Sensing from the Basement: A Feasibility Study of Unobtrusive and Low-Cost Home Activity Recognition

James Fogarty
Computer Science and Engineering
University of Washington
Seattle, WA 98105
jfogarty@cs.washington.edu

Carolyn Au†*, Scott E. Hudson*
Human Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15213
scott.hudson@cs.cmu.edu

ABSTRACT
The home deployment of sensor-based systems offers many opportunities, particularly in the area of using sensor-based systems to support aging in place by monitoring an elder's activities of daily living. But existing approaches to home activity recognition are typically expensive, difficult to install, or intrude into the living space. This paper considers the feasibility of a new approach that "reaches into the home" via the existing infrastructure. Specifically, we deploy a small number of low-cost sensors at critical locations in a home's water distribution infrastructure. Based on water usage patterns, we can then infer activities in the home. To examine the feasibility of this approach, we deployed real sensors into a real home for six weeks. Among other findings, we show that a model built on microphone-based sensors that are placed away from systematic noise sources can identify 100% of clothes washer usage, 95% of dishwasher usage, 94% of showers, 88% of toilet flushes, 73% of bathroom sink activity lasting ten seconds or longer, and 81% of kitchen sink activity lasting ten seconds or longer. While there are clear limits to what activities can be detected when analyzing water usage, our new approach represents a sweet spot in the tradeoff between what information is collected at what cost.

Author Keywords
Activity recognition, sensor-based models, sensing in the home.

ACM Classification Keywords
H5.2. Information interfaces and presentation: User Interfaces; H1.2. Models and Principles: User/Machine Systems.

INTRODUCTION AND MOTIVATION

The home deployment of sensor-based systems promises many new opportunities for human computer interaction. One such opportunity motivating our work is the use of sensor-based systems to support aging in place by monitoring activities of daily living [5, 6, 11, 12, 14, 15].

UIST'06, October 15–18, 2006, Montreux, Switzerland.

Potential benefits of in-home elder activity sensing include providing peace of mind to physically distant adult children, collecting detailed logs of daily activities for examination by a medical professional, and the detection of anomalous patterns in behavior that may suggest a fall or some other situation that warrants investigation.

A number of approaches to home activity recognition have been considered. One is to install extensive sensing infrastructure, such as vision-based systems, microphones, or strain sensors under floorboards [1, 5, 14]. While this can enable the collection of a large variety of information about the home, the cost of installing and maintaining the necessary sensing is typically very high. A second approach is to use many low-cost sensors that can be inexpensively deployed throughout a home [2, 10, 15]. But these sensors intrude into the living space (as with a contact switch taped to a kitchen cabinet), and elders may reject such sensing because it detracts from the appearance of their home or creates feelings of embarrassment related to a need for assistance [9]. A third approach is to assume a wearable device [12], but elders might choose not to wear such a device in some situations (such as when bathing).

This paper considers a new approach to sensor-based home activity recognition. Instead of deploying many sensors throughout a home, we hope to leverage a home's existing infrastructure to "reach into the home" with a small set of strategically-placed sensors. Specifically, this paper considers the use of a home's existing water distribution infrastructure. Fresh water enters a home at a single point and wastewater leaves the home at a handful of locations. We propose the use of low-cost microphone-based sensors at these critical locations. Attached to the outside of existing pipes, these sensors listen for the flow of water. Based on a model of water flow into and out of a home, we aim to provide approximately the same information that would be obtained by installing sensors on sinks, toilets, showers, and appliances throughout the home. While there are obvious limits to the activities that can be detected when only considering water usage, our proposed approach represents a sweet spot in the tradeoff between what information is collected at what cost. Using just a handful of sensors in the basement, our approach can recognize many activities that are important to elder activity sensing.

† Carolyn Au is now with Google, Inc.

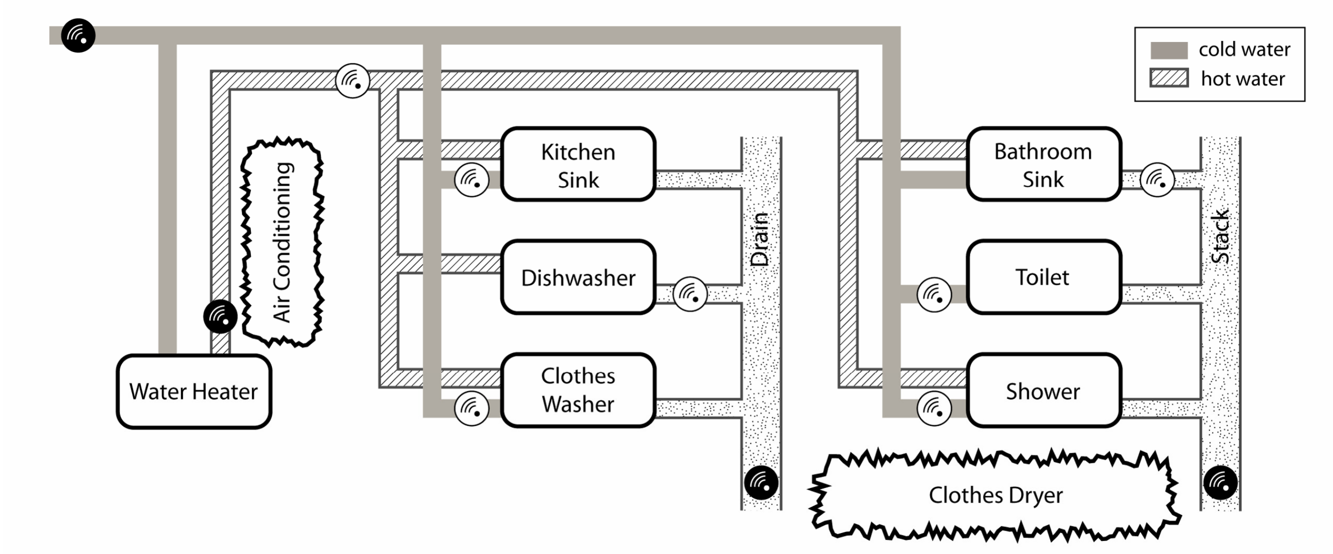


Figure 1. Schematic of the water pipes in the home where we conducted our feasibility study. The air conditioning and clothes dryer are shown because they rattled nearby pipes, introducing noise that needs to be considered in analyses. The four shaded sensors are used for modeling activities, while the unshaded sensors are included only for validating our results.

