Pervasive and Mobile Computing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

journal homepage: www.elsevier.com/locate/pmc

Fast track article

Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments

Eric Larson [a,*], Jon Froehlich [b], Tim Campbell [c], Conor Haggerty [d], Les Atlas [a], James Fogarty [b], Shwetak N. Patel [a,b]

[a] Electrical Engineering, University of Washington, Seattle, WA 98195, United States
[b] Computer Science & Engineering, University of Washington, Seattle, WA 98195, United States
[c] Mechanical Engineering, University of Washington, Seattle, WA 98195, United States
[d] Community, Environment, and Planning, University of Washington, Seattle, WA 98195, United States

ARTICLE INFO

Article history:
Received 3 January 2010
Received in revised form 24 August 2010
Accepted 30 August 2010
Available online xxxx

ACM classification codes:
H5.2
H1.2

Keywords:
Infrastructure-mediated sensing
Activity sensing
Sensors
Water usage

ABSTRACT

We present an extended analysis of our previous work on the HydroSense technology, which is a low-cost and easily installed single-point sensor of pressure for automatically disaggregating water usage activities in the home (Froehlich et al., 2009 [53]). We expand upon this work by providing a survey of existing and emerging water disaggregation techniques, a more comprehensive description of the theory of operation behind our approach, and an expanded analysis section that includes hot versus cold water valve usage classification and a comparison between two classification approaches: the template-based matching scheme used in Froehlich et al. (2009) [53] and a new stochastic approach using a Hidden Markov Model. We show that both are successful in identifying valve- and fixture-level water events with greater than 90% accuracies. We conclude with a discussion of the limitations in our experimental methodology and open problems going forward.

1. Introduction and motivation

Infrastructure-mediated sensing is based on the idea that human activities can be sensed via their manifestations in an environment's existing infrastructures. Recent work has examined how a home's water [1], electrical [2–5], and HVAC [6] infrastructures can be exploited to sense various human activities in the home such as opening a refrigerator door or walking through a doorway. Because of its practical, low-cost, and unobtrusive nature, infrastructure-mediated sensing offers significant promise as a general human activity sensing method. A home's water infrastructure is a particularly rich sensing space because of its relevancy to core activities of human life (e.g., drinking, bathing, cooking, cleaning, and gardening). As such, sensing a person's water usage in the home, like taking a shower or using the kitchen sink, can reveal important information about their activities and well-being. Sensing water usage also has significant implications for water conservation. Residential water use accounts for 50%–80% of public water supply systems and 26% of total use in the US [7]. Although most people have relied on the benefits of modern plumbing their entire lives, few actually know how much water they consume or what activities consume the most water (e.g., showering versus toilet flushing). Providing fixture-level details on water usage to consumers and utilities alike will likely reduce wasteful practices (as eco-feedback has done for energy [8,9])

* Corresponding author.
E-mail addresses: eclarson@uw.edu (E. Larson), jfroehli@uw.edu (J. Froehlich), tcampbll@uw.edu (T. Campbell), conorh@uw.edu (C. Haggerty), atlas@uw.edu (L. Atlas), jfogarty@uw.edu (J. Fogarty), shwetak@uw.edu (S.N. Patel).

1574-1192/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.pmcj.2010.08.008

Fig. 1. HydroSense can be easily installed by screwing the pressure sensor onto an exterior hose bib, water heater drain valve, or utility sink faucet (in the same way that one would attach a garden hose). If these installation points are not available, HydroSense can also be installed at the 3/8″ hose connection points of toilets, bathroom and kitchen sinks (far right image).

and allow utilities to better evaluate the success of their conservation programs. Some public utilities, for example, offer customers free low-flow shower heads, but have no means of quantifying how much water and money this saves.

This paper presents HydroSense, a low-cost and easy-to-install solution for sensing water usage from a single installation point. HydroSense is based on the continuous analysis of *pressure* within a home's water infrastructure. When a water fixture is opened or closed, its valve generates a unique pressure wave that propagates throughout the plumbing system. HydroSense observes these pressure transients or waveforms and classifies them into fixture usage events (e.g., the upstairs bathroom toilet was just flushed). HydroSense also provides estimates of the *amount of water being used* at each fixture based on the magnitude of the resulting *pressure drop* within the water infrastructure during water usage events.

HydroSense supports two key application areas of ubiquitous computing: (1) human activity sensing (e.g., aging in place elder care applications [10,11]) and (2) eco-feedback applications (i.e., applications that present information about consumption in order to reduce wasteful practices [9]). Both areas require different ways of thinking about water usage and have different sensing requirements. For assistive care applications, water usage activities serve as a rich proxy for inferring both a person's activity (e.g., cooking, using the bathroom) *and* their location in the home (e.g., kitchen, bathroom). Here, simply identifying *what* and *where* a water usage event occurred is useful. For eco-feedback applications, *how much* water is being used and whether that water is *hot* or *cold* is also relevant. HydroSense supports both of these application areas by automatically identifying water usage activity and providing estimates of flow volume and flow rate down to the individual valve.

HydroSense represents a significant advance over prior research in several regards:

First, HydroSense can be *easily installed* at any accessible location within a home's existing water infrastructure. Typical installations points include an exterior hose bib, utility sink spigot, or water heater drain valve (Fig. 1). The installation point largely does not affect the performance of HydroSense, with the exception of the hot water heater, which can make cold water valves difficult to classify in some homes. If these installation points are unavailable or not easily accessed (e.g., in an apartment unit), HydroSense can also be installed at the water connection point for a dishwasher, clothes washer, toilet, or sink (Fig. 1, far right). All of these are simple screw-on installation points, with no need for cutting into the pipe nor professional installation. Previous research in water usage event detection (e.g., [12,13]) required inline installations.

Second, HydroSense's analysis of *pressure* provides the unique capability of sensing both the *individual fixture* at which water is currently being used as well as an estimate of the *amount of water being used*. HydroSense is the first practical approach to enabling applications that require both. The sensing of pressure is also less susceptible to ambient noise, as has been encountered in previous microphone-based infrastructure-mediated systems (e.g., [1]).

Third, we provide an in-depth analysis of HydroSense in ten diverse homes and compare two different classification techniques: template matching and a Hidden Markov Model (HMM) approach. Thus, this paper provides a more *robust evaluation* than any previous work on water-related home activity sensing. In particular, we demonstrate reliable segmentation of valve pressure *events* from the surrounding sensor stream, show reliable classification of *valve open* and *valve close* events, show the successful identification of *individual fixtures* and *individual valves* with greater than 90% aggregate accuracy, and show that an appropriately located and calibrated system can estimate water usage with error rates comparable to empirical studies of traditional utility-supplied water meters. In addition, we present initial forward-looking analyses of compound event detection, a comparison of sensing at different locations, and a first look at the long-term stability of pressure event signatures. Our evaluation both validates the feasibility of our approach and provides a basis for future analyses and improvements.

The remainder of this paper discusses the theory behind our approach, presents our sensor implementation, and summarizes our in-home data collections. We then present our analyses of individual fixture identification and water flow

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

Study (year)	Goal	Flow Measurement?	Hot/Cold Disaggregation	Class. Method	Classification Level	Sensing
Ladd and Harrison (1985)	Study residential end uses of hot water	Hot water only, inline	Yes, sub-metered at water heater	Direct sensing	Hot valves only	Inline flow meters (at washer and hot water tank) Elec. current sensors
Weihl and Kempton (1985)	Study residential end uses of hot water	Hot water only, inline	Yes, sub-metered at water heater	Direct sensing	Hot valves only	Inline flow meter (at hot water tank) Temperature sensors
Lowenstein and Hiller (1996)	Study residential end uses of hot water	Hot water only, inline	Yes, sub-metered at water heater	Flow-trace analysis	Hot valves only	Inline flow meter at hot water tank
Mayer, et. al. (1999)	Study residential end uses of water	Yes, inline	No	Flow-trace analysis	Fixture category	Inline flow meter
DeOreo and Mayer (2000)	Study residential end uses of hot and cold water	Yes, inline	Yes, sub-metered at water heater	Flow-trace analysis	Fixture category	Dual inline flow meter (whole home meter plus hot water tank meter)
Fogarty, et al. (2006)	Activity inference (e.g., for elder care applications)	No	No	Machine learning	Fixture category	Acoustic sensors on intake pipes and drain pipes
Kim, et al. (2008)	Eco-feedback	Yes, inline	Yes, direct	Direct sensing	Untested (lab study only)	Inline meter accelerometers on pipe
HydroSense (2009)	Activity Inference / Eco-feedback	Yes, calibrated from pressure	Yes	Machine learning	Valve	Single pressure sensor

Fig. 2. A view of the related work in the evolution of water usage measurements. Most studies before 2006 were designed to measure end uses of water and thus focus on monitoring water at the fixture category level rather than the individual valve or specific fixture.

estimation, followed by a discussion of important directions for future work. First, however, we review the state of the art in human activity sensing and current work in disaggregated water usage and flow sensing.

2. Related work

There are four high-level areas of work related to HydroSense: (1) sensing and inferring human activity in the home, (2) sensing water flow, (3) automatically identifying water usage events, and (4) disaggregating between hot and cold water usage. We address each of these related areas in turn. Fig. 2 provides a summary of the water sensing disaggregation related work and its evolution over time.

2.1. Sensing human activity in the home

Prior work has demonstrated at least three approaches to the fundamental challenge of sensing human activity in the physical world: mobile and wearable sensing, distributed direct environmental sensing, and infrastructure-mediated environmental sensing. Promising mobile and wearable activity sensing methods include accelerometer-based activity recognition [14,15] and the detection of interaction with tagged objects via a wearable RFID reader [16]. There are many compelling applications of mobile and wearable methods, but they share a common need for a person to be willing to wear or carry the necessary device.

Environmental sensing systems take a complementary perspective, instrumenting an environment to detect activity within it. In the home, distributed direct sensing can be based on computer vision [17], microphones within the living environment [18], many simple sensors throughout the home (e.g., reed switches on cabinet doors, accelerometer-based object manipulation sensors, infrared motion and break-beam sensors [19–21]), or a smaller number of targeted direct sensors (e.g., strain sensors under floorboards at strategic locations [10]). Direct sensing provides valuable insight into home activities, but comes with practical costs. Installation and maintenance can be cost-prohibitive, direct sensing can create privacy concerns and a feeling of stigmatization (especially with cameras or microphones), and a littering of sensors throughout a home can be problematic with children and pets [22,23].

Infrastructure-mediated sensing [1,6,2,5] has therefore evolved to help address many practical obstacles to everyday deployment of home activity sensing, but is inherently limited by what information can be practically and reliably extracted from a home's infrastructure. The limitations of existing approaches are especially salient for water. Fogarty et al. used microphones pressed against the exterior of a home's major water pipes (cold water inlet, hot water inlet, waste water exit) to demonstrate recognition based on patterns of water use (e.g., the series of fill cycles associated with a dishwasher [1]). However, Fogarty et al. found they could not reliably differentiate among multiple instances of similar fixtures (e.g., multiple sinks or toilets within a home), could not reliably identify concurrent activities (e.g., a toilet flush while a person is showering), and did not attempt to estimate the volume of water being used. They also reported difficulties with ambient noise and audio-based sensors (e.g., an air conditioning unit in close proximity to a sensor placed on a home's hot water heater). By addressing these shortcomings, HydroSense significantly advances the state of the art for water-mediated home activity sensing.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

2.2. Sensing water flow

There are two basic approaches for measuring water flow: *inline direct* flow measurement and *non-intrusive flow* estimation. Inline systems typically use either positive displacement or velocity measurements to calculate flow. Positive displacement relies on water to physically displace the measuring element (e.g., an oscillating piston or rotating disc) in proportion to water flow volume. Most residential meters use positive displacement because it is generally accurate at the low to moderate flow rates found in the residential sector [24].

