Recognizing Water-Based Activities in the Home Through Infrastructure-Mediated Sensing

Edison Thomaz[1], Vinay Bettadapura[1], Gabriel Reyes[1], Megha Sandesh[1], Grant Schindler[1], Thomas Plötz[2], Gregory D. Abowd[1], Irfan Essa[1]

[1] School of Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA

[2] Culture Lab
School of Computing Science
Newcastle University, UK

ABSTRACT
Activity recognition in the home has been long recognized as the foundation for many desirable applications in fields such as home automation, sustainability, and healthcare. However, building a practical home activity monitoring system remains a challenge. Striking a balance between cost, privacy, ease of installation and scalability continues to be an elusive goal. In this paper, we explore infrastructure-mediated sensing combined with a vector space model learning approach as the basis of an activity recognition system for the home. We examine the performance of our single-sensor water-based system in recognizing eleven high-level activities in the kitchen and bathroom, such as cooking and shaving. Results from two studies show that our system can estimate activities with overall accuracy of 82.69% for one individual and 70.11% for a group of 23 participants. As far as we know, our work is the first to employ infrastructure-mediated sensing for inferring high-level human activities in a home setting.

Author Keywords
Activity Recognition, Activities of Daily Living, Machine Learning, Infrastructure-Mediated Sensing, Health, Vector Space Models

ACM Classification Keywords
I.5 Pattern Recognition; J.4 Social and Behavioral Sciences; J.3 Life and Medical Sciences

General Terms
Algorithms, Human Factors, Activity Recognition, Smart Homes, Healthcare

INTRODUCTION
Inferring user behavior in a home environment remains an interesting and challenging goal for future computing applications. There are many compelling reasons for this, such as developing home automation systems that anticipate user needs, providing feedback to inform homeowners of resource consumption, and healthcare applications that assist with aging in place and chronic disease assessment. Most efforts in inferring user behavior at home have concentrated on recognizing activities of daily living (ADLs) for health monitoring. Examples of ADLs include eating, bathing, and dressing. We are similarly motivated to recognize such domestic activities, and consider activities specifically related to water consumption in the home. We believe that focusing not only on whether individuals can perform ADLs but also on how well they do it can be a good metric for deciding whether an individual's medical condition is becoming such a limiting factor that assistance is needed.

Researchers in ubiquitous computing have attempted to address the challenge of activity recognition in the home from different perspectives. Varied approaches from instrumenting the environment with many sensors, to using body-worn sensors to infer activities in domestic settings have been proposed. We are motivated to recognize activities in the home with a real and ecologically valid deployment of sensing, that is sustainable, low-cost and easy to install. Ease of installation is the key, as we want to deploy this system in large number of homes. For this reason, we turn to the newly introduced method of Infrastructure-Mediated Sensing (IMS) [4, 6, 19, 2].

The concept behind IMS is to use the infrastructure of a home as a sensor, whereby infrastructure we mean a house's power lines, water lines, gas lines, HVAC ducts, etc. In practice, this means creating a device that attaches and monitors the chosen infrastructure with a high degree of fidelity. Our hypothesis is that by sensing some physical phenomenon (in our case, water pressure) that is linked to the chosen infrastructure, it is possible to infer what human actions and activities are being carried out on that infrastructure.

In this paper, we leverage IMS to sense human activities for health assessment purposes. We are interested in the recognition of ADLs, but not exclusively, since there are many types of activities beyond those specified in the ADL definition [11][15] that can be leveraged for health-related inferences. IMS based on water usage is an approach that is perfectly suited to our needs, since it can be installed with little

UbiComp '12, Sep 5-Sep 8, 2012, Pittsburgh, USA.

effort in a typical home and is not disruptive of the privacy and aesthetics of the physical environment. Moreover, water usage is highly correlated with a number of medically meaningful activities, including healthy eating and hydration, hygiene, and medication compliance. However, to quantify several critical health metrics, from well-being to personal hygiene, it is imperative that we have the ability to recognize what residents are doing at home beyond simply which fixture they are manipulating. To this end, we adapt the Hydrosense work [5][3] to introduce Hydrostream, a scalable system to measure water-based events in the home.

The primary research contribution of this work is *the use of water-based, single-point infrastructure-mediated sensing combined with a vector space model learning approach for high-level activity recognition in a home setting*. The results reported in this paper focus specifically on kitchen and bathroom activities, where many important activities of daily living take place. Through two user studies, where we asked individuals to perform eleven different activities in a home-like setting, we see that it is possible to recognize high-level activities using our technique with up to 82.69% for one individual and 70.11% for a group of 23 participants. To analyze the data that we gathered from our user studies, we built a Bag-of-Word (BoW) feature representation within a Vector Space Model (VSM) classification framework [7] [16].

RELATED WORK

Human activity recognition is a research problem that has been extensively studied (cf. [1] for a recent survey). In the Ubicomp domain, activity recognition is typically based on monitoring human activities by means of sensors that are either integrated into objects or the environment, or body-worn. Activity recognition using wearable sensing is potentially more accurate than environment-based sensing since the human activities are sensed more directly. However, instrumenting the environment is often more feasible than "wiring up" its occupants.

