Using Infrastructure-provided Context Filters for Efficient Fine-grained Activity Sensing

S Vigneshwaran[†], Sougata Sen[†], Archan Misra[†], Satyadip Chakraborti [‡] and Rajesh Krishna Balan[†]
[†]Singapore Management University and [‡]Indian Institute of Technology- Kharagpur
[†]{vigneshwaran,sougata.sen.2012,archanm,rajesh}@smu.edu.sg and [‡]satyadipc@gmail.com

***Abstract*—While mobile and wearable sensing can capture unique insights into fine-grained activities (such as gestures and limb-based actions) at an individual level, their energy overheads are still prohibitive enough to prevent them from being executed continuously. In this paper, we explore practical alternatives to addressing this challenge–by exploring how cheap infrastructure sensors or information sources (e.g., BLE beacons) can be harnessed with such mobile/wearable sensors to provide an effective solution that reduces energy consumption without sacrificing accuracy. The key idea is that many fine-grained activities that we desire to capture are specific to certain location, movement or *background* context: infrastructure sensors and information sources (e.g., BLE beacons) offer practical and cheap ways to identify such context. In this paper, we first explore how various infrastructure, mobile & wearable sensors can be used to identify fine-grained location/movement context (e.g., transiting through a door). We then show, using a couple of illustrative examples (specifically, the detection of 'switch pressing' before exiting a room and the identification of 'water drinking' after approaching a water cooler) to show that such background context can be predicted, with sufficient accuracy, with sufficient lead time to enable a 'triggered' model for mobile/wearable sensing of such microscopic, transient gestures and activities. Moreover, such 'triggered' sensing also helps to improve the accuracy of such microscopic gesture recognition, by reducing the set of candidate activity labels. Empirical experiments show that we are able to identify 82.2% of switch-pressing and 91.73% of water-drinking activities in a campus lab setting, with a significant reduction in active sensing time (up to 92.9% compared to continuous sensing).**

I. Introduction

The world of sensing today exhibits a decidedly dual, but rapidly converging, track. On one hand, after the recent wave of smartphone-based mobile sensing applications (providing context such as locomotive state and movement trajectories [13], [1]), attention is migrating to wearable sensing (e.g., using a smartwatch) that can capture finer-grained activity context (such as stress levels and interaction gestures). On the other hand, urban public spaces are being progressively fitted with ambient and multimedia sensors (such as BLE beacons, pressure, motion and video sensors) that help provide near-continuous and person-independent observability of physical environments. Finding effective ways to harness the combined capabilities of these two paradigms (mobile & infrastructure) is clearly a compelling and timely research goal.

This work is supported partially by Singapore Ministry of Education Academic Research Fund Tier 2 under research grant MOE2011-T2-1001, and partially by the National Research Foundation, Prime Ministers Office, Singapore, under IDM Futures Funding Initiative and International Research Centre @ Singapore Funding Initiative, and administered by the Interactive & Digital Media Program Office, Media Development Authority. All findings and recommendations are those of the authors and do not necessarily reflect the views of the granting agency, or Singapore Management University.

Energy overhead is the most formidable obstacle to the ubiquitous application of mobile/wearable sensing: while on-body sensors can provide very deep insights into individual activity and behavior, it is practically impossible to keep them active continuously throughout the day. The only plausible solution to support such fine-grained observation is *adaptive sensing*, where mobile/wearable sensors are activated only at appropriate instants. Additionally, while many fine-grained gestures and actions can be deduced from mobile/wearable sensing, such actions can often be very *transient* and similar to other unrelated daily-lifestyle based actions–e.g., classifiers may find it hard to distinguish the "pushing of a door" from "the pressing of a button on a vending machine" from pure smartwatch-based sensing.

Our primary insight is that, many such fine-grained personal activities exhibit a high-degree of location/context dependence–i.e., they are performed only under certain very-specific location or movement-based contexts. For example, (a) the act of 'switching lights on/off' is performed shortly before exiting or after entering a room; and (b) the act of 'filling up a bottle and drinking' is performed usually only at or near a water-cooler. Accordingly, if the mobile/wearable sensors could be activated, and their activity recognition vocabulary modulated, in response to such movement-based context, it should be possible to achieve both our objectives: energy efficiency (by activating sensors only during certain events) and accuracy (by restricting the set of activities/gestures to the ones relevant to the current context).

Hence, in this work we investigate the following two **Key Research Questions** :

- Can multi-modal sensing provide the vital disambiguating information (context filters) needed to improve the accuracy of the fine-grained personalized context inferred from wearables?
- Can infrastructure sensors provide cheap triggers (instantaneous/anticipatory) to activate wearable/mobile sensors, thus resulting in energy savings?

These questions are often application-dependent, and it may be hard to establish universal responses to them. Hence, this paper presents an exploration of these important questions, using three simple daily-life scenarios that can be observed in a laboratory/office environment, viz. *Door: detecting who entered/exited through a door*, *Lights: detecting whether the lights were turned off/on during such a door transition* and

978-1-4799-8033-8/15/$31.00 ©2015 IEEE

Water: monitoring the hydration habits of users during the day. These scenarios are used to provide insights on how to combine the capabilities of commercially available infrastructure sensors and low-cost mobile/wearable sensors.