Meters that do not require physical contact with water flow are called *non-intrusive*. These methods are attractive because they do not require pipe modification for installation, which can be costly and requires the expertise of a plumber. Non-intrusive flow estimation techniques use either active or passive sensing approaches. Many active sensing systems use a non-intrusive actuation probe that is distorted by water flow and perceived by an awaiting sensor. For example, Ultrasonic Doppler Velocimetry (UDV) uses the sound transit time and Doppler shift from pulsating ultrasonic waves emitted into the fluid flow to determine velocity [25]. Similar methods have been developed using lasers (e.g., Laser Doppler Velocimetry [26]). A more recent technique, called Particle Image Velocimetry, uses computer vision to record the position of tracer particles injected into the stream over time [27]. Although each of these techniques are often extremely accurate and can be used to measure flow of dangerous liquids (e.g., hot metallic melts), they are intended for industrial applications (e.g., manufacturing plants or large-scale irrigation systems) and thus are prohibitively expensive for typical residential users (with units ranging from $2000 to 8000).

Lower-cost non-intrusive techniques have been proposed and investigated that use *passive* sensing. Evans et al. showed in a laboratory environment that accelerometers mounted on the exterior of water pipes have a strong deterministic relationship to water flow rate [28], but this is highly sensitive to pipe diameter, material, and configuration. Kim et al. demonstrated the use of an aggregate water flow meter together with a network of accelerometers on pipes to infer flow rates throughout a home [29]. Both of these approaches require placement of multiple sensors along water pipe pathways that are uniquely associated with each fixture of interest (i.e., they are distributed direct sensing methods).

The next generation of resource measurement systems (often referred to as "smart meters") will soon provide real-time (or near real-time) data on electricity, gas, and water usage in homes and businesses. It is unclear, however, if this data will be open or closed (i.e., proprietary) and what temporal resolution will be available (most smart meters advertise fifteen minute interval data). For those water meters not yet upgraded to their smart meter counterparts, a popular and low-cost non-intrusive retrofit has emerged using magnetic sensors. Most residential meters use magnetic coupling to transfer data between the rotating disc, which spins proportional to flow rate, and the counter unit, which converts this spin rate to flow. Magnetic sensors, such as a hall effect sensor, can be placed at the top or bottom of a residential water meter to sense this spinning magnetic field and convert the spin rate to flow (e.g., see [30,31]). AquaCraft [32] has used this approach since 1996 to digitize traditional inline water meters and perform flow-trace analysis, a technique used to automatically identify water usage events.

2.3. Identifying water usage events

The goal of automatically identifying water usage events in the home has been pursued by utilities and water resource management scientists to better understand residential end uses of water. Only recently, with advances in technology and reductions in cost, has the ability to measure water at the fixture level from a *single installation* point been possible. In 1992, Dziegielewski, et al. [33] proposed what is now the standard water usage event identification technique called *flow-trace analysis*. Flow-trace analysis works by analyzing the aggregate flow-trace patterns off of an inline water meter to determine the source of the water usage event. Flow-trace analysis relies on the fact that most residential water uses exhibit highly consistent behavior over time (i.e., a specific toilet will flush with the same volume and flow; similarly a specific dishwasher will exhibit the same series of flow patterns each time it is run [31]).

Flow-trace analysis has since been used in a number of government- and utility-sponsored studies of residential end uses of water [34,12,13,35,36] (see also Fig. 2). The most popular of which is a large study conducted by the American Water Works Association Research Foundation (AWWARF), which used AquaCraft's flow-trace analysis toolkit called *Trace Wizard* to study residential end uses of water in 1188 households across 12 study sites in America and Canada [31]. The water data was collected using the aforementioned magnetic sensing retrofit solution. The data collection took place in the winter and the summer months from 1996 to 1998 and resulted in 1.9 million water usage events. A total of four weeks of water data per home was collected in separate two-week intervals. This study provided the canonical dataset upon which most US government statistics of residential end uses of water are based [7]. That said, as far as we are aware, the accuracy of flow-trace analysis has not been comprehensively studied. In the AWWARF study, the authors speculate that their system is 90% accurate in correctly identifying water usage events (p. 47 in [31]); however, they were not able to offer concrete accuracy statistics as they did not collect ground truth data. In addition, although flow-trace analysis is capable of classifying at the fixture category level, it cannot be used to determine the specific fixture or valve that was used (i.e., it can sense that *a* toilet was flushed, but not *which* toilet was flushed). In contrast, HydroSense can identify water usage events at the valve level.

Finally, in its current form, flow-trace analysis is not completely automated. Usage of AquaCraft's Trace Wizard, for example, is a multi-step, iterative process. The program uses the following set of statistics to help classify water usage

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

events: start time, stop time, duration, volume (gallons[1]), peak flow rate in gallons per minute (gpm[2]), the most common flow rate (gpm), and how often this most common flow rate occurs during the duration of the event. Events are classified according to their similarity with a pre-defined set of manually entered and tuned per-home parameters. For example, a toilet may be defined as using between 3.25 and 3.75 gallons per flush, the peak refill flow rate as between 4.2 and 4.6 gpm, the duration of the flush event between 30 and 50 s, and the mode flow rate between 4 and 4.5 gpm [31]. AquaCraft estimates that it takes their trained analysts approximately one hour per week of data from a home to complete a flow-trace analysis; less time after the parameter file has been properly tuned [31].

2.4. Disambiguating hot water versus cold water usage

The lack of precise measurements about the quantities of hot water used in residences has been an obstacle in the design of hot water systems (e.g., how much hot water do homes require at peak use? [37,38]) and in the analysis of conservation programs (e.g., how much hot water does a low-flow shower head save? [33]). Hot water usage also highlights the strong interconnection between energy and water conservation efforts in the home. According to the US Department of Energy, hot water heating is the fourth most energy consuming activity in the home behind air conditioning, refrigeration, and space heating [39]; however, *what* hot water usage activities contribute to this consumption are much less understood [40].

The measurement and analysis of domestic hot water consumption is typically accomplished using one of two methods: a temperature-based inference method or a flow-trace analysis method (where the inline meters are installed at the hot water heater intake line [41]). In 1985, Weihl and Kempton conducted the first quantitative study of hot water usage using automatic means [42]. They installed an inline flow meter and a temperature probe at the hot water tank as well as temperature probes at the hot water pipes leading to each fixture. Flow rate and temperature were recorded to a computerized data logger once per minute. From this data, the authors were able to deduce which fixtures were using water, how much hot water was being used, and also calculate pipe heat loss. Ladd and Harrison used a similar distributed sensing approach; instead of temperature probes, however, they used power consumption monitors on the dishwasher and clothes washer, an inline flow meter installed at the clothes washer, and an inline flow meter at the hot water tank [43].

Lowenstein and Hiller [13] simplified the above approaches by eliminating the need for multiple sensors. They collected flow-trace data at 15 s intervals at an inline meter installed at the hot water tank feed line. The volume of flow and the average flow rate for each hot water draw were used to classify the hot water usage events into their end use-category, a technique they referred to as "bin analysis", but is simply a reduced form of flow-trace analysis with fewer parameters (e.g., without time of day, event duration). Although this technique was used successfully in a year-long study of seventeen sites (and repeated by DeOreo in a study of 14 Seattle homes [37]), it was noted by the authors in a follow-up paper [44] to have two major limitations: (1) they were not able to discriminate between end uses when multiple hot water activities occurred at the same time and (2) hot water draws for different end uses can have the same flow rate and total flow volume, making it difficult to unambiguously associate specific end uses. Thus, in their follow-up study, they attached thermocouples 1–3 inches[3] downstream from where the hot water line branched, up to three thermocouples per home (depending on its plumbing layout) to mitigate these limitations.

HydroSense presents a unique opportunity for discriminating between hot and cold water usage because it is much easier to install than the above approaches. Pressure waves are shaped not just by the valve type but also by their propagation pathway through the pipe system, allowing HydroSense to disambiguate hot water events from cold water events (even though the hot and cold water valves are right next to each other at a faucet, their pipe pathway through the home is quite different).

3. Background and theory of operation

In this section, we provide background on residential water supply systems and in-home plumbing. We also introduce the basic theory of operation that motivates our approach.

Households obtain water from one of two sources: a public water supply or a private well. Public water is distributed by local utilities, relying on gravity and pumping stations to push water through major distribution pipes. Residences are connected to a water main by a smaller service line, where the water meter is typically found. A backflow valve accompanying the water meter prevents household water from flowing back into the main. Homes with private wells use a pump to draw the water out of the ground and into a small tank within the home. A pressure bladder pushes water from the tank when a valve in the home is opened. Private wells are typically unmetered.

Fig. 3 depicts a typical in-home plumbing system. Cold water enters through the service line, typically at 50–100 pounds per square inch (psi[4]) depending on such factors as the elevation and proximity to a reservoir or pumping station. Many homes have a pressure regulator that protects the home from transients (or pressure spikes) from the main and also reduces the incoming water pressure to a level safe for household fixtures.

[1] 1 US gallon ≈ 3.785 L.
[2] 1 gpm ≈ 0.063 L/s.
[3] 1 inch ≈ 2.54 cm.
[4] 1 psi ≈ 68.95 mbar.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

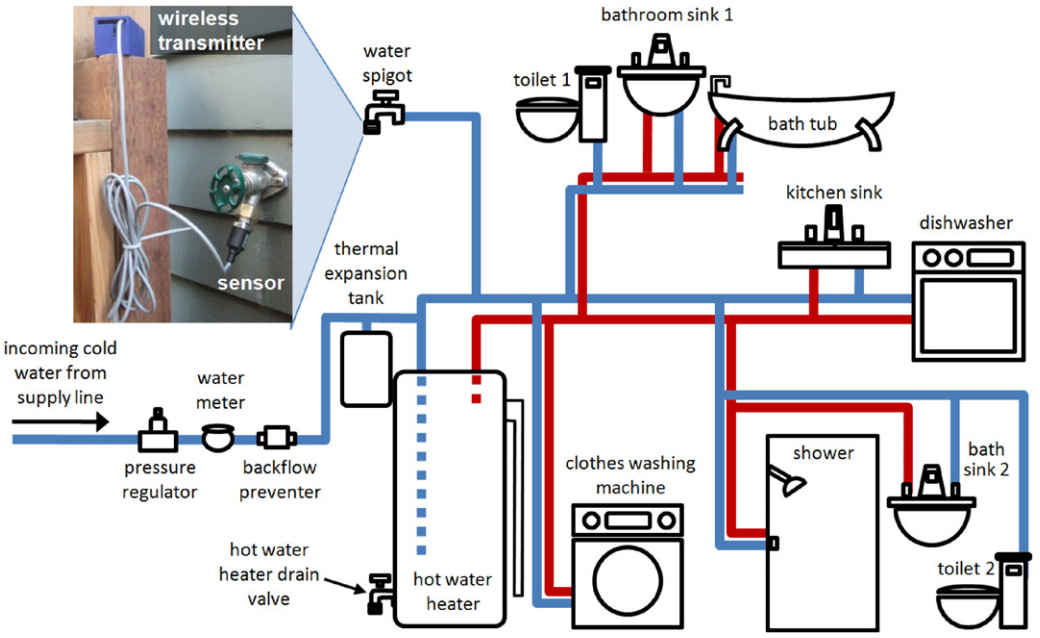

Fig. 3. An illustrative schematic of a basic plumbing layout in a two-bathroom home. HydroSense can be easily installed at any accessible location in a home's water infrastructure, with typical installations at an exterior hose bib (shown above), a utility sink spigot, or a water heater drain valve. By continuously sensing water pressure at this single installation point, HydroSense can both *identify individual fixtures* at which water is being used as well as estimate the *amount of water being used*.

After the regulator, there are two basic layouts found in typical residential piping, *series plumbed* and *branched*. Almost all multi-fixture homes have a combination. The cold water supply branches to the individual water fixtures (e.g., toilets, sinks, and showers) and into the water heater. A traditional water heater heats water in an insulated tank using electric coils or gas. When hot water is used, the pressure from the cold water supply line pushes hot water out of the tank and refills it with cold water. The cold water feed line stretches down to the bottom of the tank. Cold water remains at the bottom until it is heated (as cold water is denser than hot water) and the hot water line draws from the top of the tank. Every hot water tank has a pressure relief valve and a drain valve, which is important for maintenance as water heaters should be drained at least once a year to flush mineral deposits and increase operating efficiency. Many homes also have a thermal expansion tank connected to the water heater, providing space to store excess water as it expands during heating. Some homes instead use tankless heaters, which provide hot water on demand by circulating it through burners or electric coils. Both approaches connect cold and hot pipes of a home's plumbing system. The pressure waves leveraged in our approach travel through this connection, enabling the HydroSense unit to detect both hot and cold water activity with a single sensor.