A number of approaches for environment-based sensing for human activity recognition have been described in the literature. Smart houses, like MIT's PlaceLab [9] or the Georgia Tech Aware Home [12], are prominent examples of sensor equipped environments for human activity recognition. By employing a large number of small state-change sensors, activities of daily living (ADL), such as "toileting", "bathing" and "preparing lunch", can effectively be inferred in such environments [18]. Object-centric approaches infer activities by analyzing the way humans interact with certain objects and appliances. For example, Philipose et al. described a technique for observing activities of daily living (ADL) by identifying how residents interact with objects [21]. Based on RFID tags attached to objects of interest and an RFID-enabled glove, they show how it is possible to make activity predictions through the statistical analysis of sequences of interactions. As an alternative to RFID, in some approaches accelerometers have been integrated into objects and utensils and human activities were inferred via analyzing the way these objects were used (e.g., [13, 20]).

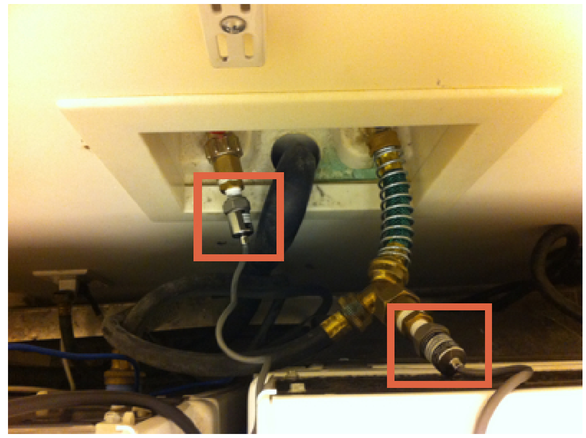

Figure 1. The two pressure sensors, for hot (left) and cold (right) water, were installed in the laundry room of the testbed home.

The most significant shortcoming of the aforementioned approaches is the large number of sensors that has to be installed for recognizing relatively simple activties. Other substantial negative implications are that these sensing approaches are perceived to be invasive from a privacy point of view, and that there are considerable installation and maintenance challenges with these techniques [8]. As an alternative, recently so-called infrastructure-mediated sensing technologies have been developed where human activities are sensed by means of single sensors attached to domestic infrastructures like the plumbing system [4], the electrical system [6, 19], or the gas supply [2]. For example, Hydrosense, a pressure-based, single-point sensing solution, has been shown to infer fixture level human activities in a home setting through water usage classification with accuracies of about 90% [3, 5]. Despite these promising results, Hydrosense succeeded primarily as a mechanism for identifying fixture-level water usage, as opposed to higher-level activities of daily living. While there is a direct mapping between certain types of ADLs and water fixture (e.g., doing laundry, running dishwasher), several important health-related activities are carried out in a single location, such as the bathroom (e.g., brushing teeth, shaving) and the kitchen (e.g., washing hands, cooking). Therefore, discriminating activities in these locations is desirable.

The work presented in this paper is the first that addresses activity recognition in a domestic setting using infrastructure-mediated sensing, which goes beyond existing, basic event detection thereby focusing on typical water-based activities.

WATER PRESSURE SENSING

Hydrostream is the name of our water-based infrastructure-mediated sensing system. It is made up of 3 components, Hydrobeacon, a hardware-software unit that probes the water system for pressure readings, Hydroserver, a cloud-based server backend for data analysis and storage, and Hydropad, an iPad-compatible application for data collection and visualization. In this work, we only employed Hydrobeacon and Hydroserver.

Hydrobeacon is architecturally identical to Hydrosense [4]

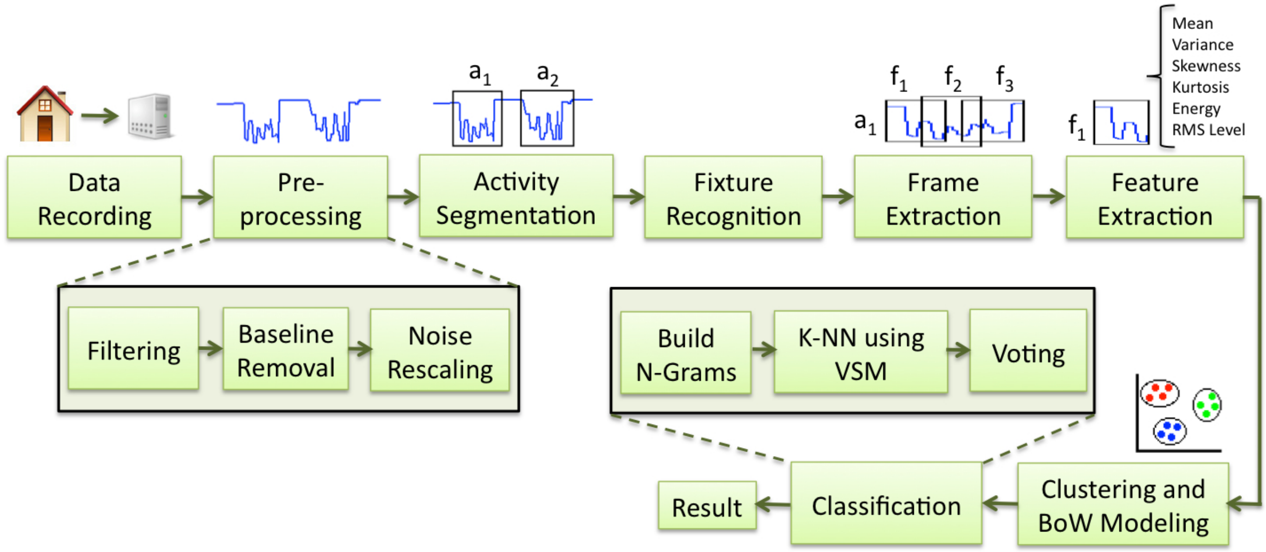

Figure 2. The activity recognition processing pipeline of our system

when it comes to water pressure sensing. At its core lies a transducer that converts pressure into voltage. The pressure sensor in our system is the MSP300 by Measurement Specialties. It can measure up to 100psi with ratiometric voltage output between 0.5v and 4.5v.