Using six weeks of data collected in an actual home shared by two adults (including one of the authors†), this paper examines the feasibility of using unobtrusive and low-cost sensing of water usage to model home activities. Figure 1 presents a schematic of the home studied in this paper. We deploy four sensors for activity recognition: one where cold water enters the home, one where hot water exits the water heater, and one on each of two pipes through which wastewater leaves the home. Another seven sensors are deployed for the sake of collecting ground truth in this feasibility study, but these would be not be necessary in resulting deployments. With approximately 3.4 million sensor readings collected from these sensors over the course of six weeks, we examine three questions:

What effect will ambient noise have on microphone-based sensors of water flow? While microphone-based sensors are inexpensive and can be installed on the outside of existing pipes, pipes are excellent conductors of sound and ambient noise could be very problematic. For example, the hum of central air conditioning or a clothes dryer might be detected by a microphone and misinterpreted as water flow.

What home activities can be reliably recognized from patterns of water usage? Some sources of water usage, such as dishwashers and other appliances, are likely to generate highly consistent mechanical patterns. In contrast, the use of water at a sink may be relatively random. Still other usage, such as the hot water heater filling, may be irrelevant in the context of the activities we are recognizing.

† While using ourselves as participants in this preliminary work presents some difficulties, it is appropriate for early feasibility testing because of the advantages it provides for critical early debugging as well as greater access for placing more intrusive ground truth sensors within the living space. In particular, it allowed us to collect data from live use in a real home much earlier than would have otherwise been possible.

How do simultaneous home activities interfere with recognition? Even if an activity can be recognized in isolation, field deployments generally require the ability to cope with simultaneous activities. This might be because multiple people live in a home or, for example, because a person uses a toilet while the dishwasher is running. Under a variety of conditions, one activity may mask another, introducing additional recognition difficulties.

This paper presents a series of analyses to explore these problems. Using our real-world data, we separate and study the impact of these problems on the recognition of an important set of home activities. Among our other results, we show that a model built with microphone-based sensors that are placed away from systematic noise sources can identify 100% of clothes washer usage, 95% of dishwasher usage, 94% of showers, 88% of toilet flushes, 73% of bathroom sink activity lasting ten seconds or longer, and 81% of kitchen sink activity lasting ten seconds or longer. While future work is needed to extend this initial result to a larger variety of homes and to consider whether more specific activities can be recognized at sinks (instead of just recognizing what sink is in use), this feasibility study demonstrates significant potential for our new approach.

The next section reviews related work, with a focus on home activity recognition and elder care applications. We then introduce our microphone-based sensor. This is followed by a discussion of the home in which we conducted our feasibility study and a characterization of the data obtained from our sensors. We then discuss our development of statistical models of water flow based on our microphone features. The output of these models is fed into a pattern-based recognition algorithm, which we then evaluate against our data. We then present a short discussion of our modeling, give an example of a high-level activity illustration based on our recognition, and conclude.

RELATED WORK

Mynatt *et al.*'s Digital Family Portrait is an excellent example of the type of an application that we believe can benefit from our approach to unobtrusive and low-cost home activity recognition [11, 14]. Intended primarily to offer piece of mind to distant family members, the Digital Family Portrait provides an overview of daily life and long-term trends. Inspecting the Digital Family Portrait, a relative can find reassurance over such concerns as whether an elder is eating enough and whether the elder is active or sedentary. If a potential problem arises, the relative can raise the issue with the elder and work with the elder to address the problem. Conversely, the relative might know that the Digital Family Portrait is reporting a very low level of activity because the elder is currently on vacation.

Initial work on the Digital Family Portrait used Wizard of Oz techniques to simulate activity detection [11]. In a more recent study, Rowan and Mynatt installed strain sensors on the underside of the first floor of an elder's home [14]. By detecting the weight of a person standing on the floor, these sensors allow the Digital Family Portrait to be based on movement throughout the first floor of the home. We note that installation of these sensors is sufficiently difficult that expert assistance would typically be required. Further, installation requires access to the underside of the floor, making it impossible to use these sensors on the second floor of an existing home. In contrast, our approach is based on sensors that are easy to install in existing homes.

In related work, Consolvo *et al.*'s CareNet Display examines the similar problem of providing information to the local members of an elder's care network [6]. The CareNet Display aims to provide more detailed information appropriate for use in the local coordination of care-related activities. This includes exactly what activities were detected and when they were detected, as well as what medications were taken and when they were taken. Because this level of sensing is extremely difficult to achieve, their evaluation is based on a Wizard of Oz approach using periodic phone conversations to collect information about an elder's activities.

Hirsch *et al.* examine the social and psychological factors that influence the design of elder care applications [9]. Among their findings is a concern that assistive technology may be rejected if it detracts from the aesthetics of the home, leads an elder to feel spied upon, or creates a feeling of embarrassment over the need for assistance. Because our sensors will typically be in a basement, utility closet, or some other secluded location, our approach enables home activity recognition without the negative stigmas that might be associated with sensors in the living space.

Beckmann *et al.* present a study of end-user sensor installation and reaction to sensors in the home [2]. They had end-users install vibration sensors, electricity usage sensors, motion detectors, cameras, and microphones. They found that end-users made a variety of errors, often due to the directional requirements of sensors or uncertainty over exactly where a sensor needs to be positioned. They also found many negative reactions to the intrusion of sensors into the living space, including objections to the potential for damage caused by the adhesive used for installation, concern that sensors were placed in locations accessible by children or pets, and objections to the placement of cameras and microphones in the home. By requiring only a few relatively easy to install sensors in the basement, our approach greatly reduces these concerns.

Tapia *et al.* discuss home activity recognition using many state change sensors, primarily contact switches [10]. These sensors were taped to surfaces in the home, logging activations of the sensor for the duration of a study. While the ability to install a contact switch nearly anywhere in the home might seem to provide more information than can be obtained from our approach, most of the sensors in this research were installed in the kitchen or bathroom, with success in detecting activities such as meal preparation and toileting. We believe our approach can provide many of the same benefits without intruding into the living space.