In summary, the plumbing system forms a closed loop pressure system, with water held at a stable pressure throughout the piping when no water is flowing. Homes with a pressure regulator have stable pressure unless the supply pressure drops below the regulator's set point. Homes without a regulator may experience occasional minor changes in water pressure depending on neighborhood water demand.

For explanatory purposes, it is useful to draw an analogy between a household water pipe network and its electrical equivalent (Fig. 4). Water pressure is analogous to voltage, water flow to electrical current, and water valves to electrical switches with variable resistance. For example, water pressure regulated down to a constant 45 psi is analogous to a 45 V DC voltage supply. Just as water flow occurs in the direction of higher pressure to a lower pressure, so does current flow from a higher voltage to a lower voltage: the greater the difference, the greater the flow. When an electrical switch is thrown, the value of the resistance controls the amount of current that escapes to ground (i.e., the same way that the size and amount that a valve is open controls the amount of water flow to a drain). In a simplified form, a water pipe is analogous to an electrical transmission line with an inductor and a capacitor (Fig. 4, far right). In this way, the network of water pipes in a home can be modeled as a collection of electrical transmission lines (i.e., a linear system with a transfer function given by the interconnection and length of pipes). Although we continue to use some electrical analogies throughout the paper, they serve only as an additional supplement to our explanations and are not necessary to fully understand the principles upon which HydroSense operates.

3.1. Identifying water fixtures

The instant a valve is opened or closed in a water fixture, a pressure change occurs and a *pressure wave* is generated in the plumbing system (Fig. 5). Transient pressure wave phenomenon results from the rapid change of water velocity in a

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

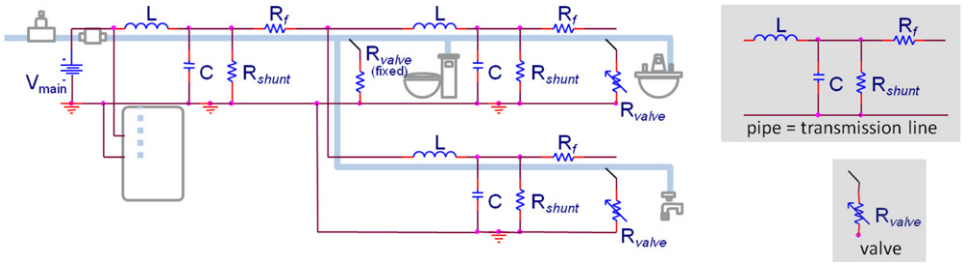

Fig. 4. The transmission line equivalent network for a common bathroom fixture configuration. Only cold water is shown. In this analogy, pressure is analogous to voltage, water flow to electrical current, and the water valve and water pipe as shown.

pipeline. For example, when a running faucet is turned off, the valve closure abruptly stops the water flow. Flowing water in a pipe, however, has momentum and is compressible. So, rather than stopping outright, the momentum compresses a finite volume of water against the closed valve, which results in a buildup of pressure at the valve head. Once this buildup reaches a critical point, the pressure at the valve head starts to decompress and the water flow is reversed until it reaches the backflow prevention valve of the house, where the cycle repeats. In the electrical analogy, this is analogous to the cyclic oscillation of charge and flow between inductors and capacitors in transmission lines.

The water pressure wave phenomenon is often referred to as a *surge* or *water hammer* and can create a loud hammering noise as the wave travels through pipes. The magnitude of the surge is dependent upon the operating pressure of the home and the flow of water through the valve. The water hammer transient can have a positive or negative rate of change depending on whether a valve is being opened or closed. Appliances such as dishwashers or clothes washers control their valves mechanically and thus often create the most pronounced water hammer. An abrupt change in flow can create dangerously high pressure transients that exceed safe operating limits for residential pipes. A thermal expansion tank and water hammer arresters offer partial dampening of these transients. Without arrestors, the flow decreases so quickly that the resulting pressure wave can burst the pipe (like the arc voltage produced by opening a switch near an inductor). However, most valve openings/closings manifest as a water hammer transients that are harmless but can be detected by a pressure sensor installed on the plumbing system. Water hammer typically lasts several seconds as the pressure wave oscillates back and forth through the pipes. We can detect this water hammer effect *anywhere* along the plumbing infrastructure (even with dampeners installed), thus enabling the single-point sensing approach.

The *unique* transient or water hammer signature that we sense for a particular fixture depends on three factors: the *valve type,* the valve's *location* in the home pipe network, and, to a lesser extent, the way in which the valve is opened or closed. Intuitively, you can think of each valve exciting the plumbing system differently along different points in its infrastructure, much like blowing a harmonica at different points produces different sounds. The physical reasons why these factors produce a unique transient, on the other hand, is best explained using the electrical analogy. The *valve type* controls the resistance of the analogous electrical switch. When the switch is closed, a low resistance path to ground is created suddenly generating an *impulse* which excites the network. Valves with lower resistances allow more water flow and create impulses with greater magnitude. The *valve location* changes where the impulse is applied. When the same impulse is applied in a different location, the path back to the sensor goes through a different set of transmission lines, altering the transfer function from the valve to the sensor and, thus, changing the shape of the transient. Moreover, each transmission line has a unique resonance that is excited by the impulse. For example, in Fig. 5 it is easy to see two resonant frequencies in each transient, the lower frequency resonance results from a long length of pipe (which excites a larger wavelength) and the high frequency resonance from a shorter pipe length. This point provides great discriminative power allowing us to distinguish between two fixtures of the exact same model such as the same model toilet located in two different bathrooms of a home because each pressure wave traverses a different path before reaching the sensor. We have even observed that the pressure transient generated between two cold sink valves located in the same bathroom (such as those found in a "his and her" bathroom sink setup) have unique transients even though there is only a small additional length of pipe between them. As this additional length of pipe becomes smaller, the transients begin to look more alike.

Finally, for manually operated valves, the way the valve is operated can affect the impulse applied to the system in two ways: the *speed* at which the valve is opened affects the slope of the excitatory impulse and the *flow rate* the valve opens to affects the magnitude of the impulse. As the speed of the valve opening becomes slower, the input excitation is not modeled well by an impulse, and instead is better modeled by a ramp function. In practice, slowly opened valves are rare occurrences, but opening a valve to different flow rates is quite common (e.g., opening a bathroom sink cold valve full stop to fill a cleaning bucket versus opening the cold valve partially to wet a toothbrush). As the magnitude of the impulse becomes smaller, the amplitude of the resonances excited in each transmission line also becomes smaller, but the resonant frequency does not change. The relationship between impulse magnitude and resonance amplitude is non-linear and system dependent. Transmission lines close to the switch are affected less than lines farther away. The algorithm outlined in this paper does not address this issue directly, but we return to this phenomenon in the discussion section.

Changes in pressure and the rate of transient onset allow us to accurately detect and identify the source of valve open and valve close events. They also allow us to estimate flow. This stems directly from our electrical analogy, where knowing the

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

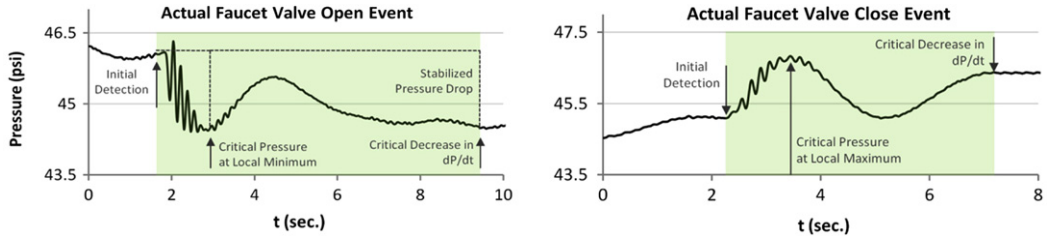

Fig. 5. Actual pressure measurements generated when a kitchen faucet is turned on and turned off, as observed by a single sensor at an exterior water bib. Our previous fixture identification work is based on segmenting the event (indicated by the highlight) and then classifying it according to its shape.

resistance and the change in voltage (i.e., pressure) allows one to determine the current (i.e., flow). Unlike current, however, water flow comes in two forms: laminar and turbulent, which affects the relationship between pressure and flow.

Laminar flow is characterized by the movement of fluid particles parallel to each other, with no transverse movement or mixing. In contrast, turbulent flow is characterized by vigorous intermixing within the flow field resulting in small, random fluctuations in flow. Physically, the two flow states are linked in that any laminar flow can become turbulent with a change in fluid velocity—think of the water flowing out of a typical bathroom faucet: when the faucet is opened partway, water flows in a clear, solid-looking stream and does not splash; however, when the faucet is opened all the way, the water flow turns more opaque, bubbly and chaotic.

The conditions under which liquids transition from laminar to turbulent flow are characterized by a dimensionless value known as the Reynolds number [45], which is dependent on the kinematic viscosity ν of the fluid, the volumetric flow rate of the fluid in a pipe Q, and the radius of the pipe r:

$$R_e = Q\frac{2r}{\upsilon}.$$

It is generally accepted that a Reynolds number less than 2300 results in laminar flow and greater than 4000 result in turbulent flow. In between there is transition between both [45]. Using the most common flow rates for residential water fixtures, it is possible to calculate the Reynolds numbers for a typical home. The Reynolds numbers for water flow in a typical home are in the range of 1000–50,000 for $3/8''$ diameter pipe segments and in the range of 400–22,000 inside the $1''$ diameter supply lines—which means fluid flow in a home can be laminar, turbulent, or a mix of both. We will address the laminar flow condition first, showing that the relationship between pressure drop and flow is linear. We then show that the non-linearity induced by turbulent flow can be sufficiently modeled as linear, with minimal loss of accuracy.

3.1.1. Laminar flow

When laminar, water flow is directly proportional to the pressure drop sensed at the HydroSense sensor. Poiseuille's Law offers a precise definition of this relationship: when the flow through a pipe is laminar, the volumetric flow rate of fluid in a pipe Q is dependent on the radius of the pipe r, the length of the pipe L, the viscosity of the fluid μ and the pressure drop ΔP from the start and end of the pipe length:

$$Q = \frac{\Delta P \pi r^4}{8\mu L}.$$

Note how increasing pipe length reduces flow (by creating more resistance) and increasing the pipe radius pipe increases flow. Pipe length has a linear relationship with flow; if you double the pipe length, the flow is halved. If, instead, you halve the pipe diameter, flow is reduced by a factor of sixteen. Poiseuille's law can be simplified by the fluid resistance formulation, which states that the resistance of flow is proportional to the drop in pressure divided by the volumetric flow rate.

$$R_f = \frac{\Delta P}{Q} \equiv \frac{8\mu L}{\pi r^4}.$$

Thus, we can use fluid resistance to abstract some of the variable complexity from Poiseuille's law, resulting in:

$$Q = \frac{\Delta P}{R_f}.$$

This is analogous to Ohm's law ($I = V/R$). HydroSense measures the change in pressure ΔP from the pressure regulator to the sensor. In order to compute Q, we must estimate the remaining unknown R_f. R_f is bounded by two factors: (1) water viscosity, which can easily be calculated according to temperature and (2) the radius of residential pipes, which are either $1/4''$ or $3/8''$ in diameter ($1''$ for supply lines in an apartment). This leaves L, the length of the pipe, as the main unknown. L will change depending on the water fixture being used, as each path from intake to fixture is different.