The analog signal from the sensor is converted to a digital representation through an Arduino Uno board, sampled at 200Hz. All digital pressure readings are forwarded to a computer using a standard USB interface. The data is validated and sent over to a web server in bursts of 10Kb. At the server level, a script written in PHP saves the data to a database, where each time-stamped data burst is stored in a new row. The script also calculates a moving average with the incoming data as a way to save a smaller, down-sampled representation of the pressure readings.

Similarly to Hydrosense, Hydrobeacon can be installed in any number of locations in a typical home. We have found the laundry room tends to be a good place, providing easy access to one or more water fixtures in addition to outlets to power sensors and other instruments. This is where we connected our sensors for the studies, as shown in Figure 1.

ACTIVITY RECOGNITION

We built a processing pipeline for water-based activity recognition consisting of eight stages: *(1)* data-recording and pre-processing; *(2)* activity segmentation; *(3)* fixture recognition; *(4)* frame-extraction for activities; *(5)* frame-based feature extraction; *(6)* clustering; *(7)* mapping of raw, continuous sensor data to symbolic representation; *(8)* activity classification. Figure 2 shows a block diagram of the pipeline.

An important element of activity recognition techniques in general, is the computational approach employed to identify activity patterns in the sensor data. Traditionally, sequential models like Hidden Markov Models (HMMs) and Dynamic Bayesian Networks (DBNs) have been used in the machine learning and computer vision literature to address activity recognition problems. However, the assumption of Markovian dynamics restricts the application of such sequential models to relatively simple problems with known structure of the data to be analyzed [23]. Similarly, syntactic methods like Parse Trees and Stochastic Context Free Grammars [17] [10] are not well suited to recognizing weakly structured activities and are not robust to errors in the data, such as the errors that could be caused due to occasional failures in the physical sensor equipment and due to errors and delays in transmission and storage.

Since the activities analyzed in this paper are weakly structured, we employ an alternative to traditional activity recognition techniques. Water-based activities are represented as Bag-of-Words (BoW), which are classified using a Vector Space Model (VSM) approach. By means of this approach, which is robust against moderate levels of error, we can analyze data with weak or unknown structure.

Data-Recording and Pre-Processing

We record water pipe pressure data continuously with a sampling frequency of 200Hz. The resulting streams of raw sensor readings represent the basis for all further processing and recognition. In the first step, the data is pre-processed by filtering using a running average with a window size of 5 followed by baseline removal and noise rescaling. Baseline removal involves subtracting the mean and noise rescaling involves dividing by the standard deviation.

Activity Segmentation

Pre-processed data are then segmented regarding water activities. Therefore, significant deviations from the pressure baseline are detected that represent the boundaries of activity segments. Since the pressure in a plumbing system largely remains constant if none of the fixtures are used, segmentation is straightforward using a threshold-based procedure. For robustness regarding potential baseline drifts we apply a hysteresis thresholding procedure with empirically determined cut-off points.

Fixture Recognition

Fixture information, i.e., the knowledge about which fixture was used for a particular water activity (e.g. flushing the toilet in the master bathroom) is of value for the overall activity recognition procedure. Previous research related to IMS has shown that high accuracy fixture identification is possible using a water pressure sensing system like Hydrostream [5]. Therefore, in this work, we assume fixture information is available and assign it to every extracted activity segment manually during the study. Fixture recognition is a very important component of our pipeline, and we plan to implement our own recognizer in future work.

Frame Extraction

For all further processing we hypothesize that the characteristics of the water pressure data remain approximately constant over short periods of time. This short-time analysis is standard in general signal processing applications assuming locally stationary data and realistic for our application. We extract frames of N consecutive samples, which corresponds to M milliseconds of data, using a sliding window procedure. For the single-participant study, the segment size was 100 with 30% overlap, and for multi-participant study, segment size was $1,000$ with 90% overlap. The large segment size was needed in the multi-participant study because we have more data to match.

Feature Extraction

For every frame we extract a compact and meaningful representation by calculating six statistical features that describe the underlying data. The features are the signal's mean, variance, skewness, kurtosis, energy and root mean square (RMS) level. In doing so we map frames to 6-dimensional feature vectors, which are the basis for classification. Table 1 gives an overview of the feature definitions.

Clustering and BoW-modeling

Extracted feature vectors are then clustered using a standard unsupervised k-means algorithm. The optimal value of k was empirically determined. For the single-participant expert study, k was set to 150 and for the multi-participant study, k was set to 300. All feature vectors, i.e., frames of activity segments, are then mapped to the k prototypes derived during clustering resulting in BoW representations of input data. Thus, we map the continuous sensor readings to sequences of symbols originating from a finite lexicon of size k.