Wilson and Atkeson examine tracking and activity recognition using motion detectors, pressure mats, break beam sensors, and contact switches [15]. This work is of interest because it tackles the problem of recognizing the activities of several people sharing a home. In contrast, most research assumes that all sensor activations are being caused by a single person. While they are able to track the locations of multiple people in the home, their approach requires the installation of many sensors in the living space and activity recognition is currently limited to movement.

Chen *et al.* examine the recognition of bathroom activities using a microphone placed near a bathroom sink, with a focus on aiding in the care of incontinence and dementia [5]. While the placement of a microphone in the bathroom raises potential privacy issues, it does allow the detection of some activities that will be invisible to our approach, such as failing to flush after using the toilet. But other activities of interest to Chen *et al*. could be detected using our approach, including repeated showering within a single day (a potential indication that a person may be confused or experiencing difficulties with incontinence).

Philipose *et al.* present the use of an RFID-enabled glove to monitor activities of daily living [12]. A person wearing the glove interacts with RFID-tagged objects, and the system recognizes activities based on interactions with objects. Even if a person is generally willing to wear the reader, they may choose to remove it in situations where it may come into contact with water, as during bathroom use and meal preparation. We see our work as complimentary to this approach, as we can recognize a variety of important activities without requiring a wearable device, but Philipose *et al.*'s approach has the potential to provide information about activities that are invisible to our approach, such as the proper self-administration of medication.

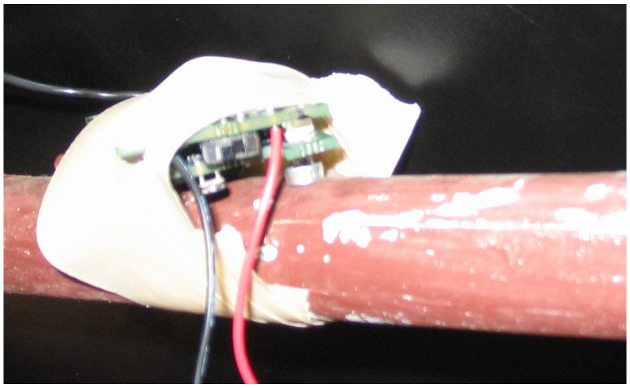
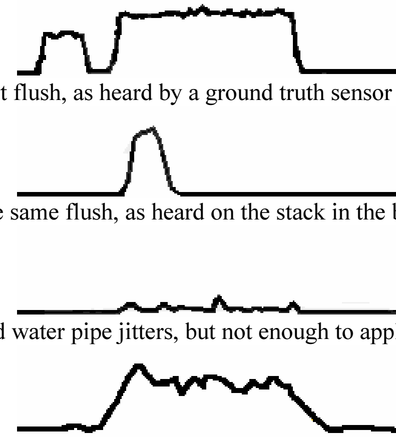


Figure 2. Our prototype sensor attached to a water pipe. The microphone is the cylinder in the center. We pressed the microphone against the pipe and secured the sensor with tape.



(a) A toilet flush, as heard by a ground truth sensor on the toilet.

(b) The same flush, as heard on the stack in the basement.

(c) The cold water pipe jitters, but not enough to apply a threshold.

(d) Applying an entropy transformation to the cold water pipe yields a feature more appropriate for use with a threshold.

Figure 3. Several signals obtained from our sensors.

PROTOTYPE SENSOR DESIGN

While we eventually envision specialized sensing hardware similar to current home alarm sensors, such hardware should not be built until we know exactly what signal to detect. To provide a flexible platform for this feasibility study, we use Mica2 series Berkeley Motes and an MTS300 sensor board, which includes a microphone [7]. Used with TinyOS [8], the Mica2 is a relatively easy to program platform that includes a microprocessor and support for wireless communication. It is more flexible and powerful than should be necessary in a low-cost sensor, but it is very appropriate for this feasibility study. As seen in Figure 2, we installed these sensors by pressing them directly against pipes and used masking tape to secure them.

In normal operation, the Mica2 can obtain several thousand 10-bit microphone samples per second. This is far too much information to transmit over the wireless connection, so we conducted informal experiments to determine what features to extract from the audio signal. Using a serial port connection, we recorded short segments of audio from sensors attached to several pipes. While recording, we turned the water on and off at known intervals. We then performed informal offline experiments to find appropriate features for identifying water flow from the recorded audio.

After considering several features, we decided to capture the zero-crossing rate and the root mean square of microphone samples. These features have different characteristics, but are both computationally inexpensive and generally increase when water is flowing in a pipe. Because power management is always important in battery-powered wireless sensors, we collect audio features only once every two seconds. The sensor spends 1.75 of every two seconds asleep. In then wakes up and captures 1000 microphone readings. It computes a feature pair for these 1000 samples (the zero-crossing rate and the root mean square) and stores the feature pair in memory before returning to sleep. After nine feature pairs have been collected, the sensor checks if any of the zero-crossing features are non-zero. If so, the nine feature pairs values are transmitted in a wireless packet. In order to make efficient use of the space available in this wireless packet, a scaling factor is applied to convert the feature values to 8-bit integers. The wireless packet is received by a Mica2 series Berkeley Mote with a serial connection to a laptop computer running logging software. To reduce data loss due to transmission failures, each packet is sent twice.

Using two D-cell batteries to power each sensor, we experienced no battery failures during our six-week feasibility deployment. These prototype sensors continued to work properly for at least another five weeks, at which point the laptop being used for data collection was accidentally unplugged. We believe custom hardware battery life could exceed two years, and so prefer this approach over, for example, the use of a sensor powered by water flow. We note that such a sensor would likely require costly professional installation, while it is a very reasonable maintenance requirement to ask a caregiver to occasionally change three to four batteries in an elder's home.

Figure 3 shows several actual zero-crossing rate signals obtained from our sensors. Signal (a) is a toilet flushing, as heard by a ground truth sensor installed on the metallic hose connected to the toilet's tank (the placement of a sensor directly on the toilet is discussed in the next section). The small peak is ambient noise prior to the toilet being flushed, and the large plateau is the sound of water filling the toilet's tank when it is flushed. Signal (b) is this same flush, as heard by a sensor in the basement on the stack (the large pipe which serves as a drain for the toilet, the shower, and the bathroom sink). Both signals are appropriate for use with a threshold, as they clearly have larger values when activity is detected. But this is not immediately true of signal (c), which shows the same flush as heard by a sensor installed on the cold water pipe in the basement. There is a jitter in signal (c) while the toilet tank is filled, but this jitter is not large enough to reliably apply a threshold. Signal (d) therefore computes the entropy of signal (c). Because signal (d) clearly has larger values when water is flowing, it is appropriate for use with a threshold.