3.1.2. Turbulent flow

When the flow of water through the pipes is turbulent instead of laminar, Poiseuille's law does not directly apply. Turbulent flow results in a non-linear relationship between pressure and flow. Instead of pressure being proportional to

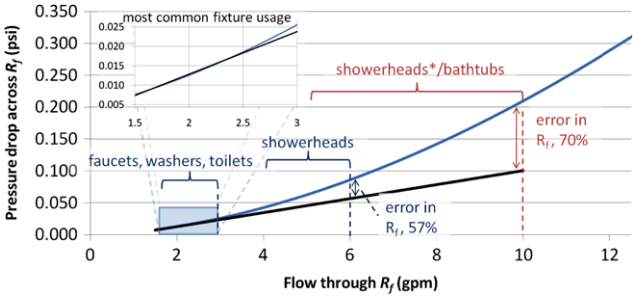

Fig. 6. The pressure versus flow in a 1″ diameter pipe of length 6.5′ using the Blasius formula (blue) and the linearization of the equation (black) around the point of most common fixture usage. Also displayed are the most common fixture flow rates. The (*) denotes a fixture built before the 1992 Energy Policy Act, requiring US manufactured fixtures to use less water. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

flow, the relationship becomes such that P is proportional to $Q^{1.75}$. The mixing induced in turbulent flow dissipates the force from the head pressure, requiring that more pressure be exerted to sustain the same flow of water. The relationship between pressure and mean flow can be determined empirically, and is given by

$$\Delta P = Q^2 \frac{L\rho 0.3164}{4r\sqrt[4]{R_e}}$$

where ρ is the density of the liquid. This is known as the Blasius formula [46] and is valid for liquid flow with Reynolds numbers less than 100,000. The pressure and flow relationship given by the Blasius formula is plotted in Fig. 6 (blue). Also shown is the typical operating range for flow in a household and the best fit linearization (black) of the Blasius formula in the most common operating range (top left inset). If we assume laminar flow, the slope of the black line is proportional to the estimated R_f from Poiseuille's law. The actual R_f is proportional to the instantaneous slope of the blue line. The worst case approximation error of R_f is 70% at a flow rate of 10 gpm. However, this only results in an error in ΔP of 0.11 psi. Because the pressure drop, ΔP, at the HydroSense unit is the sum of the pressure drop at the valve (on the order of 5–20 psi) and the pressure drop along the pipes (in this example, below 0.11 psi), the resulting error in Q is between 0.5%–2.2%. Of course, as the pipe length increases, this error increases. Even still, with a pipe length of 50 ft, the error is still less than 10%. Thus, the linear model *can* be applied with minimal loss of accuracy even when the water flow through the household is turbulent.

3.1.3. Characterizations

We note that these equations are not comprehensive. They do not account for the smoothness of the inner pipe surface, the number of bends, valves, or constrictions in pipes, nor pipe orientation (e.g., the forces of gravity and changes in barometric pressure). While many of these factors can be estimated using home size, type of plumbing (PVC, copper, etc.), and number of fixtures, we have found these effects can be treated as negligible for home pipe networks. We simply estimate R_f for each home by sampling flow rate at strategic locations (varying distances from the supply inlet).

4. Prototype sensor design

Our prototype HydroSense sensor implementation consists of a customized stainless steel pressure sensor, an analog-to-digital converter (ADC) and microcontroller, and a Bluetooth wireless radio (see Fig. 7). We built two different HydroSense prototypes: one with a pressure range of 0–50 psi and the other 0–100 psi. The higher dynamic range is useful for homes with high supply pressure or without a pressure regulator. The pressure sensor is a P1600 series manufactured by Pace Scientific™. It comes standard with a built-in 1/4″ NPT male connector, which we fitted with a 3/4″ brass adaptor and Teflon tape. This allows us to easily install our sensor at any ordinary water spigot or outlet. The sensor has an operating temperature of −40 to 257 °F and a pressure response time of less than 0.5 ms. The theoretical maximum sampling rate is therefore 2 kHz, but we found 1 kHz more than sufficient.

The pressure sensor's output is ratiometric to the 5 VDC supply voltage (the output voltage is a ratio of the supply). The sensor is connected to a 16-bit Texas Instruments ADS8344 ADC and AVR microcontroller, with a resolution of approximately 0.001 psi for the 50 psi sensor and 0.002 psi for the 100 psi sensor. The microcontroller is connected to a Class 1 Bluetooth radio implementing the serial port profile. It can reliably sample and stream pressure data over the Bluetooth channel. We use a 5 V low-drop power regulator and the entire unit operates on a single 9 V battery.

The pressure sensor has a mechanical shock rating of over 100 g, making it insensitive to pipe vibration occasionally caused by some water hammer events. Although the pressure sensor comes calibrated and tested for linearity from the factory, we confirmed the output of our entire sensor system using known pressure loads. Ten samples were taken with our sensor connected to a pressure-regulated water compressor. All measurements were well within the pressure sensor's tolerance of 0.25% at 25 °C. The entire unit is weatherproof and can be installed in damp locations. Our current

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

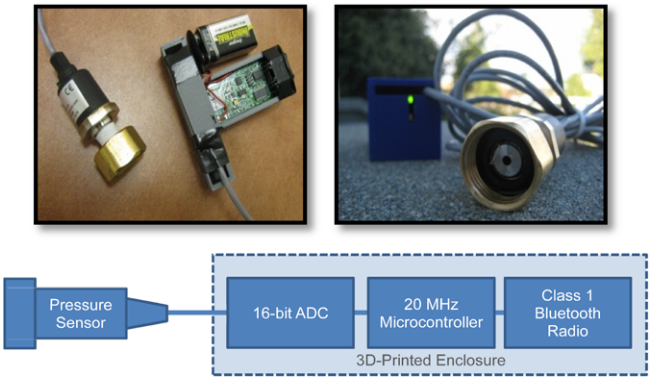

Fig. 7. Our prototype sensor implementation. The sensor twists on to a fixture and communicates wirelessly.

ID / Water Supply	Style / Built / Remodel	Size / Floors / Fixtures	Exp. Tank/ Regulator / Recirc. Pump	Water Heater / Plumbing/ Static PSI	Sensor Install Point
H1 Public Utility	Single-Family 2002	3200 sqft 2 flr + bas 12 fixture	Yes Yes No	Tank PVC 46 psi	Hose Bib
H2 Public Utility	Multi-Family 1909/96	2160 sqft 2 flr + bas 5 fixtures	No No No	Tankless Copper 46 psi	Hose Bib
H3 Public Utility	Single-Family 2003	4000 sqft 2 flr + bas 6 fixtures	Yes Yes No	Tank Copper 41 psi	Hose Bib
H4 Public Utility	Single-Family 1921	1630 sqft 1 flr + bas 4 fixtures	No No No	Tank Galvan. 43 psi	Hose Bib
H5 Public Utility	Single-Family 1913	2000 sqft 2 flr + bas 5 fixtures	No No No	Tank Copper 55 psi	Hose Bib
H6 Public Utility	Single-Family 1974/85	3100 sqft 2 flr 8 fixtures	Yes Yes Yes	Tank Galvan. 46 psi	Hose Bib
H7 Public Utility	Apartment 1927	746 sqft 1 flr 5 fixtures	No Yes No	Tank Copper+Galvan 33 psi	Water Heater
H8 Public Utility	Single-Family 1922 / 2006	3650 sqft 2 flr + bas 3 fixtures	Yes Yes Yes	Tank Copper 75 psi	Utility Sink Faucet
H9 Public Utility	Single-Family 1904 / 95 est.	1790 sqft 2 flr + bas 4 fixtures	No No No	Tank Copper 72 psi	Hose Bib + Water Heater
H10 Private Well	Resort Cabin 1950/80	900 sqft 1 flr 4 fixtures	No No No	Tank Galvan. 65 psi	Hose Bib

Fig. 8. A summary of the homes in which we collected data, including the style, size (1 sqft ≈ .093 sqm), age of the home, how many fixtures we tested, characteristics of the plumbing system, and where we installed our sensor.

implementation does not offer a pass-through solution (i.e., allowing the installation fixture to be used as normal), but this modification is trivial.

5. In-home data collection

In order to validate our general approach, our sensor implementation, and our algorithms, we collected labeled data in four cities using ten homes of varying, style, age, and diversity of plumbing systems (see Fig. 8). Homes H1–H9 used metered water from the public utility, whereas home H10 used a private well.

For each home, we first measured the baseline static water pressure and then installed the appropriate HydroSense unit (0–50 or 0–100 psi) on an available water hose bib, utility sink faucet, or water heater drain valve. Each collection session was

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

conducted by a pair of researchers: one would record the sensed pressure signatures to a laptop while the other activated the home's water valves. The pressure signatures were recorded using a graphical logging tool, which also provided real-time feedback of the pressure data via a scrolling time-series line graph. We conducted five trials per valve on each fixture (e.g., five trials for hot water and five trials for cold water). For each trial, a valve was opened completely for at least five seconds and then closed. For the toilet trials, the toilet flush and full fill cycle were logged. Note that for the faucet experiments, we did not collect data on partially opened valves nor the speed with which they were opened. We return to this issue in the discussion section.

In four of the ten houses (H1, H4, H5, and H7), we also collected flow rate information for the faucet (kitchen and bathroom) and shower fixtures. In addition to logging sensed pressure, we measured the amount of time it took to fill a calibrated bucket to one gallon (a method preferred by water utilities for accurately measuring flow). This was repeated for five trials for each valve. This in-home data collection yielded a total of 775 fixture trials and 155 flow rate trials across 76 valves and 51 fixtures.

6. Analysis of valve- and fixture-level event identification

Given our collected data, we now pursue a three-step approach to examine the feasibility of identifying individual fixture and individual valve events according to the unique transient pressure waves that propagate to our sensor. Recall that each valve event corresponds to a pressure wave when a valve is either opened or closed. We first segment each individual valve event from the stream, identifying its beginning and end to enable further analysis. We then classify each valve event as either a valve open or a valve close event. Finally, we classify the valve event at two levels of granularity: (1) according to the individual fixture that generated it (e.g., bathroom sink versus kitchen sink) and (2) according to the individual valve that generated it (e.g., kitchen sink cold water versus kitchen sink hot water). We compare two different classification protocols: the first based on matched filtering, and the second based on a Hidden Markov Model (HMM). For classification, we explicitly consider only events that occur in isolation, deferring our discussion and analysis of compound events until a later section.

6.1. Valve event segmentation

Before analyzing the characteristics of a valve event, we first segment it (i.e., isolate it) from the surrounding sensor stream. Segmentation must be effective for many different types of events, and so it is important to consider only features that are likely to be most typical of all valve events. Our approach is illustrated in Fig. 9 (and in Fig. 5). The raw signal is smoothed using two low-pass linear phase finite impulse response filters with different cutoff frequencies (1 and 13 Hz, black and blue lines in Fig. 9, respectively). Almost all the spectral energy is below 10 Hz. Thus the 13 Hz filter captures a de-noised version of the transient. Below 1 Hz the relative *decrease* or *increase* in pressure is more apparent which aids in classifying the transient as a valve open or valve close event. The derivative of the filter is also calculated using a windowed ideal differentiator filter with a cutoff of 2 Hz (i.e., the optimal differentiator in terms of squared error). The smoothed signal and its derivative are then analyzed in a sliding window of 1000 samples (one second of sensed pressure).

The beginning of a valve event corresponds to one of two conditions. The most common is when the derivative of the smoothed signal exceeds a specified threshold relative to static pressure, indicating a rapid change. For H1 (static pressure of 45 psi), this value was found empirically to be approximately 2 psi/s. For generalizing to other homes, this value was scaled by the home's actual static pressure. The less common second condition is when the difference between the maximum and minimum pressure values in the sliding window exceeds a threshold relative to static pressure, indicating a slow but substantial change (approximately 1 psi for a home with 45 psi static pressure, scaled by the actual static pressure).

Each transient represents a damped sinusoid. After the beginning of a valve event is detected, the next change in the sign of the derivative represents the time at which the extreme value of the pressure transient occurs (which may be a maximum or a minimum). This extreme value is used to assist in the calculation of the end of the transient. The end of a segmented valve event is typically detected as the first point at which an extreme of a fluctuation (a change in the sign of the derivative, dP/dt) is less than 5% of the magnitude of the first extreme following the beginning of the event. It is also possible for an event to be ended by a rapid increase in the magnitude of fluctuation. This corresponds to the occurrence of a compound event, as we will discuss in greater detail in a later section.

Applying this method to our collected in-home data yielded appropriate segmentations of 100% of our valve events from their surrounding sensor stream. Additionally, this method was effective at finding leaky flapper valves in toilets. After our initial data collection, we noticed that this method segmented multiple valve closings when a particular toilet in H2 was flushed. Further inspection revealed that the flapper valve of the toilet was loose, resulting in two concurrent valve closings.