Classification

For classification of the activities in BoW-representation we employ a Vector Space Model (VSM) approach. VSMs are an algebraic model, originally developed for representing text documents as vectors of identifiers. They are widely used in the Information Retrieval (IR) domain for text classification [22]. Since their inception they have been used extensively in the areas of text analysis, indexing and retrieval [16].

Building on the success of the BoW approaches for IR with text and images, recent efforts in human activity recognition in Machine Learning have focused on working with BoW built using local spatio-temporal features for activity recognition [24]. However, when activities are represented as BoW, there is structural information provided by the ordering of the words that is still being lost. To address this problem, as described by [24], IR researchers have used n-grams to retain some of the ordering by forming sub-sequences of n items. More recently, n-grams have been used for activity recognition by [7] to represent activities in terms of their local event sub-sequences. We use a similar technique wherein we transform our time-series water pressure data into n-gram based BoW and use VSM for classification and verification.

The BoW representation for each activity is now a vector of word frequencies where each word is drawn from a lexicon of size k. In this final stage of our recognition pipeline we classify the extracted activity vectors with respect to the relevant activities of daily living as defined in Table 2. Classification using VSM can be divided into two stages. The first stage is the weighting of the terms to enhance retrieval. The second stage involves classification using the cosine similarity measure.

Term Weighting:

Each word in the activity vector is assigned a weight in order to obtain a statistical measure of its importance. The weight assigned to each word/term depends on two factors: *(1)* The number of occurrences of each word in the vector, called the "term-frequency"; and *(2)* The number of vectors in the collection that contains the word, called the "document-frequency". The various term-frequency and document-frequency metrics are detailed in Manning et al. [16]. The final weight assigned to each term is a product of these two. For our experiments we used the logarithm term-frequency metric and the unit document-frequency metric [16].

Cosine Similarity:

After term-weighting, the cosine similarity between any two activity vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ is given by

$$\cos\Theta = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\|\|\mathbf{v}_2\|} \tag{1}$$

Given activity vectors extracted from labeled training data and from the query data, the cosine similarity of the query activity to each of the training activities is computed and a vote among the k top-scoring matches is taken (k-Nearest Neighbor classification with a cosine similarity distance measure) and the query activity is assigned the class-label to which the majority of the k matches belong.

index	feature	description	definition
1	Mean	Average value of the samples of signal $\mathbf{x}$	$\mu_{\mathbf{x}} = \frac{1}{N}\sum_{n=0}^{N-1} x_n$
2	Variance	Power of values of signal $\mathbf{x}$ with its mean removed	$\sigma_x^2 = \frac{1}{N}\sum_{n=0}^{N-1}\lvert x_n - \mu_x\rvert^2$
3	Skewness	Measure of (lack of) symmetry in data distribution	$\frac{\sum_{n=1}^{N}(x_n - x)^3}{(N-1)s^3}$
4	Kurtosis	Measure of the shape of the data distribution	$\frac{\sum_{n=1}^{N}(x_n - x)^4}{(N-1)s^4}$
5	Total energy	Sum of squared moduli of signal $\mathbf{x}$	$E_x = \sum_{n=0}^{N-1}\lvert x_n\rvert^2$
6	RMS	Square root of the average power of signal $\mathbf{x}$	$\sqrt{P_x}$, where $P_x = \frac{E_x}{N} = \frac{1}{N}\sum_{n=0}^{N-1}\lvert x_n\rvert^2$

Table 1. Feature definitions used for water-based activity recognition

Location	Activity
Bathroom	Shave
Bathroom	Brush teeth
Bathroom	Wash hands
Bathroom	Flush toilet
Kitchen	Wash hands
Kitchen	Fill up teakettle
Kitchen	Make a salad
Kitchen	Rinse a fruit
Kitchen	Take a glass of water
Kitchen	Do dishes (light load)
Kitchen	Do dishes (heavy load)

Table 2. The list of activities for the studies

STUDY DESIGN

In order to address our research question, we set up two studies centered around the performance of eleven water-based activities in the kitchen and bathrooms of a home, shown in Table 2. These particular activities were chosen for two reasons. First, to advance research in IMS-based activity recognition, it is important to study activities in which fixture identification alone is not enough to classify an activity. Therefore, we limited our attention to activities that occur in the kitchen and bathrooms. Second, we are motivated to use IMS-based activity recognition for the purposes of health monitoring, so we wanted to identify behaviors linked to eating habits, personal hygiene, and sleeping patterns.

In the first study, our goal was to evaluate the performance of our system under a realistic best-case scenario. For our purposes, best-case was characterized by low variability in the way activities were performed, with multiple examples for each activity. This was achieved by having a single individual carry out the set of pre-defined activities five times. Throughout the paper, we refer to this study as the single-participant study.

In the second study, we examined our system's ability to recognize the same group of activities when performed by twenty-three individuals, where each individual carried out each activity exactly once. This multi-participant study was designed to reflect a real-world scenario as much as possible, where an activity recognition system would be installed in a home and provided with only one training example for each activity.

The studies took place in a testbed home built specifically for the study of home technologies. The home has fully functional kitchens, bathrooms and living spaces, and offers a very realistic setting for research explorations around domestic life. For the duration of the study, we outfitted the home with Hydrostream, collecting water pressure data from the hot and cold water lines. The installation process was straightforward and took less than 15 minutes.