FEASIBILITY STUDY DEPLOYMENT

As first introduced in Figure 1, this feasibility study was conducted in a two-story, 1800 square foot house shared by a 27 year-old female and a 26 year-old male. The water pipes in this home are copper and the drain pipes are polyvinyl chloride (PVC is the most common material currently used for drain pipes). We installed four sensors intended to model activities in the home, and another seven sensors to collect ground truth data for this study.

The cold water sensor was installed in the basement immediately after the water meter, and therefore monitors all water usage. The hot water sensor was placed on the pipe coming out of the water heater (a typical 40 gallon gas water heater). A central air conditioning unit is located in close physical proximity to this hot water sensor, and the air conditioning unit introduces a rattling into the pipes. Most homes have either one or two pipes through which wastewater exits the home. In this home, the kitchen sink, dishwasher, and clothes washer all drain into one pipe (which we refer to as the *drain* throughout this paper), and the bathroom sink, toilet, and shower all drain into another pipe (which we refer to as the *stack* throughout this paper). Both the drain and the stack pass through a small utility room (joining below the floor of the basement), so our sensors were installed on the pipes as they passed through this room. The clothes dryer is also located in this room, introducing an additional noise source.

While the four locations just described are intended to be the only sensors used in a normal deployment, we installed additional sensors to provide a basis for comparison in this feasibility study. We placed a sensor on the metallic hose that fills the toilet, on the drain of the bathroom sink, on the cold water pipe feeding the shower, on the cold water pipe feeding the kitchen sink, on the drain of the dishwasher, and on the cold water pipe feeding the clothes washer. We chose these locations because they represent all of the locations in this house where water is consumed. Decisions regarding the installation of a particular sensor (whether to install it on the incoming water or the outgoing drain) were primarily based on the ease of installation and the degree to which a sensor could be hidden from view. We informally note that these six sensors were much more difficult to install than the four sensors we installed in the basement. For example, installing a sensor on the shower required removing a panel from the wall of a closet and crawling to within reach of the desired pipe. As discussed in the next section, our feasibility study uses these sensors to determine what is actually happening in the home. This provides a measure of ground truth for use in evaluating our models.

The final sensor was installed on the hot water pipe, still in the basement but far from the air conditioning unit. Because pilot data collection revealed significant noise in the hot water sensor due to the air conditioning unit, we later analyze this secondary hot water sensor instead of the sensor directly on the hot water heater. This allows us to characterize careful versus casual sensor placement.

	(Training)	Test
Dishwasher	**(4)**	**20**
Shower		1
Toilet	(2)	6
Long Bathroom Sink	(2)	2
Short Bathroom Sink		1
Long Kitchen Sink	(1)	7

	(Training)	Test
Shower	**(15)**	**66**
Dishwasher		1
Toilet	(7)	26
Long Bathroom Sink	(2)	5
Long Kitchen Sink	(2)	8
Short Kitchen Sink		2

	(Training)	Test
Long Kitchen Sink	**(87)**	**356**
Dishwasher	(1)	7
Clothes Washer	(5)	15
Shower	(2)	18
Toilet	(4)	13
Long Bathroom Sink		3

	(Training)	Test
Short Kitchen Sink	**(26)**	**168**
Dishwasher		1
Clothes Washer		4
Shower		3
Toilet		1

	(Training)	Test
Clothes Washer	**(8)**	**16**
Toilet	(2)	6
Long Bathroom Sink		4
Long Kitchen Sink	(3)	5
Short Kitchen Sink		2

	(Training)	Test
Toilet	**(100)**	**419**
Dishwasher	(4)	8
Clothes Washer	(2)	9
Shower	(7)	27
Long Bathroom Sink	(22)	83
Short Bathroom Sink		2
Long Kitchen Sink	(4)	12
Short Kitchen Sink		1

	(Training)	Test
Long Bathroom Sink	**(59)**	**262**
Dishwasher	(3)	3
Clothes Washer		12
Shower	(2)	5
Toilet	(22)	89
Long Kitchen Sink		3

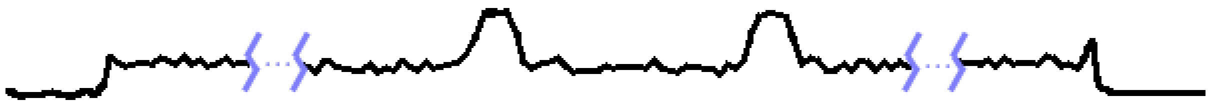
	(Training)	Test
Short Bathroom Sink	**(3)**	**42**
Dishwasher		1
Short Kitchen Sink		2

Figure 4. An overview of sensed activities in our data.

COLLECTED DATA OVERVIEW

Deployed for forty days, our thirteen sensors collected a total of approximately 3.4 million feature pairs, where each pair is a zero-crossing rate and root mean square feature, sampled at two second intervals. For our analyses, we treat the first 7 days of data as *training* data and the remaining 33 days as *test* data. The activities occurring in the test and training portions of the collected data are presented in Figure 4, with totals for the training period shown in parentheses. For example, the dishwasher was run 24 times in these 40 days, and four of these occurred during the initial training week. On one occasion, a person was in the shower at the same time the dishwasher was running. On eight occasions, the toilet was flushed at least once while the dishwasher was running (and two of these occurred during the training week). Counts are not symmetrical because multiple instances of an activity can occur during a single instance of a longer activity. Throughout this paper, a *long* sink activity is a use of a sink that lasts for more than 10 seconds, while a *short* sink activity lasts for less. We make this distinction because short uses of a sink are harder to recognize than long uses, a point we clarify in discussion.

The use of the bathroom sink reported here is lower than actually occurred, because it was generally difficult to determine whether the bathroom sink was being used while the toilet tank was being filled after a flush. Chen *et al.* report a similar difficulty, as the sound of the sink being used is faint compared to the sound of the toilet tank being filled [5]. Part of the difficulty seems to be that our ground truth sensor for the bathroom sink was attached to the pipe leading to the stack, which is also connected to the toilet. In future studies, we intend to install our ground truth sensor on the hot water pipe leading to the bathroom sink, as it has no direct connection to the toilet.