6.2. Classifying valve open and valve close events

After segmenting each valve event, we classify it as either a *valve open* or a *valve close* event. We apply a classifier that first considers the difference in the 1 Hz smoothed pressure at the beginning and the end of the segmented event. If the magnitude of this difference exceeds a threshold (approximately 2 psi for a home with 45 psi static pressure, scaled by the actual static pressure), the event can be immediately classified (a pressure decrease corresponds to a *valve open* and a

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

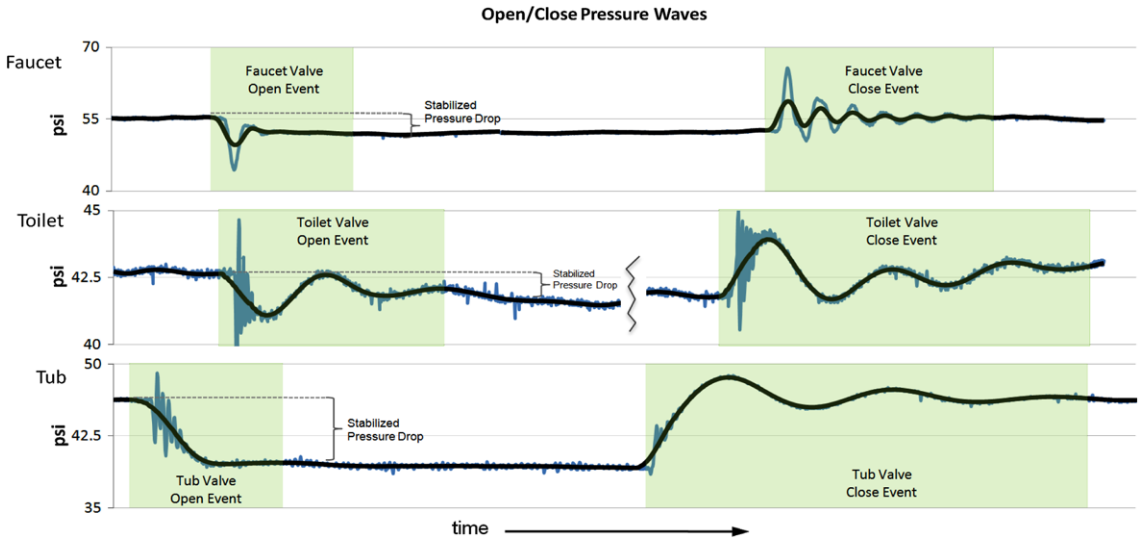

Fig. 9. Several actual sensor pressure streams from our in-home data collections. Each stream corresponds to a water valve being opened, remaining open for some amount of time, and then closed. Both raw pressure (shaded lines) and filtered pressure (black) are shown. We separately segment the valve open and valve close events from the sensor stream, as indicated by the highlighted regions of the streams. We estimate water flow to the valve based on the stabilized pressure drop while the valve remains open (the difference in pressure before the valve opens versus after it). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

pressure increase to a *valve close*). Otherwise, the event is classified according to the average value of the derivative between its beginning and its first extreme. A *valve open* creates an initial pressure decrease (a negative average derivative), while *valve close* events create an initial pressure increase (a positive average derivative). Figs. 5 and 9 illustrate this observation.

Applying this method to the segmented valve events from our collected in-home data yields 100% correct classification of valve open and valve close events. We note that this method and our event segmentation method described above require only knowledge of the static pressure in the home, which can be easily and automatically derived. HydroSense can thus be used immediately and without supervision to detect fixture openings and closings in a home. Fixture- and valve-level classification, however, require a supervised calibration process.

6.3. Fixture-level and valve-level classification

The final step in our three-step approach is to identify the valve and fixture source of the segmented open and close events. We compare two different approaches to this problem: template matching and an HMM. Before going into our analysis in detail, it is worth discussing the ways in which water is used in the home that can impact classification. First, note that valves can be divided into two groups: automatically (or electro-mechanically) controlled valves such as those found in dishwashers, laundry machines, and ice makers, and manually controlled (or human operated) valves such as those found in kitchen sink faucets, bathroom sink faucets, showers, and tubs. Toilets are a hybrid between manually and mechanically controlled valves. Manually operated valves can produce slightly different transients depending on the speed and flow rate at which they are operated, while mechanical valves produce a more distinct transient at a single flow rate. Our database of events contains both mechanically and manually operated fixtures, but our data collection did not explicitly control for different usages of hand operated valves. We address this limitation further in the discussion section.

Our first classification approach, template matching, assumes that the transient (i.e., the template) for a specific fixture or valve stays relatively constant between different trials. Our second classification approach, which uses an HMM, assumes that there is some inherent variation in the transient (a more stochastic model). The argument for using one classifier over the other is not perfectly clear. While manually operated valves can produce different transients (an argument for using an HMM), the variation in shape is not necessarily drastic (an argument for using templates). In the same way, mechanically operated valves can have highly stable transients (argument for using templates), but changes in static pressure can cause slight variations in the magnitude and duration of the transient (argument for using an HMM). Note that for both approaches, we rely solely on the pressure signal for input to our classifiers. Although incorporating contextual factors such as time of day and fixture usage duration would likely boost classification accuracies, these factors are irrelevant in our dataset because our data was collected under controlled experimental conditions.

6.3.1. Template-based classification

In this classifier, we associate *valve open* and *valve close* events with individual fixtures in a home using a template-based classifier. When classifying an unknown event, we first filter potential templates according to four complementary distance metrics.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

The first distance metric we use is a *matched filter* [47]. Very common in signal detection theory, the matched filter is the optimal detection mechanism in the presence of additive white noise. Its primary limitation is that the signals we want to differentiate are not orthogonal. It is possible to make the signals orthogonal (for example using the Karhunen Lueve expansion), but we found this type of analysis to be unnecessary. Instead, we transform the data into more orthogonal spaces.

Our second distance metric is a *matched derivative filter*. We include this because the derivatives of our events always resemble exponentially decreasing sinusoids. It is therefore reasonable to believe the derivatives are more orthogonal than the original pressure signals, and that this filter might provide value distinct from the above filter.

The third distance metric is based on the *real Cepstrum*, which is the inverse Fourier transform of the natural log of the magnitude of an event's Fourier transform. This approach attempts to approximate the original version of a signal that has been run through an unknown filter (the valve event we are trying to classify has been transformed by propagation through an unknown path in a home's water pipes). It can be shown that, for linear systems, the lower coefficients of the Cepstrum result largely from the transfer function (an event's propagation through a home's pipes) and the higher coefficients largely from the source (the original impulse at the valve) [48]. We are interested primarily in the transfer function (in part because it allows differentiating among multiple instances of identical fixtures in a home), so we truncate our Cepstrum to the lower coefficients. The resulting space is highly orthogonalized (a common property of the Cepstrum), yielding a third effective and complementary matched filter.

Finally, our fourth distance metric is the simple *mean squared error*, computed by truncating the longer of two events.

Similarity thresholds used to filter potential templates based on these distance metrics are learned from training data (filtering templates whose similarity to the unknown event are less than the minimum within-class similarity in the training data). If no template passes all four filters, the unknown event is not classified (an application might ignore the event, prompt a person to label an unrecognized fixture, or consider the possibility that the new event indicates the presence of a leak). If templates corresponding to multiple fixtures pass all filters, we choose among them using a nearest-neighbor classifier defined by the best performing distance metric, the matched derivative filter.

6.3.2. HMM-based classification

Our HMM-based classification consists of three processing steps: (1) feature set calculation, (2) model training, and (3) maximum likelihood classification.

For our feature set we chose to use a one second sliding window of Constant-Q weighted Cepstral Coefficients (CQCC). For each window we transform the time series using a frequency-based constant-Q filter bank transform [49,50]. Each filter in the transform has the exact same quality factor, Q = center frequency/bandwidth. This has the effect of logarithmically increasing the bandwidth of the filters as the center frequency increases. There are 16 filters in the filter bank spanning the frequencies of 0.1 to 20 Hz. The constant-Q filter bank transform results in 16 coefficients that are proportional to the energy in each filter. Denoting the vector of constant-Q transform coefficients of the t^{th} window as $\mathbf{q_t}$ and the vector of CQCC's as $\mathbf{s_t}$, the relationship is defined as:

$$s_t = \mathrm{DCT}_2(\ln|q_t|)$$

where DCT2 denotes the type-II discrete cosine transform. We then advance the sliding window by 100 ms and repeat the calculation. Thus our feature set is given by concatenating the CQCC's for each window, $\mathbf{S} = \mathbf{s_1 s_2} \ldots \mathbf{s_n}$. In the HMM, $\mathbf{S}$ is an observation matrix and each column of $\mathbf{S}$ contains the CQCC's calculated for a different sliding window.

These features have the advantages of reducing dimensionality (there are only 16 coefficients from the filter bank) and providing a space where it is appropriate to use Gaussian mixtures (i.e., the space has approximately diagonal covariance, a common property of the DCT). Another compelling argument for using the constant-Q filter stems from the way that resonances propagate in pipes. At higher frequencies, the bandwidth of the resonance becomes wider meaning that we should sample the energies in the frequency domain non-linearly with increasing bandwidth at higher frequencies. The constant-Q filter bank captures this aspect exactly. Furthermore, another argument for the use of Cepstra is that the electrical analogy of the plumbing system is completely linear, indicating that transformations assuming linearity may also model the non-linear hydrodynamics well. In sufficiently linear system, Cepstral coefficients tend to separate the excitation source from the system transfer function, similar to the real Cepstrum used in the template-based matching approach. The transfer function from the valve in use to our sensor provides great discriminative power, which is captured efficiently using CQCCs. Once the CQCC features are calculated, we train a set of HMMs for each valve in the home.

The goal of training is to learn a set of models, $\mathbf{V}$, for each valve from a set of observations, $\mathbf{S}$, taken during multiple trials of opening and closing the valve. If there are N valves in a home, we can train $2N$ HMMs, one for each opening and each closing of the valves in a home, $\mathrm{V} = \{V^1_{open}, V^1_{close}, \ldots, V^N_{open}, V^N_{close}\}$. We chose to use a two state HMM with four diagonal covariance Gaussian mixtures per state. This was chosen empirically based on our sparse training set. Any HMM with more than four mixtures per state resulted in decreased accuracy.

It is interesting to look at why *two* states provide an adequate modeling of each transient. In a HMM, each state should model a different *stage* (commonly called a *mode*) of the observation sequence, $\mathbf{S}$. Each mode should carry separate, distinct information about $\mathbf{S}$. One way to view the modes of a valve transient is by looking at the *resonances* of the valve and how they decay over time. Fig. 9 provides an intuition about why two states might be an appropriate model for the number of modes. Most of the transients can be divided into two sections: (1) the initial high frequency excitation which dies out quickly (in

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

Fixture-Level Classification					
		Template Based		Hidden Markov Model	
Home	Installation Point	Fixture Open Identification	Fixture Close Identification	Fixture Open Identification	Fixture Close Identification
H2 (5 fixtures)	Hose Bib	96.4%	100.0%	100.0%	100.0%
H4 (3 fixtures)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H5 (4 fixtures)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H6 (5 fixtures)	Hose Bib	100.0%	97.5%	100.0%	90.0%
H7 (5 fixtures)	Hose Bib	100.0%	100.0%	97.1%	100.0%
H8 (3 fixtures)	Utility Sink	100.0%	97.1%	100.0%	94.1%
H9A (4 fixtures)	Hose bib	97.1%	97.1%	100.0%	94.3%
H9B (4 fixtures)	Water Heater	88.6%	74.3%	82.9%	74.3%
H10 (4 fixtures)	Hose Bib	94.6%	75.0%	86.5%	75.0%
Aggregate		98.0%	94.7%	97.0%	93.4%
		96.4%		95.2%	

Fig. 10. In a cross-validation test of the robustness of learned models across multiple homes, both classification methods enable identification of the individual fixtures associated with valve open and valve close events with aggregate accuracy > 90%.

blue) and (2) the remaining low frequency excitation (where blue and black are almost perfectly overlaid). In our database, we empirically see this phenomenon in many transients. However, some transients (like at the top of Fig. 9) only display one prominent resonance (i.e., a single mode). With two states we can sufficiently model a transient with as many as two modes.

For training, each state is initialized by dividing the transient into two equal time segments, and then, for each segment, Gaussian mixtures are initialized using *k*-means clustering. Each HMM is trained according to expectation maximization [51] provided by the Matlab HMM toolbox [52].