For the first experiment, an expert (one of the authors of the paper) carried out a series of kitchen and bathroom activities five times. For the second experiment, twenty-eight participants were recruited to perform the same activities. Final data was only recorded for twenty-three participants due to reasons discussed later in the paper. Of these participants, twenty were recruited from the student population and the remaining eight came from a database of senior individuals in the Atlanta area who have agreed to be involved in research studies.

Upon arriving and checking in, we gave participants a short tour of the home and described the study and procedures. All participants in the second study carried out all activities in the same sequence, which was established by a script we read out aloud while following participants throughout the home. The script described a scenario that participants were asked to follow. It started with participants waking up and visiting the bathroom in the morning for activities such as shaving and teeth brushing. Afterwards, the script directed participants to the kitchen for breakfast preparation and all subsequent activities. The script brought consistency to the order of activities performed and injected a small dose of realism to the study. We believe that asking participants to carry out activities within a familiar context helped approximate the performance of these activities in the testbed home to how they would be performed in their homes. For activities that left significant room for interpretation, such as doing

Figure 3. A light (left) and heavy (right) load of dishes to be washed. We staged dirty dishes for every participant. Notice the presence of large pots and more plates on the heavy dishwashing side.

the dishes, we made sure to be as consistent as possible (i.e. the load of dishes was always the same for every participant as shown in Figure 3).

To ground the water-pressure data to the activities, we created a simple web-based activity-labeling tool. This tool, accessible from a mobile device, let us attach labels to the pressure readings in real-time so that we could tag exactly when activities began and ended. We used this tool while instructing participants to perform activities throughout the experiment.

DATA COLLECTION AND OBSERVATIONS

In the single-participant study, an expert carried out all eleven activities five times for a total of 52 activity samples. Three samples had to be eliminated due to logging errors. Out of the twenty-eight participants recruited for the multi-participant study, we were able to collect data for only twenty-three participants, which led to the compilation of 252 activity samples. One sample had to be discarded for the multi-participant study. The most common causes for participant data invalidation were logging errors and fluctuations in water pressure due to factors associated with the testbed home.

Challenges

We faced two issues while collecting data at the testbed home, most of them due to technical problems. First of all, even though we collected hot and cold water data continuously throughout both studies, we didn't find any difference between the two signals. This was unfortunate, since hot versus cold water usage is usually highly indicative of a high-level activity. As an example, we expected to observe only cold water pressure change for activities like toilet flushing and drinking a glass of water. We attribute this lack of differentiation in hot versus cold water pressure to the unique characteristic of the testbed home plumbing infrastructure.

Secondly, during the period the studies were taking place, we observed a water pressure phenomenon that resulted in pressure spikes at regular intervals, usually at the top of the hour. This phenomenon, which is not normally observed in a typical home, was probably the result of a faulty plumbing device. In the largest majority of cases this did not affect our study. However, in a handful of sessions we were forced to discard the data collected. Individuals on a different floor of the home who inadvertently used water while our study was in progress constituted another cause of external water pressure fluctuations.

Finally, we made mistakes logging participants' activities during the studies in some cases. This often occurred because participants engaged in an activity before telling us in advance. Data for the toilet flush activity was discarded for two participants during the multi-participant study. After flushing. they immediately proceeded to wash their hands in the bathroom sink while data was still being collected for the toilet flush.

Variability

In terms of the data collected, going from one individual in the single-participant study to twenty-eight in the multi-participant study resulted in a lot more variability in the data, since individuals tend to carry out tasks differently from each other. Moreover, the increased data variability was not matched by an increase in the number of examples for each activity. Both of these differences raised the activity classification difficulty substantially between the studies, and this difficulty is reflected in the activity classification accuracy.

As an example of the level of variability we encountered, some individuals leave the water running the whole time while doing the dishes in the kitchen. Another group first rinses the plates, pots and pans, leaves the water off until all items have been properly scrubbed, and then rinses everything again. One participant indicated that he would probably rinse and let the dishes soak for some time if the