Our **Key Contributions** are:

- We show that by combining the various sensing modalities, we can sometimes obtain contextual information that cannot be obtained from individual sensors and also improve the accuracy of classification of fine-grained activities, by eliminating irrelevant labels. We first contrast the individual capabilities of infrastructure, mobile and wearable sensors, for the *Door* scenario. While each of the individual sensors is successful to a certain extent, we find that 100% reliable door transition detection can be achieved by strategically placing 2 BLE beacons. But when we move to the *Light* scenario, we find that none of the individual sensors are able to identify this reliably and the only way is to combine the various modalities. Finally, using the *Water* scenario, we demonstrate that the accuracy of detecting a water-drinking gesture can also be improved by about 20% by eliminating activity labels that are not relevant near the water-cooler.
- For the same three scenarios, we show that by utilizing the infrastructure based sensors as a trigger to start the sensing applications on the wearables, we can achieve up to 92.9% savings on the sensing time. However, such triggering results in a trade-off between the errors due to late-triggering and false-context which are studied empirically. It is found that the triggered sensing is reliable only when the lead-times are at least 1.8 seconds.

II. Motivating Scenarios & Use Cases

To motivate the various studies and analyses performed in this paper, we consider three application examples often used in the mobile/wearable sensing literature.

Scenario 1: "Door?": Jill the building manager wants to keep accurate tabs of the occupancy statistics of precious few meeting rooms. This will allow her to go to senior management with hard data that they need to build more meeting rooms. In particular, she wants to accurately, and cheaply identify who is using each room and for how long. Besides being useful for generating operational reports, such real-time information can also be used in emergency evacuation scenarios.

Scenario 2: "Lights": Jill would also like to address another problem. Some people who use the meeting rooms don't turn off the lights when leaving, causing unnecessary utilization of electricity! Jill would like to extend her solution to also detect when the light switch is turned on and off (i.e., when a particular group leaves), and use this to prompt users to turn off the lights if they forget to do so.

Scenario 3: "Water": As the wellness manager of his office campus, Jack is bedeviled by his employees who work continuously without taking periodic drinks at the water cooler. Jack would like a simple yet accurate system that can detect if a user has approached the water-cooler and actually filled up his bottle or cup. Such information can be integrated into proactive reminders that help improve the wellbeing of employees.

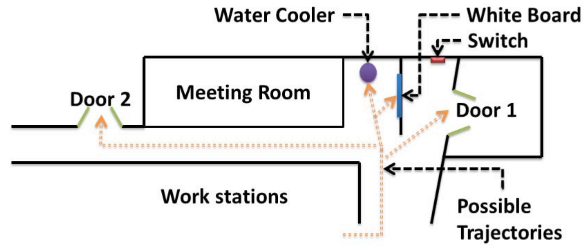


Fig. 1: Floor Layout

The three scenarios above motivate the design and evaluation in the rest of this paper; Section IV, V, VI describes how we realize the three scenarios in detail, respectively. The laboratory setup used in our studies is shown in Figure 1. The lab has two doors that can be opened only in one direction; incoming users usually pull on the door to enter, while outgoing users push on the door to exit. The doors are locked using a magnet–when a user either presses a door release button or swipes her RFID-equipped access card, the lock gets demagnetized, releasing the door. Finally, the doors are spring-mounted and self-closing–once release, the doors gradually swing back to a closed position.

To study the efficacy of various sensors at addressing the usage scenarios, we used the following four sensor platforms (2 infrastructure, 1 mobile, and 1 wearable) for all results and experiments in this paper:

1 *Door-mounted Raspberry Pi:* To capture the motion-related artifacts of a door, we mounted a Raspberry embedded platform on the door, equipped with the Adafruit 10-DOF sensor board (which includes an accelerometer, gyroscope, magnetic sensors, pressure sensor, & temperature sensor). These sensors allow us to detect when the door moved – indicating a transition.

2 *Estimote (BLE) Beacons*: We also attached BLE beacons to the doors and at locations near to the entry and exit points of the room. We use the signal strength of the beacons as observed on a client mobile device to infer door transitions. For our studies, we used the commercially available Estimote^{TM} beacon, which transmits IDs repeatedly once every 950 ms by default.

3 *Smartphone:* To capture an individual's locomotion-related context, we utilized the person's smartphone sensors – specifically, the accelerometer and the compass sensors on a representative Samsung Galaxy S4 smartphone. For door transition events, such locomotion data provides useful insights into when a person is stationary (e.g., while opening a door) vs. when the person is moving (e.g., approaching or passing through the door).

4 *Smartwatch*: Finally, we also used a Samsung Gear smartwatch, mounted on the wrist of an individual. The Smartwatch's sensors (e.g., the accelerometer and gyroscope) helps in identifying fine-grained arm movements and gestures, such as the pressing of a wall switch or drinking water.

Our goal is to understand the relative benefits of each of these sensors, in identifying various attributes related to the three scenarios (i.e., how many people and who uses the door, and in which direction, was a light switch turned on?). As we will show, Scenarios 2 and 3 *requires* both infrastructure and mobile sensors (either alone is insufficient) to achieve high detection accuracy and low energy consumption (via appropriate and timely triggering of mobile or wearable sensors).

III. Related Work

While there is obviously a tremendous amount of past work on both mobile and wearable-based activity recognition, we highlight the important contributions that are either similar to our focus on the joint use of infrastructure & mobile sensing for smart spaces, or that provide specific techniques that underpin our work.

Infrastructure+ Mobile Sensing for Smart Environments: There has been a lot of work in building smart environment and integrating them with mobile devices. MIT's PlaceLab is one such scenario where RFID tags have been placed on objects to detect daily-use activities [4]; Gu et al. [2] improved on the activity detection in smart homes. In their work they considered complex, interleaved activities and using sequential pattern mining, were able to identify these complex activities. Again they showed how a combination of wearable and infrastructure sensors could help in identifying these interleaved activities. Other works which have shown a combination of infrastructure and mobile sensing includes the STAR algorithm [12] where the authors used many binary ambient sensors for activity detection and Roy et al. [9] showed a combination of mobile and ambient sensors could be used for semantic activity recognition.

Building Automation: Our infrastructure sensor had tasks similar to Lu et al.'s work [7]. However in the scenarios mentioned by them, the accuracy could have been improved if a combination of infrastructure and wearable sensors was used. There are many other work such as [10], [6] where infrastructure sensor has been used to detect home occupancy. However similar to [7], none of them have augmented their system by using personalized sensors. Contrary to this, [5] used a combination of infrastructure and wearable sensors to detect who is utilizing what device; such approaches do not however utilize the gesture-level information that may be captured by emerging wearable devices (such as a smartwatch).

Video-based Sensing: There has been work where video based sensing has been used to detect various contexts and activities performed in an indoor space [3]. Although video feed provides rich contextual information, video-based sensing techniques have their own disadvantages. First of all, they increases the privacy concern of participants significantly. Also, the energy cost in performing continuous video capture and processing is very high. Hence, for our work, we explicitly exclude the use of video sensors.

Specific Technologies: Our approach of using BLE-based signal strength to infer movement patterns is similar to and inspired by Wang et al. [11]. Our focus on using smartwatches (and their specific sensors, e.g., accelerometer, gyroscope) adopts techniques demonstrated in approaches such as e-Gesture [8]. We feel that a combination of infrastructure and wearable/mobile sensors can help in identifying a wide range of novel and compelling context aware applications.

IV. Scenario 1 (Door): Location Transitions (Detecting Room Occupancy)

In this section, we show how we detect the occupancy of a room. To do this, we make the assumption that the room has a door that leads in and out (the solution can generalize easily to multiple doors). We thus need to detect every individual that enters or leaves through the door. In particular, we answer the following four questions (which together allow a comprehensive and accurate occupancy map to be generated) :

- *Q1: Was a specific door used?* This is the most basic question and basically provides a binary answer detailing whether (and when) a particular door was used.
- *Q2: Were they entering or leaving?* The next level of sophistication is to identify whether the door was used to enter or leave the room.
- *Q3: How many people were involved?* We next show how to count exactly how many people used the door at the same time, including the possibility that the door was opened by one person, but then used by multiple people to enter or exit.
- *Q4: Who specifically used the door?* Finally, we evaluate how to identify exactly who entered or left the room.

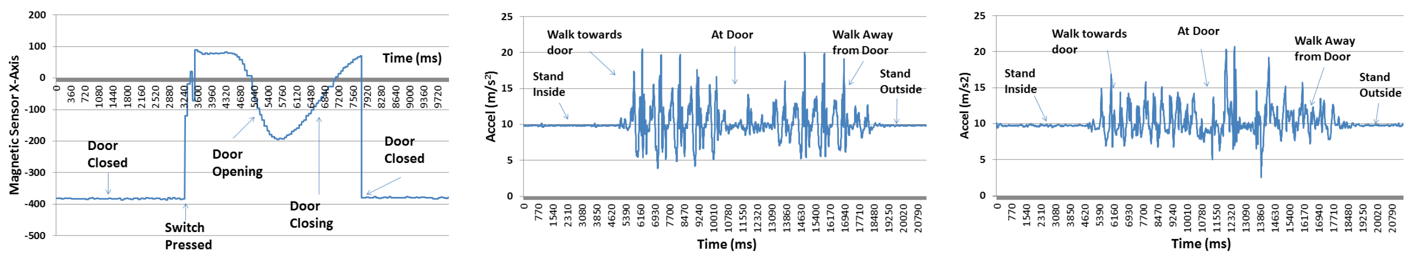
Empirical observations revealed that our doors exhibited four distinct states–(i) *Door Close*, where the magnet was activated and the door is closed at its position of rest; (ii) *Disarmed*, when the door has been demagnetized to allow the ingress or egress of an authorized person, but the door is still closed; (iii) *Door Opening*, where the door is pushed or pulled and remains in an open position, and (iv) *Door Closing*, where the door is finally returning to its initial closed position. Fig 2a shows a sample reading of the magnetic sensor on the Raspberry Pi (as well as the accelerometer readings for a user-carried smartphone (Figure 2b) and smartwatch(2c)) during these four different states of the door activity.

To run our experiments, we recruited 6 volunteers (3 males, 3 females) from our research lab to engage in multiple instances of both *leaving* & *entering* door activities – with each participant carrying and wearing a unique phone and watch and conducting between 6 to 27 separate activity instances. To train the phone and watch-based recognition models, we adopted the typical approach of using labeled training data to build a supervised activity classification model, and then utilized this model to classify unlabeled training data (all results reported use a stratified 10-fold cross validation approach, unless otherwise mentioned). For the Estimote beacons, we captured the signal strength readings (and its temporal variation) as recorded by the smartphone.

A. Q1 & Q2: Detecting Door Activity & Direction

We now present results for the first two questions listed at the start of the section – can we detect if a door was opened, and if so, did a person enter or leave the corresponding area? We first provide results for each sensor used independently. We then show how combining sensors can lead to improvements.

Door Based Infrastructure Sensor: Since the infrastructure sensor is mounted on the door, we could identify with *100%*



(a) Raspberry Pi's Magnetic Sensor (b) Smartphone's accelerometer sensor (c) Smartwatch's accelerometer reading

Fig. 2: Sensor reading patterns for the door activity

accuracy if a door activity took place, i.e. whether someone entered or exited the lab, based purely on the change in readings of the door-mounted compass sensor (Figure 2). We also found the magnetic sensor to be more reliable than the accelerometer, as seemingly random touches on the door sometimes caused fluctuations in the accelerometer readings.

Smartphone Based: To identify a door transition activity, we utilize the accelerometer sensor on the smartphone, and attempted to classify frames corresponding to a "Door Usage" activity. Multiple models of various class sizes were trained where models had class sizes of either 3, 7 or 21. Table I lists down the various class labels and the corresponding accuracy obtained using the phone (as well as the watch). To generate the model for activity recognition we used a sliding window approach with frame length of 1 second, and extracted both time and frequency domain features (as explained in [13]). The smartphone attempts to capture a sequence of motion-based artefacts, such as the user walking (up to the door), stopping to release the magnetic lock, gradually opening the door and subsequent walking again (away from the door). Based on our experiments, we obtained accuracies of *87.96%, 79.2% and 71.07%* for 3,7 and 21 class identification respectively in determining if a door activity took place. Moreover, if a door activity took place, the orientation of the phone could be used, together with the ground truth about the door's layout, to determine whether a person entered or exited a door.

Smartwatch: Similar to the smartphone, we developed a 3, 7 and 21 class label identification model for the smartwatch and obtained accuracies of *83.4%, 72.4% and 63.4%* respectively for the three classes. While the overall accuracy drops as the smartwatch is unable to correctly classify many of the other activity labels (e.g., elevator or slow-walk), its accuracy is much higher in detecting the door pushing or pulling motion (The true-positive rates for the door activity for 3-class and the 7-class classifiers are *97.2% and 85.8%* respectively.). Additionally, the smartwatch was able to distinguish *door-pull* gesture from *door-push* gesture with an accuracy of 93%.

BLE Beacon Based: Lastly, we tested the accuracy of door-activity identification using the Estimote beacons mounted on each door–the transmission signal strength and the time interval between successive advertisements are programmable and set to the 'normal' range (-8dBm) and 109ms (unless otherwise specified). The smartphone then monitors the signal strength variation of the beacons continuously—this signal strength increases rapidly as one approaches the door, and then drops again. While it is possible to detect if a door activity took place, this approach (using a single beacon per door) leads to false positives, as it cannot distinguish cases where a user approaches a door but then turns back.

Label Count	Class Labels	Accuracy (%) Phone	Accuracy (%) Watch
3	Stationary, Moving, Door Activity	87.96	83.40
7	Sit, Stand, Walk, Stairs, Elevator, Escalator, Door Activity	79.20	72.40
21	Sit, Sit-relax, Sit-Work, Sit-Write, Sit-Read, Stand, Stand-relax, Sit-Work, Stand-Write, Stand-Read, Walk, Slow-walk, Normal-Walk, StairsUp/Down, EscalatorUp/Down, ElevatorUp/Down, Door(Push/Pull), Door-Follow (Follow someone through the door)	71.07	63.40

TABLE I: Class Labels

We then tried out a slightly more involved layout, which involved the use of 2 beacons. Similar to the conveyor belt approach for asset tracking using RFID, if two beacons are placed at a little distance from each other, it turns out that we can identify both door activity as well as the direction of movement, based on the time evolution of the received signal strength indicator (RSSI) data at each phone. As mentioned earlier, we have two beacons in our experimental setup, one on each door. Both beacons are audible inside the lab. However as one approaches a door, the RSSI of the beacon at the door rises rapidly (up to -67dBm) however the RSSI of the other beacon falls and bottoms out at around -85dBm to -88dBm. After the person exits the door, the RSSI from the beacon on the door comes down in a nice slope, but the other RSSI from the other beacon becomes very weak and becomes inaudible in about 5m. Similar behavior can be observed when the person exits through the other door, with the beacons reversing their roles. (These observations are reversed for a person entering the door). We also tested out this approach in an additional campus classroom, obtaining identical results (100% classification accuracy for both door events as well as enter/exit determination).

Given the **remarkable success of the two beacon strategy**, we further investigated the performance impacts of the use of such BLE beacons. As our method relies on the use of the short-timescale evolution of the beacon signal strength measurements, we expect that our accuracy may diminish if the transmitting interval become larger (i.e., the readings become more intermittent). The transmitting interval affects its lifetime and for our experimental setup the expected lifetime is about 100 days. But for transmission interval of 50ms, the lifetime of a beacon becomes considerably short (26 to 40 days). The

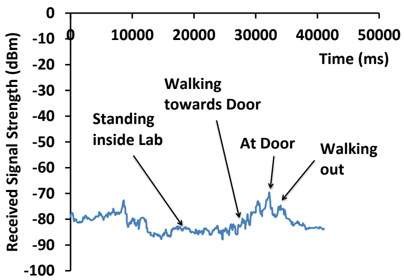


Fig. 3: RSSI of BLE beacon

DutyCycle	Transmit Interval 200ms	1000ms	2000ms	3000ms
500ms On - 500ms Off	100%	100%	93.75%	81.25%
750ms On - 750ms Off	100%	100%	87.5%	81.25%
1s On - 1s Off	100%	87.5%	87.5%	75%

TABLE II: Effect of duty-cycling on beacon accuracy

beacon-based approach also requires continuous scanning of Bluetooth signals on the smartphone, but in order to preserve battery lifetime of the phone, most experimental approaches adopt scanning duty cycles of 50% or lower (turning off the Bluetooth sensor during the off-part of the duty cycle). The impact of duty cycling (on the phone) and transmission interval (on each of the 2 BLE beacons) on the accuracy of the detection of door activity is summarized in Table II. We see that the accuracy of door detection remains 100% as long as the transmission interval does not exceed 1 second; even when the intervals is 1 sec (close to the default values of 950ms), the accuracy of door activity detection may drop to 87.5% if the phone's duty cycle is too long. Overall, these results suggest that an infrastructure-based *dual BLE beacon* strategy may be an effective strategy for capturing the door-based ingress/egress events of an individual, providing far higher accuracy as compared to mobile or smartwatch-based approaches (which also impose significantly higher energy overheads).

B. Q3: Number of People Transiting

From empirical observations, we see that people transiting a door in groups exhibit one of 3 different types of behavior: (i) One person holds the door open to let others behind him pass by to exit the door; (ii) Each person exiting the door taps the door enough to let herself or himself pass by. (iii) The lead person opens the door and continues walking, while the people trailing this person "sneaked in/out", before the door returns to its default closed position.

If an app installed on an individual's smartwatch/smartphone determines his/her transition through the door (using sensor data from smartphone, smartwatch and BLE beacons), then the problem of "people counting" is trivial. Hence, in this subsection we focus on investigating situations where the infrastructure-based approach can be used to estimate the number of people transiting, without relying on any mobile sensing or app mounted on an individual's personal device.

Door-based Infrastructure Sensor: We tried to use the duration of *how long the door remained open* to determine how many people passed through. However, in the realistic study we found that different groups had different time to pass through. Also the results were even more skewed when someone opened the door and stood there to chat with another nearby person.

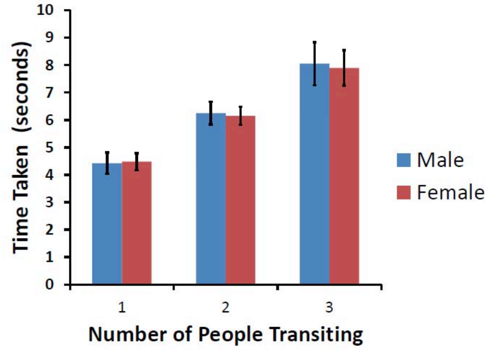


Fig. 4: Time taken for groups to pass through a door

Figure 4 plots the variation in the total transit time (i.e., how long the door was open) for different group sizes through one door. We see that there is a gradual increase in the transit time and there is difference between different group sizes. However on observing the data from the second door (not plotted), we found that the time taken for different groups to transit was different from door 1 – i.e. for a group size of 2, the transit time through that door was within the error range of group size 3's transit time through the other door. On closer inspection, we found that the duration of *door activity* time varied between the two doors, due to : (i) differences in the spring constant values for each door, resulting in different levels of force being applied to each door, and (ii) due to the door access readers being located on different sides (left wall vs. right wall), resulting in different people using either their dominant or non-dominant hand to open the door. Accordingly, it appears that pure infrastructure-based sensing (without requiring any App or cooperation from a mobile device) cannot accurately estimate the number of people crossing the door.

C. Q4: Identifying the Person(s) Transiting

As mentioned previously, using multiple BLE sensors(and the temporal pattern of the resulting RF signal), an individual phone can predict its own door transition quite accurately. However, the accuracy of detecting individual-specific door transitions using either mobile or smartwatch-based sensing is not extremely high (87% and 83% approximately). We now look at how the accuracy can be improved by correlating and fusing the data from a combination of infrastructure and mobile/wearable sensors.

Combining {Pi, beacon, smartphone, smartwatch}: To identify the person transiting a specific door (i.e., determining that person A just passed through door $D1$), our key idea is to correlate the time when the infrastructure sensor indicates a "door opened" event, with the 'door activity' labels reported by the smartphone or the smartwatch. More specifically, the data from both the infrastructure and mobile/wearable sensors were uploaded to a central server, which utilized a Spearmans Rank Correlation ranking to identify the person using the phone (the one with the highest value). We observe the following insights:

- *Phone + Raspberry Pi*: When we combine the Pi-based detection of a door-opening event with the smartphone-based classifier, we were able to improve the classification accuracy (precision) to 82%–i.e., 82% of door-opening events by individuals were correctly identified. While the improvement in precision was not dramatic, this approach helps improve recall significantly, by filtering out all the smartphone-generated values of 'door transition' that were caused by mis-classification.

- *Watch + Raspberry Pi*: When we combined the Pi-based detection of the door-opening event with the smartwatch-based classifier, we could improve the accuracy (precision) of door activity detection to 93%. Clearly, a smartwatch often generates false-positives (mis-classifying various other gestures with the push/pull motion of the door).
- *Watch/Mobile + 1 Beacon*: To complete our understanding, we also studied the effectiveness of using either a smartphone or smartwatch, in combination with a single door-mounted beacon (recall that a single beacon alone can result in false classification, especially if an individual approaches the door and then reverses her motion). In this case, we observed that a combination of mobile+beacon and watch+beacon resulted in classification accuracies (precision) of 91% in both cases.

D. Summary

Overall, our detailed investigations of various combinations of infrastructure and mobile sensing provides several important takeaways. First, as long as one does not desire ultra-long operational lifetimes on the BLE beacons, signal strength variation-based approaches that utilize readings from multiple BLE beacons can detect door transitions very precisely (with almost 100% accuracy); the accuracy numbers, however, drop to $\sim 90\%$, if the beacon intervals grow larger than 1 sec. Second, mobile/wearable sensing-based activity recognition can also be used to detect door-based transitions, but its accuracy is much lower (close to $\sim 70\%$), especially when the range of possible activities is larger.

Table III summarizes the various accuracy results obtained, in terms of context accuracy, either by using a single sensing device in isolation or by combining multiple devices. Overall, we can see that an appropriate choice of infrastructure beacons can help detect door transitions very accurately. In the subsequent section, we shall see if these transitions can be detected early enough to allow us to capture transient gesture-based activities.

V. Scenario 2 (Lights): Are The Lights Still On? (Triggered Sensing of Transient Events)

Even though infrastructure sensors can detect door activities fairly accurately, other cases such as whether the "*user switched off the light before leaving*" are either tougher to determine using infrastructure sensors or monetary cost of setting up such sensing infrastructure can be high. In this section we show how a combination of infrastructure and mobile + wearable sensors can help in easily detecting such scenarios. We have seen previously that the Beacon can help in identifying door activity based on its RSSI. In this section we show how varying the RSSI threshold from the beacon affects the *lead time* in detecting if a door activity will take

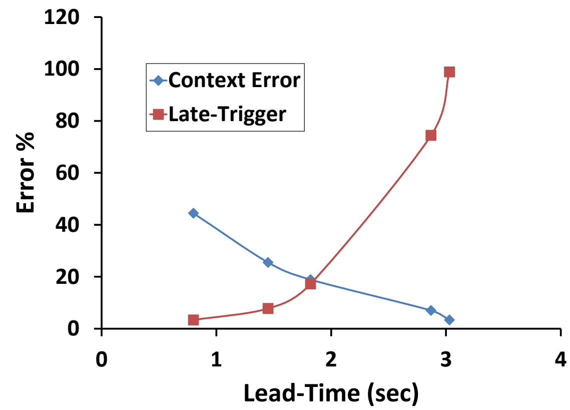


Fig. 5: Ideal Lead-time

place. If we can achieve a reasonable lead time to trigger the smartphone or smartwatch to load the model which can detect *switch-activity*, we can save on energy in the devices by turning off the sensors at times when door activity will not take place.

We again set up the experiments in our laboratory as described earlier. There is a switch near the door which the users can press just before exiting via the door(ref. Fig 1). Our use-case is to identify effectively if the user pressed the switch before exiting or not. A beacon was placed at the door and the users wore the smart watch on their wrists as well as carried a smartphone in their pant pockets during these experiments. While the smartwatch can capture the gesture made by the user when reaching out to the switch, it is highly energy-inefficient to continuously keep the sensors on the smartwatch *ON*, for this solitary purpose. Instead, we aim to use the RSSI from the beacon captured by the smartphone as a trigger to turn on the sensors in the watch. As the user nears the door, the RSSI increases and when it exceeds a threshold value we trigger the sensing mechanism on the watch. A point to note here is that it is not necessary that the user will transit through the door every time the threshold is hit and thus unnecessary sensing might happen. In our setup we have a water cooler and white board situated approximately 3 meters away from the door. Thus there is a large common path between a person who wishes to exit through the door and a person who wishes to go to the water cooler or to the white board. If we trigger the sensing mechanism too early, we may end up at the wrong context, if we trigger too late, we risk missing the switch context. We investigated this trade-off empirically on 3 users. Each user was asked to perform three activities – press the switch and go through the door, go to the water-cooler and have water, go to the white board and write a word. The users performed each of three activities 30 times. For every activity being performed, the user would start from his/her work station and go through the common path to reach their respective destination (Water-cooler/Board/Door).