To *classify* unknown transients, we calculate the observation matrix, **S**, of the unknown transient and then compute the log-likelihood of the observation given a set of HMMs in each home. The HMM with the largest log-likelihood, $\widehat{V}$, is classified as producing the observations. For a home with a set of HMMs, **V**, this is given by:

$$\widehat{V} = \arg\max_{V \in V} P(V|\mathbf{S}).$$

6.4. Fixture classification evaluation

We evaluate *valve-level* classification and *fixture-level* classification using an experimental design selected to demonstrate robustness of learned model parameters across the multiple homes in our collected data. Specifically, we conduct a cross-validation experiment that folds our data according to the home in which it was collected.

For our template-based classifier, there are ten trials in the cross-validation, with each trial using data from one home as the test data and data from the other nine homes as the training data. After learning model parameters from the test data (the four similarity filter thresholds), we classify each event in the test home using a leave-one-out method. Each test home event is classified using the other events as templates together with the model parameters learned in training.

For our HMM-based classifier, we perform a similar leave-one-out cross-validation within each home. Our database consists of five open and five close transients for each valve in the home. We train five HMMs per valve; each trained in the absence of one trial of the valve. Then, for each home and each transient in a home's database, we remove the four HMM models that use the test transient as training data, and then compute the log-likelihood of each of the remaining HMMs in the home (one HMM for the correct valve, five HMMs per incorrect valve). The test HMM with the highest log-likelihood is chosen as the best match.

Figs. 10 and 11 present the results of this evaluation for both classifiers. In particular, Fig. 11 shows the identification percentages for classifying each *valve* in a home, whereas Fig. 10 shows identification percentages for classifying each *fixture* in a home. The main advantage of *valve-level* identification is that hot and cold water usage can be disambiguated. A kitchen faucet, for example, is a single fixture that has two valves, one for hot water and one for cold. Each figure shows the accuracy of classification for *open* and *close* events within each home (and thus each test fold of the cross-validation), as well as the aggregate accuracy of classification. Accuracies for H9 are shown for two different installation points of the sensor, hose bib (H9A) and hot water heater (H9B). H1 and H3 are excluded from Fig. 11 because they contain the same number of fixtures as valves. In these homes it was impossible to isolate the valve on each fixture because the style of sinks and showers was such that both hot and cold were activated simultaneously.

Analysis of valve and fixture classification reveals that both classifiers perform similarly on H1–H8 and on H9A (>90% for fixtures, >88% for valves). However, both methods perform poorly on H10 and on H9B. The relatively poor performance in identifying valve-level events in H10 was due to noise from the eleven cabins that share the same supply line at the resort.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

Valve-Level Classification					
		Template Based		Hidden Markov Model	
Home	Installation Point	Valve Open Identification	Valve Close Identification	Valve Open Identification	Valve Close Identification
H1 (12 valves)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H2 (8 valves)	Hose Bib	96.4%	100.0%	100.0%	100.0%
H3 (6 valves)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H4 (5 valves)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H5 (9 valves)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H6 (8 valves)	Hose Bib	100.0%	97.5%	100.0%	90.0%
H7 (8 valves)	Hose Bib	100.0%	100.0%	97.1%	100.0%
H8 (6 valves)	Utility Sink	100.0%	97.1%	100.0%	94.1%
H9A (7 valves)	Hose bib	97.1%	97.1%	97.1%	88.6%
H9B (7 valves)	Water Heater	88.6%	71.4%	71.4%	62.9%
H10 (7 valves)	Hose Bib	75.7%	43.8%	70.3%	56.3%
Aggregate		96.2%	91.8%	94.2%	90.3%
		94.1%		92.3%	

Fig. 11. In a cross-validation across multiple homes, both classification methods enable identification of the individual valves associated with valve open and valve close events with aggregate accuracy $> 90\%$.

Because the cabin was not separately metered our sensor was picking up water events from a portion of these cabins during data collection. When looking at *fixture-level* performance for H10, the classification accuracies increase markedly. This is an indication that noise in a multi-unit domain (e.g., a duplex, small apartment building) may have more effect on *valve-level* identification than *fixture-level* identification. We note that the valve-level results presented here for the template-based classifier differ from Froehlich et al. [53], which presented similar results on the same database [53]. We report 75.7% and 43.8% accuracy whereas the previous reporting was 97.1 and 77.1 for H10. The reason for this is twofold: firstly, the learned model parameters have changed in the cross-validation because a new home (H9B) was added. Secondly, Froehlich et al. reported results for H10 at the fixture level, rather than valve level, a mistake which is remedied here. More work is needed to disambiguate signals in a single-meter multi-unit domain, but these results indicate a single sensor may be sufficient to sense *fixture-level* detail across more than one housing unit on a shared supply line, a somewhat different conclusion than that presented by Froehlich et al.

The poor performance on H9B is largely due to the installation point (on the water heater). Although events are transmitted *through* the water heater, many events are significantly dampened. This dampening suppresses some of the defining characteristics of the transients, making them more similar to one another. It appears as though the matched filter is slightly more robust to this dampening, but we cannot say with high confidence that the matched filter is statistically better than the HMM.

To investigate performance further, it is interesting to look at how different valves are confused in each home. Fig. 12 presents a different view on the same data, showing the accuracy of *valve-level* classification for different types of fixtures across homes. Also shown is the number of times a valve is confused according to three categories: (1) the valve is confused for a fixture of the same type, (2) the valve is confused for the hot or cold valve from the same fixture, or (3) the valve is confused for a different type of fixture altogether. In both classifiers, notice that similar appliances are rarely confused for one another (e.g., two different sinks or two different toilets). Instead, the source of confusion comes from the *proximity* of two valves in a home. For instance, the main source of confusion in the figure comes from a sink and shower in H9B that are consistently confused with each other. The valves are positioned extremely close to each other and within 10 ft. of the water heater. The valves already have similar characteristics because of their proximity, and the water heater acts like a large low-pass filter, dampening out most of the characteristics that make the pressure waves identifiable. Other sources of confusion in the figure are largely due to classification errors in H10. However, there is no pattern to valve confusion in H10, supporting the notion of noise from other cabins being the main culprit.

Fig. 13 shows the accuracy for classification of hot water valves and cold water valves separately. Again, H1 and H3 are excluded because fixtures always contain a mixture of hot and cold water. Again, the template classifier and HMM perform similarly. No real pattern presents itself from the data. Both hot and cold valves can be classified with similar accuracy. Despite these similarities, we remind the reader that our experimental data collection did not control for the speed at which manually controlled valves were opened, or when multiple valves are running in the household. More examination is needed to investigate whether the HMM could be important for classifying in the presence of these factors.

Our overall *valve-level* and *fixture-level* classification across all homes is above 90%, including a number of cases where classification accuracy is 100%. All of these results are equal to or better than prior results by Fogarty et al. with microphone-based sensors [1]. Of particular note is our ability to reliably distinguish among different sinks within a home, as Fogarty et al. found that their microphone-based sensors did not capture enough information to reliably make this distinction. Our dataset

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

Valve Level Confusability Matrix										
	Template Based					Hidden Markov Model				
Fixture Type	Accuracy	Confusion				Accuracy	Confusion			
Faucet	95.3%	424	6	6	9	93.0%	414	8	8	15
Toilet	99.3%	136	0	0	1	97.1%	133	0	0	4
Clothes Washer	100.0%	15	0	0	0	100.0%	15	0	0	0
Dishwasher	100.0%	6	0	0	0	100.0%	6	0	0	0
Tub	100.0%	36	0	0	0	100.0%	36	0	0	0
Shower	82.4%	112	0	12	12	85.3%	116	0	12	8
Confused for:		Correct	Similar Appliance	Hot/Cold	Diff. Appliance		Correct	Similar Appliance	Hot/Cold	Diff. Appliance

Fig. 12. A different view of the results, showing accuracy of identification of individual valves by fixture type, and confusability of different fixtures.

Hot/Cold Valve-Level Classification					
		Template Based		Hidden Markov Model	
Home	Installation Point	Valve Cold Identification	Valve Hot Identification	Valve Cold Identification	Valve Hot Identification
H2 (8 valves)	Hose Bib	97.3%	100.0%	100.0%	100.0%
H4 (5 valves)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H5 (9 valves)	Hose Bib	100.0%	100.0%	100.0%	100.0%
H6 (8 valves)	Hose Bib	98.3%	100.0%	96.7%	90.0%
H7 (8 valves)	Hose Bib	100.0%	100.0%	96.9%	100.0%
H8 (6 valves)	Utility Sink	97.6%	100.0%	95.2%	100.0%
H9A (7 valves)	Hose bib	100.0%	93.3%	95.0%	90.0%
H9B (7 valves)	Water Heater	75.0%	86.7%	77.5%	53.3%
H10 (7 valves)	Hose Bib	62.8%	57.7%	81.4%	34.6%
Aggregate		94.6%	92.9%	95.5%	85.0%

Fig. 13. A different view of valve-level classification separated for hot and cold valves in a home.

contains only a few instances of clothes washer or dishwater use, in part due to time constraints during data collection and in part because Fogarty et al. found these fixtures can be easily recognized by their structured cycles of water usage (an approach that can be combined with ours). However, we note that our approach is independent of the number of fill cycles (important if a dishwasher is sometimes run with a pre-rinse cycle) and allows recognition as soon as these appliances first use water (in contrast to being able to recognize them only after their pattern of fill cycles becomes apparent).

6.5. Analysis of flow estimation

As previously discussed, the volumetric flow rate Q is proportional to the change in pressure ΔP divided by a resistance variable R_f ($Q = \Delta P/R_f$). We calculate the change in pressure ΔP by measuring the difference between the pressure at the onset of a detected *valve open* event and the stabilized pressure at the end of the segmented *valve open* pressure wave impulse. The resistance variable R_f cannot be directly measured, so we instead learn it empirically by capturing ground truth flow rate information together with the corresponding change in pressure for each valve. This section considers two scenarios with regard to learning R_f. In the first, we assume a single calibration of flow for every valve of interest. In the second, we attempt to use information from the calibration of some valves to estimate R_f at valves that have not been calibrated. In both scenarios, we assume laminar fluid flow.

6.5.1. Individually calibrated valves

It is not unreasonable to imagine that the process of installing a system like HydroSense might include a single calibration of each fixture in a home. In such a scenario, each valve in the home would be labeled with a known R_f value which could be combined with the sensed pressure change ΔP to estimate water flow at those valves.

We examined the accuracy of the flow estimation that might be obtained in this scenario using a cross-validation experiment to analyze the five calibrated bucket trials collected for each of the faucet and shower fixtures in H1, H4, H5, and H7 (as previously discussed in our in-home data collection section). Each trial in the cross-validation used a single calibrated bucket trial to infer a resistance variable R_f for the valve. The inferred value of R_f was then used to estimate flow in the other four trials according to the measured change in pressure ΔP. We then noted the difference between these estimated flow

Home	Avg Error (GPM)	Stdev Error (GPM)	Avg Error (%)	Stdev Error (%)
H1 (7 valves)	0.17	0.13	7.3	6.7
H4 (6 valves)	0.19	0.17	5.6	5.3
H5 (8 valves)	0.13	0.11	4.5	5.5
H7 (8 valves)	0.67	1.47	22.2	46.0

Fig. 14. In homes H1, H4, and H5, our system is able to estimate flow at individual fixtures throughout the home with error rates comparable to that found in empirical studies of traditional utility-supplied water meters. In H7, placing the sensor on a hot water heater appears to result in a confounding of supply water main pressure with gravitational pressure due to the water in the tank.

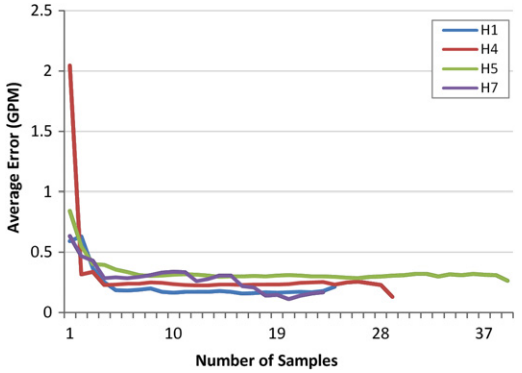

Fig. 15. Because valves share a fair amount of spatial overlap in the length and overall layout of their piping, it is possible to learn to generalize calibrations across valves.

rates (based on the inferred R_f) and their corresponding actual flow rates (obtained through the calibrated bucket trials). The results of this experiment are shown in Fig. 14.