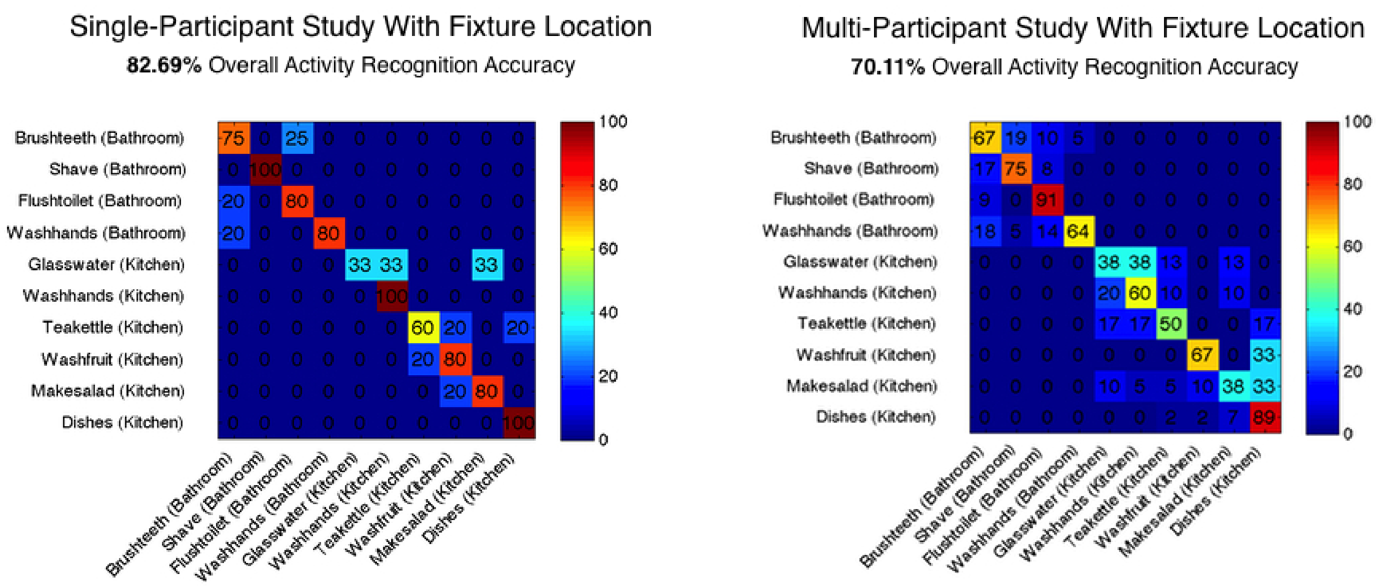

Figure 4. Confusion matrices and overall recognition results for our two studies

spaghetti sauce was stuck to the plate and needed scrubbing. We observed this type of entropy in the performance of all activities except for toilet flushes.

Another example of the variability encountered was the use of water in activities associated to the main activity. For example, rinsing occurred in a variety of forms, including rinsing the toothbrush before brushing teeth, rinsing the sponge before applying soap to the dishes, and also rinsing the sink after cleaning the dishes. Each of these associated activities were not labeled as separate activities, but instead were considered to be part of the main activity in every case.

While there was variability across participants, we observed that as individuals most participants were fairly methodical in conducting activities with multiple on/off events (e.g. dish washing, shaving, and hand washing). We also observed that certain sequence of activities appear to be second nature in a person's way of doing things at home. These observations and results motivate the next steps for this research to conduct studies in people's homes and train the machine learning algorithms focusing on the individual users.

Observations

A short exit interview was conducted with the participants after completion of the study to expand on their experience during the study, answer any questions, and gain additional insight into participants water usage at home. One participant indicated that he does not have a dishwashing machine at home and would always clean the dishes by hand. But he also noted that there was a dishwasher at his parents' home and that he would use it when he was visiting them. Another participant expressed difficulty using the kitchen sink. He said that it looked like it was designed for a left-handed person and that he might have ended up using more water than he usually does. However, another participant expressed an opposing view after the study, indicating that she had probably used less water than usual.

Participants were also keen to share other water usage scenarios based on their own experiences. One participant, identifying himself as a foreigner in the U.S., mentioned that standard devices like faucets and fixtures will be different in different countries. The same participant listed other examples of water usage that could be studied, including use of a Brita filter container, the refrigerator water dispenser, and bottled water.

STUDY RESULTS

The best recognition results we obtained assumed two conditions, *(1)* the grouping of both 'doing dishes' activities (light and heavy) into one activity, and *(2)* that fixture-level location information is available thanks to the work by Froehlich et al. on Hydrosense [5].

In the single-participant study, we obtained 82.69% in recognition accuracy for all activity classes. In this case, the participant performed all eleven activities five times. In the multi-participant study with twenty-three participants performing all activities exactly once, we obtained 70.11% in recognition accuracy for all activity classes. Figure 4 shows two confusion matrices that facilitate the visualization of the algorithm's performance for our two studies. The leave-one-out cross-validation method was used to provide an unbiased estimate of the generalization error of our classifier and minimize over-fitting.

It is worth noting that based on the raw signal, particularly poor discrimination was observed between certain types of activities such as shaving and doing dishes. The water pressure data for these two activities does look very similar, as shown in Figure 5. With fixture-level location information,