(a) The stack sensor detects significant noise created by a nearby clothes dryer, but additional peaks associated with two toilet flushes are still detectable. The stack sensor is labeled as on for the two visible peaks, and as off otherwise.

(b) The cold water sensor's view of a toilet flush occurring during a shower. The long jittering is water usage by the shower, while the higher plateau is caused by the toilet. The cold water sensor is labeled as on for the entire signal jitter, thus masking the toilet flush.

Figure 5. Two interesting cases that arise in labeling sensor activations from the microphone-based sensor readings.

LABELING SENSOR ACTIVATIONS

In order to associate meaningful labels with our training and test data, we used custom software to manually examine each collected reading and label it as either on or off. Our software presents time-aligned signals from our sensors in a format similar to Figure 3 (showing multiple sensors displayed with their output vertically aligned).

Using the output of the six ground truth sensors placed in the living environment, we manually labeled when different activities occurred in the home. For example, a signal like that in Figure 3(a) makes it very clear when the toilet flush begins and ends, so we labeled the toilet sensor as on for the entire duration of the flush (from the beginning to the end of the large plateau in Figure 3(a)). For sensors like the dishwasher, which use water only part of the time they are active, we labeled the sensor as on from the beginning of the first water-related activity until the end of the last water-related activity (either the use of water from a pipe or the release of water into a drain). Extracting ground truth from sensors in the living environment was a very straightforward process, with the one exception discussed in the last section (that activity at the bathroom sink was difficult to detect during a toilet flush).

More interesting issues arise in labeling the output of our hot, cold, drain, and stack sensors. As previously noted, the central air conditioning in this home rattles the pipe on which we placed our hot water sensor and the clothes dryer rattles both the drain and stack sensors. The effect of this noise can be seen in Figure 5(a). There is no activity at the leftmost edge of the signal. When the clothes dryer is started, the resulting noise is detected by our stack sensor. After some time passes, there are two toilet flushes in rapid succession, yielding noticeable peaks. In this case, we labeled the stack sensor as on for only the two peaks, as the noise of the clothes dryer is not associated with water flow.

Conversely, Figure 5(b) shows a toilet flush as heard by the cold water sensor when a person was in the shower. In this case, cold water is already being used by the shower when the toilet is flushed. When multiple activities used water simultaneously as illustrated here, we labeled the sensor as on for any time that water was flowing. In this example, this results in the toilet flush being masked by the shower.

LEARNED WATER FLOW SENSOR MODELS

In manually labeling our data, it became clear that a simple threshold would be inadequate for determining if water was flowing past our basement sensors. For example, the stack and drain pipes join shortly after going below the basement floor. Activity in one typically generates a detectable signal in the other. A weak signal detected by the drain sensor might therefore be a small amount of water flowing in the drain, or it might be noise caused by activity in the stack. Using our manually-specified labels and our collected microphone-based features, we learned statistical models of whether water was flowing at our hot, cold, drain, and stack sensors. For each sensor, we used Weka [16] to learn a support vector machine [13]. We learned a model based on 4000 readings randomly selected from the first 7 days of data for each sensor, 1000 readings that we had labeled as on and 3000 readings that we had labeled as off. This model was based on the original zero-crossing and root mean square features, the entropy of those features, and product and ratio of those features with the equivalent features on related sensors (considering the stack and drain to be related, as well as the hot and cold to be related).

After building these models from our training data, we evaluated them against the remaining 33 days of data. We applied the model to each sensor reading, labeling pipes as on during the time between any two readings that were both labeled as on (this has the side-effect of ignoring an isolated reading that a model labels as on). To protect our analyses against the random selection of a particularly good or bad set of training examples, we executed this process 15 times and report the mean of the recall, precision, and F_1:

Hot:	*Recall:* 91 ± 00%	*Precision:* 71 ± 01%	F_1: .80 ± .01
Cold:	*Recall:* 89 ± 01%	*Precision:* 94 ± 00%	F_1: .91 ± .01
Drain:	*Recall:* 86 ± 01%	*Precision:* 77 ± 02%	F_1: .81 ± .01
Stack:	*Recall:* 86 ± 03%	*Precision:* 89 ± 01%	F_1: .88 ± .02

where F_1 treats recall and precision equally and is defined as (2 * *Precision* * *Recall*) / (*Precision* + *Recall*). The analyses that follow use the model that performed closest to the mean of F_1 for each sensor. We do not dwell on these reliabilities, instead waiting to focus on the reliability of these sensors in the context of the activity recognition algorithm discussed next.

Clothes Washer

Cold:	**Mostly On:** 3 to 9 minutes	**Mostly Off:** 6 to 18 minutes	**Mostly On:** 3 to 9 minutes	**Mostly Off:** 90 seconds to 6 minutes	
Drain:	**Mostly Off:** 7 to 21 minutes	**Mostly On:** 40 seconds to 2 minutes	**Mostly Off:** 5 to 15 minutes	**Mostly On:** 35 to 105 seconds	**Mostly Off:** 0 to 5 minutes

Toilet

Cold:	**Mostly On:** 40 to 70 seconds		
Stack:	**Mostly Off:** 0 to 10 seconds	**Mostly On:** 3 to 20 seconds	**Mostly Off:** 30 to 55 seconds

Shower

Cold:	**Mostly On:** 5 to 30 minutes
Hot:	**Mostly On:** 5 to 30 minutes
Stack:	**Mostly On:** 0 to 30 minutes

Dishwasher

Hot:	**Mostly On:** 45 seconds to 2.5 minutes	**Mostly Off:** 2 to 3 minutes	**Mostly On:** 45 seconds to 2.5 minutes	**Mostly Off:** 5 to 7 minutes	**Mostly On:** 45 seconds to 2.5 minutes	**Mostly Off:** 45 to 90 seconds	**Mostly On:** 45 seconds to 2.5 minutes

Figure 6. Clothes washer, dishwasher, and toilet activity are matched based on the mechanical patterns in their water usage. Showers are the only activity that continuously use hot and cold water while generating waste water for an extended period of time.