We now study the effect of the RSSI threshold, on the lead-time that we can get before the user presses the switch. Due to various factors, the threshold RSSI level may not be hit at the same point in the trajectory for all the trials and hence this lead-time can vary. In Table IV we have reported the average lead-time we obtained for different values of threshold RSSI levels. As the threshold increases, the lead time becomes

	Door Event					Number of People		
	Door Opened?	Which door?	Who?	Entry or Exit ?	Switch Press?	2-people	4-people	N people walk to door, r transit
Raspberry Pi only	Yes	Yes	No	No	No	Yes	Yes	No
SmartPhone's accelerometer only	Yes. But false positives when user is just walking around	Partially. If the doors face different direction orientation sensor can be used	Yes	No	Yes. But many false positives	At Server end		
SmartWatch's accelerometer only	Yes. Better than phone but still many false positives	No	Yes	Partially. If the doors swing only one way, push Vs pull can be used.	Yes. Better than phone but many false positives.	At Server end		
Beacon only	Yes	Yes	Yes	Yes (Multiple beacons).	No	Yes	Yes	Yes
Raspberry Pi + Smartphone's accelerometer	Yes	Yes	Yes	No	Yes. Lesser false positives	Yes	Yes	No
Raspberry Pi + Smartwatch's accelerometer	Yes	Yes	Yes	Partially.	Yes	Yes	Yes	No
Beacon + Smartphone's accelerometer	Yes	Yes	Yes	Yes. With multiple beacons	Yes	Yes	Yes	Yes
Beacon + Smartwatch's accelerometer	Yes	Yes	Yes	Yes. With multiple beacons	Yes	Yes	Yes	Yes

TABLE III: Summary of capabilities of various sensing techniques

RSSI(-dBm)	Lead-time (Sec)	Late-trigger	False-context
-72	3.03	3.33%	98.89%
-70	2.87	7%	74.4%
-68	1.82	18.89%	17.22%
-67.5	1.45	25.56%	7.78%
-67	0.8	44.44%	3.33%

TABLE IV: Scenario 2: Variation of Lead-time and accuracy with RSSI

smaller since the user is closer to the door. At higher threshold levels, in some cases we found that the threshold level is hit after the switch press event and hence we cannot trigger the watch model at the right time. This can be handled by giving sufficient lead-time at a lower RSSI threshold. However, doing so could result in false-triggering since lower RSSI thresholds can be hit along the common path much before the user branches off towards the door. From such a point in the trajectory it is not possible to infer that the user will head towards the door. Hence there is a trade-off here and the ideal point can be found such that both the context and triggering errors are low enough as shown in Figure 5. For this experiment we find that at a threshold RSSI of -68dBm, the context errors and the late-triggering errors are low enough resulting in an average lead time of 1.82 seconds. Assuming that the watch recognition model was exactly triggered at the lead-time, we found that the watch model detected the switch press event 82.2% of the time.

VI. Scenario 3 (Water): Drinking (Hydration) Monitoring

In this section, we show how the same set of sensors can be used to provide accurate and energy-efficient answers to a different yet important sensing task – namely to track how often, a person goes to the water-cooler to drink water. Similar to the scenario of detecting whether a user turned off the lights before leaving through a door, this scenario involves multiple short-term gestures that is relevant to a particular location. Drinking water is a short-term gesture that can be captured by the watch. However, similar to the previous scenario, we wish to turn on the sensors closer to the water cooler so that we can conserve battery on the watch. For our experiment setup, we attached a beacon near the water cooler. There is a white board besides the water cooler where lab members often made notes. We could not use the Beacon alone to determine the context, since the user could come near the water-cooler to use the board. To ensure we could differentiate between the two activities - *filling bottle/drinking water* and *writing on notice board* we combined the data from the beacon with the data from the watch.

Similar to the previous section, we varied the RSSI threshold to determine when the smartwatch classifier should be triggered. The empirical results for this case is reported in Table V and the Fig 6. The lead-time for the hydrating activity was obtained using similar methodology as the lead time for the switch press activity. For the hydrating activity, we found that having a threshold of ≈ -68dBm gave a good lead time with not many false positives. However in this case, the overall error at this ideal lead-time is smaller than the previous experiment. This can be due to the fact that the user simply had to walk along a straight line to the water cooler and there was clear line-of-sight all the time. In the case of the door activity, users will have had to branch off in "Y-turn" and hence the door was slightly farther away than the water-cooler.

To evaluate the accuracy, we first built a classfciation model on the smartwatch that could differentiate between 4 activities - *drinking water, writing on board, door activity and walking inside lab*. We found that these activities could be identified with 71.54% accuracy. However, since the *drinking water* activity is location specific, we need to turn *ON* the classifier in the watch only when the person is near the water cooler (similar to the switch activity monitoring). This eliminated classifying the other 2 activities (door activity and walking inside lab) as they were not performed near the water cooler. Using this approach, we found that we could identify the *drinking water* activity with an accuracy of 91.735% and a precision of 89.51%. Thus we obtained a significant increase in the classification accuracy by eliminating activity labels that are not relevant at a particular location.

Energy savings: Continuous sensing severely impacts the battery life of mobile and wearable devices. In order to demonstrate the energy savings resulting from the triggered sensing paradigm, we performed a naturalistic user-study on 3 lab members. The BLE beacons were setup as mentioned earlier and the users carried the smart-phone in their front pant

RSSI(-dBm)	Lead-time (Sec)	Late-trigger	False-context
-72	3.30	1.67%	80%
-70	1.96	7.78%	23.3%
-68	1.45	8.33%	3.33%
-66	1.12	12.78%	0%
-65	0.52	20.0%	0%