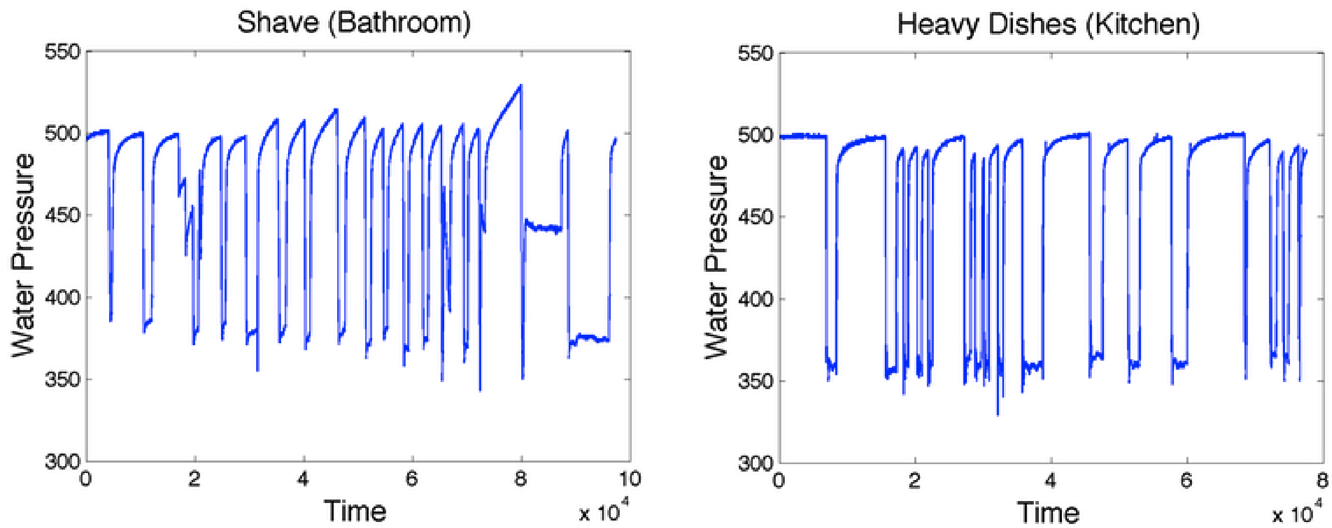

Figure 5. The water pressure signal for two distinct activities, shaving and doing the dishes.

however, our system was able to classify such similar water pressure patterns more efficiently.

DISCUSSION

In this section, we discuss three topics related to the design of the system and the two studies, namely the study script, the grouping of activities, and our choice for the number of training examples for each participant.

Study Script

Participants in the studies were asked to follow a script instructing them to perform all eleven activities in sequence. An argument against this approach is that it forces participants to carry out activities in an ordered, mechanical way that is not representative of how these tasks would be performed in everyday life. There are two reasons why we chose to format the study in this particular way.

Firstly, the script facilitated the process of collecting accurate activity ground truth data. Activity ground truth was collected manually, by means of a tool developed specifically for these studies. Having a script made it possible for the experimenters to anticipate participants' activities and log them correctly, reducing the chance of errors.

Secondly, the primary motivation of this work was to assess the feasibility of IMS combined with a vector space model learning approach for activity recognition. We were able to evaluate the system in a realistic, yet controlled setting that resulted in 'best-case' benchmark results under two studies. In the future, as we move towards studies in real homes with true ecological validity, the benchmark will serve as a reference point to help us interpret our results, and most importantly, guide us as we make improvements to the algorithms.

Activity Groupings

Our best classification results were obtained when we grouped the two dishwashing activities. However, the grouping itself accounted for an improvement of less than 5% in overall recognition accuracy. The reason why we chose to split dishwashing into two separate activities in the first place was because we hypothesized that we might be able to infer individual and family eating styles according to dishwashing patterns. For example, frequent observations of light dishwashing could be a sign that the family or individual does not eat at home often, or at least brings semi-prepared foods to eat at home. On the other hand, the recognition of heavy dishwashing could mean that a meal was prepared at home, with pots, pans and other utensils that had to be cleaned.

Training Examples

In the multi-participant study, we brought twenty-three participants to the testbed home and asked them to perform eleven activities. We collected only one example for each one of these activities. Although activity recognition systems typically require large amounts of training data, we chose to evaluate the performance of our system with a minimal number of examples. This is because in practice, collecting training data for an activity recognition system in a home setting is very disruptive for residents, and we would like to minimize this step as much as possible. Based on our results, we are hopeful that we will be able to improve the performance of the system while keeping the number of training examples low.

CONTRIBUTIONS

In this work we demonstrate that water-based, single-point infrastructure mediated sensing jointly with a vector space model learning approach can be successfully used for high-level activity recognition in a home setting. The results reported in this paper focus specifically on kitchen and bathroom activity discrimination, where many important activities of daily living take place.

Although infrastructure mediated sensing has been used before to estimate real-world water usage at the fixture level, as far as we know, our work is the first to employ this method for inferring high-level human activities. In our view, the low-cost, ease of deployment and respect to privacy that are inherent of water-based infrastructure-mediated sensing suggest that it is a viable option for population-scale activity recognition systems in the home, supporting applications in fields like healthcare and others.

FUTURE WORK

At this initial phase of research in minimally-instrumented activity recognition for the home, our focus has been on understanding what type of performance we can expect from our approach. As a result, our study was set up to be realistic but not ecologically valid. Ecological validity (i.e. evaluating the system in real homes) would have introduced uncertainties into the analysis of our system, such as whether our ground truth labels really corresponded to ground truth.

Motivated by the promising results reported in this paper, our next step we will be to collect activity data from individuals in their homes. This presents two fundamental challenges: the difficulty of collecting ground truth activity labels outside of a laboratory environment, and technical issues with infrastructure-mediated sensing.

While it was straightforward to collect ground truth activity labels in real-time in a controlled setting, doing so in a typical home environment is a much more difficult undertaking. It is critical that ground truth labels are provided to classifiers so they are trained to discriminate activities. We will most likely need participants to log their own activities to some degree, which can be disruptive, and directly affect the performance of the activity we are trying to observe and recognize. We plan to evaluate traditional survey instruments (e.g. experience sampling method [14]) and devise new techniques for collecting participant feedback and input.

As far as water-based, infrastructure-mediated sensing is concerned, further work is required to study how to adapt this sensing technique to a wider variety of homes, such as multi-family homes and apartment buildings. Even in a variety of single-family home settings, where it has been shown to produce reliable results, it would be important to confirm that the performance of our activity recognition algorithms match the results obtained in the testbed home, in light of different water pressure baselines, plumbing set ups and other unknowns.