Three of four houses tested (H1, H4, H5) have error rates below 8% (or approximately 0.16 gpm**), comparable to 10% error rates found in empirical studies of traditional utility-supplied water meters [54]. The fourth house (H7), however, had an error rate above 20%. We believe this is due to the installation location of the sensor. Whereas the first three homes had HydroSense installed on an exterior water bib, H7's installation used the hot water heater drain valve. This results in two confounding pressure sources (the supply water main and the gravitational pressure of the water in the tank). As previously discussed, our simple pressure model currently assumes a straight pipe. It is likely this situation requires a different model of R_f, and it seems that cold water valves in H7 were particularly affected. Indeed, removing H7's four cold water valves from our analysis results in an dramatically improved average error of 0.15 gpm (SD = 0.18), or 4.5% (SD = 3.8%). Because our dataset includes only one home with both hot water heater installation and flow rate information, future work is needed to investigate the feasibility of measuring cold water flow using a sensor installed at the hot water heater drain.

6.5.2. Generalizing to uncalibrated valves

In a scenario where only some of the valves in a home have been calibrated, it is reasonable to attempt to build a model of fluid resistance for the entire home from that subset of valves. The key idea here is that, although the pathway to each valve in the home is unique, those paths also share a fair amount of spatial overlap in the length and overall layout of the piping. For example, the toilet and sink in a particular bathroom share the same branch.

To examine this approach, we separated our calibrated bucket trials data into two datasets: a model and a test. The model was initially populated by a single randomly selected trial which was then used to infer a baseline R_f value. We applied this R_f value to calculate a flow estimate for each trial in the test dataset, comparing each to the corresponding actual flow. We next added a second random trial to the model (and removed it from the test dataset), then used the model to create a linear regression ($Q = R_f * \Delta P + b$). This regression was used to calculate flow estimates for the remaining trials in the test set. This process was repeated until all trials had been sampled. To avoid a particularly fortunate or unfortunate random sampling, we repeated this process five times for each home and averaged the results. Fig. 15 presents the results (note we exclude the cold water valves from H7, consistent with our prior analysis).

After sampling five trials, the average error dropped 74% to 0.27 gpm across the four homes and within 0.11 gpm of the more comprehensive R_f data from the previous analysis. This initial result indicates considerable potential for learning to generalize calibrations across valves in a home.

7. Discussion

The initial results presented in this paper show significant promise for single-point sensing of whole-home water activity via continuous monitoring of water pressure. We have presented a reliable method for segmenting valve pressure events

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

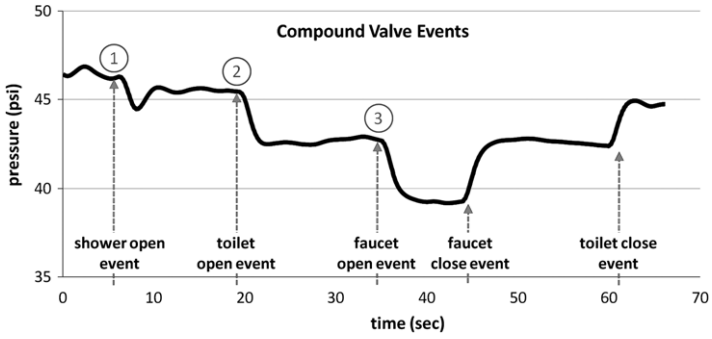

Fig. 16. Our existing event segmentation correctly identifies and segments overlapping events (shown here for overlapping shower, toilet, and faucet valve open events), and classification of such compound events is an important direction for future work. Prior work was unable to even identify the occurrence of compound events.

from their surrounding sensor stream and for determining whether a segmented event corresponds to a valve being opened or closed. Using data collected in ten homes, we have shown >90% aggregate accuracy for identifying the specific fixture and valve associated with a transient using two very different classification methods, template-based and HMM-based. Analyzing flow data collected in four of those homes, we have shown that an appropriately located and calibrated system can estimate water usage with error rates comparable to empirical studies of traditional utility-supplied water meters. Our ability to identify activity at individual fixtures using a single sensor is itself an important advance, and we are not aware of prior work even attempting single sensor estimation of the amount of water being used at individual valves throughout a home. A particular contribution of HydroSense over other methods (like flow-trace analysis) is the ability to disaggregate hot and cold water events in a home. Previously, this type of disaggregation required the use of two water meters, one connected to the water supply line into the home and another connected from the supply line of the water heater.

Although our analysis focused on identifying valve events occurring in isolation, it is clearly important to consider the case where multiple events overlap, especially for multi-family homes and apartment units with shared plumbing systems (see Fig. 16). In Fogarty et al.'s prior work with microphone-based sensing, they note an inability to even detect this situation [1]. As an initial investigation, we collected six compound events in H1 (two each of shower/sink, toilet/sink, and shower/toilet/sink overlaps). Our event segmentation algorithm correctly segments these overlapping events (ending the ongoing event when it detects a rapid increase in the magnitude of fluctuation corresponding to the beginning of another event). Preliminary experiments suggest that the magnitude and shape of the events is indeed altered by the overlap. Some aspects of the frequency domain signature remain constant (i.e., high energy harmonics) but we do not have enough compound event data to convincingly evaluate the effectiveness of a classification procedure. Furthermore, events that occur at the *exact* same instant cannot be distinguished as separate events with our current segmentation algorithm. In any case, classification of compound events is a highly important direction for building upon our current results. Bathroom activity, for example, is dominated by the usage of multiple fixtures.

An additional limitation relates to how closely our controlled experiments represent naturalistic usage of fixtures (e.g., how often do people partially open valves versus open them full stop, how is the signal affected by the speed of valve opening or closure). These concerns primarily involve fixtures with manually controlled valves (e.g., bathtubs/showers/faucets) rather than mechanically controlled valves (e.g., dishwashers/toilets/laundry machines). For mechanically controlled valves, the transient is determined based on fixture type (e.g., dishwasher versus laundry machine) and on its location in the home. For manually controlled valves, the way in which the valve is opened or closed is also a factor (e.g., the speed with which the valve is opened, how much the valve is opened). Our test dataset contains both mechanically and manually operated valves, but our data collection did not explicitly control for different usages of the manually operated valves. Thus, our classification analysis is a first step towards determining the viability of using water pressure waves to classify valves and fixtures in the home.

To address the above limitations, we have begun longitudinal deployments in multiple households collecting labeled, naturalistic water usage data to investigate these issues further. In addition, these studies should help to highlight advantages of different classification approaches. A more stochastic classifier, like the HMM presented here, may be more robust to classifying partial turn fixtures under naturalistic conditions. Furthermore, note that the contextual behaviors of water usage are important. Factors such as the time of day (e.g., showers are more likely in the mornings [31]), the duration of usage (e.g., toilet refills last between 30–60 s depending on the home's water pressure), and temporally cyclic behaviors (e.g., dishwashers and laundry machines use repeated cycles that make identification easier) provide a rich knowledge about water usage and the context of human activity surrounding water usage. To this end, we have begun work using a graphical modeling approach that allows for inclusion of such contextual variables.

We have found that reliable estimation of flow is sensitive to calibration, and we have noted that our segmentation and identification algorithms include threshold parameters that worked well in the homes we studied but are not necessarily ideal. We are interested in developing techniques for automatically calibrating our methods over the course of extended

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

usage. For example, flow estimation could potentially be automatically calibrated through occasional knowledge of whole-home aggregate water usage. Continuing deployments of wireless utility meters make this an increasingly viable approach. We also have initial evidence that system behavior is stable over time, based on a second dataset collected in H1 five weeks after our original collection. We applied our fixture classification methods to this dataset using templates from the opposite dataset (classifying unknown events using templates collected 5 weeks apart), finding no degradation in fixture identification performance. Another argument for stability comes from preliminary analysis of our longitudinal dataset (currently being collected). It shows that static water pressure and outside temperature largely do not affect generated water transients. This preliminary analysis suggests system behavior might be stable enough to apply a variety of machine learning methods for auto-calibration.

Our in-home data collection included installations at several different types of fixtures (hose bibs, utility sink faucets, water heater drain valves) with generally good results except in houses H9B and H10. There are many examples in Fig. 10 where our current approach differed in its ability to identify the fixture associated with valve open and close events. To help increase the accuracies in H9B and H10, a possible approach is to look at *both* open and close events together before classification. Because events obviously come in a series of opening and closing, it seems natural to pursue an approach that classifies *pairs* of open/close transients rather than individual transients. To investigate the implications of this pairing in H9B and H10, we looked at the number of times a valve open *and* valve close event is classified incorrectly in the same stream using the HMM classifier (an upper bound on how pairing might perform). Pairing before classification of valve-level events could increased accuracies in H9B to 83% (from 63%) and could increase accuracies in H10 to 81% (from 57%). Many other factors could be used to increase identification and flow rate accuracies. For example, we currently estimate flow independent of fixture identification, but the two are clearly related and an improved method could consider them simultaneously.

8. Conclusion

We have presented a new approach to single-point infrastructure-mediated sensing of whole-home water activity. Our initial results for two different classification methodologies both validate the effectiveness of our approach and provide a basis for future analyses and improvements. The infrastructure-mediated sensing strategy shows significant promise as a practical, low-cost, and unobtrusive approach to the broad deployment of sensing-based ubiquitous computing applications in activity inference and real-time resource monitoring. In particular, our approach validates how a single sensor can be used for disaggregated hot and cold water usage, enabling broader and more in-depth studies of the end use of water.

References

[1] J. Fogarty, C. Au, S.E. Hudson, Sensing from the basement: a feasibility study of unobtrusive and low-cost home activity recognition, in: Proceedings of the ACM Symposium on User Interface Software and Technology, UIST 2006, 2006, pp. 91–100.
[2] S.N. Patel, T. Robertson, J. Kientz, M.S. Reynolds, G. Abowd, At the flick of a switch: detecting and classifying unique electrical events on the residential power line, in: Proceedings of the International Conference on Ubiquitous Computing, 2007, pp. 271–288.
[3] S.N. Patel, E.P. Stuntebeck, T. Robertson, PL-Tags: detecting batteryless tags through the power lines in a building, in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2009, 2009, pp. 256–273.
[4] S.N. Patel, K.N. Truong, G.D. Abowd, PowerLine positioning: a practical sub-room-level indoor location system for domestic use, in: Proceedings of the International Conference on Ubiquitous Computing, UbiComp 2006, 2006, pp. 441–458.
[5] E.P. Stuntebeck, S.N. Patel, T. Robertson, M.S. Reynolds, G.D. Abowd, Wideband powerline positioning for indoor localization, in: Proceedings of the International Conference on Ubiquitous Computing, UbiComp 2008, 2008, pp. 94–103.
[6] S.N. Patel, M.S. Reynolds, G.D. Abowd, Detecting human movement by differential air pressure sensing in HVAC system ductwork: an exploration in infrastructure mediated sensing, in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2008, 2008, pp. 1–18.
[7] A. Vickers, Handbook of Water Use and Conservation: Homes, Landscapes, Industries, Businesses, Farms (Hardcover), 1st edition, WaterPlow Press, 2001.
[8] C. Fischer, Feedback on household electricity consumption: a tool for saving energy? Energy Efficiency 1 (2008) 79–104.
[9] J. Froehlich, L. Findlater, J. Landay, The design of eco-feedback technology in: Proceedings of CHI2010, Atlanta, GA, USA, April 10–15, 2010.
[10] J. Rowan, E.D. Mynatt, Digital family portrait field trial: support for aging in place. in: Proceedings of the ACM Conference on Human Factors in Computing Systems, CHI 2005, 2005, pp. 512–530.
[11] S. Consolvo, P. Roessler, B.E. Shelton, The CareNet display: lessons learned from an in home evaluation of an ambient display, in: Proceedings of the 6th International Conference on Ubiquitous Computing, Nottingham, England, Sep. 2004, pp.1–17.
[12] W.B. DeOreo, J.P. Heaney, P.W. Mayer, Flow trace analysis to assess water use, Journal of the American Water Works Association (AWWA) 88 (1) (1996) 79–90.
[13] A. Lowenstein, C.C. Hiller, Disaggregating residential hot water use: part I, in: American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Transactions: Symposia. AT-96-18-1, 1996, pp. 1019–1027.
[14] L. Bao, S.S. Intille, Activity recognition from user-annotated acceleration data, in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2004, 2004, pp. 1–17.
[15] J. Lester, T. Choudhury, N. Kern, G. Borriello, B. Hannaford, A hybrid discriminative/generative approach for modeling human activities, in: International Joint Conference on Artificial Intelligence, IJCAI 2005, 2005, pp.766–772.
[16] M. Philipose, K.P. Fishkin, M. Perkowitz, D.J. Patterson, D. Fox, H. Kautz, D. Hahnel, Inferring activities from interactions with objects, IEEE Pervasive Computing 3 (4) (2004) 50–57.
[17] B. Brumitt, B. Meyers, J. Krumm, A. Kern, S. Shafer, EasyLiving: technologies for intelligent environments. in: Proceedings of the International Symposium on Handheld and Ubiquitous Computing, HUC 2000, 2000, pp. 12–29.
[18] J. Chen, A.H. Kam, J. Zhang, N. Liu, L. Shue, Bathroom activity monitoring based on sound. in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2005, 2005, pp. 47–61.
[19] E. Munguia Tapia, S.S. Intille, K. Larson, Activity recognition in the home using simple and ubiquitous sensors, in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2004, 2004, pp. 158–175.
[20] E. Munguia Tapia, S.S. Intille, L. Lopez, K. Larson, The design of a portable kit of wireless sensors for naturalistic data collection, in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2006, 2006, pp. 117–134.
[21] D. Wilson, C.G. Atkeson, Simultaneous tracking & activity recognition (STAR) using many anonymous, binary sensors, in: Proceedings of the International Conference on Pervasive Computing, Pervasive 2005, 2005, pp. 62–79.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