ACTIVITY RECOGNITION PATTERNS

To recognize activities from patterns of activations in the hot, cold, drain, and stack sensors, we developed an algorithm to match sequences of activations over ranges of time intervals. The patterns matched by our algorithm are shown in Figure 6. While we manually crafted these patterns by examining our training data, our discussion comments on the potential automatic generation of such patterns. The structure of these patterns is based on the need to account for two aspects of our data: sensor activations might not be continuous and activities might be based on multiple activations of varying length.

To illustrate the first problem, consider that a 60-second activation of cold water associated with a toilet flush might not be detected as a single 60-second activation. Instead, the model analyzing the sensor might output, for example, a 16-second activation followed by a 4-second interval labeled as off and then another 40-second activation. An approach based completely on rules to identify toilet flushes by looking for 60-second uses of cold water would likely fail in this case. Similar problems are common due to the failure of a wireless packet transmission, which creates a short gap in the data available for a sensor. Our algorithm is therefore based on, for example, finding a 60-second interval in which a sensor is mostly active.

By itself, this first problem could have been addressed using a convolution-based approach, but a convolution cannot account for the fact that the structure of some activities varies with the length of activations within those activities. For example, the time required to fill the clothes washer in this dataset varies based on whether it is being filled with just cold water or with both hot and cold water (requiring less time to fill the washer when both water sources are being used) and may also vary based on the amount of clothes in the washer. Because the length of each fill is not fixed, no single template is appropriate for a convolution (the difference in the length of a fill in the template versus the data would create an error that would cascade through the rest of the convolution). Our algorithm therefore matches each activation independently and checks whether the sequence of activations matches an activity.

Our recognition algorithm applies the above patterns, starting with clothes washer, then the dishwasher, then the shower, and finally the toilet. Any remaining unexplained hot water usage is considered a sink-related activity. As the kitchen sink and the bathroom sink drain down different pipes, we first check for activity on the stack or the drain. If the stack is active for more of the time than hot water was being used, bathroom sink activity is recognized. If the drain is active for more of the time that hot water was being used, kitchen sink activity is recognized. In many cases, no activity is detected on either the stack or the drain. In this situation, the recognizer uses the heuristic that a person is likely in the bathroom if any bathroom activity occurs within four minutes of the unexplained hot water usage (an earlier or later toilet flush, an earlier or later shower, or earlier bathroom sink activity). Otherwise, the hot water usage is labeled as kitchen sink activity.

MODEL EVALUATION

As discussed by Tapia *et al.*, the evaluation of activity recognition is highly dependent on how a model will be used [10]. Some applications may require exact start and stop times for an activity, while others may only need an indication that an activity occurred. This section examines our algorithm's ability to recognize that an activity occurred and provide a reasonable estimate of the time interval in which the activity occurred.

For reasonably long-lasting activities (lasting more than 30 seconds), our evaluation is based on the length of the actual activity. The difference between the start time of the actual activity and the start time of the recognized activity must be less than 40% of the length of the actual activity. Similarly, the difference between the end times of the recognized and actual activities must be less than 40% of the length of the actual activity. Finally, the difference in the length of the two activities must be less than 40% of the length of the actual activity. These rules require both a useful indication of the length of an activity and a significant overlap in the actual and recognized activity times. For example, it is acceptable to recognize a 60-second toilet flush when the actual flush lasted 56 seconds, but it would be incorrect if the recognized flush was reported 30 seconds after the actual flush started.

For very short activities, the 2-second sampling rate of our microphone-based features makes the above requirements impractical. If an activity lasts for 6 seconds, reporting that it started one reading later than it did or that it lasted for 8 seconds would result in a recognition being marked as incorrect. While this might be the correct approach for an application that requires exact timing, it is not the correct approach for the elder care applications that motivate our work. For activities lasting less than 30 seconds, we consider recognition correct if the actual and recognized activity overlap or if the gap between them is less than the length of the actual activity. This applies only to kitchen and bathroom sink activities, as these are the only activities that can last for less than 30 seconds. This also means multiple recognized activities can map to the same actual activity. For example two 4-second uses of a sink could be recognized when there is a single 10-second actual use.

Recognition Results

Based on these criteria for a match, Figure 7 presents the accuracy of our activity recognition as tested against our 33 days of test data. The unshaded rows present the reliability of recognition for the full test dataset, while the shaded rows separate out potential causes of error. The *In Isolation* rows use our actual water flow sensor models, but only consider activities that do not overlap any other activity. The *Ideal Sensors* rows remove the effect of errors in our water flow sensor models by using our manually-coded labels for those sensors. Finally, the *Ideal in Isolation* rows use our manually-coded water flow labels to examine only activities that do not overlap any other activity. The *Actual* column indicates how many instances of an activity are present in the test data, and the *Found* column indicates how many of these were found by our recognizer. The *False* column indicates how many recognized instances of an activity do not match an actual instance. The precision is the ratio of these false instances to the total number of recognized instances. Note that the total number of recognized instances is not the sum of *Found* and *False*, because multiple recognized instances of an activity can match the same actual instance. We also do not present a separate precision for the isolated rows because it is not clear how a false positive can be attributed to either an isolated or non-isolated actual instance.

	Actual	Found	Recall	False	Prec
Clothes Washer	16	16	100%	0	100%
In Isolation	4	4	100%		
Ideal Sensors	16	16	100%	0	100%
Ideal in Isolation	4	4	100%		
Dishwasher	20	19	95%	0	100%
In Isolation	10	10	100%		
Ideal Sensors	20	19	95%	0	100%
Ideal in Isolation	10	10	100%		
Shower	66	58	88%	4	94%
In Isolation	28	23	82%		
Ideal Sensors	66	64	97%	1	98%
Ideal in Isolation	28	28	100%		
Toilet	419	329	79%	10	97%
In Isolation	281	243	86%		
Ideal Sensors	419	380	91%	4	99%
Ideal in Isolation	281	277	99%		
Long Bathroom Sink	262	170	65%	153	65%
In Isolation	156	93	60%		
Ideal Sensors	262	221	84%	23	91%
Ideal in Isolation	156	143	92%		
Long Kitchen Sink	356	278	78%	60	84%
In Isolation	308	263	85%		
Ideal Sensors	356	300	84%	15	95%
Ideal in Isolation	308	281	91%		
Short Bathroom Sink	42	22	52%	43	92%
In Isolation	34	16	47%		
Ideal Sensors	42	37	88%	21	75%
Ideal in Isolation	34	31	91%		
Short Kitchen Sink	168	88	52%	106	62%
In Isolation	153	84	55%		
Ideal Sensors	168	128	76%	0	100%
Ideal in Isolation	153	122	80%		

Figure 7. The reliability of our activity recognition algorithm. Unshaded rows show the reliability of models for all cases in the test data, while shaded rows examine causes for errors.

