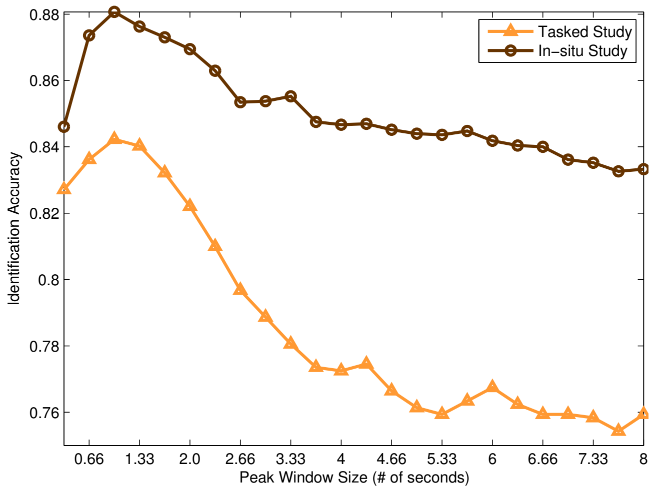

Figure 12. Accuracy was highest with a frame size of 1 second around each peak of the Linear Acceleration RSS. This may be due to the fact that time taken by people to physically interact with objects in the home is usually small

larger, we need to extract a bigger portion of sensor data from the wearables of all the occupants. The problem with this is that, If there are multiple people actively using objects in a home at the same time, the chances of identifying the wrong person increases.

According to Figure 11, extracting sensor data of window size +/-1, +/-2 and +/-3 seconds gives low accuracy because the direct sensors and the wearables had a known timestamp uncertainty of +/-3 seconds. Therefore, collecting data using a +/- 4 seconds window resulted in the highest identification accuracy. It is the largest window size that encapsulates the time difference between the two systems, as well as small enough to not have overlapping object usage event by another individual.

Once we extract sensor data from the smart watch, we perform feature extraction on sensor data segments of a fixed window size around peaks of linear acceleration. In order to identify the optimal window size of sensor data to be used for feature extraction, we varied the window size starting from 0.33 seconds (10 samples) to 7 seconds (240 samples). We observe that the highest accuracy is obtained when a frame size of 1 second (30 samples) is selected around each peak, to be be processed for feature extraction. This maybe an indication that most of interfaces of the objects present in our test homes typically required only a short interaction time.

DISCUSSION

In this paper, we investigated the hypothesis that we can accurately identify object users in a home by mining the IMU data from the wearables of home occupants. Despite a large number of simultaneous object usages by the participants (50% of object usages were less than 5 seconds apart) in our task based study, we were able to identify the correct individual with high accuracy. Based on our prior observations of how infrequently people actually use objects in homes (Figure 6), we expect accuracy of our approach to increase much more in real life.

Our system shows promise because it can start identifying individuals very accurately with only one training example per object. So a person would simply have to install an app on the wearable, use every object in the house just once to create a reference dataset, and the system would be ready to work. Since wearables are readily available in marketplaces, one of the biggest barriers to success of our approach at this point, is the commercial availability of object monitoring systems such as NILM.

Our approach is not restricted to objects within homes only. Our approach is general enough to be applied to other scenarios where there are physical interfaces, and we want to know who is using the interface, given a fixed set of people, all wearing smart wrist devices. For example, this approach can be used to detect who is currently writing on a smart whiteboard in a conference room.

Limitations

While our system shows promise, it has certain limitations too. For this approach to work, it assumes that all home occupants who use objects, wear smart wrist devices on their dominant hand. The system might identify an object user incorrectly, if the actual user is not wearing a smart wrist device. The system might also identify object users incorrectly if a person uses the un-instrumented or non-dominant hand to operate an object. Our main argument to this problem is that people mostly tend to use appliances using their dominant hand, and therefore wearing the wearable device on the dominant hand should be able to work for most objects. Finally, we concede that while our approach is expected to work accurately with objects having fixed locations in a home, we have to come up with newer ways to learn the gestures required to operate non-fixed appliances such as hand blenders and hair dryers.

CONCLUSION

In conclusion, we present our work on performing object user identification, in which we detect the presence of object hallmarks in the wearable wrist sensor to determine the person who interacted with the object. We evaluate this concept with a smart home application: recognizing who is using an object or appliance in a multi-person home by combining smart meter data and wearables. Unlike other systems, our approach does not mandate the instrumentation of any object with special sensors in the homes and can be very well integrated with NILM systems. Our results show that our approach can correctly identify the object user in 90% of the total 3378 object usage events in a 2-person scenario and 84% accuracy for 22587 objects in a 3-person scenario.

ACKNOWLEDGMENTS

We would like to thank all the participants who helped us with this study. This work is supported by the National Science Foundation under Grants 1038271 and 0845761, and the NSF Graduate Research Fellowship Program under Grant 0809128.

REFERENCES

1. Android Wear Orientation Sensor. http://developer.android.com/reference/android/hardware/Sensor.html#TYPE_ROTATION_VECTOR.
2. Fitbit Aria: Smarter scale. Better Results. https://www.fitbit.com/aria.
3. Batra, N., Kelly, J., Parson, O., Dutta, H., Knottenbelt, W., Rogers, A., Singh, A., and Srivastava, M. Nilmtk: An open source toolkit for non-intrusive load monitoring. In *Proceedings of the 5th international conference on Future energy systems*, ACM (2014), 265–276.
4. Berlin, E., Liu, J., Van Laerhoven, K., and Schiele, B. Coming to grips with the objects we grasp: detecting interactions with efficient wrist-worn sensors. In *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, ACM (2010), 57–64.
5. Chang, K.-h., Hightower, J., and Kveton, B. Inferring identity using accelerometers in television remote controls. In *Pervasive Computing*. Springer, 2009, 151–167.
6. Cheng, Y., Chen, K., Zhang, B., Liang, C.-J. M., Jiang, X., and Zhao, F. Accurate real-time occupant energy-footprinting in commercial buildings. In *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, ACM (2012), 115–122.
7. Garmin Navigation System. http://www.garmin.com/en-US.
8. Hamilton, M. T., Healy, G. N., Dunstan, D. W., Zderic, T. W., and Owen, N. Too little exercise and too much sitting: inactivity physiology and the need for new recommendations on sedentary behavior. *Current cardiovascular risk reports 2*, 4 (2008), 292–298.
9. Hart, G. W. Nonintrusive appliance load monitoring. *Proceedings of the IEEE 80*, 12 (1992), 1870–1891.
10. Hay, S., and Rice, A. The case for apportionment. In *Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, BuildSys '09, ACM (2009), 13–18.
11. Ho, B.-J., Kao, H.-L. C., Chen, N.-C., You, C.-W., Chu, H.-H., and Chen, M.-S. Heatprobe: a thermal-based power meter for accounting disaggregated electricity usage. In *Proceedings of the 13th international conference on Ubiquitous computing*, ACM (2011), 55–64.
12. UX90-002M: HOBO UX90 Light On/Off Logger with Extended Memory. http://www.onsetcomp.com/products/data-loggers/ux90-002m.
13. Hodges, M. R., and Pollack, M. E. An object-use fingerprint: The use of electronic sensors for human identification. In *UbiComp 2007: Ubiquitous Computing*. Springer, 2007, 289–303.
14. Kim, Y., Schmid, T., Charbiwala, Z. M., and Srivastava, M. B. Viridiscope: design and implementation of a fine grained power monitoring system for homes. In *Proceedings of the 11th international conference on Ubiquitous computing*, ACM (2009), 245–254.
15. Lee, S., Ahn, D., Lee, S., Ha, R., and Cha, H. Personalized energy auditor: Estimating personal electricity usage. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, IEEE (2014), 44–49.
16. Leeb, S. B., Shaw, S. R., and Kirtley Jr, J. L. Transient event detection in spectral envelope estimates for nonintrusive load monitoring. *Power Delivery, IEEE Transactions on 10*, 3 (1995), 1200–1210.
17. LG G Watch. http://www.lg.com/us/smartwatch/g-watch-r.
18. Marceau, M. L., and Zmeureanu, R. Nonintrusive load disaggregation computer program to estimate the energy consumption of major end uses in residential buildings. *Energy Conversion and Management 41*, 13 (2000), 1389–1403.
19. Nest Thermostat. https://nest.com/thermostat/.
20. Philipose, M., Fishkin, K. P., Perkowitz, M., Patterson, D. J., Fox, D., Kautz, H., and Hahnel, D. Inferring activities from interactions with objects. *Pervasive Computing, IEEE 3*, 4 (2004), 50–57.
21. Ranjan, J., Griffiths, E., and Whitehouse, K. Discerning electrical and water usage by individuals in homes. In *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, ACM (2014), 20–29.
22. Ranjan, J., Yao, Y., and Whitehouse, K. An rf doormat for tracking people's room locations. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, ACM (2013), 797–800.
23. Ruzzelli, A. G., Nicolas, C., Schoofs, A., and O'Hare, G. M. Real-time recognition and profiling of appliances through a single electricity sensor. In *Sensor Mesh and Ad Hoc Communications and Networks (SECON), 2010 7th Annual IEEE Communications Society Conference on*, IEEE (2010), 1–9.
24. Google's Self Driving Car. http://www.google.com/selfdrivingcar/.
25. UX90-001: HOBO UX90 State Logger. http://www.onsetcomp.com/products/data-loggers/ux90-001.
26. Wolff, M. Behavioral biometric identification on mobile devices. In *Foundations of Augmented Cognition*. Springer, 2013, 783–791.
27. Zoha, A., Gluhak, A., Imran, M. A., and Rajasegarar, S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. *Sensors 12*, 12 (2012), 16838–16866.