TABLE V: Scenario 3: Variation of Lead-time and accuracy with RSSI

pocket and wore the smart-watch as well. The phone and the watch were charged fully before starting the experiment and they collected accelerometer and gyroscope sensor readings at 100Hz. We evaluated the number of times the classification models on the smart-watch (for door switch detection and hydration detection) would have been triggered based on the RSSI from the beacons, as the user moved about the lab performing their regular activities. For uniformity, the experiment started at 12:00 noon and ended at 4:00 PM, over 2 days. By analyzing the RSSI of beacons we found that, on average, the door switch detection model was triggered 7.4 times while the hydration detection model was triggered 4.3 times, resulting in an average sensing time of ≈ 17 minutes, which is a drastic reduction of 92.9% compared to keeping the sensors *ON* throughout. In our study, we have considered only two transient event scenarios. As the number of transient events increases, the sensing time will also increase due to more frequent triggering, thus resulting in lesser energy savings. We also found that if we kept the accelerometer and gyroscope in the smartwatch *ON*, to sample at 100Hz, the device would run out of charge in ≈ 3 hours while according to specifications, it could last for ≈ 24hours when no sensor is turned on.

VII. Conclusion

In this paper we showed how various infrastructure, wearable and mobile sensors can be used to identify fine grained user context which would have been either not possible or expensive if the these sensors tried predicting individually. We also showed that using triggers from cheaper infrastructure sensors can be used to save energy on wearable or mobile devices while not compromising too much on the accuracy. In our experiments we found that with 100% accuracy we could detect using infrastructure sensor if a person transited a door. However, it was not possible to detect whether a person performed a short gesture such as a switch-press, while exiting using existing infrastructure sensors alone. However with a combination of infrastructure, mobile and wearable sensors, we achieved an accuracy of 82.1% in detecting switch press cases. We tried another similar scenario involving multiple gestures involving water drinking habits of the users. We could identify the multiple gestures effectively and achieve 91.7% accuracy for this scenario. All this improvement is on top of the energy savings on the wearables due to the 92.9% reduction in sensing time resulting from the triggered sensing paradigm.

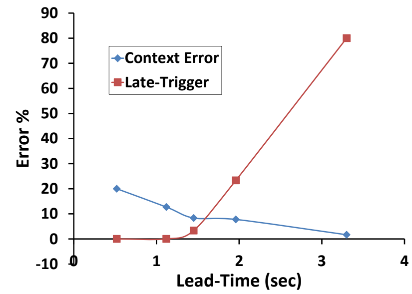


Fig. 6: Ideal Lead-time

References

[1] Chon, Y., Talipov, E., Shin, H., and Cha, H. Mobility prediction-based smartphone energy optimization for everyday location monitoring. *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*, SenSys '11, pages 82–95, New York, NY, USA, 2011. ACM.

[2] Gu, T., Wu, Z., Tao, X., Pung, H. K., and Lu, J. epsicar: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition. *Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on*, pages 1–9. IEEE.

[3] Hongeng, S., Nevatia, R., and Bremond, F. Video-based event recognition: activity representation and probabilistic recognition methods. *Computer Vision and Image Understanding*, 96(2):129–162, 2004.

[4] Intille, S. S., Larson, K., Tapia, E. M., Beaudin, J. S., Kaushik, P., Nawyn, J., and Rockinson, R. Using a live-in laboratory for ubiquitous computing research. *Pervasive Computing*, pages 349–365. Springer, 2006.

[5] Lee, S., Ahn, D., Lee, S., Ha, R., and Cha, H. Personalized energy auditor: Estimating personal electricity usage. *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, pages 44–49, March 2014.

[6] Li, N., Calis, G., and Becerik-Gerber, B. Measuring and monitoring occupancy with an rfid based system for demand-driven hvac operations. *Automation in construction*, 24:89–99, 2012.

[7] Lu, J., Sookoor, T., Srinivasan, V., Gao, G., Holben, B., Stankovic, J., Field, E., and Whitehouse, K. The smart thermostat: Using occupancy sensors to save energy in homes. *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, SenSys '10, pages 211–224, New York, NY, USA, 2010. ACM.

[8] Park, T., Lee, J., Hwang, I., Yoo, C., Nachman, L., and Song, J. E-gesture: A collaborative architecture for energy-efficient gesture recognition with hand-worn sensor and mobile devices. *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*, SenSys '11, pages 260–273, New York, NY, USA, 2011. ACM.

[9] Roy, N., Misra, A., and Cook, D. Infrastructure-assisted smartphone-based adl recognition in multi-inhabitant smart environments. *Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on*, pages 38–46. IEEE, 2013.

[10] Scott, J., Bernheim Brush, A., Krumm, J., Meyers, B., Hazas, M., Hodges, S., and Villar, N. Preheat: Controlling home heating using occupancy prediction. *Proceedings of the 13th International Conference on Ubiquitous Computing*, UbiComp '11, pages 281–290, New York, NY, USA, 2011. ACM.

[11] Wang, Y., Yang, J., Chen, Y., Liu, H., Gruteser, M., and Martin, R. P. Tracking human queues using single-point signal monitoring. *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, MobiSys '14, pages 42–54, New York, NY, USA, 2014. ACM.

[12] Wilson, D. H. and Atkeson, C. Simultaneous tracking and activity recognition (star) using many anonymous, binary sensors. *Pervasive computing*, pages 62–79. Springer, 2005.

[13] Yan, Z., Subbaraju, V., Chakraborty, D., Misra, A., and Aberer, K. Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach. *Wearable Computers (ISWC), 2012 16th International Symposium on*, pages 17–24. Ieee, 2012.