[22] C. Beckmann, S. Consolvo, A. LaMarca, Some assembly required: supporting end-user sensor installation in domestic ubiquitous computing environments, in: Proceedings of the International Conference on Ubiquitous Computing, 2004, pp. 107–124.
[23] T. Hirsch, J. Forlizzi, E. Hyder, J. Goetz, C. Kurtz, J. Stroback, The ELDer project: social and emotional factors in the design of eldercare technologies, in: Proceedings of the ACM Conference on Universal Usability, CUU 2000, 2000, pp. 72–29.
[24] Z. Satterfield, V. Bhardwaj, Water meters tech brief, Water Meters 4 (2) (2004), National Environmental Services Center.
[25] Y. Takeda, Velocity profile measurement by ultrasonic doppler method, Experimental Thermal and Fluid Science 10 (1995) 444–453.
[26] J.W. Foreman, E.W. George, R.D. Lewis, Measurement of localized flow velocities in gases with a laser doppler flowmeter, Applied Physics Letters 7 (4) (1965) 77–78.
[27] R.J. Adrian, Twenty years of particle image velocimetry, in: 12th International Symposium on Applications of Laser Techniques to Fluid Mechanics, Lisbon, July 12–15, 2004.
[28] R. Evans, J. Blotter, A. Stephens, Flow rate measurements using flow-induced pipe vibration, Journal of Fluids Engineering 126 (2) (2004) 280–285.
[29] Y. Kim, T. Schmid, Z.M. Charbiwala, J. Friedman, M.B. Srivastava, NAWMS: non-intrusive autonomous water monitoring system, in: Proceedings of the ACM Conference on Embedded Network Sensor Systems, SenSys 2008, 2008, pp. 309–322.
[30] E. Cheung, Municipal water meter monitor, 2003 http://www.edcheung.com/automa/water.htm (last accessed on 11.12.09).
[31] P.W. Mayer, W.B. DeOreo, J. Kiefer, E. Opitz, B. Dziegieliewski, J.O. Nelson, Residential End Uses of Water, American Water Works Association, Denver, Colo, 1999.
[32] Aquacraft Corporation, http://www.aquacraft.com (last accessed on 11.12.09).
[33] B. Dziegielewski, E. Opitz, J. Kiefer, D. Baumann, Evaluation of Urban Water Conservation Programs: A Procedures Manual, Prepared for California Urban Water Agencies by Planning and Management Consultants, Ltd., Carbondale, IL, Feb 1992.
[34] W.B. DeOreo, P.W. Mayer, Project Report: A Process Approach for Measuring Residential Water Use and Assessing Conservation Effectiveness, City of Boulder Office of Water Conservation, Boulder, Colorado, 1994.
[35] P.W. Mayer, Residential water use and conservation effectiveness: a process approach, Master's Thesis, University of Colorado, Boulder, 1995.
[36] W.B. DeOreo, P. Lander, P.W. Mayer, New approaches in assessing water use, in: Proceedings of Conservation 1996, AWWA and AWWARF, Orlando, FL, 1996.
[37] W.B. DeOreo, P.W. Mayer, The end uses of hot water in single family homes from flow trace analysis, Aquacraft Inc. Report, 2000, http://www.aquacraft.com/Download_Reports/DISAGGREGATED-HOT_WATER_USE.pdf (last accessed on 11.12.09).
[38] C.C. Hiller, New hot water consumption analysis and water-heating system sizing methodology, in: American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Transactions: Symposia. SF-98-31-3, 1998, pp. 1864–1877.
[39] US Department of Energy, US household electricity report, energy information administration, US Department of Energy, 2001, http://www.eia.doe.gov/emeu/reps/enduse/er01_us_tab1.html (last accessed on 11.12.09).
[40] G.P. Henze, D.K. Tiller, M. Fischer, M. Rieger, Comparison of event inference and flow trace signature methods for hot water end use analysis, in: ASHRAE Transactions, vol. 108, part 2, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2002, pp. 467–479.
[41] C. Aguilar, D.J. White, D. Ryan, Domestic water heating and water heater energy consumption in Canada, in: Canadian Building Energy End-Use Data and Analysis Centre, CBEEDAC 2005–RP-02, 2005.
[42] J.S. Weihl, W. Kempton, Residential hot water energy analysis: instruments and algorithms, Energy and Buildings 8 (3) (1985) 197–204.
[43] G.O. Ladd, and J.L. Harrison (Eds.), Electric water heating for single-family residences: group load research and analysis, EPRI EA-40006, 1985.
[44] A. Lowenstein, C.C. Hiller, Disaggregating residential hot water use: part II, in: American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) Transactions: Symposia. SF-98-31-2, 1998, pp. 1852–1863.
[45] O. Reynolds, Philosophical Transactions 174 (1883) 935–982; Scientific papers, 2, p. 51, Cambridge University Press, London.
[46] H. Blasius, Das Ahnlickkeitsgesetz bei Reibungsvorgangen in Flussigkeiten Forshungheft 131, Berlin, 1911.
[47] D.O. North, An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems, Proceedings of the IEEE 51 (7) (1963) 1016–1027.
[48] A. Oppenheim, R. Schafer, From frequency to quefrency: a history of the cepstrum, IEEE Signal Processing Magazine 21 (5) (2004) 95–106.
[49] Judith Brown, Calculation of a constant Q spectral transform, Journal of Acoustical Society of America 89 (1) (1991) 425–434.
[50] Judith Brown, S.Puckette Miller, An efficient algorithm for the calculation of a constant Q transform, Journal of Acoustical Society of America 92 (5) (1992) 2698–2701.
[51] L. Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition, Proceedings of the IEEE 77 (1989) 257–286.
[52] Kevin Murphy, HMM toolbox for Matlab, 1998. http://people.cs.ubc.ca/~murphyk/Software/HMM/hmm.html.
[53] J. Froehlich, E. Larson, T. Campbell, C. Haggerty, J. Fogarty, S. Patel, HydroSense: infrastructure-mediated single-point sensing of whole-home water activity, in: Proceedings of UbiComp 2009, Orlando, FL, USA, September 30th–October 3rd, 2009.
[54] F.J. Arregui, C.V. Palau, L. Gascon, O. Peris, in: E. Cabrera, E. Cabrera (Eds.), Evaluation of Domestic Water Meter Accuracy: A Case Study, 2003, pp. 343–352.

Eric Larson is a senior Ph.D. student in the laboratory of Ubiquitous Computing at the University of Washington and is advised by Shwetak Patel. His main research area is supported signal processing in ubiquitous computing applications, in particular, for healthcare and environmental sustainability applications. He received his B.S. and M.S. degrees in electrical engineering from Oklahoma State University, Stillwater in 2006 and 2008 respectively where he specialized in image processing and perception under the advisement of Damon Chandler. He is active in signal processing education and is a member of IEEE and HKN. Contact him at eclarson@uw.edu.

Jon Froehlich is a Ph.D. candidate and Microsoft Research Graduate Fellow in Human Computer Interaction (HCI) and Ubiquitous Computing (UbiComp) at the University of Washington co-advised by James Landay and Shwetak Patel. He was recently selected as the UW College of Engineering Graduate Student Innovator of the Year for 2010. His dissertation is on promoting sustainable behaviors through automated sensing and feedback technology. Jon received his MS in Information and Computer Science specializing in analytic visualizations from the University of California, Irvine under the advisement of Paul Dourish. Contact him at jonfroehlich@gmail.com.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008

Tim Campbell is a Senior undergraduate student in Mechanical Engineering at the University of Washington. His research focus is in Mechatronics and combines mechanical devices with self-powered sensing and feedback techniques. He is a member of the Ubiquitous Computing Laboratory under the advisement of Dr. Shwetak Patel. He recently had the honor of speaking on behalf of the Mechanical Engineering student body at the department's Scholarship Luncheon. Contact him at tim.b.campbell@gmail.com.

Conor Haggerty graduated from the University of Washington in 2010 with a degree in Community Environment and Planning. His studies focused on land use and its effect on waterways. Advised by Dr. Robert Edmonds, he received a minor in Streamside Studies as well as an ISA Certification in arboriculture. Conor worked in the development stages of sensor technology under the advisement of James Landay. He competed in the UW Environmental Challenge and the Business Plan Competition showcasing the HydroSense technology he helped to develop. He now owns a small business where he works as a contracting arborist. Contact him at conorhaggerty@gmail.com.

Les Atlas received his M.S. and Ph.D. degrees in Electrical Engineering from Stanford University in 1979 and 1984, respectively. He joined the University of Washington in 1984, where he is a Professor of Electrical Engineering. His research is in digital signal processing, with specializations in audio and acoustic analysis, and time frequency and modulation frequency representations. Prof. Atlas received a NSF Presidential Young Investigator Award and a 2004 Fulbright Senior Research Scholar Award. He is a Fellow of the IEEE and was General Chair of the 1998 IEEE ICASSP, founding Chair of the IEEE Signal Processing Society Technical Committee on Theory and Methods and a Member-at-Large of the Signal Processing Society's Board of Governors. He is currently a member of the IEEE Signal Processing Society Technical Committee on Speech and Language Technology. His current research sponsors are the AFOSR and the ONR.

James Fogarty is an Assistant Professor of Computer Science & Engineering at the University of Washington. His broad research interests are in Human–Computer Interaction, User Interface Software and Technology, and Ubiquitous Computing. His focus is on developing, deploying, and evaluating new approaches to the human obstacles surrounding widespread everyday adoption of ubiquitous sensing and intelligent computing technologies. He received his Ph.D. in 2006 from the Human–Computer Interaction Institute at Carnegie Mellon University, where he worked with Scott E. Hudson. He received his B.S. in Computer Science in 2000 at Virginia Tech, where he worked with John Carroll and Mary Beth Rosson.

Shwetak N. Patel is an Assistant Professor in the departments of Computer Science and Engineering and Electrical Engineering at the University of Washington. His research interests are in the areas of Human–Computer Interaction, Ubiquitous Computing, and User Interface Software and Technology. He is particularly interested in developing easy-to-deploy sensing technologies and approaches for location and activity recognition applications. Shwetak is also the co-founder of Usenso, Inc., a demand side energy monitoring solution provider. Shwetak received his Ph.D. in Computer Science from the Georgia Institute of Technology in 2008 and B.S. in Computer Science in 2003. He was also the Assistant Director of the Aware Home Research Initiative at Georgia Tech. Dr. Patel was recently a TR-35 award recipient for 2009.

Please cite this article in press as: E. Larson, et al., Disaggregated water sensing from a single, pressure-based sensor: An extended analysis of HydroSense using staged experiments, Pervasive and Mobile Computing (2010), doi:10.1016/j.pmcj.2010.08.008