The clothes washer pattern successfully matches all occurrences of the clothes washer, with no false positives. Similarly, the dishwasher pattern successfully matches 19 of 20 occurrences, with no false positives. The single failure is the case where a person is showering while the dishwasher is running. The continuous hot water usage of the shower interferes with the detection of a delay between two fills associated with the dishwasher pattern.

The shower pattern matches continuous simultaneous hot and cold water usage lasting more than 5 minutes, together with any stack activity occurring during that time. This detects 58 of 66 showers in our test data, with four false positives. The near-perfect performance obtained when matching patterns with the ideal sensors suggests that our errors are due to the models that determine whether water is flowing in each pipe, and an inspection of the errors confirms this. Of the eight failures, seven are caused by a failure of the cold water sensor. In each case, cold water is flowing but the model for the cold water pipe does not detect the flow. In one of these cases, the hot water model also fails. The eighth error is a case where the length of the recognized shower does not correspond to the length of the actual shower, again due to a failure of the cold water sensor. Note that this case is also one of the false positives. The other three false positives are caused by the erroneous detection of water flow by the hot water sensor (caused by the central air conditioning rattling the pipe) at the same time as cold water is being used for an extended period of time (likely by an outside faucet used to water the garden).

Our toilet matching pattern correctly identifies 329 of 419 actual toilet flushes, with 97% of recognized flushes corresponding to an actual flush. There again is a noticeable improvement when matching the ideal water sensors instead of the actual sensors, indicating that noise in the low-level models is causing errors in the higher-level recognition. The recall also rises when examining toilet flushes that occur in isolation, suggesting that showering and other simultaneous activities are a significant cause of recognition error. This is consistent with the example presented in Figure 5(b), where the use of cold water by a shower masks the use of cold water by the toilet.

While noise in the low-level water flow models introduces some problems for the highly-structured activities discussed so far, it is especially problematic for sink activity. For example, random noise from a low-level model is unlikely to take a form resembling a toilet flush. But sink activity lacks a high-level structure that can be used to distinguish it from noise. While we are still able to recognize 65% of bathroom sink activity lasting 10 seconds or longer and 78% of kitchen sink activity lasting 10 seconds or longer, we can see that the reliability of recognition is much higher when considering the ideal water flow sensors. This suggests that improving the reliability of the low-level water flow sensors would improve these results, an issue we discuss next. That the ideal accuracies are also not close to 100% indicates room for improvement in our recognizer algorithm. For example, many kitchen or bathroom sink activities are classified as the wrong type of sink activity. Our four-minute threshold (sink activity within four minutes of any bathroom activity is recognized as bathroom sink activity) was chosen arbitrarily, and is probably not optimal. A recognizer could also consider the time of day when classifying sink activity, or patterns in how long a person stays in the bathroom after a shower or toilet flush. Anecdotally, kitchen sink activity seems to occur in bursts of many sink usages, while bathroom sink usage seems to occur once or twice in close proximity to another bathroom activity. The rapid occurrence of several sink activities could suggest to a recognizer that they are more likely to be taking place at the kitchen sink.

Analysis of the Secondary Hot Water Sensor

As previously noted, our pilot data collection showed that the central air conditioning unit introduced significant noise into the nearby hot water sensor. We therefore deployed a secondary hot water sensor, still in the basement but far away from the central air conditioning unit. While all of our previous analyses have been based on the hot water sensor attached to the hot water heater, we now further examine the impact of this systematic noise on our models by constructing a model using this secondary hot water sensor.

We first constructed new low-level models for the hot and cold water sensors (because these models are considered related, they include features based on each other and both need to be rebuilt). The new models perform as follows:

Hot:	*Recall:* 92 ± 01%	*Precision:* 95 ± 01%	F_1: .94 ± .01
Cold:	*Recall:* 93 ± 01%	*Precision:* 95 ± 00%	F_1: .94 ± .01

Most notably, the precision of the hot water sensor increases from 71% immediately beside the air conditioning unit to 95% at our secondary sensor.

	Actual	Found	Recall	False	Prec
Clothes Washer	16	16	100%	0	100%
In Isolation	4	4	100%		
Dishwasher	20	19	95%	0	100%
In Isolation	10	10	100%		
Shower	66	62	94%	2	97%
In Isolation	28	27	96%		
Toilet	419	368	88%	7	98%
In Isolation	281	266	95%		
Long Bathroom Sink	262	192	73%	28	83%
In Isolation	156	93	60%		
Long Kitchen Sink	356	288	81%	18	92%
In Isolation	308	270	88%		
Short Bathroom Sink	42	26	62%	26	90%
In Isolation	34	21	62%		
Short Kitchen Sink	168	98	58%	16	93%
In Isolation	153	84	55%		

Figure 8. The reliability of a model based on our secondary hot water sensor, which was placed further away from the systematic noise created by the central air conditioning unit.

Figure 8 presents the result of applying our recognition algorithm based on the secondary hot water sensor. While still not as reliable as a model based on our manual labels, there are many improvements over the models based on the noisier hot water sensor. The recall of shower recognition improves from 88% to 94%. The recall of toilet recognition improves from 79% to 88%. Long bathroom sink activity recognition improves from a recall of 65% to 73% and from a precision of 65% to 83%. Long kitchen sink activity recognition improves from a precision of 84% to 92%. As a whole, the improved results obtained with our secondary hot water sensor further illustrate the importance of accounting for systematic noise in recognition.

DISCUSSION

The modeling presented in this paper is compatible with a scenario in which a home monitoring service, such as those currently providing home alarm monitoring, uses an expert interface to manually label a small amount of data collected in the initial week that sensors are deployed. Our experience in labeling the data collected in this study leads us to believe that a expert could label the structured activities in a dataset without a need for ground truth sensors in the living space. For the unstructured activities, such as sink usage, an unobtrusive sensor (such as a motion detector) could be deployed for this training.