ACKNOWLEDGEMENT

This work was funded in part by the Intel Science and Technology Center for Pervasive Computing (ISTC-PC).

REFERENCES

1. L. Atallah and G. Yang. The use of pervasive sensing for behaviour profiling — a survey. *Pervasive and Mobile Computing*, (5):447–464, 2009.

2. G. Cohn, S. Gupta, J. Froehlich, E. Larson, and S. Patel. GasSense: Appliance-level, single-point sensing of gas activity in the home. *Pervasive Computing*, pages 265–282, 2010.

3. J. Fogarty, C. Au, and S. E. Hudson. Sensing from the basement: a feasibility study of unobtrusive and low-cost home activity recognition. In *UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology*. ACM Request Permissions, Oct. 2006.

4. J. Froehlich, E. Larson, T. Campbell, C. Haggerty, J. Fogarty, and S. Patel. HydroSense: infrastructure-mediated single-point sensing of whole-home water activity. *Proceedings of the 11th international conference on Ubiquitous computing*, pages 235–244, 2009.

5. J. Froehlich, E. Larson, E. Saba, T. Campbell, L. Atlas, J. Fogarty, and S. Patel. A Longitudinal Study of Pressure Sensing to Infer Real-World Water Usage Events in the Home. *Pervasive Computing*, pages 50–69, 2011.

6. S. Gupta, M. Reynolds, and S. Patel. ElectriSense: single-point sensing using EMI for electrical event detection and classification in the home. *Ubicomp '10: Proceedings of the 12th ACM international conference on Ubiquitous computing*, Sept. 2010.

7. R. Hamid, S. Maddi, A. Johnson, A. Bobick, I. Essa, and C. Isbell. A novel sequence representation for unsupervised analysis of human activities. *Artificial Intell.*, 173:1221–44, 2009.

8. T. Hirsch, J. Forlizzi, E. Hyder, J. Goetz, C. Kurtz, and J. Stroback. The ELDer project: social, emotional, and environmental factors in the design of eldercare technologies. *Proceedings on the 2000 conference on Universal Usability*, pages 72–79, 2000.

9. S. S. Intille, K. Larson, E. M. Mungia-Tapia, J. S. Beaudin, P. Kaushik, J. Nawyn, and R. Rockinson. Using a live-in laboratory for ubiquitous computing research. In *Proc. Int. Conf. on Pervasive Computing*, pages 349–365, 2006.

10. Y. A. Ivanov and A. F. Bobick. Recognition of visual activities and interactions by stochastic parsing. *PAMI*, 22(2):852–872, August 2000.

11. S. Katz, A. B. Ford, R. W. Moskowitz, B. A. Jackson, and M. W. Jaffe. Studies of illness in the aged. *Jama The Journal Of The American Medical Association*, 185(12):914–919, 1963.

12. J. Kientz, S. Patel, B. Jones, E. Price, E. D. Mynatt, and G. D. Abowd. The Georgia Tech aware home. In *Proc. CHI*, 2008.

13. M. Kranz, A. Schmidt, A. Maldonado, and R. B. Rusu. Context-aware kitchen utilities. In *Proc. Int. Conf. Tangible and Embedded Interaction*, 2007.

14. R. Larson and M. Csikszentmihalyi. Experience sampling method (esm). *In H T Reis Ed Naturalistic Approaches to Studying Social Interaction New Directions for Methodology of Social and Behavioral Science*, 15:41–56, 1983.

15. M. P. Lawton and E. M. Brody. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist*, 9(3):179–186, 1969.

16. C. Manning, P. Raghavan, and H. Schütze. *Introduction to Information Retrieval*. Cambridge Univ. Press, 2008.

17. D. Moore and I. Essa. Recognizing multitasked activities from video using stochastic context-free grammar. In *AAAI*, 2002.

18. E. Munguia-Tapia, S. Intille, and K. Larson. Activity recognition in the home using simple and ubiquitous sensors. *Proceedings of PERVASIVE*, pages 158–186, 2004.

19. S. Patel, M. Reynolds, and G. Abowd. Detecting Human Movement by Differential Air Pressure Sensing in HVAC System Ductwork: An Exploration in Infrastructure Mediated Sensing. *Pervasive '08: Proceedings of the 6th International Conference on Pervasive Computing*, Mar. 2009.

20. C. Pham and P. Olivier. Slice & Dice: Recognizing Food Preparation Activities Using Embedded Accelerometers. In *Proc. Europ. Conf. Ambient Intelligence*, 2009.

21. M. Philipose, K. Fishkin, M. Perkowitz, D. Patterson, D. Fox, H. Kautz, and D. Hahnel. Inferring activities from interactions with objects. *IEEE pervasive computing*, 3(4):50–57, 2004.

22. G. Salton. *The SMART retrieval system: Experiments in automatic document processing*. Prentice-Hall, Upper Saddle River, NJ, 1971.

23. P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea. Machine recognition of human activities: A survey. *IEEE Trans. Circuits Systems for Video Tech.*, 2008.

24. H. Wang, M. M. Ullah, A. Kläser, I. Laptev, and C. Schmid. Evaluation of local spatio-temporal features for action recognition. In *BMVC*, 2009.