Moving forward, we intend to pursue a variety of future work. Our results show that the informed placement of sensors can greatly reduce the effects of systematic noise sources (as with our secondary hot water sensor), but we are also interested in unsupervised approaches to detecting and accounting for systematic noise. Systematic noise likely affects only a single sensor, so analyzing patterns in the co-activations of sensors should provide insight into detecting and accounting for systematic noise. We are also generally interested in unsupervised approaches to learning the types of activity patterns used in this work, though it

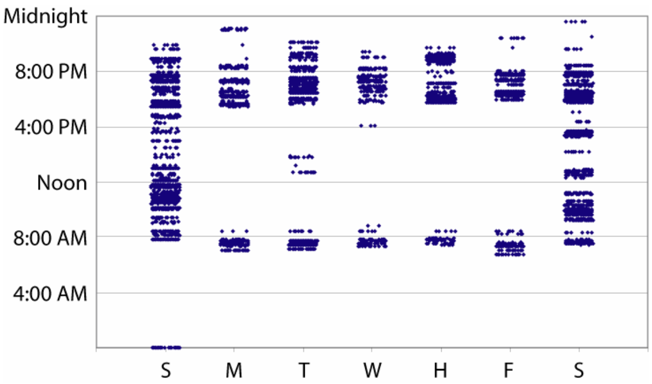


Figure 9. Kitchen sink activity, by day of week. Meal preparation times can be seen in the dense bands, and morning activity starts later on the weekend.

would be premature to use a single dataset to pursue such work. We therefore intend to collect additional data in a variety of homes to further examine our approach.

ILLUSTRATING ACTIVITY TRENDS

The focus of this paper is on examining the feasibility of our new approach to home activity recognition, but it is interesting to consider how applications might use this recognition. Figure 9 shows all of the kitchen sink activity recognized in our test data, plotted using a visualization similar to those developed by Begole *et al*. [3, 4]. With kitchen sink usage grouped by day of week, it is clear that the occupants prepare a morning meal before leaving for work and another meal after returning from work. Activity starts later on Sundays, indicating that both people sleep later than usual. This visualization, and others like it, could be very useful in applications like the Digital Family Portrait [11, 14] and the CareNet Display [6]. Even though our recognizer makes occasional recognition errors, trends in behavior are clearly visible in the recognized activities.

CONCLUSION

This paper demonstrates the feasibility of a new approach to home activity recognition, using unobtrusive and low-cost microphone-based sensors that "reach into the home" from the basement. Because we use sensors that can be attached to the outside of the pipe, there is no need for a professional plumber during installation (as would be required if using flow sensors installed within the pipe). Because we use just a handful of sensors in the basement, a variety of information important to elder care applications can be collected without intruding into the living space. Our approach therefore represents an important and interesting new point in the design space surrounding home activity recognition for elder care applications.

ACKNOWLEDGMENTS

We would like to thank Joonhwan Lee for creating the schematic in Figure 1. This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. NBCHD030010, and by the National Science Foundation under grants IIS-0121560 and IIS-0325351.

REFERENCES

1. Abowd, G. and Mynatt, E.D. (2000) Charting Past, Present, and Future Research in Ubiquitous Computing. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *7*(1). 29-58.
2. Beckmann, C., Consolvo, S. and LaMarca, A. (2004) Some Assembly Required: Supporting End-User Sensor Installation in Domestic Ubiquitous Computing Environments. *Proceedings of the International Conference on Ubiquitous Computing (UbiComp 2004)*, 107-124.
3. Begole, J.B., Tang, J.C. and Hill, R. (2003) Rhythm Modeling, Visualizations, and Applications. *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST 2003)*, 11-20.
4. Begole, J.B., Tang, J.C., Smith, R.B. and Yankelovich, N. (2002) Work Rhythms: Analyzing Visualizations of Awareness Histories of Distributed Groups. *Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW 2002)*, 334-343.
5. Chen, J., Kam, A.H., Zhang, J., Liu, N. and Shue, L. (2005) Bathroom Activity Monitoring Based on Sound. *Proceedings of the International Conference on Pervasive Computing (Pervasive 2005)*, 47-61.
6. Consolvo, S., Roessler, P. and Shelton, B.E. (2004) The CareNet Display: Lessons Learned from an In Home Evaluation of an Ambient Display. *Proceedings of the International Conference on Ubiquitous Computing (UbiComp 2004)*, 1-17.
7. Crossbow Technology. *http://www.xbow.com*
8. Culler, D.E., Hill, J., Buonadonna, R. and Woo, A. (2001) A Network-Centric Approach to Embedded Software for Tiny Devices. *Proceedings of the International Workshop on Embedded Software (EMSOFT 2001)*.
9. Hirsch, T., Forlizzi, J., Hyder, E., Goetz, J., Kurtz, C. and Stroback, J. (2000) The ELDer Project: Social, Emotional, and Environmental Factors in the Design of Eldercare Technologies. *Proceedings of the ACM Conference on Universal Usability*, 72-79.
10. Munguia Tapia, E., Intille, S.S. and Larson, K. (2004) Activity Recognition in the Home Using Simple and Ubiquitous Sensors. *Proceedings of the International Conference on Pervasive Computing (Pervasive 2004)*, 158-175.
11. Mynatt, E.D., Rowan, J., Jacobs, A. and Craighill, S. (2001) Digital Family Portraits: Supporting Peace of Mind for Extended Family Members. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2001)*, 333-340.
12. Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H. and Hahnel, D. (2004) Inferring Activities from Interactions with Objects. *IEEE Pervasive Computing*, *3*(4). 50-57.
13. Platt, J.C. Fast Training of Support Vector Machines using Sequential Minimal Optimization. In Schölkopf, B., Burges, C. and Smola, A. eds. *Advances in Kernel Methods: Support Vector Learning*, MIT Press, 1999, 185-208.
14. Rowan, J. and Mynatt, E.D. (2005) Digital Family Portrait Field Trial: Support for Aging in Place. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2005)*, 521-530.
15. Wilson, D.H. and Atkeson, C.G. (2005) Simultaneous Tracking and Activity Recognition (STAR) Using Many Anonymous, Binary Sensors. *Proceedings of the International Conference on Pervasive Computing (Pervasive 2005)*, 62-79.
16. Witten, I.H. and Frank, E. (1999) *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann.