

It Takes Two: Exploring Interactions between Smart Objects and Wearables to Implicitly Identify and Authenticate Object Users

A Dissertation

Presented to

the Faculty of the School of Engineering and Applied Science

University of Virginia

In Partial Fulfillment

of the requirements for the Degree

Doctor of Philosophy (Computer Science)

by

Juhi Ranjan

December 2016

Abstract

In today's context, the term 'Smart Object' is used to refer to any object that incorporates computing and communication in some capacity to enhance the functionality and/or interaction experience for the end user. People interacting with objects have different personalities and preferences, and therefore different requirements from the same objects. Therefore, personalization of a Smart Object's behavior will soon become an important function. Often when people interact with objects that are shared with other people, they need to re-configure the objects to suit their personal requirements. For example, while showering, people prefer different water temperatures, and therefore have to set the hot and cold water mixer to their preferred configuration every time they shower.

In the current state of the art, if the Smart Objects need to adapt themselves to the person using them, they perform an explicit identification process for the user. However as Smart Objects are expected to permeate every aspect of people's daily lives, this approach will not be scalable. In order for objects to perform personalized functions, they must solve what we refer to as the Implicit Object User Identification (IOUI) problem: understanding who is actually using a given object, and being able to validate their identity without expecting the user to explicitly participate in an identification process.

In this dissertation, we explore the use of wearable devices in performing IOUI. There are two main reasons why wearable devices are an attractive technological solution that can assist Smart Objects in solving this problem: a) wearables adoption is growing at a rapid pace, and b) they are embedded with sensors that can monitor the location context and hand motion of

the wearer. However, while sensors in wearables are great for making approximate measures of a person’s activities, the imprecision of their sensing systems makes them challenging for use in applications such as IOUI, which require high precision and accuracy.

In this work, we explore the following hypothesis: Despite the coarse granularity of its location sensing, and imprecision in sensing the trajectory of hand’s motion, data from sensors in wearable devices, when augmented with data from Smart Objects, can be used to identify users interacting with Smart Objects. We explore different levels of information shared by the Wearable device with a Smart Object, and explore how each level of data abstraction affects the user identification accuracy.

Approval Sheet

This dissertation is submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy (Computer Science)

Juhi Ranjan

This dissertation has been read and approved by the Examining Committee:

Kamin Whitehouse, Adviser

Jack Stankovic, Committee Chair

James Scott

Gabriel Robins

Laura Barnes

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, Dean, School of Engineering and Applied Science

December 2016

Dedicated to maa, papa, dida, dadu and my husband

Acknowledgments

Many people have contributed to the making of this work - either directly, or indirectly, and words cannot do justice to their love and support.

Undoubtedly, the biggest force behind this work is my advisor, Kamin Whitehouse. When I first came to grad school I was very naive and untrained. Kamin has been very patient with me while teaching me the ropes of the scientific method. He has taught me how to think critically, broaden the scope of ideas, and communicate effectively. I have been very lucky to have worked with such an eloquent person. Communication was probably one of my biggest weaknesses, and over the past five years, Kamin has made me practice how to communicate my ideas in various forms - be it writing, presentations, or whiteboard discussions. He is a very creative person. His perspective on research contributions never fails to amaze me. I hope someday I can learn to look at the same problem with as many lenses as he can. He has always helped me improve by expecting me to be more strategic about my work and time. He has shared with me countless tips about research, networking, writing, etc., and these tips have served me well throughout the last five years. I cannot thank him enough for all that he has done for me.

Most people are lucky to find one good mentor during their PhD; I have been incredibly lucky to have found two. James Scott was my supervisor when I interned at MSR Cambridge, and I continued working with him until recently. James has been an amazing mentor, and a friend, and I will always be thankful to him for his advice and guidance. Working closely with him during the internship was a great learning experience in designing experiments and

conducting research. No matter how busy his schedule, James has always taken the time to meet with me. His feedback has been honest and direct, which made it very easy for me to understand things I was doing right, and things I needed to improve. His passion and excitement for research ideas is inspiring, and it kept me driven through the low phase of our research. Although I have enjoyed annoying him with my sub-optimal presentation making habits (read: complex builds on the same slide), thanks to his relentless efforts to teach me to do better, I have improved. His drive for perfection, and attention to detail, has shown me how little things can make a big difference between good and great. Working with him has given a big boost to my confidence as a researcher.

I would also like to thank other members of my committee for their contribution to my research career at UVa. Even though I haven't worked with Prof. Stankovic directly, I have absorbed some of his guidance indirectly, while interacting with his students. I really appreciate his support and thoughtful suggestions during my quals proposal, defense and Ph.D proposal. He has always supported me in my endeavors and fellowship applications, and has given me good career advice. I'd also like to thank Prof. Robins. He has been very accommodating in making time for all of my quals and PhD presentations. His feedback and perspective on my research has been very unique and enriching. I have also enjoyed taking two of his classes - Theory of Computer and Design and Analysis of Algorithms. His teaching methodology really broadens the perspective of solving problems, and I have really enjoyed attending his lectures. I would also like to thank Prof. Laura Barnes, with whom I had also taken my first Machine Learning class. Her use of ML tools in the class and assignments help me understand how to apply a concept that was earlier very unclear to me. Things I learned in her class have served me well in many of my research projects.

There are other faculty members in UVa to whom I am very grateful. I'm very thankful to Prof. Soffa, who was a part of my quals defense for her helpful advice on my masters thesis. I would also like to thank Prof. Cohoon for being so kind to me during my tenure as his teaching assistant. By observing him, I have learnt a great deal about what caring for

students truly means. I would also like to thank Prof. Skadron for his encouragement and kindness.

I have also learned a great deal from people in my research group: Erin, Yu Yao, Dezhi, Elahe, Devika, and Avinash. Their critical feedback on my work, and presentations, as well as thoughtful questions, has been pivotal in developing my research ideas. Our research group can sometimes be a tough crowd to please, and that is good thing because preparing my presentations with them has helped me pick up techniques on addressing a skeptical audience. I have had the opportunity to work alongside and learn from them in different capacities, and I admire different qualities in each of them. I would also like to acknowledge some of the senior students of our group - Tim Hnat and Jiakang Lu for their advice. I would like to thank Tamim Sokoor, for always being available to answer any and all questions that I had about navigating grad school. He was very supportive when I was applying for NSF GRFP, and I really appreciate his encouragement. I would also like to acknowledge the undergrads that I have mentored at UVa - Tahiya, Paul, Eugene and Rahim. Working with enthusiastic and hard working students like them has been an energizing experience. Working with them has helped me understand how to structure projects better, and it has helped me better understand the process of developing new ideas.

I'd like to thank our department's system staff - Brian Donato, Scott Ruffner and Essex Scales, who have always been available and prompt in helping me with systems issues, and poster printouts. I'd also like to thank our department's office staff - Kim, Christine, Debbie, Jan and Brenda. I'd like to thank Wendy for always taking the time to help me sort the finances of my trips and purchases, and for being so nice to me.

I would also like to thank the National Science Foundation for funding my research as well as providing me with a generous fellowship for three years during my Ph.D.

I have been incredibly blessed to have made a great set of friends at UVa. They have cooked for me, let me plan our trips, tolerated my (awesome) jokes/pranks, participated in my research studies, given me rides to the airport, and pushed me to try new things. When

you have friends like I do, having your family 13,500 km away doesn't bother you that much.

Here are some of the people who have been a big part of my life at Charlottesville: Tanima - She gave me a place to stay when I arrived at Charlottesville. She cooked for me, and drove me to stores. Over time we also became gym buddies, and she has been like a sister to me; Enamul - From him, I learnt how to stay calm no matter how difficult the situation. I jokingly call him and Tanima, dad and mom, because of how caring they both are; Vidhya - We became very good friends, very fast. In fact, she called me her soulmate. But then, friendships like that don't last too long. She met her husband, and I was reduced to the position of a soulsister - I don't even know what that means. In fact now, she denies having ever called me a soulmate; Angad - He is the reason that Vidhya demoted my significance in her life. Angad is one of the most well informed people that I know, and extremely articulate. He was also the conductor/pseudo-priest at my wedding renewal ceremony and he did a fantastic job; Anjana - She kept me at an arm's length till she thought I was Miss Goody Two Shoes. Of course, once she lived with me for three months, and then moved to the apartment upstairs, she realized how deeply flawed I was. After that she readily inducted me into her circle of close friends, and we have been BFFs ever since. Now we buy matching clothes, purses and toothbrushes; Anubhav - I didn't know that an anesthesiologist was a real doctor till I met him. He is a very caring person, and he's made sure that I took my medicines on time when I was sick; Kaushik - He was responsible for my love for hiking. I would have never realized the joys of rock scrambling, had he not insisted that I go to the Old Rag Mountain; Nivi - One of the sweetest people I know, she makes it very hard for me to even jokingly be nasty to her. She is very accommodating, and the perfect companion for long chats; Samyukta - She is one of the most generous people I have ever come across. We became friends through the dinner parties that she would cook for the entire department. Eventually, we became neighbors, and she took care of me like a mother. She knew how much I hated eating out during my week long conference trips, and I would always come back

to great home cooked meals by her; Avinash - He was just a regular acquaintance until we became neighbors when he moved upstairs. He turned me into a monster with his laundry room (in)abilities and I had to come up with the most ridiculous house rules to make sure he didn't use the dryer as a closet. He's been a great friend and his Mafia playing skills make me think that he's a mind reader. He goes on the most incredible (and risky) vacations, and inspires me to take a break from my work and go for a vacation; Piyush - I have no idea why every time my husband plays tennis with him, Piyush brings him back to me in a broken condition. He is a smart person, and one of the funniest people that I know; Heather - My respect for nurses has increased after meeting her. She is such an incredible person - very calm and resilient; Sid - He was in Charlottesville for only three months, but in that short time he taught me so much. He pushed me to run longer and faster, and I give him credit for my my first 5K run. We also made a short horror movie together, in which I played the part of a ghost, and Avinash played my body double - it was a riot; Devika - She is my adopted daughter. Her dedication to work and her attention to detail is inspiring. She has been a true friend to me; Divya - I really admire her, not only because she's smart and sensitive, but also because of how she pushes her physical endurance in activities such as rock climbing; Ritambhara - She is another girl I really admire. Very well spoken and polite, she knows how to aim high and reach for the stars; Essex and Liza - They have almost been like second parents to me and my husband. We've enjoyed spending many weekends together, going to the river, shooting at the range, and riding horses; Yamina - She and I started at UVa together. While she was at UVa, every year she would cook amazing food for me during Karwachauth (Indian festival); Anindya and Jisa - They have been very kind to me, and going to their home for a chat has always felt like therapy to me; Emi - She has been a great friend to me. Always helpful and adjusting. Once, she even treated me to a surprise grand dinner; Asif - He is humility personified. He has been very encouraging, and I really admire his photography'; Sarah - Although, I haven't spent a lot of time with her, I really appreciate her honesty and how dedicated she is to her work; Samee - We became great friends when we

started learning how to ice skate on our own. We bonded over YouTube skating wisdom, and board games; Yan - He and I have deep conversations about life, work, and family. He's very adjusting, caring and thoughtful; Kirti - He was one of my first friends at UVa. When he came to know how frustrated I was at my inability to judge a used car, he went out of his way to make sure I got a good one, and his choice has served me well over the last five years; Ashwin - He was also one of my first friends at UVa. We bonded over meals that we'd cook together, and songs that we sang; Anup and Katie - They were part of the Friday movie club group. Trips on weekends were so much fun with the two of them around. Other than UVa, I would like to mention some of my old friends at Bangalore - Baba, Neeraj, Pragnya, Daddu, and Keshu. We've had a lot of fun times together. Leaving you guys behind to work on my research was not easy. I would also like to acknowledge the support of some of my other friends whom I've known since my school and undergrad days: Jayanti, Richa, Mansi, Pallavi, Sonal, Komal, Lakshmi and Nilesh. I have very fond memories of college and school, thanks to all the fun I've had with them.

PhD is not an easy journey to embark upon. It would have been impossible without the love and support from my family - especially my father, mother, maternal grandparents (dida and dadu), and my husband. My father was my first research mentor. He never told me that I couldn't do something. He never brought up my gender as being related to what I could or could not do. This may seem trivial to some, but knowing the society I grew up in, in India, this is huge. He has always been very supportive of all my career decisions, and never pressurized me into doing something just because it was good for the family's optics. My father as a researcher is also very inspiring. He never cared about creating publication records, or being part of the rat race. He has always looked for real problems that he wanted to solve, and actually focused on solving them. He is always looking for learning new things, and his oratory skills are legendary. My love for presentations come from my memories as a child when I would listen to my father describe how a tokamak worked during Open House, and then when he would take a break, I would take over explaining it to the groups of visitors.

I would also like to acknowledge my mother. She's probably upset at this point that I talked about my father first :P. My mom's attitude towards her work is very inspiring, and she is passionate about learning new computing techniques. But more than her work attitude, I really appreciate my mom's support throughout my academic career. She has always put me first, ahead of her own comfort, her career, and choices. I know everyone thinks that their mom is the best, but my mom really IS THE BEST. She never gave up on me, even when I was intolerable. In fact I have no idea how she tolerated me in my rebellious teenage years. But because she is awesome, she did, and I am a more secure person because I know she's there for me no matter what. Every time she visits me, she goes through a lot of effort to make sure I am comfortable, working hard, and eating well. She takes great joy in the smallest things that I do. For instance, the other day I babysat someone's dog for a weekend, and she told me gushingly how proud she was of me. I cannot imagine how happy this degree will make her. I'd also like to mention my maternal grandparents, who have been a big part of my life. They both passed away two years back. My grandmother (dida) was a very strong woman. Till the very end she lived life on her own terms. She was very accomplished, and was even offered a fully funded PhD at Germany, which she declined to raise my mother. I am sure wherever she is, she is very happy that I completed my doctoral degree. My maternal grandfather (dadu) would have also been very proud of me. He was a very loving person. He was the real doctor in the family. My love for learning comes from him. Even when he was 90 years old, he stayed up to date with the last medical techniques. I'd also like to acknowledge my husband, who's been a great partner during this long PhD journey - Who knew that marrying the first guy I went on a date with, would turn out to be the best decision that I made. My husband is the reason I can never be complacent about my career. He pushes me out of my comfort zone and challenges me to do something good. He is probably the only person who knows how to make me laugh, no matter how stressed I am. He also lets me go through all sorts of crazy ideas - like buying a shiny new red convertible. Seeing him work with dedication and build a company, also inspires me to work harder. He is also a great

travel partner. During the last five years, we've been on many trips together. Weekends with him have helped reset and refresh my mind. Its been 10 years since we met, and my life seems to keep getting better with him in it. Also, thanks to him, I have two other wonderful people in my life - his brother Naman and my sister-in-law, Ankita. They've given me a lot of love and support when I needed it the most. I'd also like to acknowledge the support of my cousins - Nupur, her husband Prashant, Mayank, his wife, Shruti, Sonam, Himanshu and Diksha. I know they'll be thrilled to bits when I finally graduate with a Ph.D.

I would like to acknowledge some of the MeetUp groups I was a part of in Charlottesville: The Good Reads book club - Until I joined them, I never realized how fun it can be to discuss about a book at great lengths with other people. Thanks to Jen, Sam, Carrie and others, I read books I had never even heard of; Photography club - I learnt how to take better pictures thanks to photo walks they organized, and the monthly themed photo shoots I'd participate in. I'd also like to thank Terry Dean's ballroom dancing group for organizing amazing classes and events. I have enjoyed dancing at their studio every single time.

I do want to mention my car (WSNGAL), who was a quintessential companion in my research endeavors at Charlottesville. She never failed me on the road; no matter how long the journey, or how harsh the weather, be it driving on early winter mornings, or escaping Hurricane Katrina, she has been a faithful and steady friend. I'm also very grateful to have found a home that was very conducive to my quality of life during Ph.D. A big part of my research was conducted at my home at Piedmont Ave, which was my haven with its peace and serenity. Living in a house with such a gorgeous backyard made it very hard to go back to a 6x6 office cubicle.

I am pretty sure I have unintentionally missed acknowledging many other people who've been part of my life and my research. To all those I have missed mentioning - A BIG THANK YOU.

Contents

Contents	xiii
List of Tables	xvi
List of Figures	xvii
1 Introduction	1
1.1 What are Smart Objects?	1
1.2 What is the State of the Art in Smart Objects?	2
1.3 What is Implicit Object User Identification?	5
1.4 How can User Identity Transform Smart Objects?	6
1.5 Why is Implicit User Identification Challenging?	8
1.6 What Are Wearables, and Why Are They a Promising Solution?	8
1.7 What are the Main Challenges in Using Wearables?	10
1.8 Thesis Statement	11
1.9 How can Wearables and Smart Objects Interact to Identify and Authenticate Users?	12
1.10 Contributions	16
1.11 Limitations	18
1.12 Thesis Outline	19
2 Background	21
2.1 Indoor Positioning System	21
2.2 Activities of Daily Living	23
2.3 Wearable Devices	25
2.4 Inertial Measurement Unit	26
2.5 Personal Energy Feedback	27
3 Related Work	29
3.1 Indoor Positioning System	29
3.2 Object User Identification	33
3.3 Object User Authentication	35
4 Level 1 - Using Location to Identify Object Users	39
4.1 Introduction	39
4.2 Related Work	43
4.3 Approach	45

4.3.1	Step 1 - Object Usage Detection	45
4.3.2	Step 2 - Location Tracking	47
4.3.3	Step 3 - Object User Identification	49
4.3.4	Step 4 - Applying Heuristics	50
4.4	Case Studies	52
4.4.1	Lights Energy Analysis	52
4.4.2	Habit Monitoring	53
4.5	Experimental Setup	54
4.5.1	Participants	54
4.5.2	Sensing Infrastructure	55
4.6	Results	57
4.7	Discussion	63
4.8	Limitations	64
4.9	Conclusion	65
5	Level 2 - Using Gesture Matching to Identify Object User	66
5.1	Introduction	66
5.2	Related Work	69
5.3	Approach	70
5.3.1	Sensing Object Hallmarks	70
5.3.2	System Requirements	71
5.3.3	Overview of Object User Identification	72
5.4	Experimental Setup	77
5.4.1	Study Design	78
5.5	Results	83
5.6	Discussion	91
5.7	Limitations	92
5.8	Conclusion	92
6	Level 3 - Using Raw Sensor Data to Authenticate Object User	93
6.1	Introduction	93
6.2	Related Work	97
6.3	Approach	99
6.3.1	System: Hardware and Sensor	100
6.3.2	Features used in Authenticating User	101
6.3.3	Prototype	104
6.4	Feasibility Study	105
6.4.1	Procedure	105
6.4.2	Analysis	106
6.5	Prototype Interaction Study	112
6.5.1	Procedure	113
6.5.2	Analysis	116
6.6	Adversarial Study	119
6.6.1	Procedure	120
6.6.2	Analysis	120

6.7	Limitations	121
6.8	Conclusion	122
7	Conclusion	124
7.1	Summary	124
7.2	Contributions	125
7.3	Limitations	126
7.4	Future Research Directions	126
	Bibliography	128

List of Tables

4.1	Analysis of room location based object assignment: The confusion matrix represents the object assignments for individuals when the room is singly occupied. Ambiguous object usage (when multiple people are present in same room) is the main reason why room location tracking has lower object assignment accuracy than coordinate system. Group 4 has 0 ambiguous events, and therefore 100% room location based object assignment accuracy	58
4.2	Improved room-location based object assignments: The use of simple heuristics with room location tracking resulted in object assignment accuracy of 87% in group 1 and 92% in group 2. This is comparable to 96% and 98% accuracy achieved respectively in coordinate level tracking system	60

List of Figures

1.1	First level of interaction - Is to figure potential object users based on proximity. Sometimes only one user may be detected. In this case, the lone user is identified as the object user. This level of interaction is insufficient if there are multiple users.	13
1.2	Second level of interaction - Is to query the devices nearby for how well their motion data matches the gesture required to interact with the object. This can help disambiguate cases when multiple users are detected near the object. However, this method is vulnerable to being spoofed, as devices may maliciously respond with a higher matching score. Therefore this method cannot be used for sensitive and privacy critical applications.	14
1.3	Third level of interaction - Is reserved for applications where it is important to validate that a user is the actual object user. In this stage, the identified user is queried for their raw sensor data, which is analyzed and correlated with the object's sensor data at a millisecond level accuracy	15
2.1	While more number of sensors can detect a richer set of ADLs, it is more practical in terms of cost and maintenance to have systems that can infer ADLs from fewer sensing points (figure from [1])	24
2.2	Wearable devices come in many different form factors, although currently wrist worn devices such as smartwatches and fitness trackers are leading in the Wearables market	26
2.3	In the PowerHouse, the game player manages a simulated domestic environment with seven characters. The meters in the lower pane displays a specific characters mental and physical state. In the upper right corner are the money and power meters that show the accumulated points and how much energy is being consumed.	27
3.1	Energy-unaware behavior uses twice as much energy as the minimum that can be achieved (figure from [2])	34
3.2	In PhoneTouch [3], surface and phones detect touch events independently. The device-level observations are communicated over a wireless network, correlated in time, and combined to associate a touch with both surface position and phone identity	37

4.1	Energy disaggregation (of one of the participant groups): Per-person per-fixture disaggregation makes it evident that Person B consumes most of the electrical and water fixture energy in the house except for TV and Dryer, which are used only by Person A	40
4.2	Sensors Deployed: (a) All major electrical and water fixtures in the test home were instrumented, including lights (with light sensor), refrigerator (with reed switch on door), toaster (with power meter), and stove (with reed switch inside knob). (b) Sensors included power meters for plugs, in-line water flow meters, door open/close sensors, and fiber optic light pipes attached to light sensors. (c) Onset's data loggers, such as 4-channel pulse data logger were used	41
4.3	Sensor layout: the first floor of our test home had at least one instrumented light object, water object, or electrical appliance in every room. All doorways were instrumented with an RF Doormat.	46
4.4	Case study of Lights energy: (a) Aggregate energy analysis reveals that person B uses light energy comparable to person A, but wastes three times more energy than A, (b) Energy disaggregated per person, per room indicates that Lights energy in dining room and hallway are mostly wasted, (c) Energy disaggregated per person, per object in the kitchen shows that the stove and banquet lights are used more efficiently	52
4.5	Person A showers more frequently than B. However A's total shower water usage (916 L) is lower than B's (1327 L). This is because B tends to take longer showers compared to A	53
4.6	User identification can be used to track hygiene habits, for e.g. how many times do people wash their hands after using the toilet flush	54
4.7	Evaluation results: House level tracking has an average object assignment accuracy of 12%. Room location level tracking alone has 70% accuracy. Heuristics augmented room location tracking has an average accuracy of 87%. Coordinate based tracking has an average of 97% accuracy across all groups	57
4.8	Heuristics analysis: (a) House Level Tracking - Heuristics achieve about 75% accuracy in groups 2 and 3, which have 23% and 40% accuracy without any heuristic. In group 1, 4 & 5, heuristics obtain 50% accuracy indicating that none of them worked better than random assignment. (b) Room Location Tracking Heuristics achieve over 85% accuracy in groups 1, 2 and 3 due to a high number of unambiguous object assignments. In group 4, heuristics achieve 72 % accuracy because of a small percent (45%) of unambiguous object assignments	59
4.9	object level breakdown of event assignment accuracy, for different heuristics for all groups, shows potential for developing complex heuristics which are sensitive to typical usage pattern of individual objects	61
5.1	When an object is used, the object user is identified as the person making a hand movement containing the object hallmark. The hand is monitored using wearable devices having IMU sensors, such as fitness trackers	67

5.2	Total acceleration across the x, y and z axis is decomposed into gravity and linear acceleration along the three axes. The gravity component indicates how a hand is oriented, and the linear acceleration defines the movement of the hand	72
5.3	Segmentation: Linear acceleration across x, y and z axis are filtered and combined using Root Sum Square (RSS). Peak detection is performed on RSS to determine bursts of acceleration. Sensor data of a fixed window size around each peak is used for feature extraction	74
5.4	Experimental Setup: (L-R) The fridge, freezer and microwave face the same direction, and have doors that open the same way. The dishwasher door is hinged at the bottom. The bathroom hot and cold faucets pull out from left to right, and right to left respectively. The kitchen sink hot and cold faucets both turn in counter-clockwise direction. Light switches for kitchen and living room are right next to each other. The x, y, and z axis of a LG android smart watch are annotated	75
5.5	Object Layout: All the major appliances, light and water fixtures of a two-bedroom apartment were instrumented for the study. Quite often, two objects with similar interfaces are co-located, such as the kitchen sink hot and cold faucets	78
5.6	CDF of time differences between object usage: 50% of the object usage in the tasked study happened within 5 seconds of each other	82
5.7	Gesture classification: Confusion Matrix of 10-fold cross validation in scripted study show (a) 95% accuracy for a person dependent gesture model, (b) 88% accuracy for person independent gesture model	83
5.8	Classifying between objects with similar interfaces: Different features are evaluated to determine how gestures are differentiated between sets of co-located and similar objects	84
5.9	Different sensing features are evaluated to determine person identification accuracy. Linear Acceleration (L) based features do not perform as well as the orientation features (Gravity (G) and 3-D orientation (O))	86
5.10	An average accuracy of 80% is achieved with just a single training example for each object gesture. More training examples bring incremental improvements in the accuracy	87
5.11	Highest accuracy is obtained when a sensor data size of +/- 4 seconds around the object usage timestamp, is selected. The likelihood of an overlapping event from another person increases with a larger temporal range of sensor data. Accuracy is low for smaller sensor data size because the object monitoring hardware had a 1-3 seconds difference with the time on the watch	89
5.12	Accuracy was highest with a frame size of 1 second around each peak of the Linear Acceleration RSS. This may be due to the fact that time taken by people to physically interact with objects in the home is usually small	90
6.1	The third level of interaction correlates motion and orientation of the phone and the wearable to authenticate smartphone user. Features from the sensor data are extracted separately before and after the time when the touch event is initiated	94

6.2	Our approach works in different hand positions - the wearable may be worn in the same hand as holding the phone, on the hand touching the phone, and the hand holding and touching the phone. The only position in which our approach doesn't work is if the non-wearable hand holds as well as touches the phone	96
6.3	Wearable's linear acceleration patterns: A snapshot of 10 seconds compares the acceleration signatures of touch interactions as well as other activities. The touch signal consists of 8 phone touches which can be observed by the 8 peaks in the signal	101
6.4	Classifier evaluation: Random Forest and Boosted JRip classifiers have the highest true acceptance rate, as well as the lowest false acceptance rates, making them better suited for an authentication system. While Random Forest slightly outperforms the Boosted JRip classifier, we use the latter in our Android based prototype, for faster response time	107
6.5	Per participant analysis: Training per participant performs with an average f-score value of 96.4%, while leave-one-out (LOO) method performs with an average f-score value of 92.3%. LOO indicates that learning valid touch features from other participants may be suboptimal for some participants such as P2, P4, P8 and P12	108
6.6	Sensitivity analysis of training data: When the number of samples in the training dataset is varied from 12 to 120, we can observe that for some participants such as P7, classifier is able to achieve very high accuracy with very few samples. The classifier achieves an average of 95% accuracy after training on an average of 84 samples for all participants	110
6.7	Prototype interaction accuracy: The maximum TAR per session indicates that even though initially some users may initially not be able to interact easily on a system that is not trained on them, eventually most of them intuitively figure out a good way to interact with the phone. The adversarial case shows the low false acceptance rate of the system	111
6.8	Per Session Accuracy: Validation accuracy improves with every session of tasks indicating the participant starts to figure out how to use the system . .	112
6.9	Device Response: For the baseline case, almost all participants felt they were able to get the device to respond to their touches. Once the authentication system kicked in, there was a temporary dip in their response, although eventually it settled higher at an average perception of 'Agree'	114
6.10	Noticeability of device's lag in response: The baseline response was 'Disagree', which is a result of Android's inherent system delay. When the authentication system kicked in, the average response increased to 'Agree', though it eventually settled to 'Neutral' by the end of the task sessions	115
6.11	Acceptability of device's lag in response: The baseline response was 'Agree'. Once the authentication system kicked in, it dropped to slightly below 'Neutral', and improved and fell a couple time before settling at 'Neutral' by the end of the task sessions	116

- 6.12 Median number of touches that were allowed, before the system displays a strong password screen upon detecting **k** number of consecutive invalid touches. The average number of touches across all participants for the values of **k** from 1 to 5 are 1.1, 1.4, 1.8, 2 and 2.55 118
- 6.13 Maximum number of touches that were allowed, before the system displays a strong password screen upon detecting ‘k’ number of consecutive invalid touches. The average number of maximum touches the adversaries were able to achieve for the values of **k** from 1 to 5 are 5.7, 11.8, 12.4, 13.5 and 15.3 . . . 119

Chapter 1

Introduction

1.1 What are Smart Objects?

The quality of human life has been greatly improved by the functionality provided by different objects¹ that serve in multitude of roles, such as security, communication, entertainment, cleaning, cooking, storage, computation etc. Leaps in technological capabilities have transformed the designs of objects over time. The improvements in manufacturing techniques and factory line production, enabled the mass production of objects with mechanically moving parts. The availability of electricity and water, further enabled these objects to be operated in a semi-autonomous fashion. Another transformation was brought about by the advent of embedded computing, which then allowed these devices to be programmed to operate in different modes. The current revolution in communication techniques is allowing these objects to communicate with each other, and learn different contexts about their environment. This allows them to function with greater autonomy and improves the type of functionality that they offer to end users.

In today's context, the term 'Smart Object' is used to refer to any object that incorporates computing and communication in some capacity to enhance the functionality and/or interac-

¹In this dissertation, we use the term 'object' to refer to devices, gadgets, and other entities that people can physically interact with, to fulfill certain tasks and activities.

tion experience of the end user. One definition [4] states that *The concept smart for a smart physical object simply means that it is active, digital, networked, can operate to some extent autonomously, is reconfigurable and has local control of the resources it needs such as energy, data storage, etc.* Another definition [5] states that *The combination of the Internet and emerging technologies such as nearfield communications, real-time localization, and embedded sensors lets us transform everyday objects into Smart Objects that can understand and react to their environment.*

1.2 What is the State of the Art in Smart Objects?

Smart Objects are not just a technological vision for the future. They have arrived, and in fact, have already started pervading the commercial market. Some examples of objects that are currently available commercially and have ‘smartness’ embedded in them are:

- **TVs** have Ethernet ports, or WiFi capability which allow them to stay connected to the Internet. These TVs run a specific type of an operating system which allows them to run apps that can stream different content [6] directly from the Internet. They allow for the interface to be customized so that a user can customize their favorite apps, games, live TV channels to be viewed on the home screen.
- **Thermostats** are equipped with WiFi capability. This allows them to be controlled remotely via smartphones. They can also use motion sensors in the home to detect if a home is occupied, or if the occupants are sleeping, in order to reduce the heating/cooling of the home, and save energy [7]. Some thermostats can even learn the occupancy patterns of the home, and predictively control the HVAC [8].
- **Water Heaters** can operate in multiple modes, including an occupant behavior learning mode in which the system intelligently reacts to a homeowner’s hot water usage patterns, in order to reduce stand-by heat loss [9]. Other modes such as vacation mode reduces

the standby heat loss when occupants are away and the normal mode in which the water heater operates like a static water heater. These are also connected to the Internet, allowing people to control temperature settings remotely from their smartphones.

- **Speakers** connected to WiFi can not only access music content, but also access unlimited computing via the cloud [10]. This boosted computational power makes them capable of voice interaction, music playback, making to-do lists, setting alarms, streaming podcasts, playing audiobooks, and providing weather, traffic and other real time information.
- **Automobile** technology is currently a very swiftly progressing area, with car companies either being in the development or testing phases of autonomous vehicles. Vehicles are already able to park autonomously, change lanes and maintain a constant speed set by the driver. Cars now also feature an attention monitoring system that watches the driver's steering patterns to determine if they are paying sufficient attention while driving [11]. Other cars have sensors that can detect lane markings and keep the car centered within the lane. Radars mounted on the car can detect when other obstacles or vehicles get too close to the car [12]. Automatic lighting control can change the luminosity of headlights by gradually increasing or lowering the light distribution based on the distance of approaching traffic. They also enable parental controls to limit the highest speed at which the car can operate.
- **Eyewear** can now act as fitness trackers using sensors that are embedded into the frame including a gyroscope, accelerometer and magnetometer. The technology is used to monitor steps, distance, calories burned and active time [13]. Other glasses pack in a small heads up display enabling cyclists to glance at a host of useful data in real time including speed, cadence, heart rate. and power zones [14]. Other glasses use bio-sensing technology to detect changes in a wearer's eye and body movements to track and alert on safety, health and fitness [15]. They can track fatigue and alert drivers

who may be about to nod off. EOG electrodes built into the frames and in the nose pads can detect blinks and eye movement in eight directions, and the accelerometer and gyroscope sensors along the earpiece monitor body axis and walking pattern to improve posture.

- **Clothing** augmented with sensors can be used for monitoring heart rate, breathing and movement [16]. The realtime data can be used to provide insights on a range of sporty metrics including intensity and recovery, calories burned, fatigue level and sleep quality. Some workout clothes are woven with micro-EMG sensors that detect which muscles are working and transfer this workout data to a smartphone via Bluetooth [17]. Muscle effort, heart rate and breathing are all tracked and the app provides insights to help exercise correctly and avoid injury. Socks augmented with textile pressure sensors can measure the pressure placed on the foot during running [18]. Infant socks outfitted with pulse oximetry technology can monitor heart rate to make sure the baby is breathing and sleep has been uninterrupted [19].
- **Locks** with Bluetooth and WiFi connectivity allow them to be remotely locked or unlocked using a smartphone [20]. Some other locks allow people to unlock them using voice commands given to a smartphone's digital assistant [21]. Some smart locks also come equipped with touchscreen or other physical interfaces, which allow them to be operated even without the smartphone [22].
- **Electrical Sockets** have access to power lines which makes it easy for them to constantly remain connected to WiFi. People can remotely turn the sockets on or off from their smartphones [23]. These can also be easily programmed to operate on a schedule. Some of them can also track how they are being used, in order to provide a detailed energy usage feedback [24]. These can also be connected to the personal digital assistants, and therefore be controlled using voice commands [25].

- **Lights** can also be controlled remotely using smartphones. People can choose to turn them on/off, dim them, or even change their color [26]. Some lights can even connect to other Smart Objects, such as cars, or locks [27]. One can program the lights to turn on when the smart door lock opens, to turn off when the connected car drives away, and to flash when the security camera detects movement outside.

These are only a few examples of Smart Objects that are available in the market. There are many others which are much better known, such as fitness trackers, smart watches, smartphones, laptops, etc., and therefore these are not discussed in this section. To summarize, most of the body-worn Smart Objects focus on tracking a person's motion and health. While some objects help people remotely control them from smartphones (locks, TV, lights), others allow people to perform certain tasks more easily (vehicles).

1.3 What is Implicit Object User Identification?

Identification is the process of answering the question “Who are you?”. In the information security world, this is analogous to entering a username. Authentication is the process of ensuring that the identity presented is actually true, and answering the question, “How can you prove that its you?”. Typically, systems use a password for this, which is based on some information which is a secret between the user and the system. Entering a username and password is an explicit method of identifying and authenticating a user.

In order for objects to perform personalized or contextual functions, they must solve what we refer as the *Implicit Object User Identification* problem: understanding who is actually using a given object, and being able to validate their identity without expecting people to explicitly participate in an identification procedure.

1.4 How can User Identity Transform Smart Objects?

To support personalized Smart Object behavior, it is important to obtain the identity of the object's user, especially for objects that are shared between people. This is because people have different personalities and preferences, and therefore require different outputs from objects. Therefore when people interact with objects, they often need to re-configure it to suit their needs. For example, while showering, people prefer different temperatures, and need to set the hot and cold water mixer to their preferred configuration every time they take a shower. In addition to configuring objects to their preferences, people also use objects to track their health - such as measure their blood pressure using a portable automatic blood pressure instrument. However, they need to keep track of their own measurements.

The ability to infer user identity is not only transformative for Smart Objects' design in terms of convenience and ease of use, but it can also create new functionality that was previously not possible.

Let us take a look at some examples of how objects could potentially use the user identity information.

- **TV** - People in the same household may like to watch different content on TV. If a TV could know the person using the remote, it could show their preferred content on the home screen, making it easier and faster for them to navigate to their regularly watched shows. Also, a TV could learn the viewing preferences of different individuals, and use that information to suggest new content to them. Additionally, there may be children in the household, and some content could be restricted for them. A TV could automatically set parental controls, when it detects that a user who is designated as a child is using the remote.
- **Touch Lamp** - Touch lamps can begin to learn people's preference for lighting levels as they start to associate user ID with the level of brightness that they set the lamps. After a lamp has sufficient data about a user, the next time that the person touches the

lamp, it can start with their preferred level of lighting, which can be further adjusted if the user wants to.

- **Car** - People have different heights, which makes them adjust the mirror positions, and the car seat's distance from the steering wheel, every time they use a shared vehicle. If a car could detect the person unlocking the door it could customize the seat and mirror positions based on the driver's identity. It could further adjust the height of the car seat and the angle of reclining based on the person's preference. In addition to customization, the car could also monitor the person's driving habits. This can be used in providing timely feedback to a driver to ensure they correct their driving behavior and prevent accidents. Insurance companies could further incentivize their customers by allowing drivers to share their personal driving information with them for discounts to their policies.
- **Lock** - The obvious use case for door locks is to unlock the door if the user is authorized to enter. The user's experience will be just like people opening any regular lockless door: they will simply walk up to a door, and open it without engaging in any explicit identification step. This may seem like a trivial convenience for opening the main door in a home, because only one door needs to be opened. But in many commercial buildings, employees have to scan their ID through multiple doors in order to get access to a secure room. In this case, the convenience offered by this implicit security mechanism will definitely seem more obvious.
- **Whiteboard** - Whiteboards can currently use stylus-based markers, and electronically change ink color on boards with screens. If they could identify the user writing on the whiteboard, they could automatically annotate the sections of the board written by them. The board could also automatically save the notes and share with the users participating in writing the notes. It can allow also users to automatically pull up previously written notes and display them on the board.

1.5 Why is Implicit User Identification Challenging?

User identification is not a new problem, and different solutions have been proposed. An obvious solution to this problem seems to be the use of biometric sensing, such as facial scanning, retina scan, fingerprinting etc. In the current state of art, each of these solutions are either a) not implicit (for example, finger needs to be placed on a specific sensor), or b) robust to different situations (such as different lighting conditions). Another obvious solution seems to be the use of proximity sensing techniques on devices typically carried by users, such as smartphones. Unless all devices are instrumented with beaconing devices, the resolution of location provided by infrastructure-based indoor positioning systems is almost room-size. This coarse granularity can cause confusion when there are multiple users detected in the proximity of the object when it is used. Other techniques try to learn the specific ways in which a person interacts with an object. When the object is used, it observes how it is being used to identify the individual. This technique may be used for monitoring applications where object usage needs to be attributed to the user post-facto. But for cases where the user's identity needs to be established for the user to actually interact with the object, this method fails.

1.6 What Are Wearables, and Why Are They a Promising Solution?

Wearables are a special class of Smart Objects. Any device that is designed to be worn on the human body is referred to as a 'Wearable' device. These devices are not just regular gadgets that are worn on the body, such as headphones. These devices are usually augmented with sensing systems, and have wireless connectivity such as Bluetooth and WiFi. While most Wearables are wrist worn, there are many devices that can be clipped to the body, or hung around the neck. Wearables are also being integrated to jewelry. Implantables are a group

of Wearables that are surgically attached somewhere under a person's skin, for e.g. insulin pumps, or contraception. In the context of this dissertation, we use the term 'Wearables' to refer to wrist worn devices such as smartwatches and fitness trackers.

There are two main reasons why wearable devices are an attractive technological solution that can assist Smart Objects in performing implicit user identification: a. these devices are becoming very popular, and b. they are embedded with sensors that can monitor location context and hand motion of their wearer.

According to a recent survey, as many as 20% Americans are known to own a Wearable device, and the adoption rate is at par with tablets. Some of the reasons why Wearables have been popular is that these have started providing people with a perspective about their habits that were previously unknown or un-quantified to them. Wearables have become especially good at quantifying ambulation and sleep. Another reason why they are popular is that it provides people with a simplified interface to access important information from smartphones.

Wearable are also promising in that they pack a lot of sensing power. These sensing system allow Wearables to monitor biological parameters of the wearer, as well as sense other context about them. Research has shown that the biometric information sensed by these devices can be used to authenticate the user wearing the device. This can be used to authenticate that the wearer of the device is the actual owner of the device. Additionally, Wearables are typically embedded with wireless communication capabilities such as Bluetooth and WiFi. These radios also allow the devices to be located indoors using existing wireless infrastructure in buildings. They also have IMU sensor, which is a combination of accelerometer, gyroscope and magnetometer. The IMU sensor can help sense the motion of the person.

While Wearables are an exciting technological design that is finding an increasing acceptance from the consumers, they are currently not integrated completely with the lives of its users. While physical health is only one measurable attributes, the Wearables can be used to sense many other aspects of a person's daily life, such as the high level activities

that they perform, the energy that they use, etc. In this dissertation, we explore the idea that Wearables can be used in conjunction with Smart Objects to provide a much richer set of functionality to the end user. For Wearables to be most valuable to the consumer, it needs to create a simplified user experience and an easier means to achieve goals. The integration of Wearables and Smart Objects can provide a simplified means to personalize a user's experience with objects on a daily basis.

1.7 What are the Main Challenges in Using Wearables?

Wearables are getting good at measuring certain aspects of people's health and ambulation. They can record heart rate, body fat composition, perspiration, health, temperature and muscle activity by just touching the skin, as well as movement, distance and speed using GPS, and IMU sensors - accelerometers, gyroscopes and magnetometers. However, it is well known that Wearables suffer from imprecision. There are basically two main causes for inaccuracies in their measurements - 1. imprecision caused due to the location of the device, and 2. imprecision caused due to the sensors in the device.

Consumers seem to prefer placing the Wearable devices on the wrist. Since ambulation takes place due to motion in the legs, the motion transferred to the arms, which is sensed by these devices is often not accurate. Also, since people move their arms around a lot - while talking, working with objects, etc., it can result in erroneous measurements.

The errors caused by the sensors is arguably a bigger problem. The component in the devices that takes the measurement is an accelerometer, which records movement in three dimensions: up and down, side to side, and forward and backward. However, it records the rate of speed of motion. This means, to get the distance that a person has moved, one needs to temporally double integrate the acceleration value. Any offset error in the acceleration will cause the velocity calculation error to rise linearly with time. A position calculation will have errors rising quadratically with time. Gyroscopes, or gyros, measure angular velocity, or

speed of rotation around an axis. The gyro bias shows itself after integration as an angular drift, increasing rise linearly over time. Magnetometer measures magnetic field. Earth's magnetic field is relatively small, so measuring it is a pretty delicate process. Magnetic sensors are subject to cross-axis misalignment and temperature-varying scale factors and biases. In addition, the local magnetic field around the magnetometer can be distorted by magnetic and ferrous metal objects.

The sensors in a wearable device that can help locate the position of a person within a space, are the wireless communication components. Wearables are usually equipped with Bluetooth, or WiFi radio, and some have both. While Bluetooth based localization requires additional augmentation of the indoor environment, WiFi leverages the existing infrastructure of access points within a space. The technique of Wi-Fi based indoor localization of a device consists in determining the position of client devices with respect to access points using either received signal strength indication (RSSI), fingerprinting, angle of arrival (AoA), or time of flight (ToF) based techniques. The main limitation is that the WiFi based techniques have a coarse granularity of tracking.

There are inaccuracies in the other sensors used for measuring a person's vital signs. However, we are skipping those details as they are not relevant to the topic in this dissertation.

To summarize, while sensors in Wearables are great for approximate measures of a person's activities, the errors caused due to imprecision of sensors makes them challenging for use in applications that require precision, such as object user identification and authentication.

1.8 Thesis Statement

This dissertation investigates the following hypothesis:

Despite the coarse granularity of its location sensing, and imprecision in sensing hand motion trajectory, data from sensors in wearable devices, when augmented with data from objects, can be used to identify and authenticate users interacting with Smart Objects.

1.9 How can Wearables and Smart Objects Interact to Identify and Authenticate Users?

As discussed in Section 1.4, Smart Objects can provide more personalized, and even new functionalities if they know the identity of the user interacting with them. To this effect, we expect the number of Smart Objects in our daily lives to be so pervasive, that explicit identification will no longer remain a practical solution. Therefore, it is important to design implicit user identification solution.

The consequence of hundreds of Smart Objects surrounding us is that there are several different considerations in the design of the solution. One consideration is that the amount of information a Wearable reveals about itself makes it vulnerable for side channel attacks. Another consideration is that the more the Wearable needs to communicate with each Smart Object, the more energy it consumes. This becomes a serious limitation when a Wearable starts to interact with hundreds of Smart Objects on a daily basis.

In this dissertation, we explore different levels of information shared by a Wearable device with a Smart Object, and explore how that level of data abstraction affects the user identification accuracy. In the first level on interaction, Wearables only reveal their location to the object. The object uses its own location information to figure out the user close enough to use it. However, this information is not sufficient if there are multiple people near the object. In the second level of interaction, the object queries the users nearby to see which of them had a hand motion that most closely resembles the gesture required to interact with the object. Finally, which the information from the first two levels of interaction can be spoofed, by a malicious user, when applications need to authenticate a user, we investigate a third level of interaction in which the Wearable. In this level of interaction, the object request the raw data from the wearable, and correlates its own sensor data with that of the Wearable at a millisecond level.

We make two main assumptions about the system. First, we assume there exists a

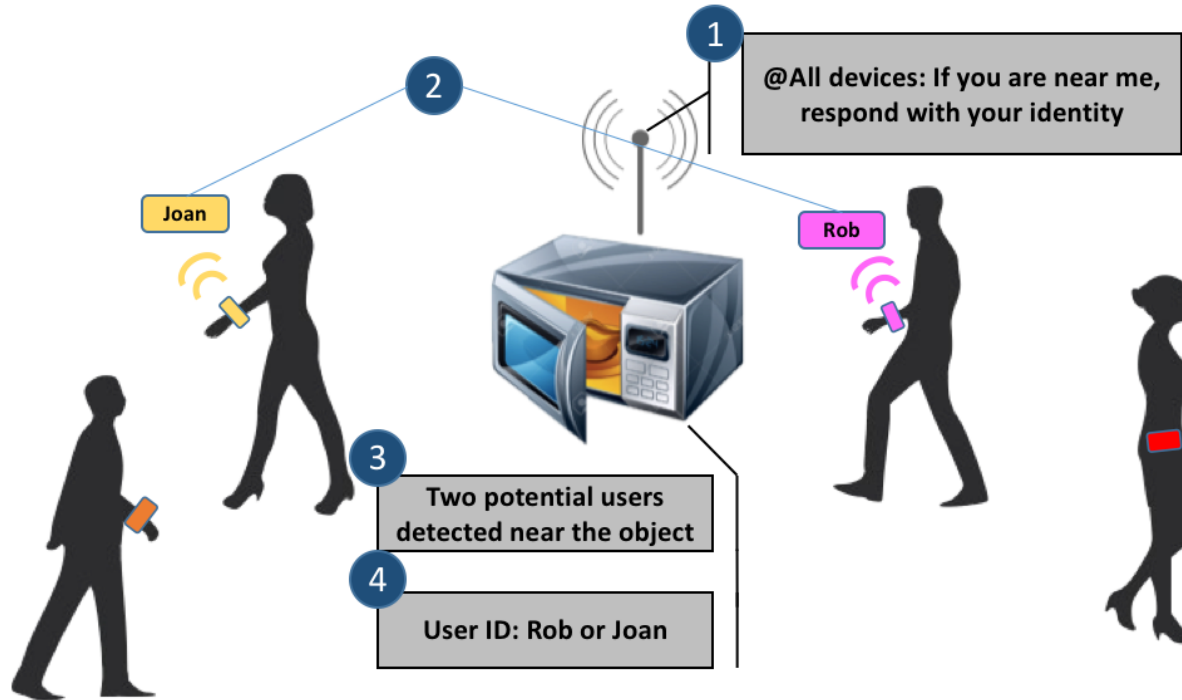


Figure 1.1: First level of interaction - Is to figure potential object users based on proximity. Sometimes only one user may be detected. In this case, the lone user is identified as the object user. This level of interaction is insufficient if there are multiple users.

framework that allows Smart Objects to communicate with Wearables using Bluetooth or WiFi. Second, we assume that objects have some sense of their own location. For example, if the object is within a building, it may know the WiFi fingerprint at its position. If its outside a building, it may know GPS coordinates at its location. We also assume that Wearables have a similar sense of location.

When an object senses its being used, it needs to first identify the set of candidate users who are close enough to physically access it. It broadcasts a message querying the devices that are in proximity of the object to respond with the identity of their user. If there is only one user detected near the object when its used, then clearly the user is also the object user. The main challenge is when there are multiple users detected in vicinity of the object. However, if there is a history of past object usage by the users, one could use heuristics to make an intelligent guess as to who might be more likely to use the object. We explore

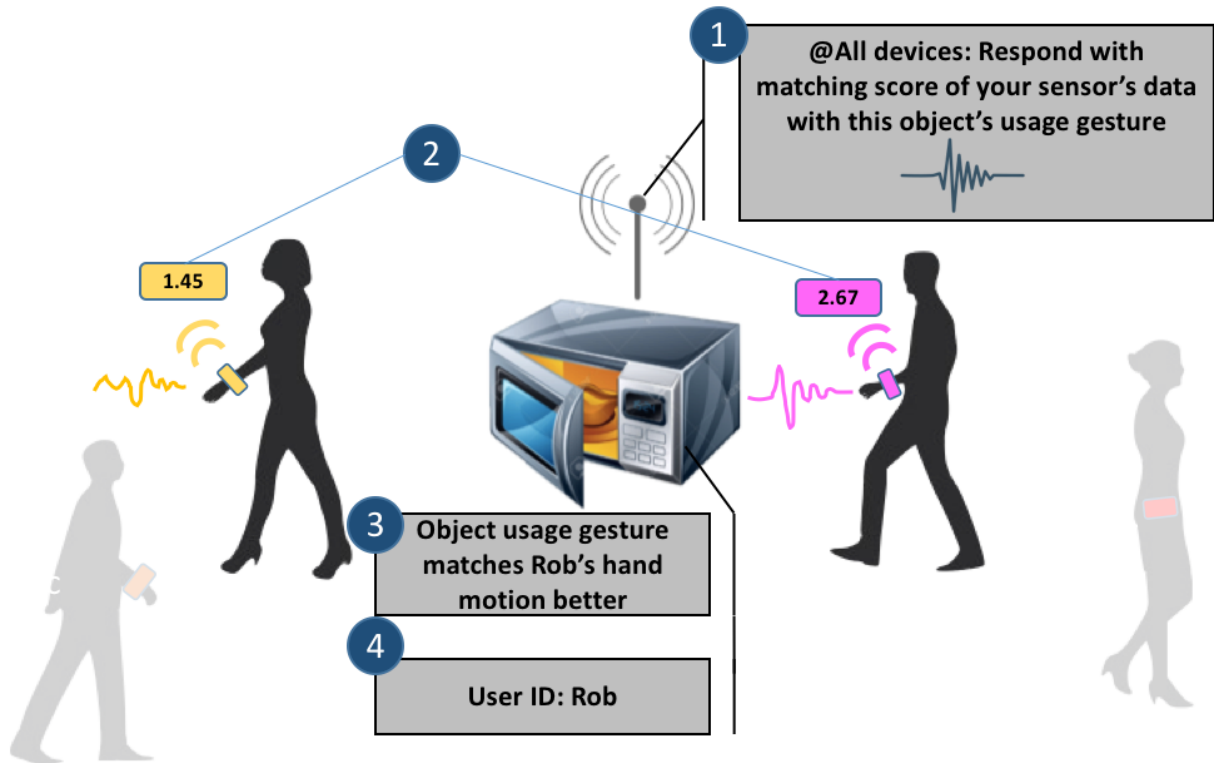


Figure 1.2: Second level of interaction - Is to query the devices nearby for how well their motion data matches the gesture required to interact with the object. This can help disambiguate cases when multiple users are detected near the object. However, this method is vulnerable to being spoofed, as devices may maliciously respond with a higher matching score. Therefore this method cannot be used for sensitive and privacy critical applications.

some heuristics in an in-situ experiment, where we asked five pairs of participants to live in an instrumented home. Although level one currently looks like a query response but could actually be just BLE beacon advertisements from the watch that are passively received by the object. A benefit of this level is that it can be only a one way protocol, i.e. it takes no setup at all, including Bluetooth pairing.

Even though we explore the use of heuristics with location context for cases where multiple people are detected near an object, this method is not robust across different situations. If the object does not have the history of usage of every user it has to choose from, or if certain objects are used equally or randomly by individuals (e.g. lights), then the use of heuristics does not help in guessing who used the object. At this point, the object initiates the next level of interaction with the Wearable.

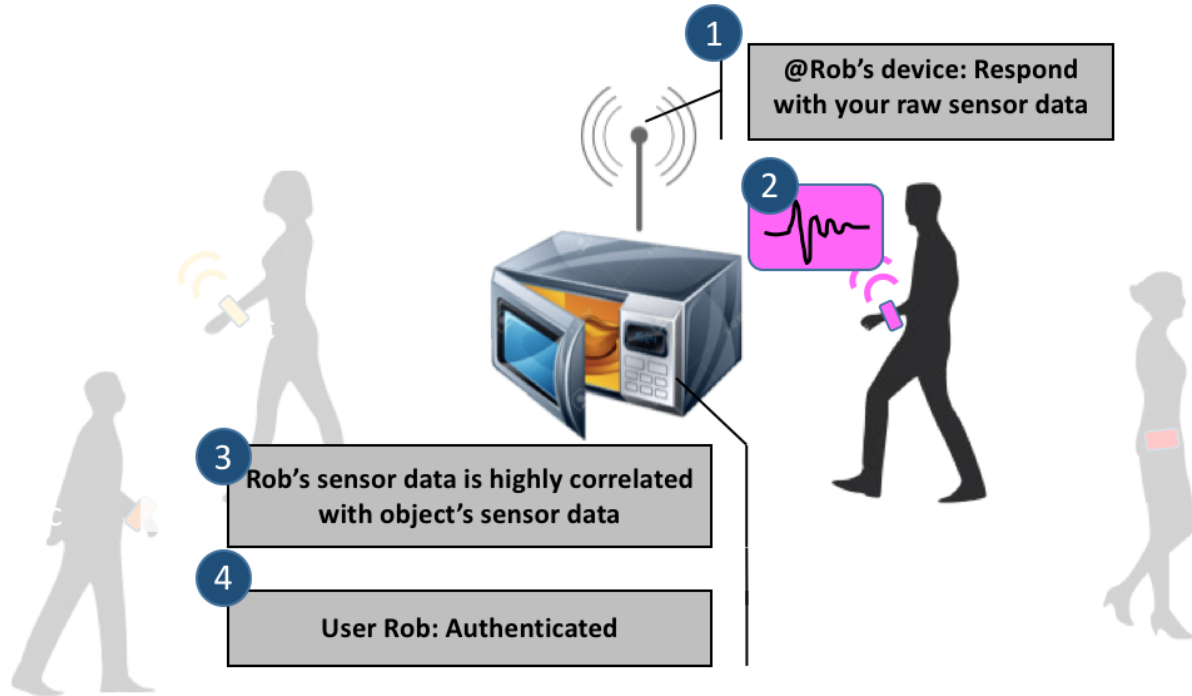


Figure 1.3: Third level of interaction - Is reserved for applications where it is important to validate that a user is the actual object user. In this stage, the identified user is queried for their raw sensor data, which is analyzed and correlated with the object's sensor data at a millisecond level accuracy

Once the Object has a list of candidate users, who could have used in, by virtue of being in proximity to the Object, it queries the candidates again. This time, it asks the devices to match the motion of their wearer's hand to the motion required to interact with the objects. The device that returns the highest matching score is the Object user. In this work, our hypothesis is that objects have unique hallmarks that are imprinted in the hand gestures of its users. By detecting the presence of an object's hallmark in the wrist sensor data, we can identify who used the object. Although it would be very helpful to sense the path of the hand's motion when it moves, the IMU sensor in the Wearables do not sense displacement. Therefore, we characterize the motion of the hand in terms of statistical features about the hand's linear acceleration, orientation, and angle of the hand w.r.t. gravity. We evaluate this concept with a smart home application: recognizing who is using an object or appliance

in a multi-person home by combining smart meter data and Wearables. We conduct three different studies with 10 participants: 1) a study with scripted object use 2) a study with high-level tasked activities and unscripted object use, and 3) a 5-day in-situ study.

Level two of interaction preserves the privacy of the user by not sharing raw IMU data, and is also energy efficient because raw IMU data does not need to be transferred wirelessly. However, one of the main concerns which this approach is that for applications concerning security, privacy or confidentiality of information, this method may not be sufficient. A device can be spoofed to send a higher gesture matching score to get identified as the object user. Therefore, for applications where identity of a user has to be validated, the object initiates a third level of interaction with the Wearable that has been identified as the object user.

In this case, the object asks the Wearable to return the raw IMU sensor data from a time window that extend from before the Object was used to just after it was used. The object computes certain features of the raw data, as well as, correlates it with its own IMU data to authenticate the object user. Level three is more privacy invasive and less energy efficient but allows the object to make authentication decisions, rather than relying on the user. We evaluate this approach with Smartphones, where we intend to authenticate the owner of the phone, vs. other phone users. Offline analyses of our approach using a 13 participant feasibility study, show very promising results. We also developed and tested a prototype with 10 other participants to validate the true acceptance rate of the system. We also conducted a 10 participant adversarial study to determine the false acceptance rate of the system.

1.10 Contributions

The contributions of this dissertation are the following:

1. To evaluate the first level of interaction, using location, we conducted a study with 5 groups of 2 participants each, who lived together for 7-12 days in a test home. We instrumented 39 different fixtures including light fixtures, water fixtures (for e.g. sinks,

showers, and toilets), and electrical appliances (for e.g. the oven, stove, and fridge). Coordinate-level tracking was deployed for 15 of those 39 fixtures.

2. We evaluated different heuristics to learn past object usage behavior in order to perform user identification when multiple people are in proximity of an object. While only 70% of object usage was unambiguous (i.e. only a single person was in the room), we found that user identification could be performed with 87% average accuracy using our heuristics.
3. To evaluate the efficacy of using IMU based features for user identification, we perform multiple studies where participants wore smartwatch on their wrist and used different objects. For this evaluation, we instrumented 16 objects in a home, including as lights, water fixtures, and major appliances. Some of the research questions we intended to answer with these studies are: whether IMU based object usage features are unique and identifiable, if these features are person-independent, and how performing real world tasks affected the user identification accuracy.
4. We performed a scripted study in which 10 study participants performed specific object usage. Our results show that we can classify object hallmark with 95% accuracy when the objects are used in isolation.
5. We asked 5 groups of two individuals (the same 10 participants) to complete a list of 80 real-world tasks within two hours. Despite the high frequency of object usage, we were able to identify the correct individual in 85% of the total 986 object usage events.
6. To evaluate a more realistic scenario in which people spend a large fraction of the day resting or still, we collected 30 hours of in-situ data from a single participant who lived in the test home for 5 days. We used this data to emulate a multi-person home and the results show that our approach can correctly identify the object user in 90% of the total

3378 object usage events in a 2-person scenario and 84% accuracy for 22587 objects in a 3-person scenario.

7. For the user authentication study, we designed algorithms for determining smartphone usage using IMU sensor on phone and a smartwatch paired
8. We conducted a study (with 5000 touch samples spanning 13 individuals) and analysis of features differentiating click vs. other activities
9. We also conducted a prototype use-case study with 10 individuals and analysis of the learnability of the approach
10. We conducted an adversarial analysis with 10 individuals evaluating the false acceptance rate of the approach

1.11 Limitations

There are many limitations in our evaluations:

1. In the study for location based user identification, the population of users recruited for the study were not representative of all possible types of multi-person households, such as those with grandparents. kids or pets living together as well. We were limited by the number of bedrooms in the house. Recruiting participants who were willing to leave their homes, and live somewhere else for 2 weeks was also a big challenge.
2. We were able collect data from each participant for only 7-12 days. This further limited our ability to test more complex heuristics which would include richer features that encapsulate more about people's behavior in using objects.
3. In the studies pertaining to user identification using gestures, we assume that all home occupants who use objects, wear smart wrist devices on their dominant hand.

4. The system might also identify object users incorrectly if a person uses the un-instrumented or non-dominant hand to operate an object.
5. While our approach was tested to work accurately with objects having fixed locations in a home, we have to come up with newer ways to learn the gestures required to operate non-fixed appliances such as hand blenders and hair dryers.
6. In the studies pertaining to user authentication, it would be interesting to evaluate our approach with a larger set of users.
7. It would also be interesting to actually implement a prototype that could actively learn how the owner uses a phone, and use that to authenticate them.
8. Our proposed approach currently does not work in a particular hand position, when the wearable device is worn on the hand that is neither used to hold the phone or touch the phone. Our proposed approach also does not work in the scenario where a phone user is holding the phone in the hand wearing the Wearable, and someone else touches the screen.

1.12 Thesis Outline

The rest of the dissertation is organized as follows:

- Chapter 2 describes some basic applications related to contextual applications using identity, and well as the definitions of terminology which are used in the rest of the dissertation.
- Chapter 3 presents the state-of-the-art in technologies related to the user identification and authentication.
- Chapter 4 describes the evaluation of our user identification based work that uses location.

- Chapter 5 describes the evaluation of our user identification based work that uses IMU based gesture matching.
- Chapter 6 describes the the evaluation of our user authentication based work that uses raw sensor data.
- Chapter 7 concludes the thesis by summarizing the contributions and describing the future work.

Chapter 2

Background

In this section, we describe some of the key applications and contexts that are relevant to this thesis.

2.1 Indoor Positioning System

An indoor positioning system (IPS) is a system that can locate objects or people inside buildings, and is typically achieved using radio waves, ultrasonic, or infrared.

In order to develop smart building systems that respond to its occupants and physical environment, the physical location of objects and people indoors, becomes key information. Sensing position gives an application a wider perspective of the user in the context of the environment, including how they interact with objects and other people. Locating people outdoors is mostly a solved problem, indoor position can be expressed on a spectrum of granularity, depending on the requirement of the application. The ability to position people and objects indoors is crucial for identifying potential object users that are in proximity to an object when it is used.

IPS can roughly be divided into two main thrusts: exact location and presence detection. Finding the exact location of an individual or object implies an inherent knowledge of the layout or floor plan of the indoor space.

Location information on humans is easily obtained from video data. Often motion detection algorithms, such as detecting the change in pixels of a video over time or performing background subtraction, are used to identify people in video data. By understanding where the camera is in the environment and what its viewing range covers, the exact location of a person can be inferred. One of the main problems with this approach is occlusion- when one person blocks the view of another, and applications mistakenly detect only the one occupant. Approaches to counteract this often include using multiple cameras monitoring the same space from different angles [28]. Video-based localization can also detect when a person approaches an object in the environment by identifying and tracking that object. By using targeted microphones, audio information can also be used to localize people in an environment as they speak or otherwise make noise. Additionally, audio and video sensing systems can be combined to increase the accuracy of the detected location [29].

Wearables are very useful in locating the position of people indoors and outdoors. This is because wearables have systems that can transmit and receive information wirelessly, which can be used to locate the device, and therefore its wearer. While wireless based technologies, such as, WiFi, Bluetooth, Infrared, Ultrasound etc. are more commonly explored, other methods such as Pedestrian Dead Reckoning uses inertial sensors in wearable devices. Another way to locate indoors is to determine proximity of people to objects that have a fixed location. This is done typically using RFID and magnetic beaconing methods. While there are methods to track a person indoors using Wi-Fi fingerprinting, and other radio based methods, there are new techniques that automatically learn the characteristics of an indoor floorplan, and can determine where the person is located [30]. The key intuition behind the fusion approach is that certain locations of indoor environment pose a distinct signature for the various sensors on a smartphone. For example, an elevator has a unique accelerometer pattern, corridors may have a distinct set of visible WiFi AP and signal strengths etc.

Some examples of current technologies can be found in [31], which includes systems that sense from the environment, both with and without carried devices. An example of this is

the SmartFloor, where the entire floor of a space is covered with pressure sensors to perform localization anywhere in that space [32]. However, the large install cost makes this and similar solutions too expensive for practical use. Other systems look at a more limited understanding of location, such as the Doorjamb system which localizes a person to a room by monitoring only the doorways in a home with ultrasound sensors to detect when a person moves from room to room [33].

2.2 Activities of Daily Living

An important consequence of being able to implicitly identify object users, is the ability to track their activities. That is because objects are highly indicative of the type of activity that a person is doing. Monitoring the activities of the elderly is important to ensure that they are able to function well independently, and it can also help track changes to their cognitive functionality. Apart from ambulation related activities, the activities that are particularly of interest are the Activities of Daily Living.

Activities of Daily Living or ADLs are activities that people routinely perform without assistance. Some examples of basic ADLs are bathing, eating, continence, dressing, toileting, transferring etc. A person's ability to perform some or most of these ADLs determines the level of care that they need in order to lead a normal life. The assessment that evaluates a person's ability to perform ADLs is referred to as the functional behavioral assessment, and this is typically performed for elderly. There is a high emphasis to support the concept of aging-in-place, where the elderly are encouraged to live in their homes for as long as possible. However, aging-in-place needs adequate supervision and support to ensure the safety of the elderly in their residence. The safety of an elderly person in a home is highly correlated to their ability to perform ADLs. Many elderly people experience problems in ADLs because of illnesses or health-related disabilities. For example, people suffering from heart failure or lung infections may lack the physical stamina to manage household tasks like cleaning,

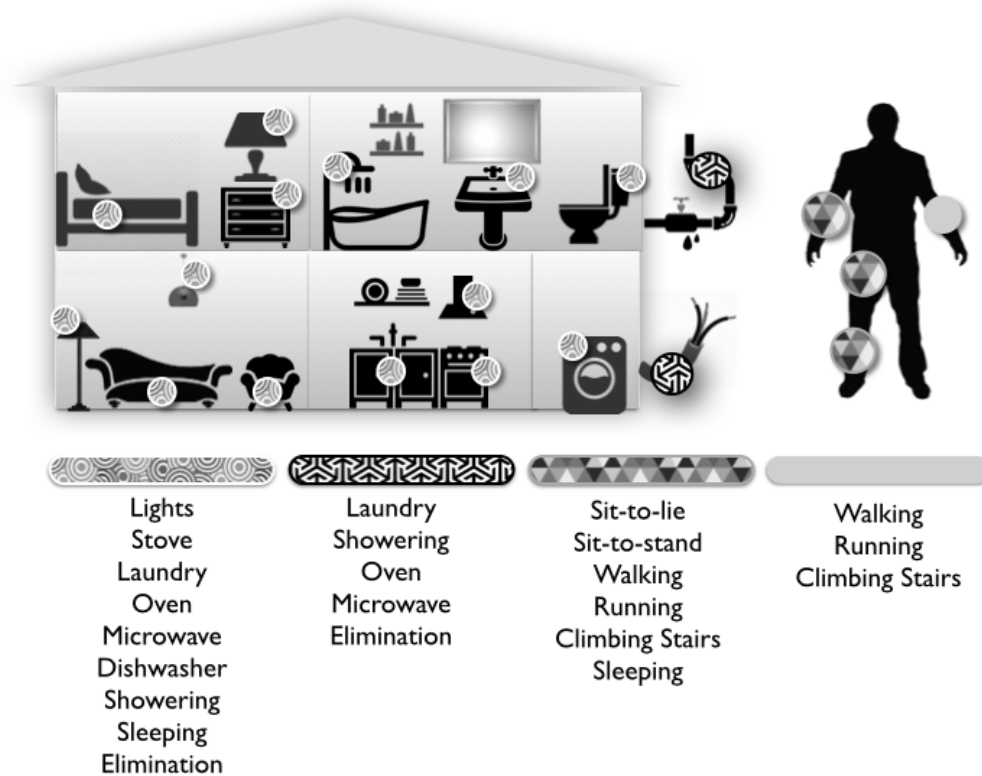


Figure 2.1: While more number of sensors can detect a richer set of ADLs, it is more practical in terms of cost and maintenance to have systems that can infer ADLs from fewer sensing points (figure from [1])

cooking, and laundry on their own. A functional assessment can identify if an elderly person needs outside help, such as home care, moving to an assisted living facility, etc.

The two main subcategories of ADLs are Basic ADLs, such as ‘Dressing’ which refer to the ability of a person to live within a home, and Instrumental ADLs such as ‘Driving’, which refer to a person’s ability to live within a community. However, within these subcategories, each clinical scale has its own set of labels, some of which may specify activities very specifically such as ‘brushing teeth’, and others just refer to broad categories, such as ‘grooming’. The instruments of functional assessment intend to track the ability to perform certain basic and instrumental ADLs.

Changes in functional abilities of an elderly person are usually the manifestation of the changes in the cognitive and physical abilities of the elderly, such as changes in how well the elderly are able to perform tasks such as meal preparation, grooming, ease of transferring

between bed and chair etc. Since these tasks are performed in a routine manner by the individuals on a regular basis, they may not notice the changes in their own abilities. Therefore it becomes important to monitor the functional status of an individual.

2.3 Wearable Devices

The terms ‘Wearable Devices’, ‘Wearable Technology’ and ‘Wearables’ are all synonymous. Wearable devices are a category of technology that include all gadgets that are meant to be worn on the body as part of the clothing or other accessories. Wearable devices come with computing power comparable to smartphones. Typically, wearable devices are optimized to perform certain sensor based tasks, such as tracking sleep. These devices offer an unique opportunity to sense its user at a physical level unlike that offered by smartphones or laptops.

Generally, wearable technology will have some form of communications capability and will allow the wearer access to information in real time. Data-input capabilities are also a feature of such devices, as is local storage. Examples of wearable devices include watches, glasses, contact lenses, e-textiles and smart fabrics, headbands, beanies and caps, jewelry such as rings, bracelets, and hearing aid-like devices that are designed to look like earrings. There are also more invasive versions of the concept as in the case of implanted devices such as micro-chips or even smart tattoos.

One of the most common uses of wearable devices is in health and fitness tracking. The goal of the fitness devices is to monitor certain aspects of the person’s health or activities throughout the day. Wearables technology also promises great potential in gaming and entertainment. Augmented reality and wearable technology can combine to create a much more realistic and immersive environment in real time.



Figure 2.2: Wearable devices come in many different form factors, although currently wrist worn devices such as smartwatches and fitness trackers are leading in the Wearables market

2.4 Inertial Measurement Unit

One of the key motion sensing systems used in Wearables is an IMU sensor. An Inertial Measurement Unit or IMU is a sensor that integrates three different sensors that can measure acceleration, angular velocity and magnetic forces. The accelerometer in the IMU is used to sense both static (e.g. gravity) and dynamic (e.g. sudden starts/stops) acceleration. Because they are affected by the acceleration of gravity, an accelerometer can tell how a device is oriented with respect to the Earth's surface. An accelerometer can also be used to sense motion. Finally, an accelerometer can also be used to sense if a device is in a state of free fall. Gyroscopes measure angular velocity, how fast something is spinning about an axis. Unlike accelerometers, gyros are not affected by gravity, so they make a great complement to each other. In an IMU gyroscopes are used alongside accelerometers for applications like motion-capture and vehicle navigation. IMUs also contain magnetometers, which can measure the magnetic force along a certain axis. This allows better performance for dynamic orientation calculation in attitude and heading reference systems which base on IMUs.



Figure 2.3: In the PowerHouse, the game player manages a simulated domestic environment with seven characters. The meters in the lower pane displays a specific characters mental and physical state. In the upper right corner are the money and power meters that show the accumulated points and how much energy is being consumed.

2.5 Personal Energy Feedback

Although not directly researchers developed a prototype game to influence a set of target activities in the home using several persuasive techniques [34]. Employing the format of a reality TV show (docu soap), the game informs implicitly and explicitly about various energy-efficient actions. Figure 3.2 shows a screenshot of this game. Another game based approach used a virtual energy pet [35]. With the use of a pet avatar, learners were encouraged to use home energy conservatively in a playful and engaging way. The aim is not simply to teach learners how to reduce energy use, but to engage them in adopting appropriate energy conservation measures. An empirical study was conducted to examine if the system could promote learners' an understanding of energy conservation. The results demonstrated that the system significantly promoted learners' self-awareness, learning motivation, as well as

willingness to conserve energy.

Energy usage feedback in residential buildings is considered an effective means to induce energy saving behavior in home residents [36]. The concept of energy usage at home disaggregated at fixture level, has been incorporated in many eco-feedback systems, such as the Aware-Cord [37], which displays fixture level ambient energy usage, and WattBot [38] which is a mobile phone based fixture and room level home energy feedback system. Researchers developed a prototype to promote water conservation in the shower [39]. The prototype consists of a flow meter, a microcontroller, and LEDs. The LEDs visualize an imaginary water level that would rise with the continuous water flow, if the drain was closed. One LED lights up for every five liters.

Another study compared the effect of numerical and abstract water usage feedback [40]. Sensors first logged baseline water usage without visualization. Then, two display styles, ambient and numeric, were deployed in random order, each showing individual and average water consumption. Quantitative data along with participants feedback contrast the effectiveness of numeric displays against abstract visualization in this very important domain of water conservation and public health.

Chapter 3

Related Work

This chapter presents the state of the art in technologies related to the problem of implicit object user identification and authentication. In Section 3.1, we discuss technologies that can be used to perform indoor localization. In Section 3.2, we discuss some of the existing technologies that can be used to perform object user identification. In Section 2.5, we discuss some of the work that has in been done in disaggregating energy used by electrical objects, which is an application of object user identification. And finally, in Section 3.3 we discuss user authentication techniques proposed for smart objects.

3.1 Indoor Positioning System

Classic work in this area includes the Cricket [41] and Active Bat systems [42]. The Cricket location support system leverages a group of ultrasonic beacons installed throughout the building. Applications running on mobile and static nodes learn their physical location by using listeners that hear and analyze information from these beacons. This system enables room-size granularity in terms of location accuracy. Cricket uses a combination of RF and ultrasound to provide a location-support service to users and applications. For every beacon, RF advertisements with location information are broadcasted at the same time as ultrasonic pulse. The listeners receive these RF and ultrasonic signals, correlate them to each other,

and infer the space they are currently in. The active bat based system works in the reverse manner. Receivers are installed throughout the ceilings of buildings, while RF transmitters are attached to objects that we want to track. Multilateration based techniques are then applied to determine the location of the objects.

The DOLPHIN system consists of distributed wireless sensor nodes which are capable of sending and receiving RF and ultrasonic signals [43]. These nodes are attached to various indoor objects. Using a distributed positioning algorithm in the nodes, DOLPHIN enables autonomous positioning of the objects with minimal manual configuration. Another proposed positioning system called COMPASS, was based on 802.11-compliant network infrastructure and digital compasses [44]. On the mobile device, COMPASS samples the signal strength values of different access points in its communication range and utilizes the orientation of the user to preselect a subset of the training data. The remaining training data is used by a probabilistic positioning algorithm to determine the position of the user.

SpotFi is a technique that only uses information that is already exposed by WiFi chips [45]. First, SpotFi incorporates super-resolution algorithms that can accurately compute the angle of arrival (AoA) of multi-path components even when the access point (AP) has only three antennas. Second, it incorporates novel filtering and estimation techniques to identify AoA of direct path between the localization target and AP by assigning values for each path depending on how likely the particular path is the direct path.

Horus is an RF-based location determination system that uses the signal strength observed for frames transmitted by the access points to infer the user location [46]. The Horus system identifies different causes for the wireless channel variations and addresses them to achieve its high accuracy. It uses location-clustering techniques to reduce the computational requirements of the algorithm.

WiFi based distance estimation uses received signal strength indicator (RSSI) is easily affected by the temporal and spatial variance due to the multi-path effect, which contributes to most of the estimation errors in current systems. Researchers explore the frequency

diversity of the subcarriers in OFDM systems and propose a novel approach called FILA [47], which leverages the channel state information (CSI) to alleviate multi-path effect at the receiver

While most WiFi fingerprinting based localization techniques require a training set of signal strength readings labeled against a ground truth location map, Ferris et. al propose a novel technique for solving the WiFi SLAM problem using the Gaussian Process Latent Variable Model (GPLVM) to determine the latent-space locations of unlabeled signal strength data [48]. They demonstrate that GPLVM, in combination with an appropriate motion dynamics model, can be used to reconstruct a topological connectivity graph from a signal strength sequence which, in combination with the learned Gaussian Process signal strength model, can be used to perform efficient localization.

A research prototype, CUPID, utilizes physical layer (PHY) information to extract the signal strength and the angle of only the direct path, thereby avoiding the effect of multipath reflections [49]. This was based on the observation that natural human mobility, when combined with PHY layer information, can help in accurately estimating the angle and distance of a mobile device from an wireless access point (AP).

Researchers proposed UnLoc, an unsupervised indoor localization scheme that certain locations in an indoor environment present identifiable signatures on one or more sensing dimensions [30]. They speculate that these kind of signatures naturally exist in the environment, and can be envisioned as internal landmarks of a building. Mobile devices that “sense” these landmarks can recalibrate their locations, while dead-reckoning schemes can track them between landmarks.

Another calibration free technique was proposed by Chintalapudi et. al [50]. The mobile devices record Received Signal Strength (RSS) measurements corresponding to APs in their view at various (unknown) locations and report these to a localization server. Occasionally, a mobile device also obtains and reports a location fix, for e.g. by obtaining a GPS lock at the entrance or near a window. The key intuition is that all of the observations reported to the

server, even the many from unknown locations, are constrained by the physics of wireless propagation. They model these constraints and then use a genetic algorithm to solve them.

Zee is another system that uses crowdsourcing of location-annotated WiFi measurements in indoor spaces, using the mobile phones carried by users [51]. A key attribute of crowdsourcing with Zee is that it does not require any active user intervention in terms of location input, placement of the phone, or other aspect. Zee employs a set of techniques to resolve ambiguity in location during crowdsourcing, using inertial and WiFi measurements, and a map of the indoor space as the inputs. The data thus gathered can help train existing WiFi-based localization algorithms.

Yang et al investigated sensors integrated in modern mobile phones and leveraged user motions to construct the radio map of a floor plan [52]. In their proposed approach, the calibration of fingerprints is crowdsourced and automatic.

Another technique used algorithms for reliable detection of steps and heading directions, and accurate estimation and personalization of step length [53]. Their indoor localization system was comprised of several modules, the step detector, step length estimator, heading estimator, particle filter, and a personalization module for adapting a step model to an individual user.

Ravi et al proposed that a smart phone be worn by the user as a pendant and images periodically captured and transmitted over GPRS to a web server [54]. The web server returns the location of the user by comparing the received images with images stored in a database. As opposed to earlier solutions for indoor localization, this approach did not have any infrastructure requirements. However, it did have the cost of building an image database.

Some other researchers propose an audio based technique for determining a mobile phone's indoor location even when Wi-Fi infrastructure is unavailable or sparse [55]. Their technique was based on ambient sound fingerprint called the Acoustic Background Spectrum (ABS). They hypothesized that ABS would serve well as a room fingerprint because it is compact, easily computable, robust to transient sounds, and distinctive. An experiment conducted

involving 33 rooms, yielded 69% correct fingerprint matches meaning that, in the majority of observations, the fingerprint was closer to a previous visit fingerprint than to any fingerprints from the other 32 rooms.

Another audio based technique uses off-the-shelf audio speakers (already in place) to provide fine-grained indoor position data to modern mobile devices like smartphones and tablets [56]. The mobile devices can localize themselves by applying a Time-Difference-of-Arrival (TDOA) pseudo-ranging technique. The main contribution of their work was a modulation mechanism that sends data utilizing ultrasonic frequencies in such a manner that is imperceptible to humans.

3.2 Object User Identification

One of the applications for performing object user identification is to be able to attribute energy usage to every person.

Hay et al explored breaking down electricity consumption for occupants of a commercial building working at different shifts in the day [57]. They used access logs to infer the occupancy of the building, and explored the usage of different policies to allocate shared resource usage. The static apportionment policy divides the entire energy consumption equally among all the people who have desks allocated in the building, regardless of how much they use the building. Dynamic occupants policy divides a building's instantaneous power consumption only amongst the current occupants of the building. These policies had to share the electricity equally between building occupants because the location information of the employees was only at the building level, and therefore the ambiguity of object user's identity extended to everyone in the building who could have access to the object.

Hay et al present their concept of a personal energy meter (PEM) which can record and apportion an individual's energy usage [58]. Their concept not only includes a person's energy footprint inside a home, but also the energy footprint outside of home - such as that of

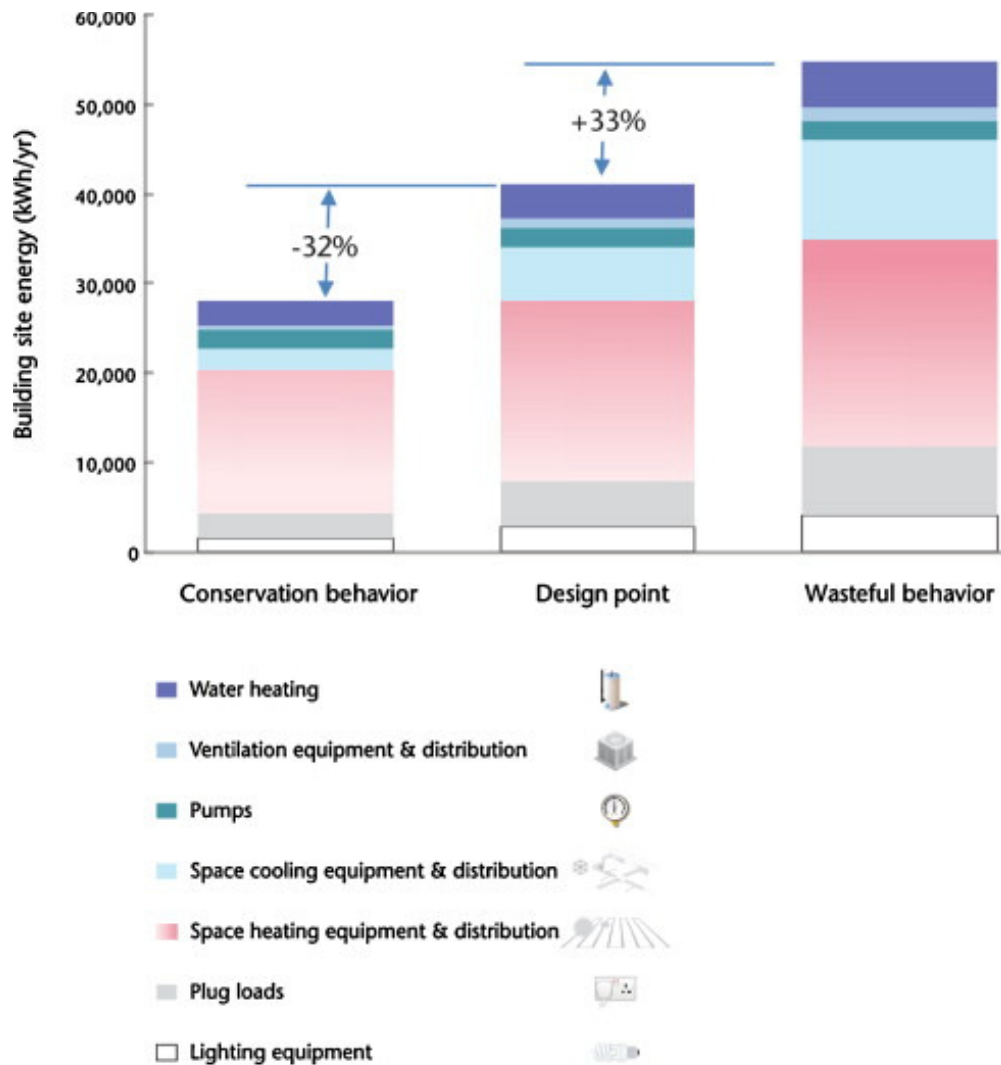


Figure 3.1: Energy-unaware behavior uses twice as much energy as the minimum that can be achieved (figure from [2])

transportation. To pursue this concept, Hsu et. al conducted a study with a mobile application that provides estimations of personal electricity consumption in a research laboratory [59]. In this study, metadata about each appliance owned by a user such as computer, printer and phone was stored in a central repository, and the energy consumption for these appliances tracked using their proposed system architecture. The total energy consumed everyday by all these devices is apportioned and displayed on the user's smartphone. They propose that this method gives users the alternative to operate specific appliances which are more energy efficient. Here the assumption is that personal objects are only used by their owner, and

therefore by default the object user's identity is set to be their owner.

Cheng et al use coordinate level tracking in a study and apportion energy usage for users in an office building [60]. To determine which person is using an appliance, they use a proximity detection system where users carry a magnetic beaconing system which are detected by special receivers near appliances. The disaggregation policy is simple - assign the energy used to the nearest person detected.

Lee et al propose tracking people's movements in the rooms of a house to disaggregate electrical energy [61]. They use people's most frequent movements in the rooms of a home to identify individual bedrooms, and assign all electrical usage in those rooms to the person. In a shared space, they use room location tracking to identify the possible person present in the room while an electrical appliance is being used. If multiple people are present in the room when an electrical fixture is used, they simply split the energy usage equally between all people present in the room.

3.3 Object User Authentication

SilentSense is a technique that leverages the user touch behavior biometrics and integrated sensors to capture the micro-movement of the device caused by users screen-touch actions [62]. By tracking the fine-detailed touch actions of the user, they build a touch-based biometrics model of the owner, and then verify whether the current user is the owner or guest/attacker.

Fiberio [63] proposes a rear-projected multitouch table that identifies users biometrically based on their fingerprints during each touch interaction using a large fiber optic plate that diffuses light on transmission, thereby allowing it to act as projection surface. This work was challenging as the surface had to fulfill two contradicting requirements with respect to the screen material. On the one hand, the screen had to reveal fingerprints, i.e., produce contrast between the ridges and valleys of the fingerprint. On the other hand, to be used as a display, the screen had to allow the rear-projection to produce a visible image, which requires

the screen material to be diffuse. Unfortunately, specular and diffuse are contradictory requirements for such a surface. Fiberio used a 3mm thick, 4233dpi fiber optic plate which (1) diffuses light on transmission, and (2) the correct illumination setup, the fiber optic plate creates a specific type of specular reflection.

Capacitive fingerprinting [64] proposes a novel sensing approach based on Swept Frequency Capacitive Sensing, which measures the impedance of a user to the environment across a range of AC frequencies. Different people have different impedance profiles, based on which they can be identified. Different impedance profiles can be attributed to the fact that every human body has varying levels of bone density, muscle mass, blood volume and a plethora of other biological and anatomical factors. Also, users wear different shoes and naturally assume different postures, which alters how a user is grounded. As the signal flows through the body, the signal amplitude and phase change differently at different frequencies. These changes can be measured in real time and used to build a frequency-to-impedance profile

The use of a trusted smart glove [65] was proposed where individual fingers and parts of a glove are tagged with fiducial tags. Fiducial tags are 2x2 reflective markers with a unique 8-bit identification pattern encoded. A tag-aware surface, such as Microsoft Surface, tracks the location and orientation of these tags, while identifying the wearer of the glove. The glove allows multiple aspects of the hand's posture, finger position, and gestures to be recognized in order to make interactions with objects more intuitive.

PhoneTouch [3] is a technique which uses phones to interact on a touch surface. The phone is used to select targets on the touch surface, and then it can be used to move the target on the screen, by dragging the phone along the screen. PhoneTouch is based on separate detection of phone touch events to identify multi-user interaction on a surface, which determines location of the touch, and by the phone, which contributes device identity. The device-level observations are merged based on correlation in time. The main system design involves all the devices independently detecting the touch events. Every time that a device detects an event, it transmits that information to a server. The individual surface and phone

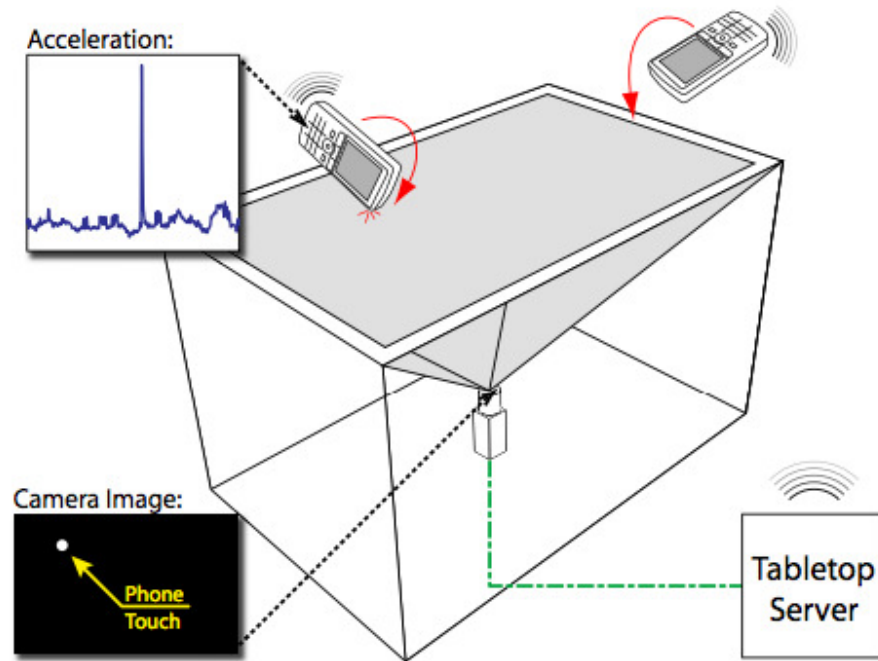


Figure 3.2: In PhoneTouch [3], surface and phones detect touch events independently. The device-level observations are communicated over a wireless network, correlated in time, and combined to associate a touch with both surface position and phone identity

events are matched based on their time-stamps, in order to determine PhoneTouch events.

YouTouch [66] proposes the use of a commodity RGB + depth camera in front of the wall, to track users on a common collaborative surface and correlate them with touch events. Whenever the user's ID is occluded from the camera, and the tracker loses the user's ID, they propose re-identifying the person by means of a color histograms of body parts and skeleton-based biometric measurements. The re-identification component of their work is most challenging, as re-identification can be hindered by factors such as varying lighting conditions, camera color calibration and the need to work with low-resolution images containing significant clutter. They use person descriptors which consist of anthropometric features (human biometric measurements such as height calculated from the skeleton data) and color features (histograms of person-specific image regions such as the torso).

Some proposed methods attempt to use the phone's app usage and calling behavior to authenticate users [67, 68]. For example, [68] look at the movement a person performs,

from the moment they press ‘start’ (to initiate a call), until they take the phone to the ear. They treat this movement as a biometric feature, and demonstrate that there are differences between different users, so that the movement can effectively be used for identification. In this way, they hypothesize that as soon as a call is answered (or placed), the phone can promptly evaluate if the user is authorized to perform this action, and block the system in case of unauthorized users.

Another thrust in implicit authentication attempts to use gait based features to authenticate users as they are walking [69, 70]. [70] normalize the accelerometer data from the phone carried by the person and then process it using correlation, frequency domain methods and data distribution statistics. Correlation is used because of the assumption that the shape of the signal during gait is unique for every person. Frequency domain based methods are used because of the assumption that there is a characteristic distribution of frequency components for each person in the walking signal. Data distribution statistics is used because of the assumption that characteristics of the signal shape affect on the data distribution.

Chapter 4

Level 1 - Using Location to Identify Object Users

4.1 Introduction

Location in an indoor space is a fundamental context around which many other concepts revolve: activities, furniture, heating, lighting, and noise. Therefore location is closely tied to the type of objects - furniture, devices, lights etc., that may be present at different parts of a building. Information about the location of objects and building residents can be used to perform object user identification based on the intuition that when an object is used, the person closest to the object must be the object user.

When an object senses its being used, it needs to first identify the set of candidate users who are close enough to physically access it. It broadcasts a message querying the devices that are in proximity of the object to respond with the identity of their user. If there is only one user detected near the object when its used, then clearly the user is also the object user. The main challenge is when there are multiple users detected in vicinity of the object. However, if there is a history of past object usage by the users, one could use heuristics to make an intelligent guess as to who might be more likely to use the object.

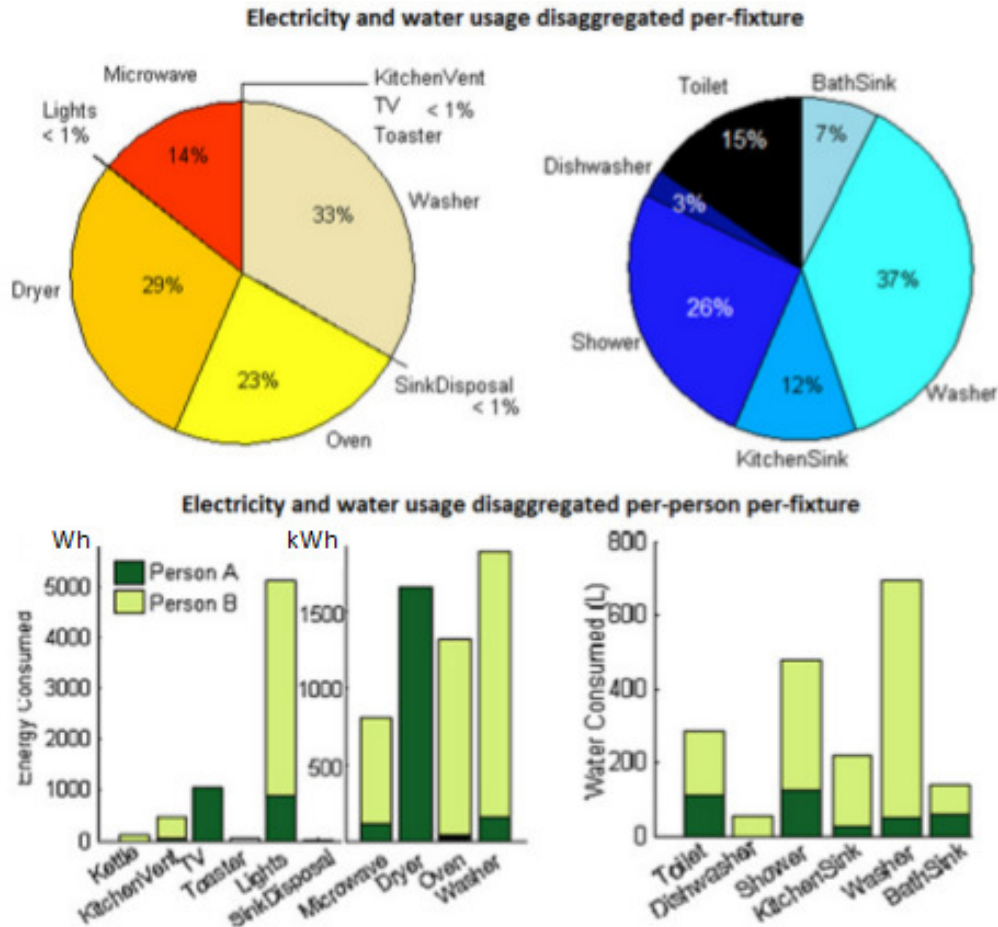


Figure 4.1: Energy disaggregation (of one of the participant groups): Per-person per-fixture disaggregation makes it evident that Person B consumes most of the electrical and water fixture energy in the house except for TV and Dryer, which are used only by Person A

To determine the location of people within buildings, because of the lack of reliable GPS signals inside a building, researchers have explored the use of WiFi beacons, magnetometer, vision, or ultrasound etc. Most technology can locate people with a room size granularity within the space. Since wearable devices have WiFi and/or Bluetooth radio, they can be used to locate their wearer with a room level granularity as well.

One of the key applications where location has been used to perform object user identification, is energy disaggregation at individual level. It is believed a system that disaggregates a home's total energy usage at a *per-person*, *per-object* level, will allow individuals to discover and reduce the energy footprint of their object usage behavior in homes [58].

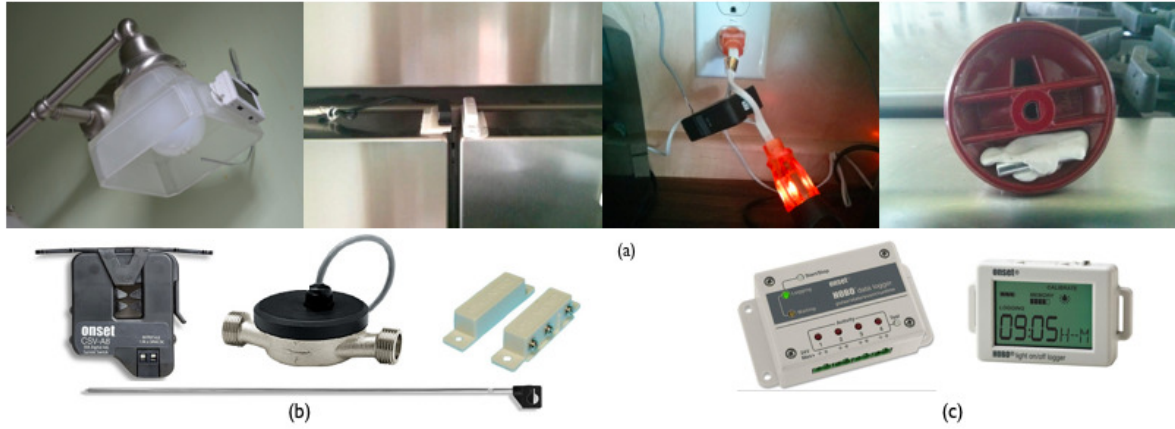


Figure 4.2: Sensors Deployed: (a) All major electrical and water fixtures in the test home were instrumented, including lights (with light sensor), refrigerator (with reed switch on door), toaster (with power meter), and stove (with reed switch inside knob). (b) Sensors included power meters for plugs, in-line water flow meters, door open/close sensors, and fiber optic light pipes attached to light sensors. (c) Onset’s data loggers, such as 4-channel pulse data logger were used

In this work, with sufficient object usage history we demonstrate that we can perform object user identification based on coarse-grained room-level location tracking – even when two people are present in the room where an object is used. Many WiFi fingerprinting based, cellphone, and emerging Bluetooth Low Energy (BLE) systems, such as the Apple’s iBeacon system can already achieve room-level location tracking. Our results offer promise that wearables based coarse-grained location systems can be used to perform user identification, even if fine-grained coordinate level indoor location tracking remains impractical and expensive in the current state of the art.

Our basic approach to user identification is to first find unambiguous object usages, where an object is used and only one person is detected in the same room. These unambiguous cases are then used to learn object usage patterns that can help disambiguate cases where two or more people are in the room. The patterns can be general trends (e.g. Person A does all the cooking) or can be similar to object-use fingerprints [71] (e.g Person A likes long, hot showers) and can even include multi-object fingerprints (e.g. Person A uses the sink often while cooking).

Recent advances in sensing technology have been able to disaggregate the coarse-grained aggregate energy consumption in homes, to an object level; it is possible to discover which objects are present in the home [72], recognize when they are used [73, 74], and discern how much energy or water was consumed during each use [75, 76]. However, unless every object is instrumented with an RFID tag [77] or a biometric sensor [78], the objects themselves can not detect who is using them. By using location information along with the knowledge that a particular object has been used, we can identify the person using the object.

In this work, we also apply the user identity information to electrical and water object usage data to demonstrate a few case studies of how the disaggregated information can be useful in monitoring energy usage and object usage habits of home residents.

To test our approach, we performed a study with 5 groups of 2 participants each, who lived together for 7-12 days in a test home. We equipped the test home with an RFID-based tracking system that tracked the location of each participant with three different granularities: house-level (home vs. not home), room-level (kitchen vs. bathroom), and coordinate-level (at the oven vs. at the fridge). We also instrumented 39 different objects including light fixtures, water fixtures (for e.g. sinks, showers, and toilets), and electrical appliances (for e.g. the oven, stove, and fridge). Coordinate-level tracking was deployed for 15 of those 39 fixtures. Participants were required to wear RFID ankle bracelets [79] for up to 12 days, and were otherwise asked to follow their normal daily routines.

While only 70% of object usage was unambiguous (i.e. only a single person was in the room), we found that object user identification could be performed with 87% accuracy on average by using our learning algorithms. In comparison, user identification based on house-level tracking achieved only 12% accuracy due to objects being used when multiple people are home at the same time. Coordinate-level tracking achieved user identification accuracy of 97%. In 3% of cases, two people were standing immediately next to the object when it was used. Thus, user identification performed with room-level tracking had comparable accuracy as coordinate-level tracking, but at potentially much lower cost.

4.2 Related Work

Hay et al explored using building level location information for breaking down electricity consumption for occupants of a commercial building working at different shifts in the day [57]. They used access logs to infer the occupancy of the building, and explored the usage of different policies to allocate shared resource usage since the building level location information was too vague to be used to pinpoint the exact object user. The static apportionment policy divides the entire energy consumption equally among all the people who have desks allocated in the building, regardless of how much they use the building. Dynamic occupants policy divides a building's instantaneous power consumption amongst only the current occupants of the building. For this, they estimated the number of people in the building using the entry and exit access logs of the building. But this approach penalizes the people who are present in the building at odd hours, such as early morning, as the entire building's energy load including the static energy consumption is divided only between a few people. They envisioned a personal load policy in which they hope to track a person's individual withdrawal from the energy bank in terms of the energy consumed by switching on the lights and computer monitors when they enter the building. Keeping in mind that they used building level occupancy data, in our evaluation we compare this level of granularity with the room level and coordinate level tracking systems.

While these general policies might work for a commercial buildings where a large percentage of the building's energy consuming systems, such as the heating and ventilation system are equally shared by its occupants, homes are different. In homes, individuals are in control of the all fixtures and appliances that contribute to the total home energy consumption. Therefore, the need for discerning who is using a fixture becomes very relevant in homes. In this direction, Hay et al. present their concept of a personal energy meter (PEM) which can record and apportion an individual's energy usage [58]. Their concept not only includes a person's energy footprint inside a home, but also the energy footprint outside of home - such as that of transportation. Discovering who uses which electrical/water fixture becomes an

essential step toward apportioning total energy between occupants.

Hsu et. al conducted a study with a mobile application that provides estimations of personal electricity consumption in a research laboratory [59]. In this study, metadata about each appliance owned by a user such as computer, printer and phone was stored in a central repository, and the energy consumption for these appliances tracked using their proposed system architecture. The total energy consumed everyday by all these devices is apportioned and displayed on the user's smartphone. They propose that this method gives users the alternative to operate specific appliances which are more energy efficient. Here the assumption is that the owner of the object is also the object user. This approach does not apply to devices which are commonly shared between home occupants, in which case such an assumption does not apply.

Cheng et. al use coordinate level tracking in a study and apportion energy usage for users in an office building [60]. To determine which person is using an appliance, they use a proximity detection system where users carry a magnetic beaconing system which are detected by special receivers near appliances. The disaggregation policy is simple - assign the energy used to the nearest person detected. We recognize that coordinate level localization is very accurate in determining the individual who used an appliance. However, it is also an expensive technique which requires carrying specialized wearables at home, in the current state of the art. We explore the idea that we can use a more coarse-grained room level localization system and achieve comparable results. Room location tracking in homes can be done much cheaper by utilizing the existing technology infrastructure at homes, such as WiFi fingerprinting.

Lee et al. propose tracking people's movements in the rooms of a house to disaggregate electrical energy [61]. They use people's most frequent movements in the rooms of a home to identify individual bedrooms, and assign all electrical usage in those rooms to the person. In a shared space, they use room location tracking to identify possible person present in the room while an electrical appliance is being used. If multiple people are present in the room when an

electrical fixture is used, they simply split the energy usage equally between all people present in the room. Our work differs from this study, in that, even in the presence of multiple people in a room, we attempt to guess the exact individual who is actually responsible for a fixture event. This does not imply that all of them are using the appliance. In our study we observed that people spent up to 60% of their time together in the same room. The energy used when people are in the same Dividing all the energy usage equally masks individual contributions to energy usage and makes it harder for individuals to recognize energy impact of their individual fixture usage. We explore learning based policies of fixture usage assignment that intend to use history of appliance usage to make an intelligent guess about which individual might be using the appliance.

4.3 Approach

We aim to compare the performance of room level granularity of location tracking systems (wearables) with coordinate level and home level granularity, to identify the people using electrical and water objects in a home. We fuse location information of individuals with electrical and water object usage events, to assign object events to residents within a home. The main challenge is disambiguating who used a object. This occurs when two people are detected within the same proximity of a object at the time of its use. We test the use of different heuristics to learn past object usage behavior in order to disambiguate such events.

4.3.1 Step 1 - Object Usage Detection

In the scope of this paper, we assume that we have prior knowledge about a object's presence, identity and location within a home. All the object sensors in the house, report object usage status at a second's granularity. For every electrical object sensor i in the system that produces a time series $E^i = e_1^i, e_2^i, e_3^i, \dots, e_t^i$, we apply DBSCAN algorithm [80] to find clusters of object usage events. For each of the sensor data streams generated by a

sensor i , DBSCAN produces a series of events ev^i . Let EV^i be the ordered set of all events $ev^i = (t_{start}^i, t_{end}^i, m^i)$, where t_{start}^i and t_{end}^i are the start and end timestamps of the electrical object usage event, and m^i is the total electrical energy (watts) consumed in the event.

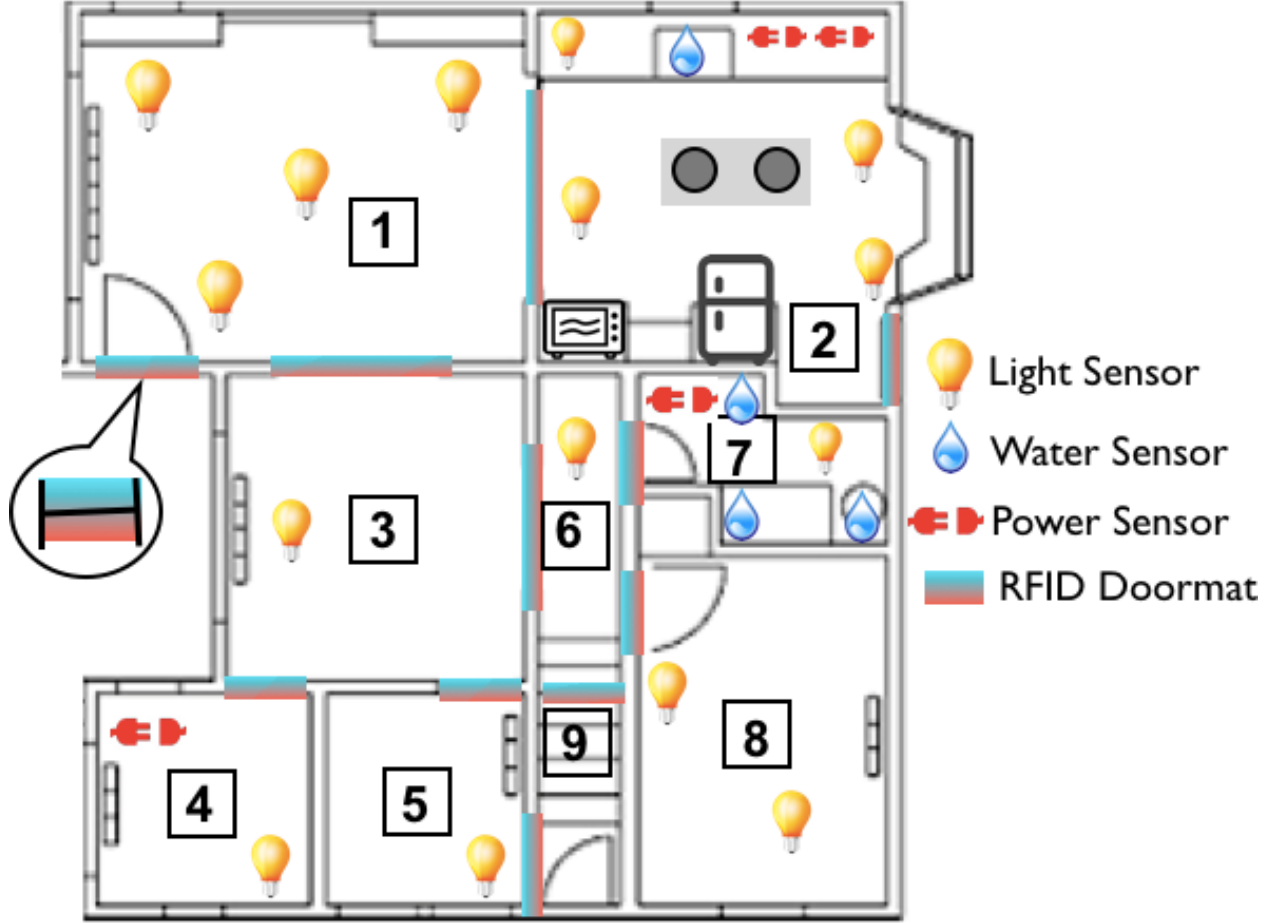


Figure 4.3: Sensor layout: the first floor of our test home had at least one instrumented light object, water object, or electrical appliance in every room. All doorways were instrumented with an RF Doormat.

For each water object sensor i , we have two streams of sensors reporting hot water volume (wh^i) and cold water volume (wc^i) at the object at a second's granularity. Therefore every water object sensor i generates a time series $W^i = (wh_1^i, wc_1^i), (wh_2^i, wc_2^i), (wh_3^i, wc_3^i), \dots, (wh_t^i, wc_t^i)$. We apply DBSCAN on W^i to generate water object usage events wv^i . Let WV^i be the ordered set of all events $wv^i = (t_{start}^i, t_{end}^i, m^i)$, where t_{start}^i and t_{end}^i are the start and end timestamps of the water object event end, and m^i is the total water (liters) consumed in the

event.

For appliances such as fridge and freezer which do not register an immediate change in appliance power, sensors on the appliance report back appliance door open events a^i . Let A^i be the ordered set of all door open events $a^i = (t_s^i, t_e^i, 0)$, where t_{start}^i is the timestamp of opening the appliance door, and t_{end}^i is the timestamp of closing the appliance door. F is the set of all object usage events, i.e., $\{EV, WV, A\} \in F$.

4.3.2 Step 2 - Location Tracking

In terms of tracking residents in homes, we compare three granularities of location inferring systems: 1. House occupancy level, 2. Room location level, and 3. Coordinate level. In order to have one system that could track people at all three levels of granularity, we used RFID based tracking systems. RFID anklets were worn by participants at home, and RFID detection zones created at pertinent locations, were used to perform these three levels of location in homes. We used a RFID tracking system, evaluated in prior work [79], to obtain highly accurate location information for our studies.

House Level Tracking

Tracking individuals at a house level implies detecting when the individuals are at home or away. This level of tracking is at a very coarse granularity, and is one of the most inexpensive level of tracking. The object user identification accuracy using this level of tracking depends on how often individuals stay at home alone. This is an inaccurate tracking mechanism when all the individuals within a home have similar hours of stay.

Let *HouseOccupancy* be the set of all home occupancy sessions (t^i, t^j, p_X) , where a person p_X occupies the house from time t^i to t^j . To detect home occupancy sessions of people, we track their movements across the entry doorways of the house. For each entry doorway ed^i , there are two RFID zones $(rf_{Exterior}^i, rf_{Interior}^i)$ on either sides of the doorway: $rf_{Exterior}^i$ lies on the outer part of the house, and $rf_{Interior}^i$ lies on the inner part of the house. We

know a person p_X enters a house at time t^i , when we observe a doorway crossing $dc^i = (p_X, ed^i, rf_{Exterior}^i, rf_{Interior}^i, t^i)$, where $rf_{Exterior}^i$ detects p_X first, followed by $rf_{Interior}^i$, at time t^i . We know that p_X was present in the house from t^i to t^j , when p_X is detected at a doorway crossing $dc^j = (p_X, ed^j, rf_{Interior}^j, rf_{Exterior}^j, t^j)$, where $rf_{Interior}^j$ detects p_X first, followed by $rf_{Exterior}^j$, such that $t^j \succ t^i$. More details of the algorithm to detect a doorway crossing can be found in [79].

Room Location Tracking

Room location tracking goes one step beyond what house level tracking does. In this level of tracking we have precise knowledge of when a person occupies individual rooms in a house. Room location is a very important level of tracking in homes, because a ‘room’ is a logical unit of space within a house which encompasses certain categories of activities which people perform within that space. These activities define the type of appliances and objects present and used in the room.

Let *RoomOccupancy* be the set of all room occupancy sessions $(p_X^i, r_x^i, t_{enter}^i, t_{exit}^i)$, where r_x^i is the room occupied by the person p_X^i from time t_{enter}^i to t_{exit}^i . To sense when people occupy rooms in a house, we detect them as they walk across the indoors doorways, to move in and out of rooms. For each doorway id^i , between rooms r_x^i and r_y^i , there are two RFID zones (rf_x^i, rf_y^i) on either sides of the doorway. RFID zone rf_x^i lies on r_x^i ’s side of the doorway, and rf_y^i lies on r_y^i ’s side of the doorway. We know a person p_X walks across a doorway id^i into the room r_y^i from room r_x^i , when a doorway crossing event $dc^i = (p_X, id^i, rf_x^i, rf_y^i, t^i)$ is generated in the doorway. Here rf_x^i is the first RFID zone that detects the person, rf_y^i is the second RFID zone that detects the person, and t^i is the time when p_X is detected at doorway. To detect how long a person stays in the room r_y , we look for the subsequent doorway crossing event made by the same person out of the room $(p_X, id^{i+1}, rf_y^{i+1}, rf_z^{i+1}, t^{i+1})$. Based on these two doorway crossing events, we can infer that person p_X was in room r_y from t^i to t^{i+1} .

Coordinate Level Tracking

Coordinate level tracking can locate a person at a precise spot with sub meter level accuracy within a house. Given this ability, we can locate a person in the spot right next to where the object is used. Although this may seem as ‘the most ideal level’ of tracking, it is still prone to errors. Sometimes two people may stand very close next to each other while one person using a object, and may even share the use of the object, for e.g. two people may cook together. In such cases, co-locating people with objects even within sub-meter range would result in ambiguous object assignment.

To track people at a coordinate level, as they stand next to objects, we created RFID detection zones $(rz_1, rz_2, rz_3, \dots, rz_n)$, with sub-meter radius. Each object i associated with a coordinate specific RFID zone rz_i . Each RFID zone i , produces a stream of RFID tags reads $rz^i = (r_1^i, r_2^i, r_3^i, \dots, r_n^i)$ when a person wearing RFID anklets stands near the corresponding objects. We use DBSCAN to cluster these tag reads into coordinate detection events $cd^i = (rz^i, p^i, t_{steppedOn}^i, t_{steppedOff}^i) \in CoordinateEvents$, where rz^i is the the RFID zone where a person was detected standing, p^i is the person identified at rz^i , $t_{steppedOn}^i$ and $t_{steppedOff}^i$ are the timestamps when p^i was first and last detected at rz^i in this cluster of RFID tag reads.

4.3.3 Step 3 - Object User Identification

House Level

For all events $f^i = (t_{start}^i, t_{end}^i, m^i) \in F$, we detect which individuals are present in the house. If $\exists(p^j, t_{enter}^j, t_{exit}^j) \in HouseOccupancy$, such that $t_{exit}^j > t_{start}^i > t_{enter}^j$, then person p^j was present in the house when the object was used. If only one person is detected in the house when a object event f^i takes place, the object usage is assigned to that person. *objectAssignment* is the set of all object assignments (f^i, p^i) . If there are multiple people in

the house when a object f^i is used, then it is declared an ambiguous object event, and added to the set *AmbiguousEvents*.

Room Level

For all object usage events $f^i = (t_{start}^i, t_{end}^i, m^i) \in F$ that takes place in a room r_{fi} , we detect which individuals are present in the room. If $\exists(r_{fi}^j, p_y^j, t_{enter}^j, t_{exit}^j) \in RoomOccupancy$, such that $t_{exit}^j > t_{start}^i > t_{enter}^j$, then person p^j was present in the room when the object was used. If only one person p^i is detected in the room r_{fi} , when a object event f^i takes place, the object usage is assigned to the same person. *objectAssignment* is the set of all object assignments (f^i, p^i) . If there are multiple people in the room r_{fi} , when a object f^i is used, then it is declared an ambiguous object event, and added to the set *AmbiguousEvents*.

Coordinate Level

For all object usage events $f^i = (t_{start}^i, t_{end}^i, m^i) \in F$, we detect which individuals are present at the associated RFID zone rz_{fi} . If $\exists(rz_{fi}^i, p^i, t_{steppedOn}^i, t_{steppedOff}^i) \in CoordinateEvents$, such that $t_{steppedOff}^i > t_{start}^i > t_{steppedOn}^i$, then we know person p^i was located right next to the object when it was used. If only one person p^i is detected at the RFID zone rz_{fi} , when a object is used f^i , the object usage is assigned to the same person. *objectAssignment* is the set of all object assignments (f^i, p^i) . If there are multiple people standing at the RFID zone rz_{fi} , when a object f^i is used, then it is declared an ambiguous object event, and added to the set *AmbiguousEvents*.

4.3.4 Step 4 - Applying Heuristics

For all object events $f^i \in AmbiguousEvents$, we apply different heuristics to make an intelligent guess as to who might have used a object.

Naive heuristic

This is a simple heuristic, which does not require any preconditions to operate. Given a set of multiple people (p_1, p_2, \dots, p_n) detected in the same proximity, where a object event f^i takes place, this heuristic randomly assigns the object usage to any one of the people present nearby.

$$Naive(f^i) = RandomSelection(p_1, p_2, \dots, p_n).$$

Blame ‘X’ Myopic heuristic

This heuristic keeps a short term memory (a day) of all the unambiguous object usage assignments made for each person. $(f^i, p_x^i) \in ShortTermMemory$. When a object is used, this heuristic refers to *ShortTermMemory* for the person p_x who has used this object most on that day. If there are no records for the object f^i in *ShortTermMemory*, i.e., $\nexists (f^i, p_x^i) \in ShortTermMemory$, then we use $Naive(f^i)$ to make the object assignment.

LastPersonXmin heuristic

If an unambiguous object assignment is made to a person within a certain time frame of an ambiguous object usage, this heuristic assigns the ambiguous object usage to the same person. To assign an ambiguous object usage f^i at time t_{fi} , this heuristic searches for any instance of an unambiguous object assignment in *objectAssignments* within a time window from $t_{fi} - X_{minutes}$ to $t_{fi} + X_{minutes}$. If $\exists (f^j, p_x^j)$, such that $f^j = (t_{start}^j, t_{end}^j, m^j)$ and $t_{fi} - X_{minutes} > t_{start}^j > t_{fi} + X_{minutes}$, then the ambiguous object usage f^i is assigned to p_x^j .

Blame ‘X’ Hyperopic heuristic This heuristic keeps a long term memory of all the unambiguous object usage assignments made for each person. $(f^i, p_x^i) \in LongTermMemory$. When a object is used, this heuristic refers to *LongTermMemory* for the person p_x who has used this object most to date.

If there are no records for the object f^i in *LongTermMemory*, i.e., $\nexists (f^i, p_x^i) \in LongTermMemory$, then we use $Naive(f^i)$ to make the object assignment.

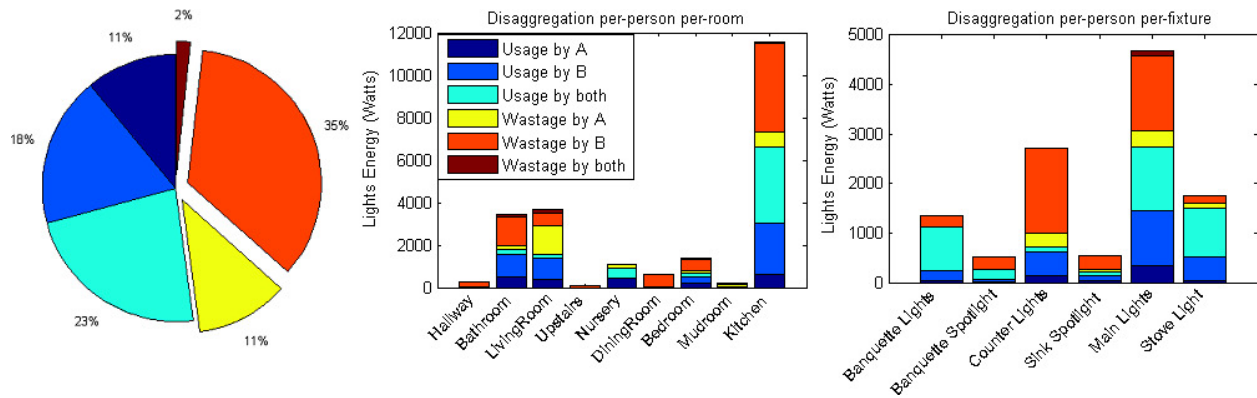


Figure 4.4: Case study of Lights energy: (a) Aggregate energy analysis reveals that person B uses light energy comparable to person A, but wastes three times more energy than A, (b) Energy disaggregated per person, per room indicates that Lights energy in dining room and hallway are mostly wasted, (c) Energy disaggregated per person, per object in the kitchen shows that the stove and banquet lights are used more efficiently

4.4 Case Studies

In this section, we take a look at some case studies demonstrating the potential for developing energy feedback applications that leverage user identification to encourage people to save energy.

4.4.1 Lights Energy Analysis

Personalized light usage feedback can reveal a lot of potential for savings. Quite often, lights once switched on, remain on even when rooms are unoccupied. We can detect exactly which person forgets to switch the light off when they leave the room. Figure 4.4 shows the breakdown lights energy usage and wastage by individuals from one of our studies. To discover wastage, for every minute that a light object is on, we detect who is present in the room. If one or more persons are present in the room, then its energy is assigned to the set of people present in the room. If the room is empty, and the light is on, then it is labeled as wastage. This wastage is attributed to the last person who left the room without switching off the light. In this case study, we observe that 50% of the total lights energy in the house is wasted by the individuals. Although person A and B use lights in the rooms almost equally,

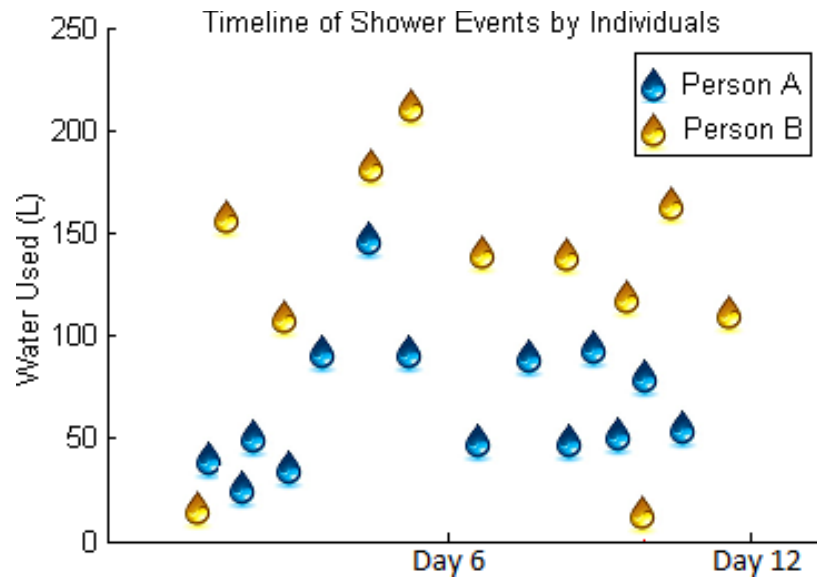


Figure 4.5: Person A showers more frequently than B. However A's total shower water usage (916 L) is lower than B's (1327 L). This is because B tends to take longer showers compared to A

person *B* wastes about 35% of the total lights energy, which is three times more than what *A* wastes. A room level breakdown further reveals that person *B* usually leaves the lights on in some rooms such as the dining room and hallway without ever using the lights in those rooms.

A wastes more light in the living room than *B*. Most of the light usage and wastage takes place in the kitchen. A further analysis of individual light objects in the kitchen reveals that certain lights objects are used more efficiently than others. For e.g., there is hardly any wastage detected for stove light and banquet light. However, most of the light energy of sink spotlight and counter light is wasted by *B*.

4.4.2 Habit Monitoring

A 'per person per object' assignment can be helpful in learning individual habits and preferences of people in a house. In Figure 4.5, we can detect non-obvious trends in shower water usage. Although person A showers 1.2 times more than person B, B uses 1.4 times

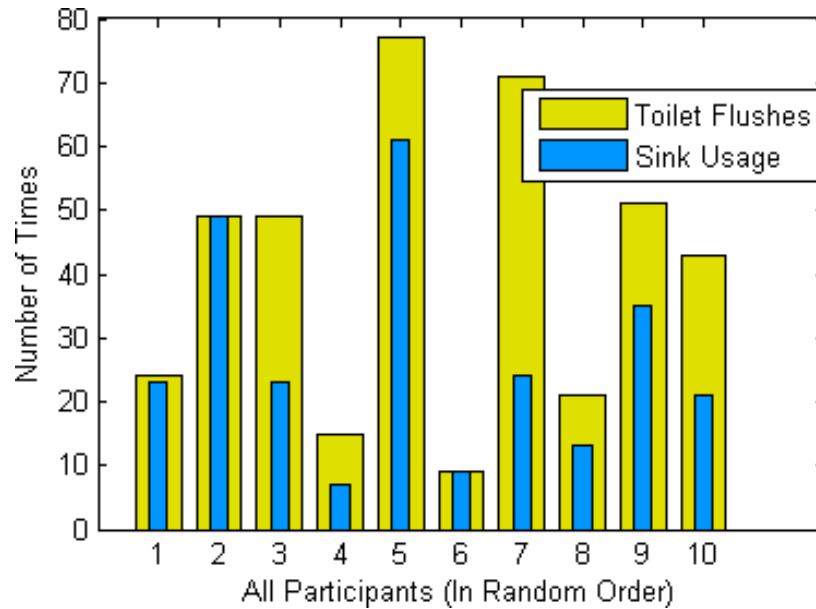


Figure 4.6: User identification can be used to track hygiene habits, for e.g. how many times do people wash their hands after using the toilet flush

more shower water than A. We can also monitor if people are following hygienic habits. In Figure 4.6, we can see that for 65% of the toilet flushes made by participant 7, there were no sink events detected within 45 seconds of the toilet being flushed.

4.5 Experimental Setup

To evaluate our hypothesis for performing a 'per person per day' disaggregation, we used a living lab model for data collection. We instrumented lights, appliances and water objects with sensors in a residential house and rotated five sets of two participants each, who lived in the same house for 7-12 days each. Compared to controlled experiments, this approach gave us in-situ data of the participants, who lived their lives normally in the house.

4.5.1 Participants

Five sets of two participants were recruited to live in the house for 12 days each. Each set had only two participants living in the house at the same time for two reasons - some of the

participants were married/dating, and for others we were limited by the number of bedrooms in the house. While most participants lived in the house for 12 days, one participant in study set 2 had to leave the study for work obligations after 7 days. Although the participants in study sets 3 and 4 lived in the house for 12 days each, we lost 5 days data in set 3, and 2 days data in set 4 due to a network failure in the living lab.

4.5.2 Sensing Infrastructure

Lights

We used light on/off sensing HOBO UX90-002M data loggers [81] to sense the lights in the house. These loggers are installed right next to the light bulbs, and have a programmable threshold for light intensity to detect when a light is on. We used off-the-shelf optical fiber pipe attachments (as shown in Figure 4.2 b.) with the light sensors to filter out external light and to make sure the sensor received light only from the object being sensed.

Appliances

We used different sensing methods for different types of appliances. *Appliances with hinged doors* such as fridge, dishwasher, washing machine etc. were instrumented with state sensing HOBO UX90-001 data loggers [82]. A rare earth magnet was attached to the door of the appliance, and the reed sensor was attached on the fixed part of the appliance. We also used the magnetic reed switches to instrument the stove, by fixing rare earth magnets inside the stove knobs, and reed switches next to the knob from inside the frame of the stove. *Small appliances plugged to wall receptacles* such as electric kettle, toaster etc. were instrumented with CSV A-8 current sensor [83] to sense when these appliances were powered on and off (as shown in Figure 4.2 a.) *Major power consuming appliances* such as microwave, oven, dryer etc. had individual power lines connecting to the circuit board, which was monitored using a TED Energy Monitor [84]. This monitored the power consumption on all circuits and reported data to a central repository via ethernet.

Water Fixtures

We instrumented all the water fixtures in the house such as sink faucets, toilet flush tanks, and showers, and obtained water flow information for each object at a second's granularity. To do so we installed in-line water flow sensors FTB4700 and FTB4707 [85] just before the outlet valve of each object. The data from these sensors were logged using UX120-017M 4-channel pulse data logger [86].

Location Sensing

The ability to conduct this study relied mainly on being able to sense the locations of residents in the house at different granularity levels. Each participant in this study wore a pair of RFID anklets with unique RFID tag numbers. We tracked the location of the participants in the house by tracking the RFID anklets worn by them using under the floor RFID antennas. We tracked participants in the house at three levels -

1. *Coordinate Level* - We created binary-state RFID detection spots near many of the objects in kitchen and bathroom. This ensured that we detected a person only when they stood at a given RFID detection spot in the house.
2. *Room Location Level* - In this level of tracking, we sensed the room locations of the participants by detecting their RFID anklets as they crossed the doorways, using the RF-Doormat room location sensing system proposed in [79]. This sensing system uses proximity based RF threshold systems: two RF thresholds placed on either sides of a doorway, determined if a person crosses a door to change his/her room location.
3. *House Occupancy Level* - In this level of tracking, we sense when people are present at house or away, by monitoring their movement into or out of the house at the entry doorways of the house using the RF Doormats.

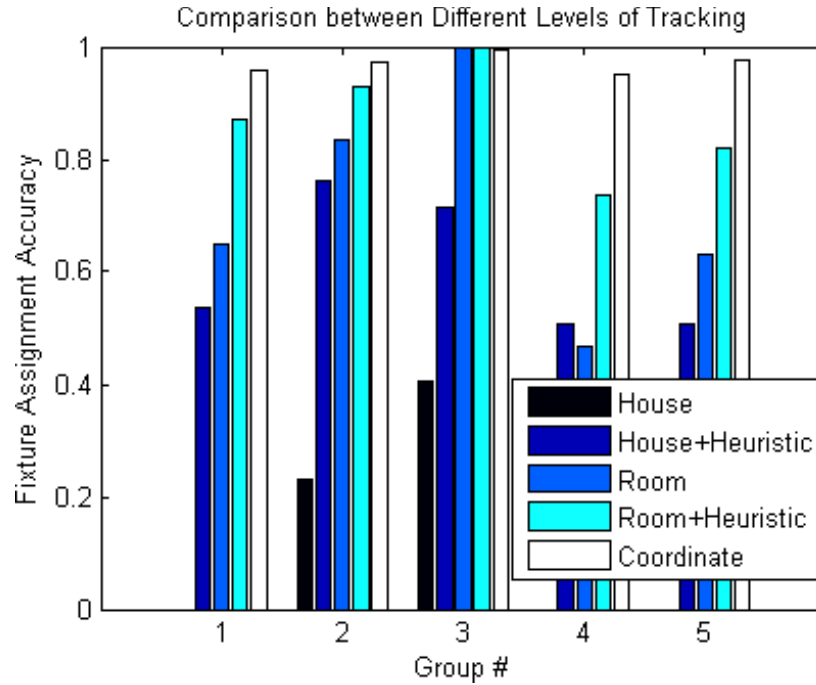


Figure 4.7: Evaluation results: House level tracking has an average object assignment accuracy of 12%. Room location level tracking alone has 70% accuracy. Heuristics augmented room location tracking has an average accuracy of 87%. Coordinate based tracking has an average of 97% accuracy across all groups

4.6 Results

To evaluate our hypothesis, we used object events obtained from ‘monitored’ appliances which have coordinate level tracking systems installed near them. These include toilet, bathroom sink, hair dryer, fridge, freezer, microwave, oven, toaster, coffee maker, electric kettle, coffee grinder, kitchen sink faucet, stove, range hood light, and dishwasher. Post study analysis revealed that none of the participant groups had used the coffee maker or the coffee grinder, and therefore these got excluded from evaluation. To test the object assignment of different tracking systems - house, room, coordinate and heuristics augmented house and room level tracking, we used the metric ‘*object assignment accuracy*’: percentage of object usage events assigned to the right person. To evaluate coordinate tracking, we used the number of events logged by the sensors on monitored objects as the ground truth. A video recording system would have been ideal for getting ground truth. However, privacy concerns of ‘being watched’

Group 1	P1	P2	Ambiguous	Total #
P1	91	1	47	139
P2	14	256	126	396

Group 2	P1	P2	Ambiguous	Total #
P1	248	0	34	282
P2	1	58	26	85

Group 3	P1	P2	Ambiguous	Total #
P1	64	0	0	64
P2	0	65	0	65

Group 4	P1	P2	Ambiguous	Total #
P1	62	1	56	119
P2	5	60	78	143

Group 5	P1	P2	Ambiguous	Total #
P1	196	11	90	297
P2	13	230	134	377

Table 4.1: Analysis of room location based object assignment: The confusion matrix represents the object assignments for individuals when the room is singly occupied. Ambiguous object usage (when multiple people are present in same room) is the main reason why room location tracking has lower object assignment accuracy than coordinate system. Group 4 has 0 ambiguous events, and therefore 100% room location based object assignment accuracy

inside rooms, and specially the bathroom, were the main reasons why we did not use cameras. For evaluating object assignment accuracy of all other tracking systems, we used the data from coordinate tracking system as the ground truth.

To get object assignments for coordinate tracking system, we incrementally searched for a person detected by the coordinate tracking system near the object starting from 5 secs to 15 seconds of the event timestamp. This flexibility in the search time window is to deal with the time sync differences between the different logging systems. If a single person is found standing near the object used, then the event is assigned to the same person. An event remains unassigned (ambiguous) if two people are detected within the search timeframe of the event. As can be seen from Figure 4.7, coordinate tracking system has an average accuracy of 97% across all groups of participants. The 3% of ambiguous object assignments

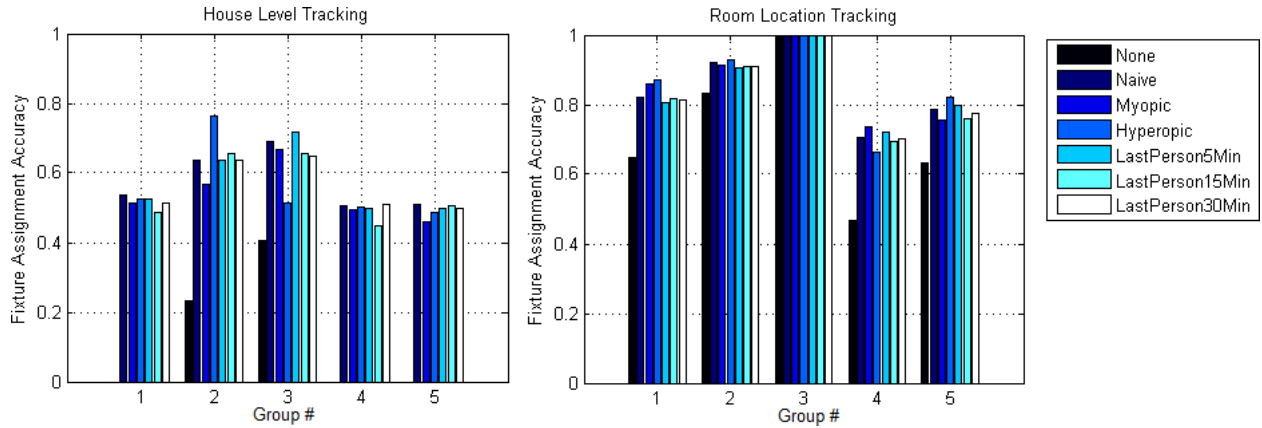


Figure 4.8: Heuristics analysis: (a) House Level Tracking - Heuristics achieve about 75% accuracy in groups 2 and 3, which have 23% and 40% accuracy without any heuristic. In group 1, 4 & 5, heuristics obtain 50% accuracy indicating that none of them worked better than random assignment. (b) Room Location Tracking Heuristics achieve over 85% accuracy in groups 1, 2 and 3 due to a high number of unambiguous object assignments. In group 4, heuristics achieve 72 % accuracy because of a small percent (45%) of unambiguous object assignments

could have been caused by several reasons, such as one person operating the object as the other person walks by, or two people jointly operating a object, such as sharing a stove while cooking a meal together.

In the room location tracking system, if a single person is detected in the room when a monitored object is used, it is assigned to the same person. A object event remains unassigned if multiple people are present in the same room at the time it is used. The room location system achieves an average object assignment accuracy of 71.60%. It achieves 100% accuracy in group 3, where there are no ambiguous events detected. Errors in detecting people entering a room, or exiting the room, may lead to mistakes in assigning the object to the right person. As can be seen in Table 4.1, these cases are few and limited to 2% of the total monitored object usage. Adding heuristics to room location tracking, improves the average object assignment accuracy to 87%.

In the house level tracking system, if a single person is detected in the house during object usage, the system assigns it to the same person. If multiple people are present in the house when a object is used, the event is unassigned. House level's object assignment accuracy is

Group 1	<i>BlameX Hyperopic</i>	P1	P2	Total #
	P1	91	48	139
	P2	21	375	396

Group 2	<i>BlameX Hyperopic</i>	P1	P2	Total #
	P1	282	0	282
	P2	26	59	85

Group 4	<i>BlameX Hyperopic</i>	P1	P2	Total #
	P1	85	34	119
	P2	35	108	143

Group 5	<i>BlameX Myopic</i>	P1	P2	Total #
	P1	246	51	297
	P2	113	264	377

Table 4.2: Improved room-location based object assignments: The use of simple heuristics with room location tracking resulted in object assignment accuracy of 87% in group 1 and 92% in group 2. This is comparable to 96% and 98% accuracy achieved respectively in coordinate level tracking system

0% for groups 1, 4 and 5 - which means both participants in these groups were present in the house every time a monitored object was used. The object assignment accuracy is higher for groups 2 and 3 at 20% and 40% respectively. For these study groups, the corresponding room location system's accuracy is also higher at 92% and 100%. Later in our exit interviews we came to know that in group 3, one participant had left town for two days. As a result, the house was singly occupied for 30% of the study, resulting in a higher object assignment accuracy at the house level.

House level tracking system when augmented with heuristics, has an average of 60% accuracy. This is a significant improvement over the average of 12.5% in unassisted house level tracking. In fact, group 4's heuristic assisted house level accuracy is higher than that of unassisted room location based system. Group 2 achieves 76% accuracy which is comparable to the 83% achieved by the unassisted room location tracking system. Groups 1, 4 and 5 achieve an accuracy of 50% which is what we expect if we were to randomly assign all objects usages between any two people. Details of heuristics assisted house level tracking will be

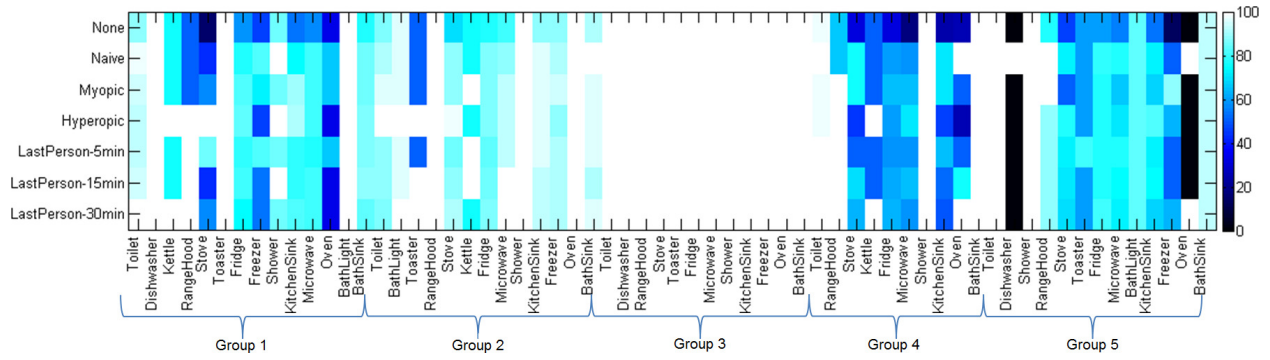


Figure 4.9: object level breakdown of event assignment accuracy, for different heuristics for all groups, shows potential for developing complex heuristics which are sensitive to typical usage pattern of individual objects

discussed later in this section.

Analysis of Heuristics

In Figure 4.8, we compare the performance of all heuristics for house level and room location tracking systems. All heuristics except for Naive, refer to unambiguous usage of objects at different timeframes to determine object assignment in ambiguous cases. In the case of house level tracking, all heuristics in study groups 1, 4 and 5, have object assignment accuracies of 50%. This is expected since all heuristics other than Naive, depend on a history/database of unambiguous object assignments, which is 0% for these groups. In the absence of any unambiguous history, they all resort to the Naive heuristic to make the object assignment determination. Group 2 has 22% unambiguous object assignments, therefore enabling the heuristics to use the history to learn object usage behavior. Hyperopic heuristic, which uses long term history to disambiguate works best for this group. This implies that these participants typically follow a consistent pattern of object usage over time. In the case of group 3, we have 40% unambiguous object assignments, however Hyperopic performs even worse than Naive. This might indicate that the unambiguous object assignments were made during a time when another person was not present. Therefore the history of appliance usage

could be skewed and a short sighted heuristic such as LastPerson5 min has higher accuracy.

In heuristics assisted room location tracking, Hyperopic heuristic performs the best in three of the five study groups. In group 1, Myopic and Hyperopic perform better than Naive, whereas the LastPerson heuristics are at par with Naive. In groups 2 and 5 all heuristics are comparable to Naive, with Hyperopic performing slightly better than Naive. The problem with Hyperopic is that the most dominant user is assigned all the ambiguous object usages, even those made by the minority user. Looking at the confusion matrix for group 2 in Table 4.1 and Table 4.2, we observe that P2 is the dominant user of all objects, and gets assigned all ambiguous object events (even those of P1). This indicates the need for insightful heuristics which even consider assigning ambiguous object events to a minority user. In group 3, the unassisted room location system achieves 100% object assignment accuracy, indicating that there were no ambiguous events at this tracking level. Therefore, all heuristics performed at par with the unassisted room location system. In group 4, Myopic heuristic performs better than the other heuristics. This indicates that in this group, the dynamics of who uses a object varied over time.

Analysis of Heuristics for Individual objects

One possible way to improve object assignment accuracy, is to apply different heuristics for different objects. In Figure 4.9, we show a detailed breakdown of event assignment accuracy per object for all heuristics across all studies. While BlameX Hyperopic achieved the highest accuracy across most groups of participants, we can clearly see that for many appliances other heuristics worked better. And this seems intuitive for general purpose appliances such as fridge and freezer. Here historical data gives little information in predicting future object usage, for e.g. predicting who would feel thirsty and open the fridge for a glass of juice. This suggests that complex heuristics will be better at performing event assignments. In future work, we want to develop systems that can apply a medley of different heuristics for individual objects, based on object usage patterns.

4.7 Discussion

In this work, we investigated the hypothesis that we can perform object user identification based on coarse-grained room-level location tracking even when two people are in the same room when a fixture is used. We compare the user identification accuracy at three different granularities of tracking - at the house level, room location level and coordinate level. While we recognize that coordinate tracking is the best level of tracking required to assign fixture usage to individuals, we can achieve comparable results using a more coarse grained room location level tracking system. We observe that room-location tracking alone cannot assign fixture usage when multiple people are present in the same room, and has a 70% average user identification accuracy. However, the use of fixture usage history based heuristics, improve the average accuracy to 87%. These heuristics look at short term and long-term usage history of a fixture to determine which person should be assigned the fixture event. In three sets of participants, a long term usage history based heuristic 'BlameX Hyperopic' had 89% user identification accuracy. These results are indicative of the potential that simple heuristics have in overcoming the limitations of a room location tracking system.

Anecdotal data in Figure 4.9 reveals that when it comes to applying heuristics to objects, there's more to it than meets the eye. Any single heuristic does not perform the best assignment for all appliances. This is because different appliances in multi-person household are 'shared' by residents in different ways. For example, one person may cook for everyone in the household, and in this case a stove's usage can be better disambiguated by a long term history based heuristics. The fridge on the other hand, may be used by everyone in the house, with no obvious bias by any one individual, and therefore a random or naive approach in disambiguating fridge usage might be the best approach.

Trying to develop a heuristic with a comparable accuracy to coordinate tracking based system is challenging: this is a modeling problem to understand how people use objects when there are other people present in the same room. Since this modeling would require more longer study periods than two weeks, in ongoing work we are generating more extensive data

sets. Our goal for future work is to test whether the system can learn personal ‘signatures’ of every object’s usage by individuals to improve user identification when multiple people are in the room simultaneously.

4.8 Limitations

We recognize that our study sets of two people are not representative of all possible types of multi-person households, such as those with grandparents, children or pets living together as well. We were limited by the number of bedrooms in the house. Recruiting participants who were willing to leave their homes, and live somewhere else for 2 weeks was also a big challenge. We were a bit concerned about whether participants would behave naturally while living in another house (Hawthorne effect). However, in the exit surveys, all the participants said that their lifestyle in the living lab was very similar to what they had at their own residence, except that some participants did not have certain appliances at their own homes. A participant said, “Yes, it is indicative of use at home, except for the fact that we do not have a dishwasher or laundry appliances. We probably used the washer and dryer more frequently than we normally would due to the fact that it was a rare opportunity to do all our laundry for free :)”. The time of study for one participant set coincided with a month long winter break at the university, where the participants said that they were at home longer than the usual work week - “no deviation from use in own house except maybe hours of occupancy since this was winter break period”. We believe the participants would have also had longer hours of home occupancy at their own homes during the winter break, and therefore we do not consider this a serious deviation from natural behavior. Participants also said that the RFID based location tracking system did not distract their behavior - “RFID receivers beneath the floor did not bother us at all - it’s easy to forget they are there”.

4.9 Conclusion

In this work, we aim to evaluate object user identification accuracy for individuals living in a home and using common household objects. We use the fusion of information from location tracking system and object usage to demonstrate, that object user identification in a home using a coarse grained room level location tracking system can perform with 87% accuracy, even when multiple people are present in the same room where the fixture is used. We apply heuristics that learn how people use objects in unambiguous assignments (when only one person was present in the room), and make an informed guess to assign ambiguous object usage to individuals. While we recognize that user identification can also be performed using coordinate level tracking, we believe that a coarse granularity of tracking would have lower cost implications.

Chapter 5

Level 2 - Using Gesture Matching to Identify Object User

5.1 Introduction

The biggest drawback of location-based object user identification is that it does not work when multiple users are detected in proximity to the object. The heuristics we explored are after all a guess, albeit a learned one. If the object does not have the history of usage of every user it has to choose from, or if certain objects are used equally or randomly by individuals (e.g. lights), then the use of heuristics does not help in guessing who used the object. At this point, the object initiates the next level of interaction with the Wearable. This time, it asks the devices to match the motion of their wearer's hand to the motion required to interact with the objects, or in other words it initiates gesture recognition for the object.

Gesture recognition is a popularly known concept in the computing community. The term 'gesture' can be used to refer to any motion of the hand or the body, or some other form of non-verbal communication that is typically intended as a means to express certain messages. Gesture recognition is the use of sensing technology to interpret the human body or hand motion using mathematical algorithms. Understanding gestures is important because

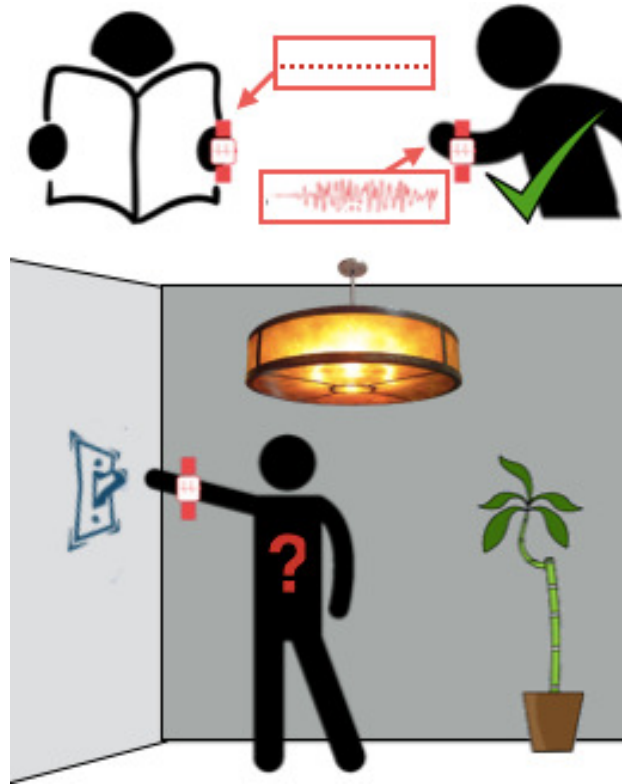


Figure 5.1: When an object is used, the object user is identified as the person making a hand movement containing the object hallmark. The hand is monitored using wearable devices having IMU sensors, such as fitness trackers

they can be used to create more natural object interactions and detect a person's activity. Typically a gesture represents the trajectory of motion undertaken by the part of body being tracked.

Traditionally, camera-based systems have been successful in tracking a predetermined set of gestures made by a person. However, the same is not true for wearable based gesture sensing. The main sensor that is used to sense motion in wearables devices is the IMU sensor. This sensor has an accelerometer, gyroscope and magnetometer, which measure acceleration, angular velocity and orientation respectively.

Objects that have embedded sensors can recognize the object user based on the unique way in which an object is touched [71] or held [?], which is referred to as 'Object Use Fingerprinting'. Most of the solutions rely on instrumenting objects with additional sensors. Therefore, until every object in the world is fully instrumented, there is not likely to be a

“silver bullet” that solves this problem for all object types.

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

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

To test our hypothesis, we use the concept of apportionment in a home - where the identity of an object user is important from the perspective of attributing the object’s energy usage to the user. We instrument 16 objects in a home, including lights, water fixtures, and major appliances. We perform a series of three feasibility studies with decreasing levels of

constraints, ending with a single person in-situ study. Some of the research questions we intend to answer with these studies are: whether object hallmarks are unique and identifiable, if object hallmarks are person-independent, and how real world tasks affect the object user identification accuracy.

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

5.2 Related Work

Object Use Fingerprint research shows that it is possible to recognize an individual based on how they touch objects by equipping them with pressure sensors or other smart surfaces [71]. Another case of Fingerprinting was demonstrated with a button. A pressure-sensitive button was built containing an Force Sensitive Resistor sensor. Springs provide return force and push resistance. The work explored the hypothesis that different users have different button pressing habits. User recognition via button pressing is particularly suitable in one-off interactions, where users are not engaged for long and where, at the extreme, the interaction

itself can be reduced to one button press. Our work explored the inverse of this hypothesis. We explored the idea that the use of an object imprints a unique and identifiable hallmark in the hand movement of its user.

Chang et. al use accelerometers embedded in a television remote control or mobile device to identify household members, based on the unique way each person uses the remote [?, ?]. They show that accelerometers, touch screens and software keyboards, can be used to differentiate different test subjects based on the unique interaction characteristics of each subject. These approaches work well when an object contains the sensors required to uniquely differentiate object users but does not extend as easily to other objects such as home appliances, light switches, or water fixtures.

Other proposed methods recognize an object user in a house by instrumenting all objects or fixtures with RFID tags[?, 77]. A specialized wrist device consisting of an accelerometer and/or a RFID reader on it is worn by the occupants. The RFID reader tells us which person is touching an object, and the accelerometer characterizes how the object is being used.

5.3 Approach

5.3.1 Sensing Object Hallmarks

Since objects have a unique combination of interface, orientation and location within a space, the hand gestures for using them in terms of the hand's acceleration signature, tilt and compass direction are also unique hallmarks. In order to sense these three parameters of a hand's motion, we use the nine-axis Inertial Measurement Unit (IMU) sensor of a wrist wearable device. The triaxial accelerometer in an IMU reports values composed of linear acceleration of the hand, as well as acceleration due to earth's gravity. The tilt of a hand can be inferred by measuring the amount of gravity acting upon each of the three axes. In order to separate gravity and linear acceleration, we apply a low pass filter on the accelerometer values. As illustrated in Figure 5.2, if we have a acceleration vector $acc_{(t_1, t_2)}^{x, y, z}$, and a gravity

vector $gra_{(t_1, t_2)}^{x, y, z}$, then linear acceleration vector $linA_{(t_1, t_2)}^{x, y, z}$ can be obtained by

$$linA_{(t_1, t_2)}^{x, y, z} = acc_{(t_1, t_2)}^{x, y, z} - gra_{(t_1, t_2)}^{x, y, z}$$

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

Let the sensor data between time t_1 and t_2 from the wearable, be represented as

$$sen_{(t_1, t_2)} = (acc_{(t_1, t_2)}^{x, y, z}, gra_{(t_1, t_2)}^{x, y, z}, orient_{(t_1, t_2)}^{x, y, z})$$

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

5.3.2 System Requirements

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

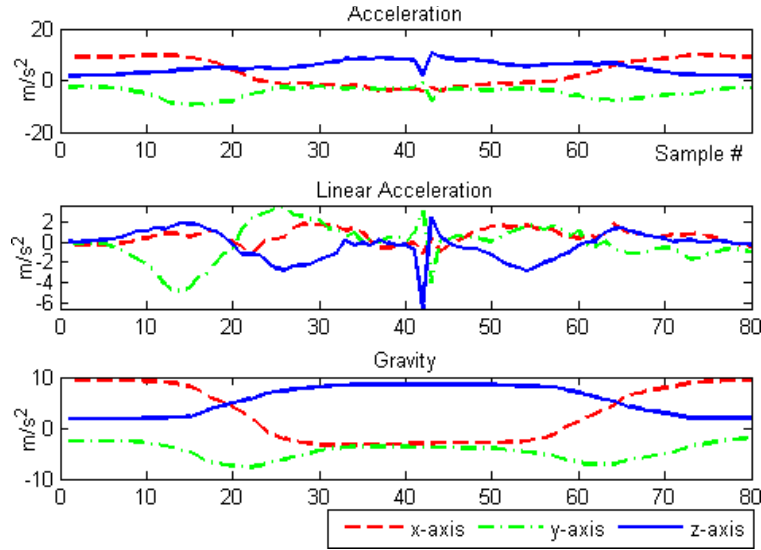


Figure 5.2: Total acceleration across the x, y and z axis is decomposed into gravity and linear acceleration along the three axes. The gravity component indicates how a hand is oriented, and the linear acceleration defines the movement of the hand

5.3.3 Overview of Object User Identification

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

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

1. *Creating Reference Dataset*: A dataset of object hallmarks of each possible object user is recorded for reference.
2. *Determining Possible Object Users*: A set of possible object users are identified by determining people whose hands were moving at the time of an object usage.
3. *Identifying Object User*: If more than one person is active during an object usage, we

process the sensor data further to determine which person made a hand movement that most closely resembles the object gesture in the training dataset.

1. *Creating Reference Dataset*

Every person has their own ‘fingerprint’ of operating an object [71]. Conversely, the object hallmarks imprinted in the hand gesture of every person will be different as well. In the current version of the system, we reference object hallmarks of every person for all monitored objects from a pre-recorded dataset to perform object user identification.

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

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

$$training_{p_x}^O = (hallmark_{p_x}^{o_1}, hallmark_{p_x}^{o_2}, ..., hallmark_{p_x}^{o_n})$$

2. *Determining Possible Object Users*

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

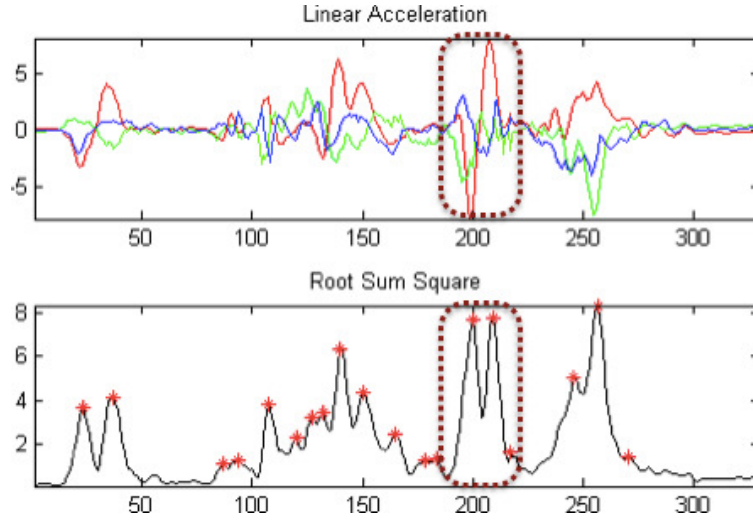


Figure 5.3: Segmentation: Linear acceleration across x, y and z axis are filtered and combined using Root Sum Square (RSS). Peak detection is performed on RSS to determine bursts of acceleration. Sensor data of a fixed window size around each peak is used for feature extraction

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

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

Having extracted the linear acceleration for all the people at home, we now determine the people who are active between the time $t - td$ to $t + td$. To determine if a person is at rest at



Figure 5.4: Experimental Setup: (L-R) The fridge, freezer and microwave face the same direction, and have doors that open the same way. The dishwasher door is hinged at the bottom. The bathroom hot and cold faucets pull out from left to right, and right to left respectively. The kitchen sink hot and cold faucets both turn in counter-clockwise direction. Light switches for kitchen and living room are right next to each other. The x, y, and z axis of a LG android smart watch are annotated

the time that an object is used, we measure the standard deviation of the linear acceleration, and see if it exceeds a certain threshold value:

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

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

3. Identifying Object User

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

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

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

$$RSS_{(t_1,t_2)} = \sqrt{(linA_{(t_1,t_2)}^x)^2 + (linA_{(t_1,t_2)}^y)^2 + (linA_{(t_1,t_2)}^z)^2}$$

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

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

Figure 5.3 shows the linear acceleration data of a person opening a microwave door. To open the door, a person places the hand on the door and pauses slightly before pulling the door. When the door opens, he holds his hand firmly to stop the door, resulting in an equal and opposite force. This gesture shows up as adjacent peaks of equal magnitude in the RSS vector.

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

For each peak $pk \in peaksRSS$ in the sensor data of a potential object user p_x , we determine the distance $dist_{pk}$ of the peak's features $features_{pk}$, to the object o_i 's hallmark $hallmark_{p_x}^{o_i}$ in the reference dataset $training_{p_x}^O$.

$dist_{pk} = distance_function(features_{pk}, hallmark_{p_x}^{o_i})$, where $distance_function$ is a euclidean distance function. Further, for each potential object user, we select a candidate object usage hand gesture as the peak pk with least $dist_{pk}$ to the object hallmark. From the set of candidate hand gestures from all potential object users, the individual having a hand gesture with the smallest value of $dist_p$ or, the best resemblance to the object hallmark, is identified as the object user.

5.4 Experimental Setup

To test our hypothesis, we used systems that could monitor the use of objects in homes, and smart wrist devices to sense hand gestures. We instrumented 16 major appliances, light and water fixtures in a two-bedroom apartment. The relative locations of all objects are shown in Figure 5.5.

In the absence of any commercially available Non Intrusive Load Monitoring systems at the time this study was conducted, we placed direct sensors on all objects to monitor their usage. We used magnetic reed sensors on appliances with hinged doors, such as Fridge (F), Microwave (M), Freezer(Fr) and Dishwasher(D) (Figure 5.4). We also used magnetic reed sensors to instrument the bathroom hot (BH) and cold (BC) faucets, as well as the kitchen hot (KH) and cold (KC) faucets. The reed sensors were plugged into HOBO UX90-001 data loggers [82]. We used light on/off sensing HOBO UX90-002M data loggers [81] to sense the Lights(L) in the house. These loggers are installed right next to the light bulbs, and have a programmable threshold for light intensity to detect when a light is on.

To sense the hand gestures, we used LG G watch [?], which runs on Android Wear platform. We wrote an app for the watch which collected IMU based sensor data (accelerometer, gravity and orientation vector) at 33 Hz and transmitted it to its paired smartphone. The smartphone had a listener app which logged the received data on the phone. For each sensor data sample, the app recorded the timestamp, the sensor type, and the set of raw sensor values along x, y and z axes. The app assigned a sequence number to each received sample as well to detect packet loss. The sequence number helped us detect a 30 second (~3000 data points) data loss during one of our trials.

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

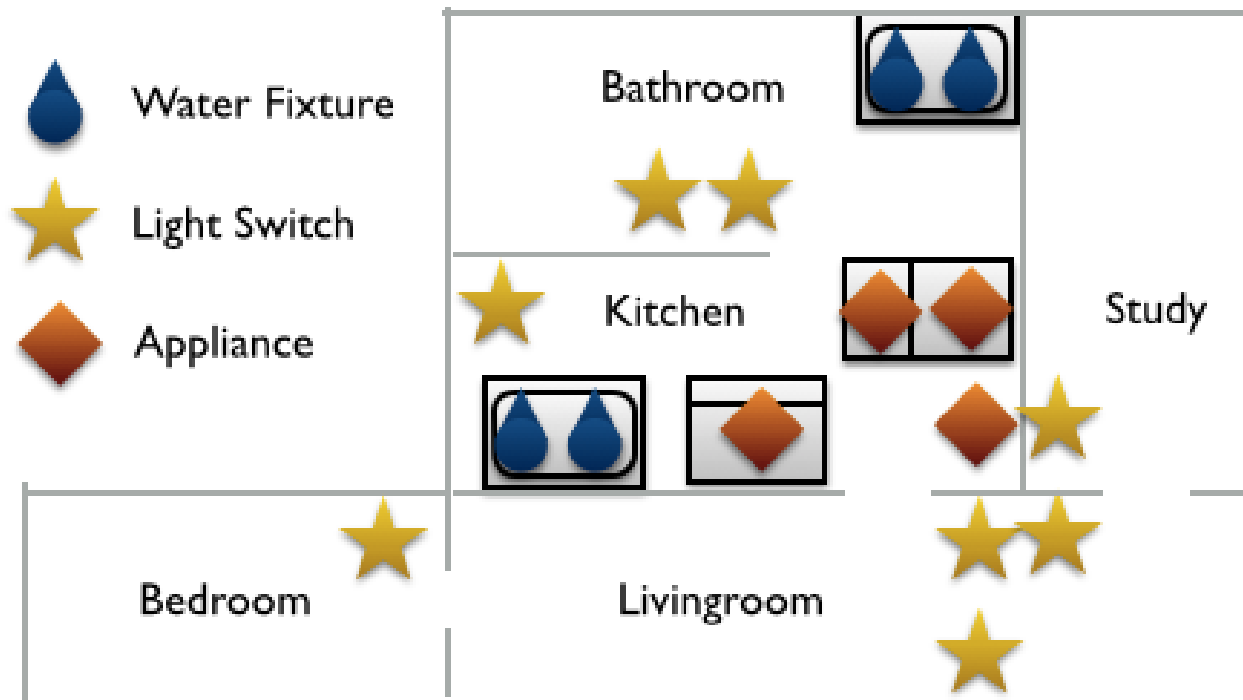


Figure 5.5: Object Layout: All the major appliances, light and water fixtures of a two-bedroom apartment were instrumented for the study. Quite often, two objects with similar interfaces are co-located, such as the kitchen sink hot and cold faucets

5.4.1 Study Design

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

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

Scripted Study

The main premise of our approach relies on the ability to detect and differentiate between hallmarks of different objects. Therefore, to ensure feasibility of our approach, we first

evaluated our approach in a scripted study. In this study the participants were instructed exactly which objects to use in a predetermined, scripted order. The aim of this study was to answer the following questions:

1. Do people use objects using gestures that are distinguishable and learnable?
2. Is it possible to learn the object hallmarks in a person-independent manner?
3. Can we differentiate between hallmarks of objects with the same interface? For e.g. the fridge and the freezer doors have similar handles.

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

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

Real World Task based Study

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

real world tasks.

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

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

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

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

For ground truth, we had an observer making real time entries about the object usage of each individual, using a logging tool that was time-synced with the sensor loggers. To perform object user identification for the tasked study, we used the annotated gestures from

the scripted study for the reference dataset. Since, the script was repeated ten times, we obtained ten reference object hallmarks for each object in the home.

We aimed to answer the following questions with this study:

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

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

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

On an average, ~200 object events were performed within 2 hours by each group. To put this into perspective, we compare this frequency to the object usage frequency from an in situ dataset from one of our previous studies [?]. The dataset contained object usage timestamps of 5 groups of 2 individuals, who lived in an instrumented home for 10-12 days. As can be seen in Figure 5.6, this actually resulted in a high frequency of object usage in our study, with about 50% of the object usages happening within 5 seconds of each other. Whereas, according to the previous dataset, only 15-20% of the object usages happened within 5 seconds of each other.

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

In-Situ Study

Although the task based study encapsulated a lot of noise from real life scenarios, we

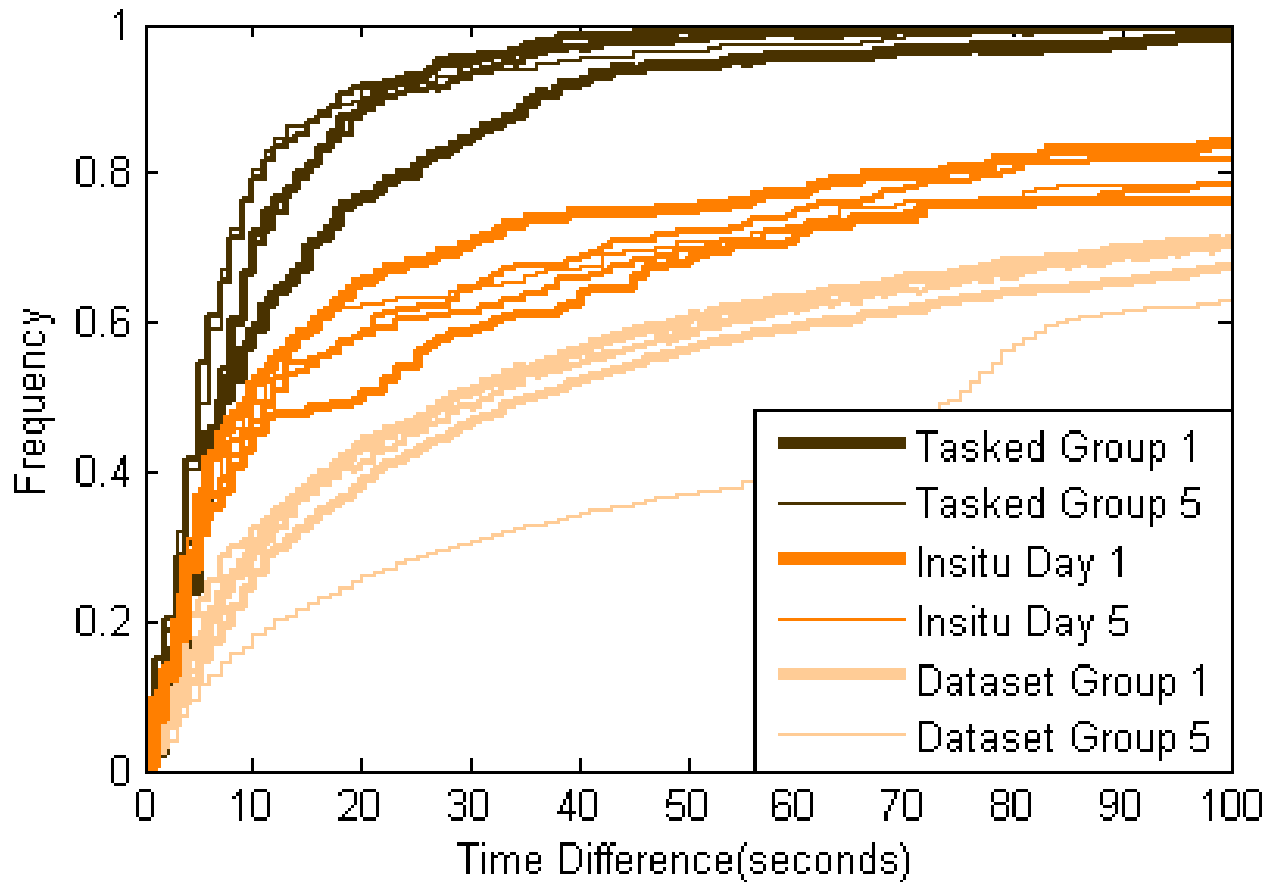


Figure 5.6: CDF of time differences between object usage: 50% of the object usage in the tasked study happened within 5 seconds of each other

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

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

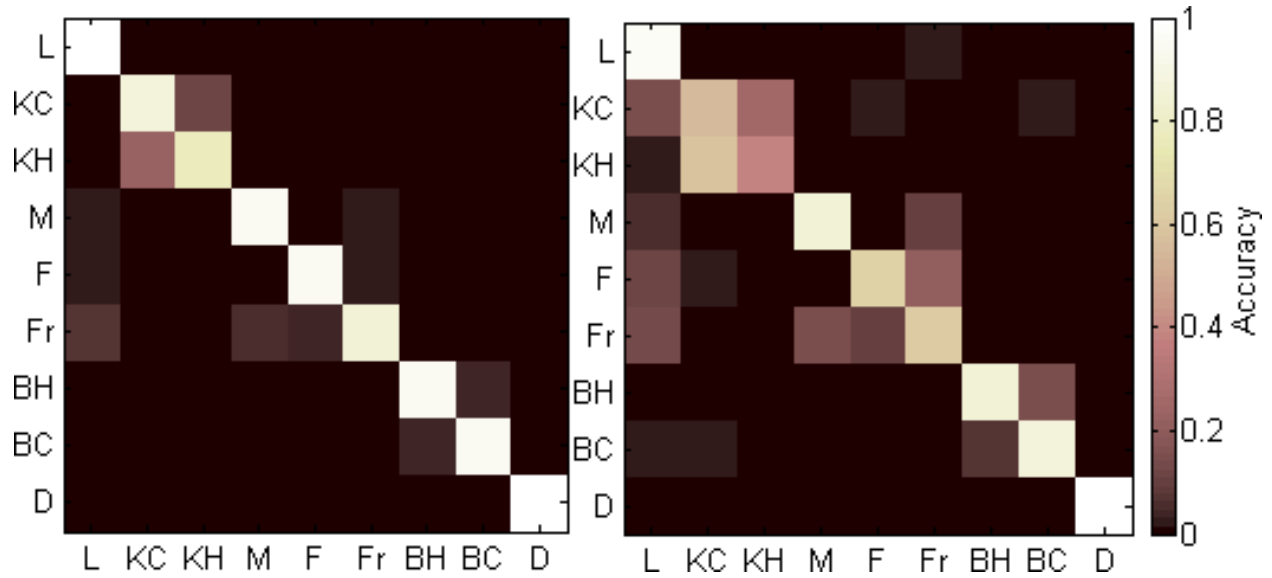


Figure 5.7: Gesture classification: Confusion Matrix of 10-fold cross validation in scripted study show (a) 95% accuracy for a person dependent gesture model, (b) 88% accuracy for person independent gesture model

of 515 in-situ object usages in total. The main research question that we aimed to answer using this study was:

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

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

5.5 Results

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

Feasibility Study

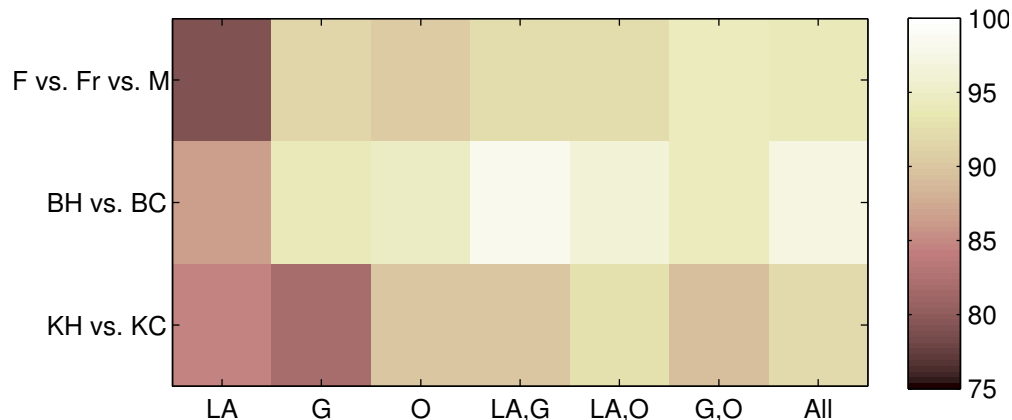


Figure 5.8: Classifying between objects with similar interfaces: Different features are evaluated to determine how gestures are differentiated between sets of co-located and similar objects

We used the annotated gesture data from the scripted study to validate the feasibility of using object gestures as hallmarks for object user identification. The metric used in this section is *Classification Accuracy*, which is defined as the sum of correct classifications divided by the total number of classifications. The annotated gesture dataset has an object ID for every object gesture, and the mean, median, standard deviation, 25th percentile data point and 75th percentile data point of the three axes of linear acceleration, gravity and orientation as the feature set. For each of the 10 participants' data, we performed 10-fold cross validation on the object gestures, using Nearest Neighbor method. As shown in Figure 5.7(a), the per-person based classifier achieved 95% classification accuracy using the 15 features calculated on the annotated sensor data. This suggests that gestures used to operate objects can be differentiated. This also suggests that for a given object, the gestures required to operate the object are consistent for the same person.

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

To determine if object hallmarks are person dependent or not, we created test cases, where we trained using annotated gesture data, from 9 people and tested on the 10th person.

As shown in Figure 5.7(b), on an average, this method was able to differentiate between gestures with 88% accuracy. In Figure 5.7(b), the precision and recall for kitchen hot and cold faucets (KC and KH) was quite low, with one getting classified as the other, quite often. This was also true for the fridge and freezer, and the bathroom hot and cold faucets. The dishwasher, microwave and lights performed the best in the person-independent classification. Given that certain object hallmarks performed well in the person independent analysis, while others performed poorly, it is hard to conclude whether object hallmarks can be learned in a person independent manner.

Figure 5.7(a) showed that a person's gestures for co-located and similar objects are differentiable as hallmarks. We further analyze how each of the three IMU sensor's features: Linear Acceleration (LA), Gravity (G) and Orientation (O), and their combinations contribute to the classification between the following sets of similar objects:

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

The accuracy of using different features in classifying between sets of similar objects are shown in Figure 5.8. Results show that orientation (O) features are the most accurate and linear acceleration features are the worst for differentiating between F/Fr/M. This is because all three objects have similar interface - handle that needs to be pulled outward in the same way. Thereby resulting in similar linear acceleration signatures.

In differentiating between bathroom hot and cold faucets, the *Linear Acceleration, Gravity* (LA,G) feature set performs the best. Since the hot faucet handle moves from left to right,

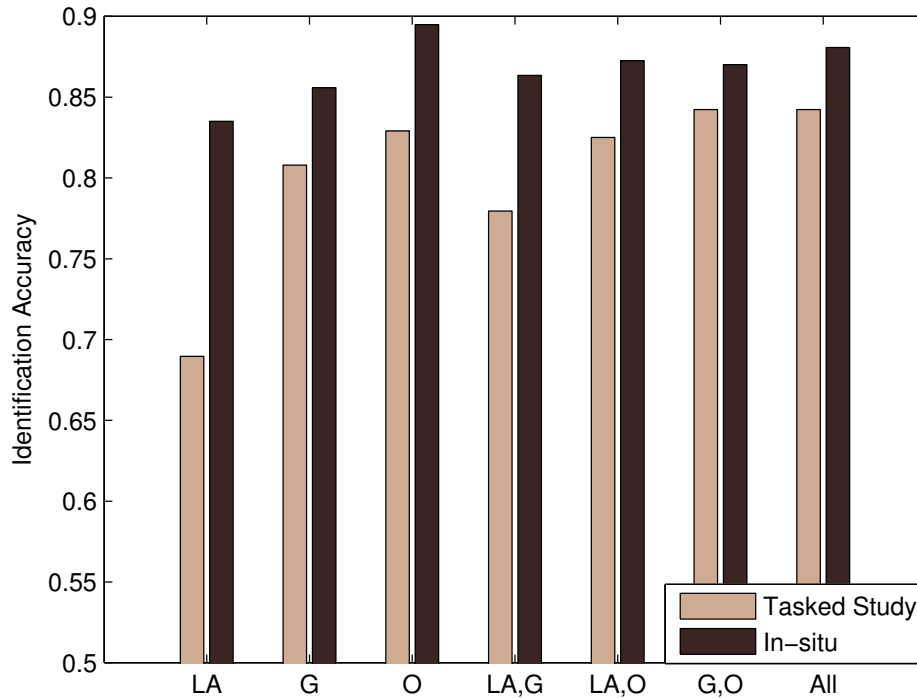


Figure 5.9: Different sensing features are evaluated to determine person identification accuracy. Linear Acceleration (L) based features do not perform as well as the orientation features (Gravity (G) and 3-D orientation (O))

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

Identifying the Fixture User

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

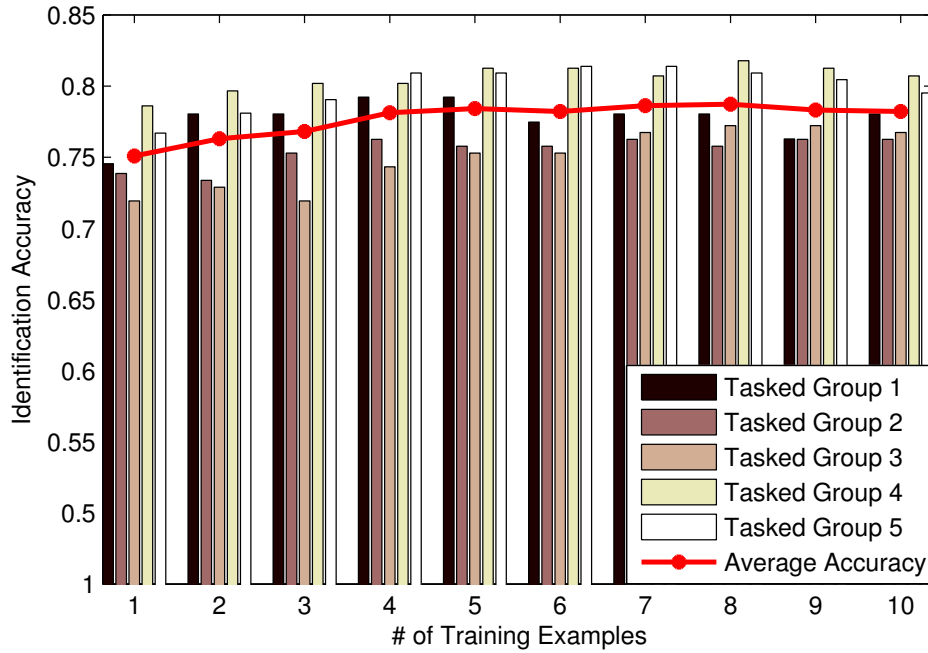


Figure 5.10: An average accuracy of 80% is achieved with just a single training example for each object gesture. More training examples bring incremental improvements in the accuracy

which is defined as the number of correct object user identified divided by the total number of object usages.

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

Since our approach uses features from the three sensors in an IMU (accelerometer, gravity and orientation), we varied the sensor features in the distance function to study which features performed the best in terms of identification accuracy. In Figure 5.9, we can see that our approach achieves 90% accuracy in the in-situ study based simulations using only orientation parameters in the distance function. We achieve 85% accuracy for the tasked study, using gravity and orientation parameters in the distance function. The linear acceleration features in the distance function, perform the worst for both the in-situ and the tasked study. This

may be explained in part that there might be variations in the force with which a person uses the same object at different times. In general, the orientation features achieve higher accuracy. This is possibly because the hand's orientation, in terms of the gravity and the compass direction in space, remains relatively constant, as the hand is constrained to face a certain way to use an object.

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

Figure 5.10 shows our system's identification accuracy on varying the number of training examples used. With just one set of training data for all the objects in the home, our system can identify the correct person 80% of the time in the tasked study. With additional training data, the identification accuracy improves incrementally. This is possibly because the training data was used from the scripted study which probably did not include all possible variations of people's object usage gestures.

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

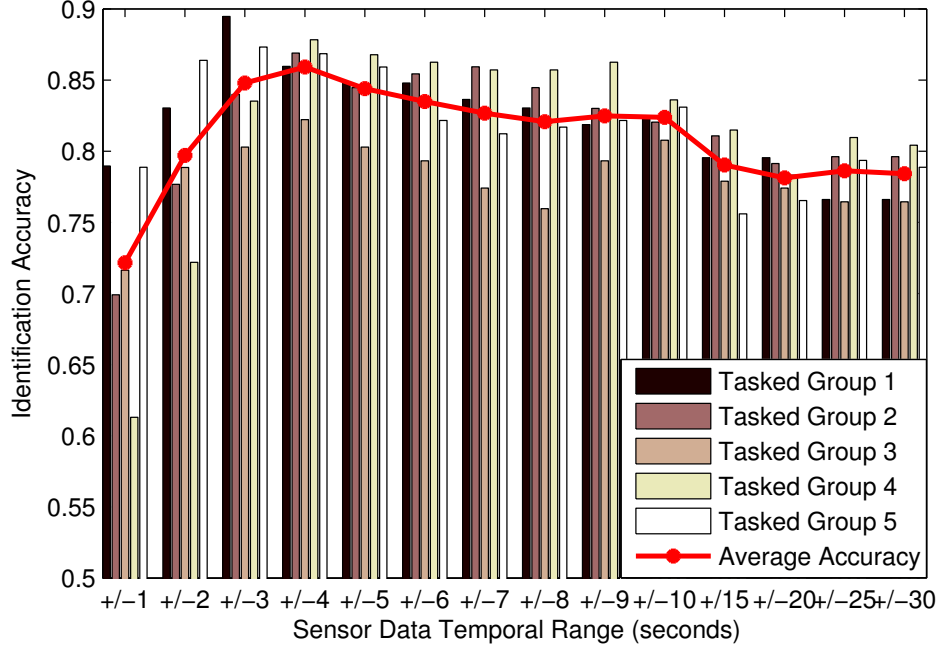


Figure 5.11: Highest accuracy is obtained when a sensor data size of ± 4 seconds around the object usage timestamp, is selected. The likelihood of an overlapping event from another person increases with a larger temporal range of sensor data. Accuracy is low for smaller sensor data size because the object monitoring hardware had a 1-3 seconds difference with the time on the watch

and 84%. For the 4 person scenario, the average accuracy came down to 65% and was similar for the five person scenario at 63.8%.

Sensitivity Analysis

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

In our approach, we assume that there exists a system which records and reports every time that an object is used in a home. To know how robust our system is to the differences in time stamps between object monitoring and the wearable, it is important to know how this time difference affects identification accuracy. Researchers have explored many systems for performing Non Intrusive Load Monitoring of appliances in a home. These systems are often

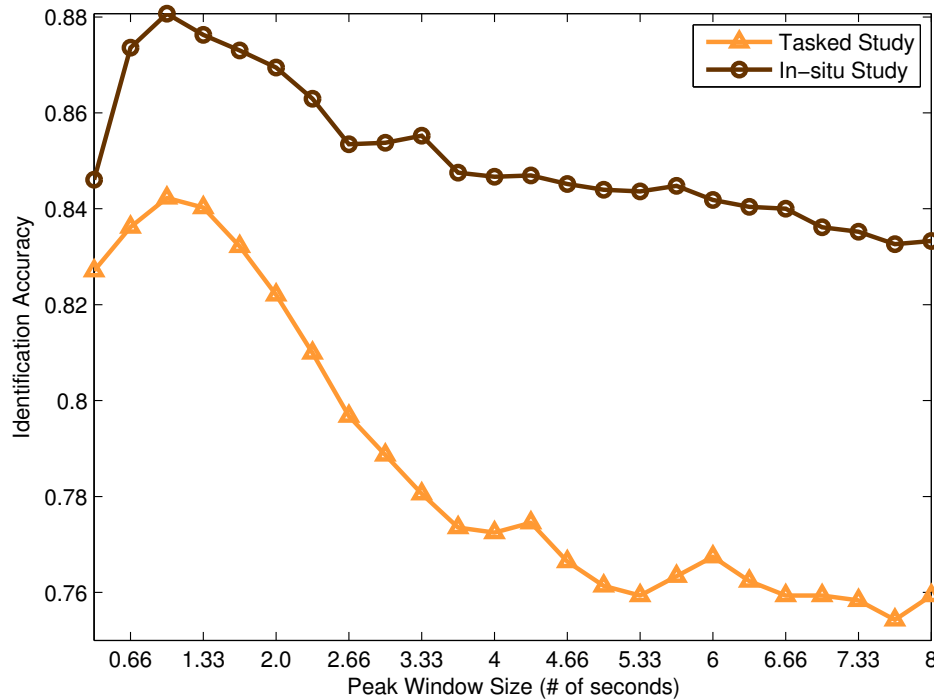


Figure 5.12: Accuracy was highest with a frame size of 1 second around each peak of the Linear Acceleration RSS. This may be due to the fact that time taken by people to physically interact with objects in the home is usually small

not very accurate in terms of the reported time of object event. In such a scenario, we evaluate how well our system can perform with increasing uncertainties about the actual time of event. If the uncertainty is larger, we need to extract a bigger portion of sensor data from the wearables of all the occupants. The problem with this is, if there are multiple people actively using objects in a home at the same time, the chances of identifying the wrong person increases.

According to Figure 5.11, extracting sensor data of window size ± 1 , ± 2 and ± 3 seconds gives low accuracy because the direct sensors and the wearables had a known timestamp uncertainty of ± 3 seconds. Therefore, collecting data using a ± 4 seconds window resulted in the highest identification accuracy. It is the largest window size that encapsulates the time difference between the two systems, as well as small enough to not have overlapping object usage event by another individual.

Once we extract sensor data from the smart watch, we perform feature extraction on

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

5.6 Discussion

In this work, we investigated the hypothesis that we can accurately identify object users in a home by mining the IMU data from the wearables of home occupants. Despite a large number of simultaneous object usages by the participants (50% of object usages were less than 5 seconds apart) in our task-based study, we were able to identify the correct individual with high accuracy. Based on our prior observations of how infrequently people actually use objects in homes (Figure 5.6), we expect accuracy of our approach to increase much more in real life.

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

Our approach is not restricted to objects within homes only. Our approach is general enough to be applied to other scenarios where there are physical interfaces, and we want to know who is using the interface, given a fixed set of people, all wearing smart wrist

devices. For example, this approach can be used to detect who is currently writing on a smart whiteboard in a conference room.

5.7 Limitations

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

5.8 Conclusion

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

Chapter 6

Level 3 - Using Raw Sensor Data to Authenticate Object User

I

6.1 Introduction

One of the main concerns for the gesture based object user identification approach is that for applications concerning security, privacy or confidentiality of information, these methods may not be sufficient. A device can be spoofed to send a higher gesture matching score to get identified as the object user. Therefore, for applications where identity of a user has to be validated, the object initiates a third level of interaction with the Wearable that has been identified as the object user. In this case, the object asks the Wearable to return the raw IMU sensor data from a time window that extend from before the object was used to just after it was used. The object computer certain features of the raw data, as well as, correlates it with its own IMU data to authenticate the object user. We evaluated this approach with Smartphones, where we intend to authenticate the owner of the phone, vs. other phone users.

According to a recent survey [87], existing authentication procedures are considered cumbersome by as many as 30% of smartphone users who leave their phones unlocked.

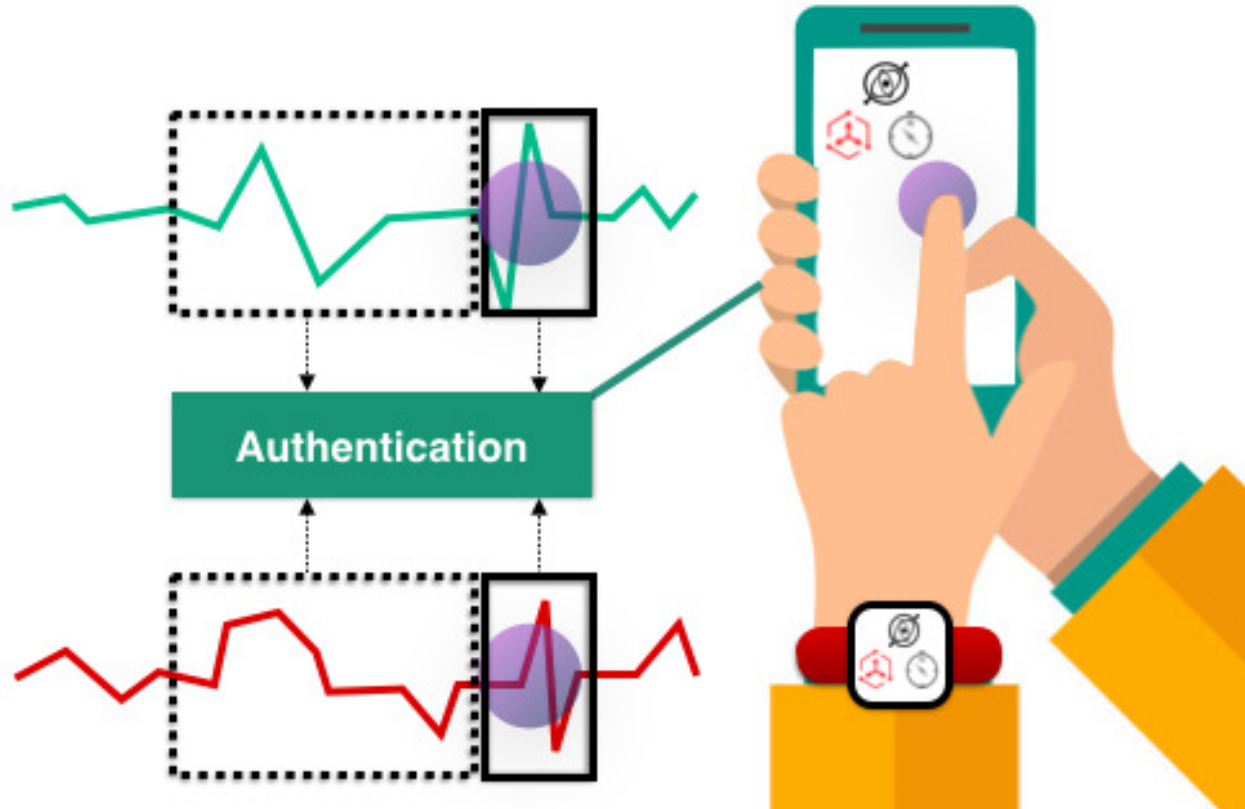


Figure 6.1: The third level of interaction correlates motion and orientation of the phone and the wearable to authenticate smartphone user. Features from the sensor data are extracted separately before and after the time when the touch event is initiated

As a result, smartphone manufacturers have developed different types of authentication mechanisms [88] such as patterns, pin, facial recognition etc.

Most of the commercially available authentication mechanisms fall under one of two categories:

1. Easy to use - for e.g., using presence of a nearby bluetooth device to unlock. In this, easier the unlocking mechanism, higher the security risk.
2. One-time secure unlock - for e.g. patterns or pin. In this, once the device is unlocked, it is not locked until the user shuts the screen off. Sharing phone with friends, for e.g., can allow them access all the data on the phone. In fact sharing a photo with a person nearby, while making sure they cannot swipe and access other photo, is a commonly recognized problem.

Apps such as Focus [89] have been created that allows the user to restrict a friend's access to a predetermined set of photos.

In this work, we propose a different type of approach - one that not only offers ease of use, but can also continuously guarantee the security of the phone's usage by making sure that every touch interaction on the phone belong to the phone's owner. We propose to do so using a paired wearable device being worn by the owner. The intuition behind our system is that a wearable is usually present on the body of the phone owner, and its sensors can be used to determine if the owner is actually interacting with the phone.

Using a wearable device to authenticate a phone may appear trivial, and the solution simple, but there are many challenges that need to be addressed to design a robust solution. An example of an obvious wearables based method is the use of Received Signal Strength Indicator, a value that indicates the level of proximity of the wearable to the phone. If the phone is detected to be close enough to the owner then it is unlocked. However, a major caveat of this method is that proximity of the phone does not guarantee possession of the phone. A phone could be close proximity of its owner (for example, on a coffee table), and yet be accessible for use by others in the same location (other guests in the living room).

Another, obvious wearables based authentication appears to be the use of temporal correlation the touch event in the phone with a motion event on the wearable device. Temporally co-occurring motion and touch event does not imply that the device is in the possession of its owner, as the owner could be performing some other activity resulting in a motion at the same time that another person uses the phone.

In this work, we hypothesize that it is possible to correlate the motion and orientation of the wearable and the phone, in order to authenticate the user, while ensuring a low false acceptance rate. We evaluate the performance of a set of phone and wearable IMU sensor features through three different user studies. We first perform a feasibility study with 13 participants who use the phone in different hand positions (Figure 6.2), where we test the performance of different classifiers, as well as, how well the system works when trained with a

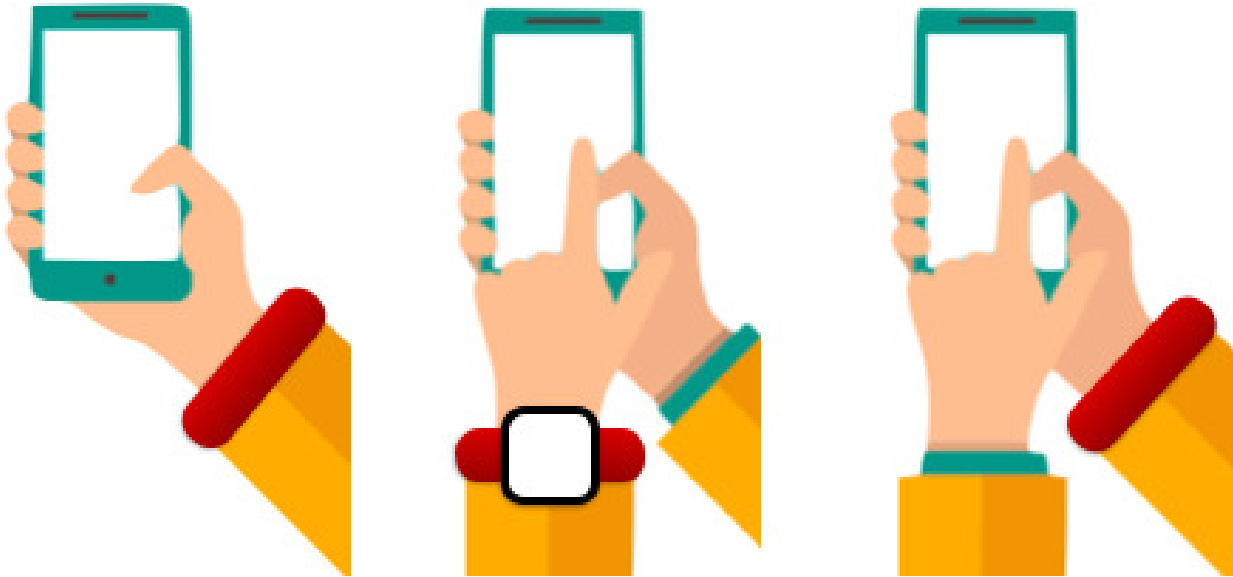


Figure 6.2: Our approach works in different hand positions - the wearable may be worn in the same hand as holding the phone, on the hand touching the phone, and the hand holding and touching the phone. The only position in which our approach doesn't work is if the non-wearable hand holds as well as touches the phone

person's data, and when trained on other persons' data. We then created a proof-of-concept system, and evaluated its performance with 10 participants using 9 sets of phone interactions tasks. The system was not trained on any of their data previously to test if the users naturally learned how to interact with the system. We also conducted an adversarial study with 10 users. In this study, we gave the phone to an adversary who tried to use the phone as the actual phone user was engaged in performing other activities. A prototype of this concept was recently demonstrated [90] at UbiComp'16 Demo Session, where many visitors who were initially skeptical about the performance of this system were convinced that it worked when they tried to use the phone as an owner, as well as adversary.

The main contributions of this work:

1. Algorithms for determining smartphone usage using IMU sensor on phone and a smart-watch paired
2. A study (with 5000 touch samples spanning 13 individuals) and analysis of features differentiating click vs. other activities

3. A prototype use-case study with 10 individuals and analysis of the learnability of the approach
4. An adversarial analysis with 10 individuals evaluating the false acceptance rate of the approach

6.2 Related Work

Commercial smartphone manufacturers have included a wide range of device unlocking mechanisms [88]. Some of these are - patterns, pin code, facial recognition, etc. Some LG devices even have a unique knock code to unlock the device. Android has released a new SmartLock feature [91] that does not lock the phone when it is within range of another bluetooth device, e.g. a smart watch. While this is convenient, any other person in the same location can also unlock the phone.

There have been many research prototypes that aim to implicitly authenticate users on smart objects. Fiberio [63] proposes a rear-projected multitouch table that identifies users biometrically based on their fingerprints during each touch interaction using a large fiber optic plate that diffuses light on transmission, thereby allowing it to act as projection surface. Capacitive fingerprinting [64] proposes a novel sensing approach based on Swept Frequency Capacitive Sensing, which measures the impedance of a user to the environment across a range of AC frequencies. Different people have different impedance profiles since they have different bone densities and muscle mass yields.

The use of a trusted smart glove [65] was proposed where individual fingers are tagged separately. A tag-aware surface tracks the location and orientation of these tags, while identifying the wearer of the glove. Phone touch [3] is a technique which is based on separate detection of phone touch events to identify multi-user interaction on a surface, which determines location of the touch, and by the phone, which contributes device identity. The

device-level observations are merged based on correlation in time. YouTouch [66] proposes the use of a commodity RGB + depth camera in front of the wall, to track users on a common collaborative surface and correlate them with touch events. Whenever the user's ID is occluded from the camera, they propose re-identifying the person by means of a color histograms of body parts and skeleton-based biometric measurements.

Specifically, when applied to smartphone authentication, many research prototypes attempt to use different touch or behavioral biometrics based methods.

SilentSense [62] proposes a framework to authenticate users implicitly using the user touch behavior biometrics and the micro-movement of the device caused by user's screen-touch actions. A touch-based biometrics model of the owner is developed by extracting some principle features, and then used to verify whether the current user is the owner or guest/attacker. Another work proposes [92] to authenticate users mainly based on how they input, using distinguishing features such as finger velocity, device acceleration, and stroke time. Some other work has also been done to authenticate a user using their tapping behavior [93, 94, 95]. It is however interesting to note that a mimicry attack based study [96] has demonstrated that touch based implicit authentication fails against shoulder surfing and offline training attacks.

Some proposed methods attempt to use the phone's app usage and calling behavior to authenticate users [67, 68]. However, this methodology takes a significant amount of time to establish authentication, and therefore although it is implicit based, it cannot be used to continuously authenticate users on smartphones.

Another thrust in implicit authentication attempts to use gait based features to authenticate users as they are walking [69, 70]. However since the gaits of people are different on different types of surfaces such as grass, road, snow, wet surface, and slippery surface, these techniques have low true acceptance rates.

In a work that is closely related to ours, researchers [97] propose authenticating a user by initiating a challenge-response protocol on-screen, requiring the user to perform a particular

sequence of gestures. While interacting with this control, the server is continuously reading IMU sensor data from connected wrist devices. It compares this sensor data with the touch data transmitted by the client device and then calculates the correlation between them to determine if the user is authenticated. This approach differs from ours in that it requires the user to explicitly engage with the system for authentication.

6.3 Approach

Our approach is based on mainly the following three observations. When people wearing a wearable device, use a phone, both the devices are present on the hand of the same person, and therefore subject to similar patterns of linear acceleration. These devices also register similar angular velocity patterns as the person's hand makes small adjustments while interacting with the phone. Furthermore, as human hands are constrained to face a similar direction during phone usage, these devices are also coupled tightly in terms of their absolute orientations. As a consequence, when a person with a wearable device also uses a phone, we expect strong correlation in the IMU features of a phone and the wearable. Therefore, if we measure appropriate IMU features of the two devices, it should be possible to authenticate the owner of the phone.

IMU sensor is a set of three individual sensors - accelerometer, gyroscope and magnetometer, that allow it to capture different aspects of a device's motion and position in space. Since we want to ensure minimal delay in authentication process, in this paper, we focus on time-domain IMU features. The delay introduced due to a continuous authentication procedure such as ours, is perhaps more noticeable than a traditional one-time unlock scheme, and therefore it needs to be minimal so that the end user does not feel interference with their ability to interact with the device.

While it is important to explore the correlation of the smartphone's IMU data with that of the watch, we want to make sure that the smartphone's IMU data is not correlated to a

smartwatch's IMU when the user is performing other tasks - such as walking, cleaning dishes. This would prevent a malicious user with physical access to the phone from being able to interact with it while the owner of the phone is unaware, and performing other activities.

6.3.1 System: Hardware and Sensor

In this paper, we explore the idea of implicitly authenticating every touch on a smartphone using a wearable device. Our approach relies on time-domain IMU features on both the wearable and the phone. Specifically, we are looking at linear acceleration, angular rotation, and orientation in the world coordinate system. While angular rotation is directly obtained from the gyroscope sensor in the IMU, linear acceleration and orientation are derived using sensor fusion.

We use a Shimmer [98], which is a IMU device that calculates quaternions using IMU sensor fusion algorithms. Using the quaternions, we calculate the gravity acting on the device's frame of reference, and therefore calculate linear acceleration by subtracting gravity from absolute acceleration. We also use the quaternions to calculate the device's absolute orientation, expressed in terms of Euler Angles. For the smartphone, we use a Nexus 6 (Android). Android API offers multiple software sensors, which allowed us to directly obtain the phone's linear acceleration and orientation.

We sampled the IMU sensor on both devices at 100 Hz. This sampling frequency is required in order to respond to every touch within time bounds of a person's visual perception. We don't expect this to remain a static sampling frequency. In fact, when the devices are not being used, they can be sampled at a much lower frequency in order to conserve battery. In the current state of work, we set up the wearable (Shimmer) to continuously transmit all the raw data to the phone, where each sample was marked with the received timestamp. The received timestamp removed the need to synchronize the clocks of the two devices, and also included the latency in the Bluetooth communication.

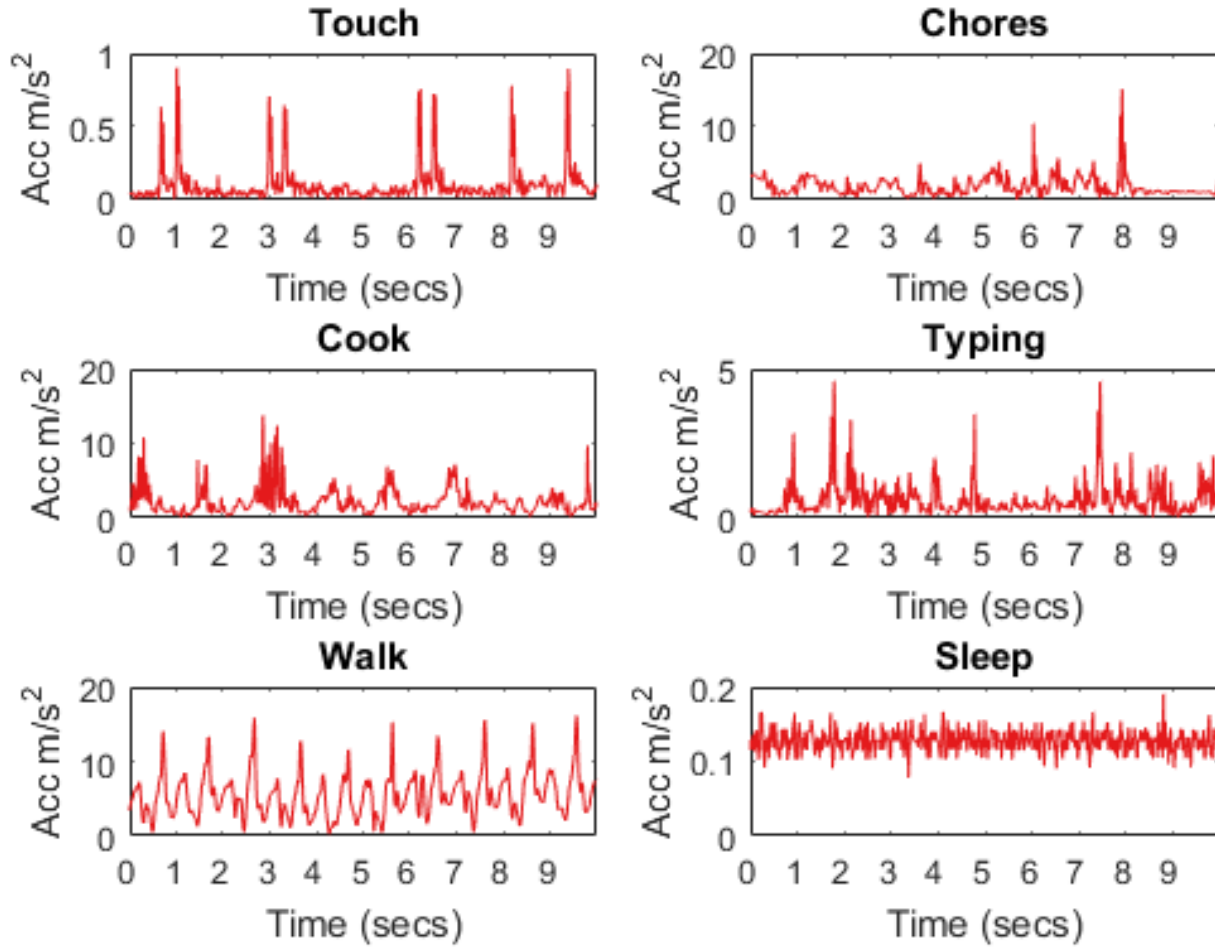


Figure 6.3: Wearable’s linear acceleration patterns: A snapshot of 10 seconds compares the acceleration signatures of touch interactions as well as other activities. The touch signal consists of 8 phone touches which can be observed by the 8 peaks in the signal

6.3.2 Features used in Authenticating User

To authenticate a person on the smartphone, our approach co-relates the IMU data from the paired wearable and the phone. Every time that a touch is sensed on the smartphone screen, the system initiates the authentication method. The authentication method takes into account the IMU data from both the devices from a time period before the touch (*beforeT*), and waits for some time to get the data after the touch (*afterT*).

It then calculates individual features for the data obtained from before and after the touch, and uses a rule based classifier to disambiguate whether to permit the touch.

In this approach, we want to correlate not only the IMU data during and just after the touch is initiated, but also before. This is because of the observation that when a person is about to use the phone, they typically hold the phone for some time while locating the region of the screen where they want to touch. We hypothesize that this results in a certain amount of correlation between the phone and the wearable even before the touch.

The following sensor data is used for feature extraction - Root Mean Square of Linear Acceleration (*Acc*), Root Mean Square of Angular Velocity (*Gyro*), Yaw (*Yaw*), Pitch (*Pitch*), and Roll (*Roll*). This data is obtained for both the wearable (*W*) and the phone (*P*). Therefore, every time a touch is initiated, the data streams obtained are:

- Linear Acceleration before touch - $Acc_{beforeT}^P, Acc_{beforeT}^W$
- Angular Velocity before touch - $Gyro_{beforeT}^P, Gyro_{beforeT}^W$
- Yaw before touch - $Yaw_{beforeT}^P, Yaw_{beforeT}^W$
- Roll before touch - $Roll_{beforeT}^P, Roll_{beforeT}^W$
- Pitch 500ms before touch - $Pitch_{beforeT}^P, Pitch_{beforeT}^W$
- Linear Acceleration 120ms after touch - $Acc_{afterT}^P, Acc_{afterT}^W$
- Angular Velocity 120ms after touch - $Gyro_{afterT}^P, Gyro_{afterT}^W$
- Yaw 120ms after touch - $Yaw_{afterT}^P, Yaw_{afterT}^W$
- Roll 120ms after touch - $Roll_{afterT}^P, Roll_{afterT}^W$
- Pitch 120ms after touch - $Pitch_{afterT}^P, Pitch_{afterT}^W$

Different features are computed on this data to encapsulate different intuitions about the data.

1. Wearable signal characterization - The most fundamental insight is that the motion signature of the wearable device is different from motion signature during other activities (as can be seen in Figure 6.3). To characterize the raw signal, we compute the following features on the following data streams:

1. Maximum value of raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W, Yaw_{beforeT}^W, Yaw_{afterT}^W, Roll_{beforeT}^W, Roll_{afterT}^W, Pitch_{beforeT}^W, Pitch_{afterT}^W$

2. Minimum value of raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W, Yaw_{beforeT}^W, Yaw_{afterT}^W, Roll_{beforeT}^W, Roll_{afterT}^W, Pitch_{beforeT}^W, Pitch_{afterT}^W$

3. 25th, 50th and 75th percentile value of raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W, Yaw_{beforeT}^W, Yaw_{afterT}^W, Roll_{beforeT}^W, Roll_{afterT}^W, Pitch_{beforeT}^W, Pitch_{afterT}^W, Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W, Yaw_{beforeT}^W, Yaw_{afterT}^W, Roll_{beforeT}^W, Roll_{afterT}^W, Pitch_{beforeT}^W, Pitch_{afterT}^W, Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W, Yaw_{beforeT}^W, Yaw_{afterT}^W, Roll_{beforeT}^W, Roll_{afterT}^W, Pitch_{beforeT}^W, Pitch_{afterT}^W$

4. Average of raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W$

5. Standard deviation of raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W, Yaw_{beforeT}^W, Yaw_{afterT}^W, Roll_{beforeT}^W, Roll_{afterT}^W, Pitch_{beforeT}^W, Pitch_{afterT}^W$

The wearable's linear acceleration and angular velocity also registers peaks, which are points where the motion abruptly changes direction. The number of peaks within a time frame indicates the number of times the motion changes direction. The average height of the peaks indicate the magnitude of the motion that is changing, and the prominence characterizes the relative height of the peak with respect to the lowest contour line. We characterize some of the peaks observed in the signal of the following data:

6. Number of peaks in raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W$

7. Average height of peaks in raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W$

8. Average prominence of peaks in raw signal:

$Acc_{beforeT}^W, Acc_{afterT}^W, Gyro_{beforeT}^W, Gyro_{afterT}^W$

In addition to characterizing the wearable's motion and orientation, we also correlate the IMU sensor data of the wearable and the phone. We correlate the signals in place, as well as by shifting the two signals. This is to account for the latency in the wearable data that may cause similar events to register at slightly different times on the phone. We shift one of the signals from -50 ms to 50 ms, and calculate the maximum correlation resulting between the two signals. The correlation features are computed on the following data:

9. Correlation in raw signal:

$Acc_{beforeT}^W$ and $Acc_{beforeT}^P$, Acc_{afterT}^W and Acc_{afterT}^P , $Gyro_{beforeT}^W$ and $Gyro_{beforeT}^P$, $Gyro_{afterT}^W$ and $Gyro_{afterT}^P$, $Acc_{beforeT}^W$ and $Gyro_{beforeT}^W$, Acc_{afterT}^W and $Gyro_{afterT}^W$

10. Max correlation in shifted raw signal:

Acc_{afterT}^W and Acc_{afterT}^P , $Gyro_{afterT}^W$ and $Gyro_{afterT}^P$, Acc_{afterT}^W and $Gyro_{afterT}^W$

Finally, the difference in the orientations of the wearable and the phone is computed, before and after the touch event.

11. Difference in orientations of devices:

$Yaw_{beforeT}^W$ and $Yaw_{beforeT}^P$, Yaw_{afterT}^W and Yaw_{afterT}^P , $Roll_{beforeT}^W$ and $Roll_{beforeT}^P$, $Roll_{afterT}^W$ and $Roll_{afterT}^P$, $Pitch_{beforeT}^W$ and $Pitch_{beforeT}^P$, $Pitch_{afterT}^W$ and $Pitch_{afterT}^P$

6.3.3 Prototype

We created a proof-of-concept of our approach using an Android app. The app had a set of multiple choice questions, and text for people to scroll and read. The app was built on top of an existing app that received Shimmer's IMU data continuously on an Android phone and logged it. When a person touches one of the UI components in the app, such as the radio button, if the system authenticates the touch, then the UI components react the touch normally, otherwise, the UI responds with a deselected component to indicate that the touch was not validated.

6.4 Feasibility Study

To evaluate the feasibility and inform the selection of classifier for our approach, we conducted a 13 person study.

6.4.1 Procedure

Our approach is mainly evaluated on its the ability to validate every touch event with the IMU sensor data that is generated when the wearable device is used to either hold the phone or interact with the phone. The touch event should not be validated by the IMU sensor data generated when the phone's owner is engaged in other activities.

The aim of this study was to answer the following questions:

1. What is the performance of different classifiers in differentiating between valid and non valid touch events?
2. How well does the system perform when it is trained on a user's touch data?
3. How well does the system perform when it is not trained on a user's touch data?
4. What is the effect of the amount of training data on the accuracy of a system, w.r.t. a new user?

To evaluate the features in differentiating between these validating and non-validating features, we collected data of valid use of phone by 13 participants, as well as IMU sensor data from five activities (cooking, typing, walking, household chores, sleeping) being performed in-situ for an hour each by a participant. The acceleration patterns can be seen in Figure 6.3.

To collect valid IMU sensor data, we asked participants to wear the wearable device, and use an app developed by us. The task given to the participants was to answer a few multiple choice questions, and scroll through some text. The participants were asked to repeat this task six times. Each time, the participant changed the hand positions of the devices. The participants used the devices in different combinations of - a. which hand holds the phone, b. which wears the smartwatch, and, c. which hand is used to interact with the phone. The

combinations for the hand holding the phone (H_p), the hand wearing the watch (H_w), and the hand touching the phone screen (H_t) are expressed in tuples of (H_p, H_w, H_t) , where the hand H can be Left(L) or Right(R). For eg., LLL means. the phone is held in the left hand, watch worn on the left hand, and the left hand is used to interact with the phone. There are eight combinations - LLL, RRR, LLR, RRL, RLL, LRR, LRL and RLR. In our evaluation, we included six of the eight combinations. Our approach cannot detect phone usage in configurations LRL and RLR, as the watch is worn on a non-moving hand, that cannot detect phone usage. Overall, this study resulted in 5000 touches across all participants.

For each touch, we compute the features discussed earlier, and generate a valid touch dataset. To compute the negative examples, we use the phone's IMU data and correlate it with IMU sensor data obtained from the insitu activities (such as cooking, walking etc.) in order to generate an invalid touch dataset. We then evaluate the ability of the classifiers in differentiate between the touches in the valid and invalid datasets.

6.4.2 Analysis

We trained different different classifiers on the dataset containing the valid and invalid touch features, and evaluated them using 10-fold cross validation. Each classifier was evaluated on two metrics - True Acceptance Rate (TAR), and False Acceptance Rate (FAR). True Acceptance Rate is defined as the ratio of the number of true acceptances divided by the number of identification attempts. False Acceptance Rate is defined as the ratio of the number of false acceptances divided by the number of identification attempts. Ideally, an authentication system should have a high TAR and a low FAR.

The first question that we aimed to answer was, how well can we classify and differentiate between the validated touch features and the invalid touch features. We tested several popularly used Machine Learning algorithms in Weka. Since the dataset contained unbalanced classes, we first balanced the classes using SpreadSubsample [99], which randomly samples and reduces the majority class to match the size of the minority class. The dataset was

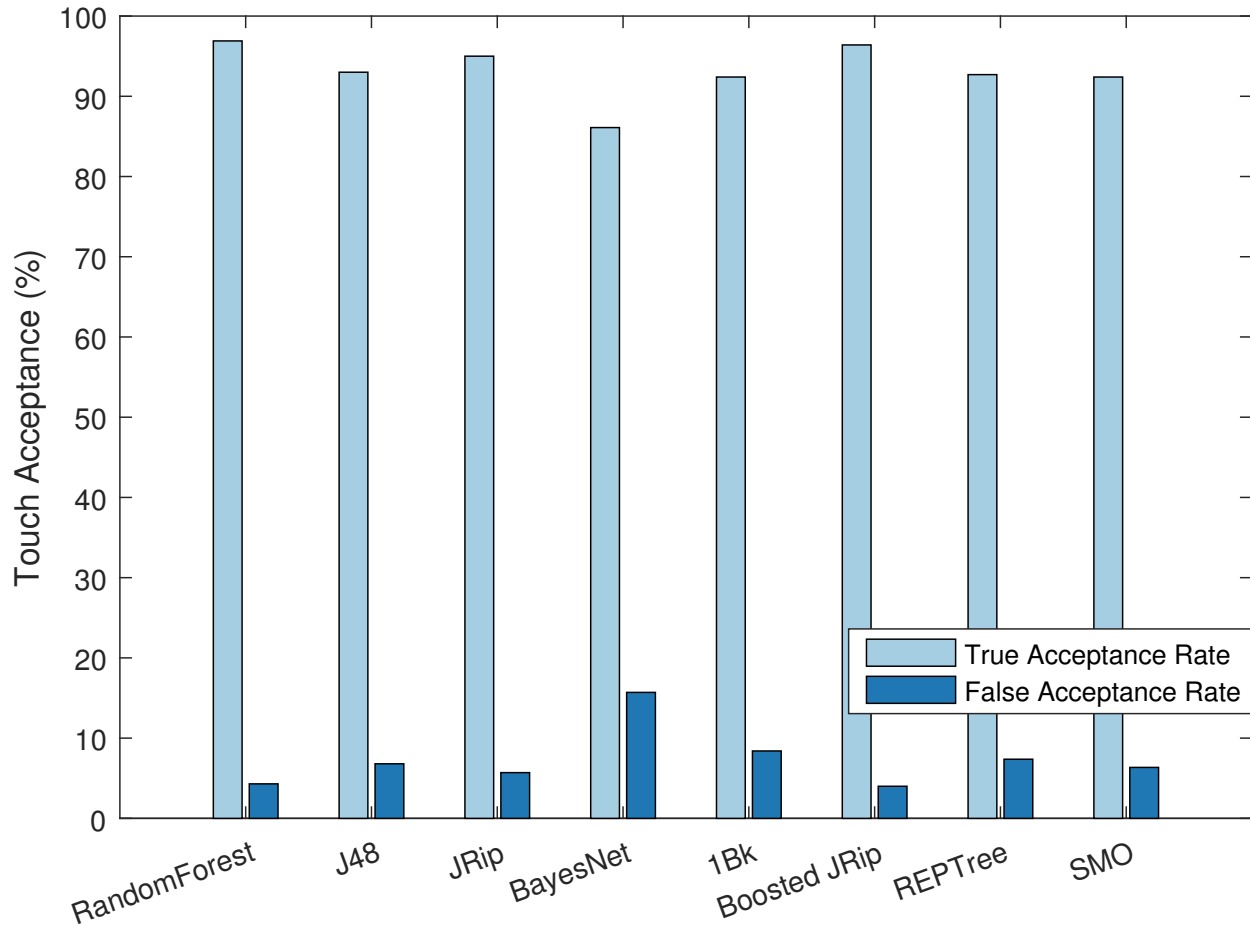


Figure 6.4: Classifier evaluation: Random Forest and Boosted JRip classifiers have the highest true acceptance rate, as well as the lowest false acceptance rates, making them better suited for an authentication system. While Random Forest slightly outperforms the Boosted JRip classifier, we use the latter in our Android based prototype, for faster response time

evaluated using 10-fold cross validation. The classifiers that we tested were - Random Forest, J48, JRip, BayesNet, 1Bk, Ada1 Boosted JRip, REPTree, and SMO. The results can be seen in Figure 6.4. In terms of TAR, Random Forest and Boosted JRip perform the best with 97% and 96% accuracy in identifying true touch features. These two classifiers also have the lowest FAR at 4% and 4.5%. While Random Forest performs slightly better than Boosted JRip, we selected the Boosted JRip classifier for developing our proof-of-concept system. This was because Weka outputs the classifier model generated by Random Forest as a file that needs to be executed with a Random Forest classifier JAR. In contrast, boosted JRip output a set

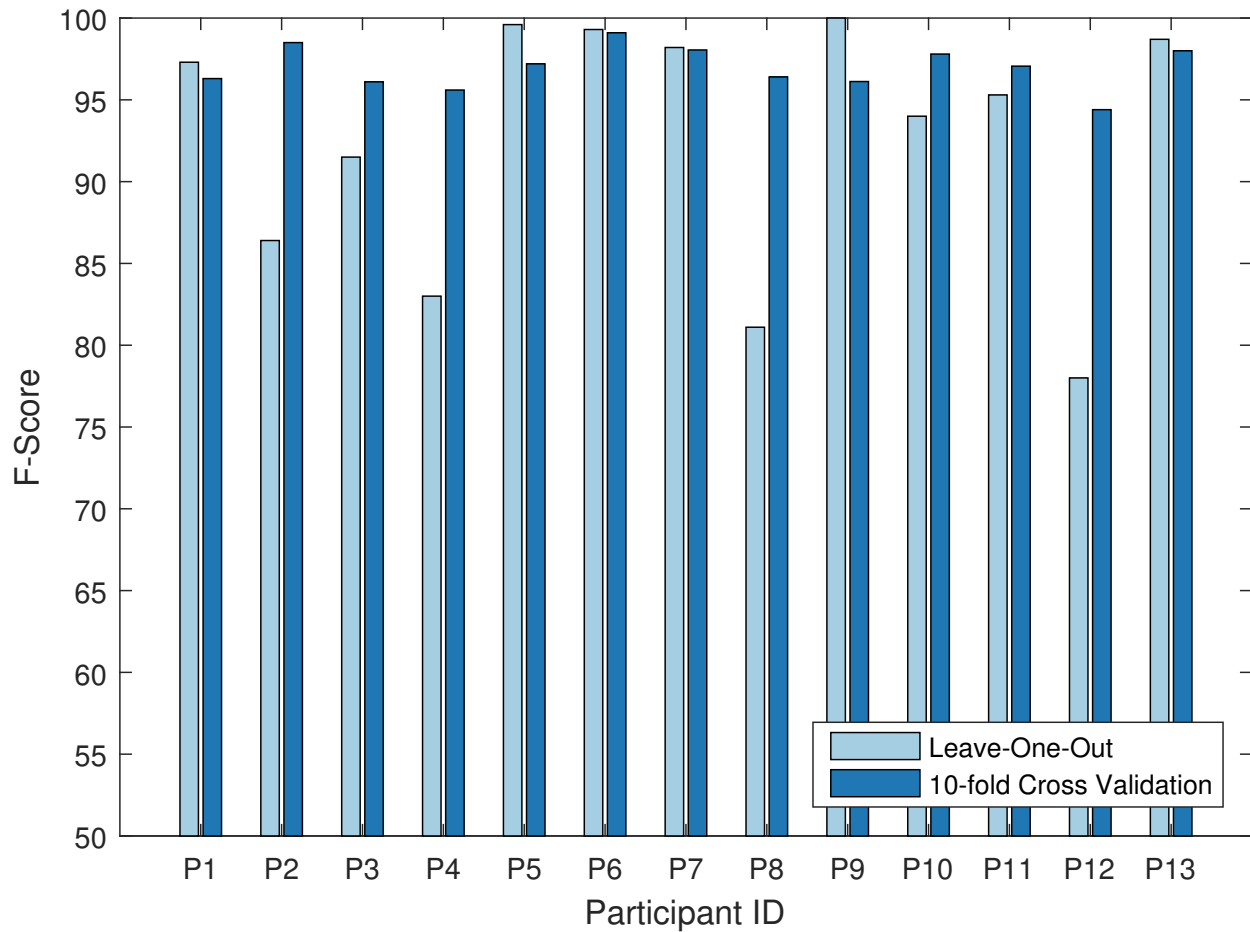


Figure 6.5: Per participant analysis: Training per participant performs with an average f-score value of 96.4%, while leave-one-out (LOO) method performs with an average f-score value of 92.3%. LOO indicates that learning valid touch features from other participants may be suboptimal for some participants such as P2, P4, P8 and P12

of rules for classification. These rules can be easily imported to Java code, and incorporated directly in the Android App's code. In a time critical application such as authentication, having a classifier directly embedded in code is much faster than having to invoke a jar to classify data. We continue to use a boosted JRip for the rest of the analysis in this paper.

We performed sensitivity analysis of the look back time window from the time of touch (*beforeT*), and varied it from 300 ms to 1.5 seconds. We also varied the look ahead window from the time of touch (*afterT*), from 120 ms to 240 ms. Although, there didn't appear to be a significant difference in varying these time windows, the results indicated a slightly

higher accuracy at a 600 ms look back time, and 120 ms look ahead time. We used these values in generating in the feature extraction stage.

The next question that we wanted to answer was how well does the classifier perform when trained on individual participants. For this analysis, we created 13 different datasets, one for each participant. Each dataset contained valid IMU features from a single participant, and the invalid IMU features from the insitu data of other activities. The classes were balanced using SpreadSubsample. We used a boosted JRip classifier and measured the classifier's performance in terms of F-Score using 10-fold cross validation. F-Score is defined as the harmonic mean of precision and recall of a classifier's output. As can be seen in the Figure 6.5, the classifier achieved an average of 95% accuracy per participant. For participants P2, P6, P7, P10 and P13, it actually performs at 97% accuracy or higher. The worst f-score is obtained for P12, at 94%.

While the previous analysis established that it was indeed possible for a classifier to learn every individuals valid touch IMU data features, an important question in an authentication system such as ours is, how well does training the system on a population translate in terms of performance to a person completely new to the system. This is important because it is a measure of how well the system will work out-of-the-box for a new user. For this analysis, we performed a leave-out-out validation study. In this for every participant, we created a test dataset using the participant's valid touch IMU features, and a training dataset using the valid touch IMU features from all other participants, and invalid touch features from the other insitu activities. Figure 6.5 shows the results of the Leave-One-Out analysis. For 5 out of the 13 participants, the classifier performed with an f-score of more than 98%. For four of the participants, the f-score was greater than 90%. However for four of the participants- P2, P4, P8, and P12, the classifier performed poorly with an f-score ranging from 75%-86%. Therefore, while the system was able to account for the way 9 out of 13 participants used the phones, it wasn't able to do so as accurately for 4 participants. However, our earlier analysis did indicate that when we trained the system on these participants' individual data, it was

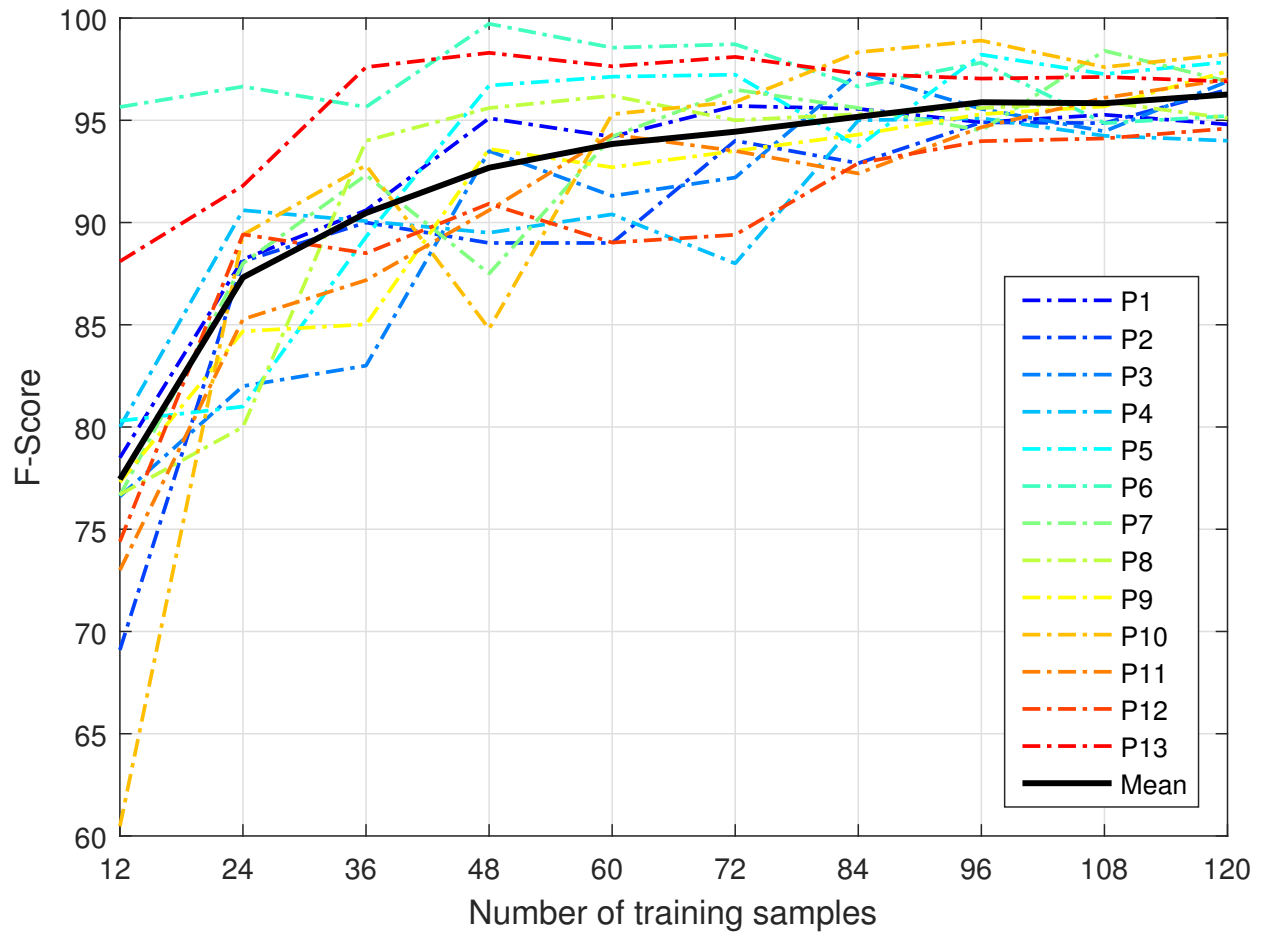


Figure 6.6: Sensitivity analysis of training data: When the number of samples in the training dataset is varied from 12 to 120, we can observe that for some participants such as P7, classifier is able to achieve very high accuracy with very few samples. The classifier achieves an average of 95% accuracy after training on an average of 84 samples for all participants

able to achieve an average of 95% f-score. This indicates that when the system is installed on a new user's phone, while there may be a learning phase required for some users, the system would eventually adapt to the person's individual style of interacting with the phone. If we compare the LOO and the per-person analysis in Figure 6.5, we make an interesting observation - the LOO classifier performs much better than the personalized classifier, for some participants - P5, P6, P7, P9, and P13. This indicates that personalized classifier sometimes over-trained, which the LOO was able to mitigate with a wider variety of phone usage styles incorporated due to data obtained from a larger population.

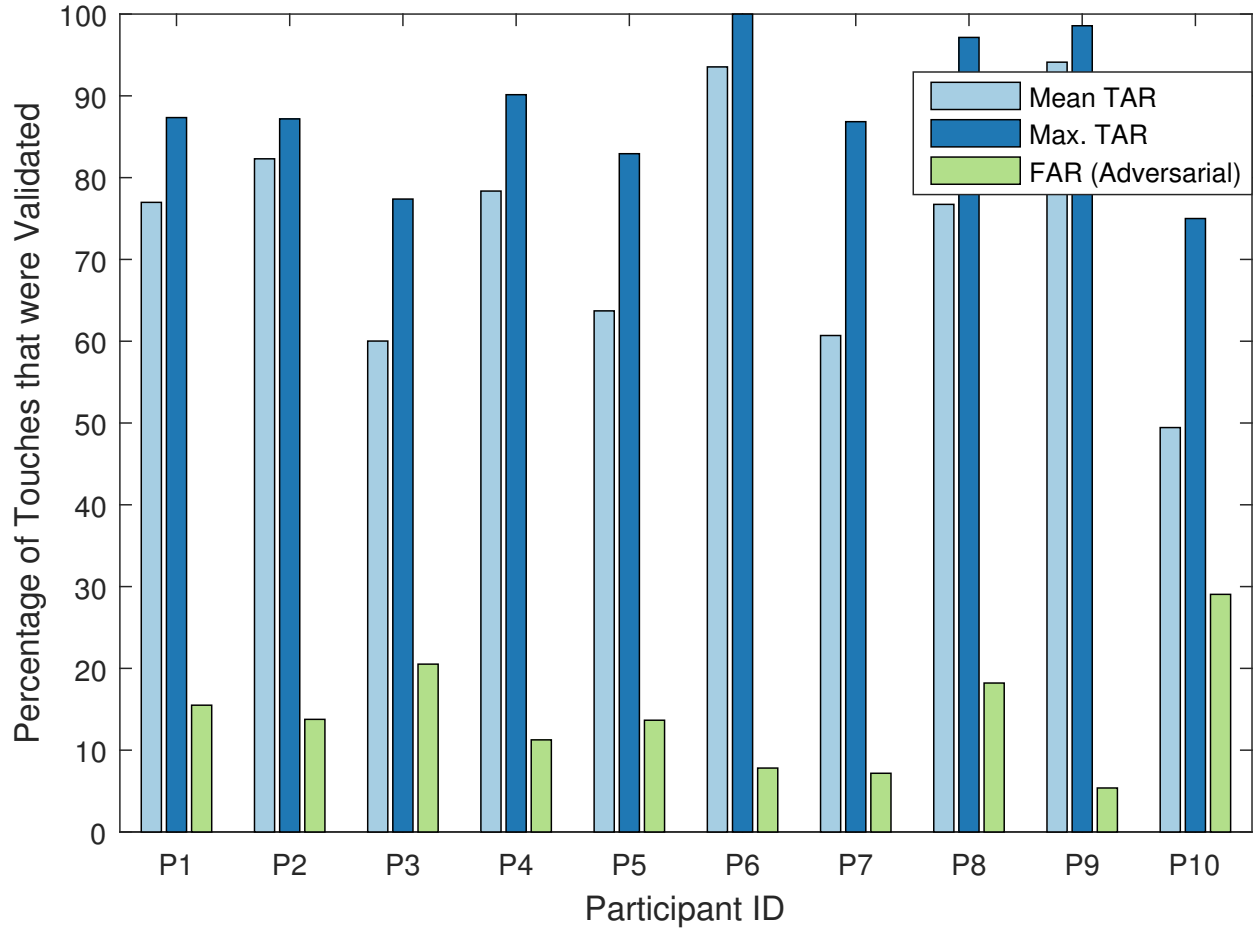


Figure 6.7: Prototype interaction accuracy: The maximum TAR per session indicates that even though initially some users may initially not be able to interact easily on a system that is not trained on them, eventually most of them intuitively figure out a good way to interact with the phone. The adversarial case shows the low false acceptance rate of the system

Since, the system may need to train on an individual’s style of using the phone, we performed a sensitivity analysis of the effect of training data on the accuracy of the system. For every iteration we removed 12 data samples from the person’s set of valid touch samples and added it to the training dataset. The remaining touch data samples were used to create a testing dataset. We varied the number of valid touch IMU data samples from 12 to 120. The classifier was evaluated using f-score, which is the harmonic mean of precision and recall. The results of the sensitivity analysis can be seen in Figure 6.6. On an average, with just 12 samples, the classifier was able perform with a 77% f-score, while with 84 touch samples, the

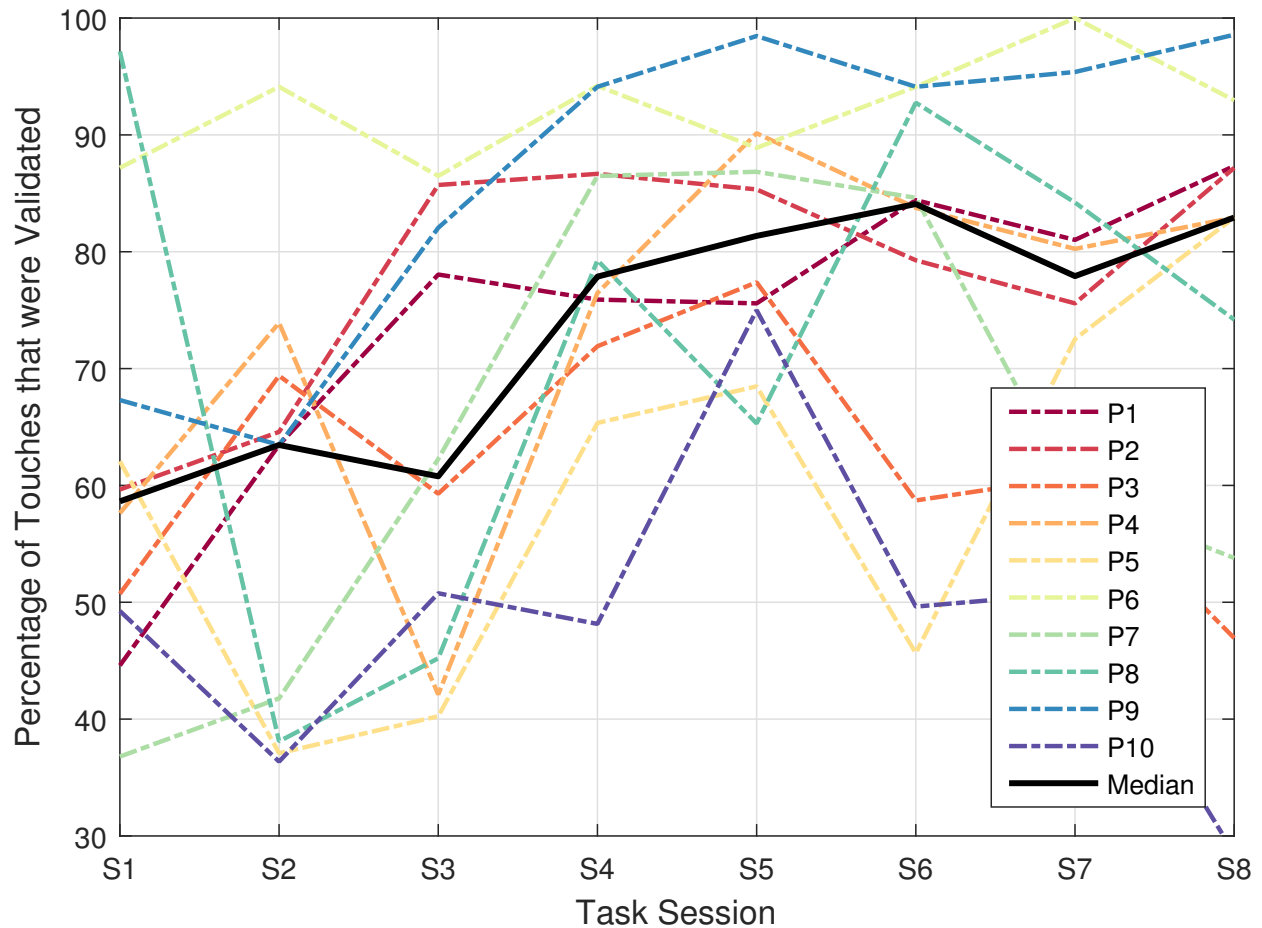


Figure 6.8: Per Session Accuracy: Validation accuracy improves with every session of tasks indicating the participant starts to figure out how to use the system

classifier was able to perform with more than 95% accuracy.

6.5 Prototype Interaction Study

One of the main take-away messages from our feasibility study was that our approach works well when the system trains for the individual using the phone. The leave-one-out analysis revealed that there were users for whom the system does not perform accurately when it didn't train on them. Given this result, we anticipated that when a user encounters a system that is not trained for them, they might learn to adapt to the system instead.

6.5.1 Procedure

Having conducted offline analysis, we decided to actually build a prototype based on the classifier resulting from the previous study. We decided to test this system with users that the classifier was not trained on.

The aim of this study was to answer the following questions:

1. What is the true acceptance rate of the system, when it is used by people that the system is untrained for?
2. What is the perception of the users for a non-personalized system in terms of their ability to get the system to respond, the noticeability of the device's lag in response to their touch, and how acceptable is the lag to them?
3. Do users learn to adapt to the system intuitively, based on the system's authentication response to their touches?

To observe the users' interactions with a non-personalized authentication system, we implemented the Boosted JRip classifier trained on the data from the 13 participants from the feasibility study, in our Android app. The participants were asked to perform touch interactions in our app.

While we had asked our participants to use each of the six hand configurations explicitly in our feasibility study, in this study we let the participants wear the wearable on their non-dominant hand (where people usually prefer wearing a watch), and let them hold the phone as they liked. The app's basic design remained the same - multiple choice questions, where every touch on the radio button or the 'Submit' button was validated using our pre-trained classifier. Each participant had to answer a set of MCQ as a task, and they had to perform a total of 9 sessions of tasks. At the end of each session, we asked the participant to give the phone back to the researcher, and then use their experience interacting with the smartphone to respond to a questionnaire. The questionnaire had three statements:

1. I was able to get the device to respond to my touches.
2. The lag in the device's response to my touches is noticeable to me.

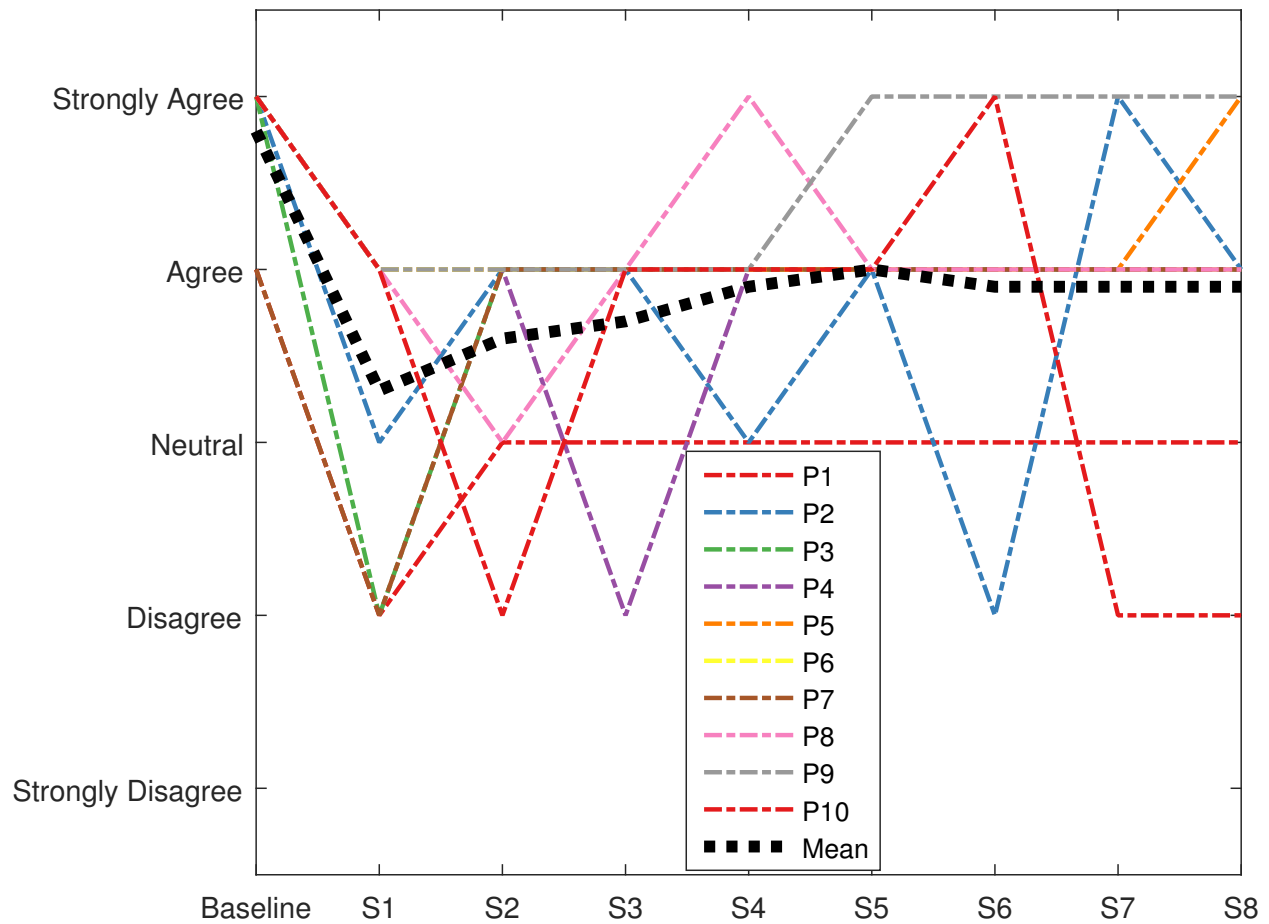


Figure 6.9: Device Response: For the baseline case, almost all participants felt they were able to get the device to respond to their touches. Once the authentication system kicked in, there was a temporary dip in their response, although eventually it settled higher at an average perception of ‘Agree’

3. The lag in the device’s response to my touches is acceptable to me.

The participants were asked to mark their responses to the statements on a five-point Likert scale:

Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree

After responding to the survey, they were handed the phone back by the researcher, and they started to work on a new session of task. Out of the nine sessions of task, the first session did not have the authentication system active. Therefore, all touches made by the user in the first session were authenticated by default. We considered this session as the *baseline*, and compared the participants’ responses to the first session to the responses received in the

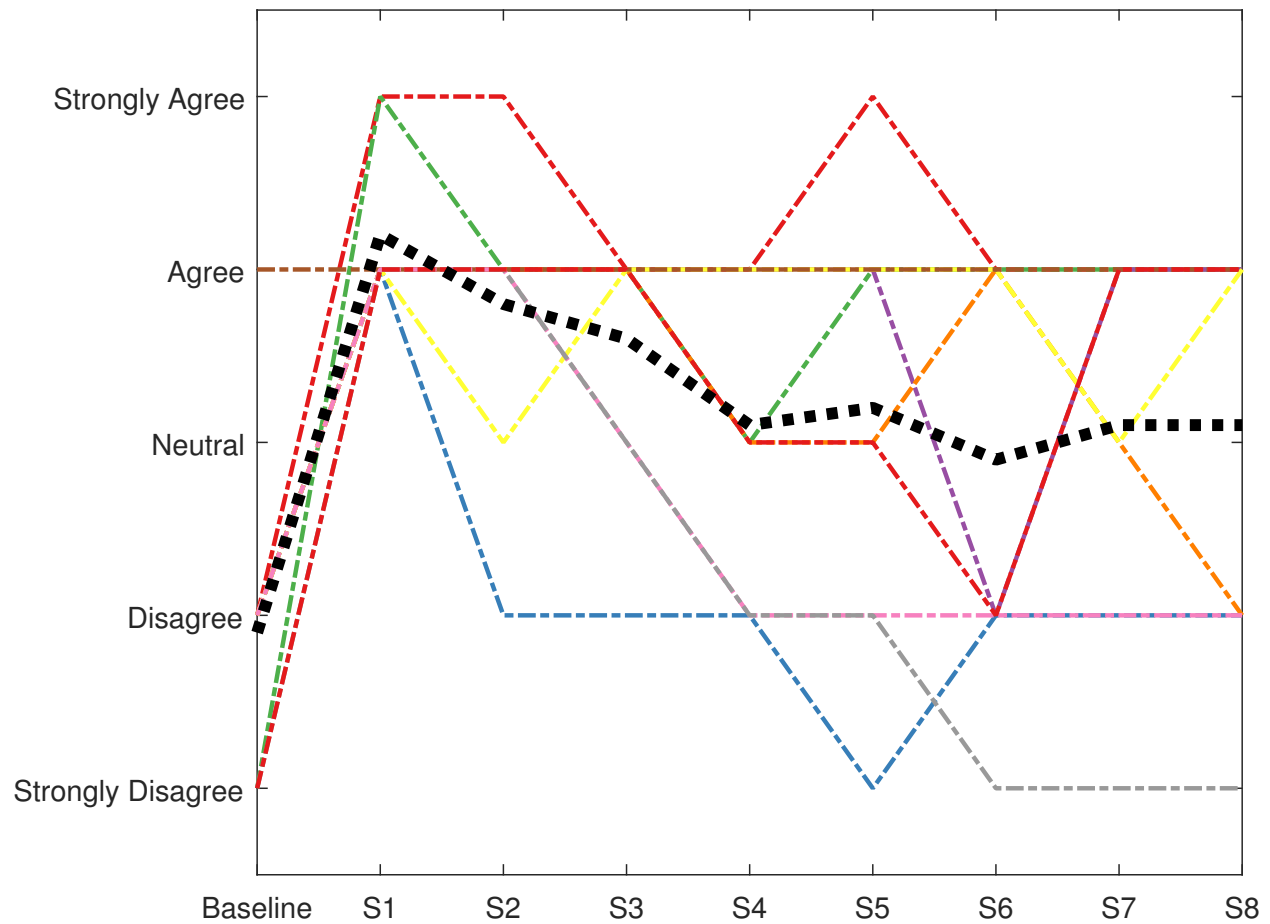


Figure 6.10: Noticeability of device’s lag in response: The baseline response was ‘Disagree’, which is a result of Android’s inherent system delay. When the authentication system kicked in, the average response increased to ‘Agree’, though it eventually settled to ‘Neutral’ by the end of the task sessions

other eight sessions. However, none of the participants were informed of this difference. The participants were informed that in every task, the system would try to authenticate every touch made by them. Their objective was to try to get the device to respond to their touches. If the phone did not seem to respond to their touch, they were free to change how they interacted with the phone to make it work for them. We were interested to see if they were able to retain an idea of the configuration that worked for them, intuitively over different sessions, when they picked up the phone from the researcher and started a new task.

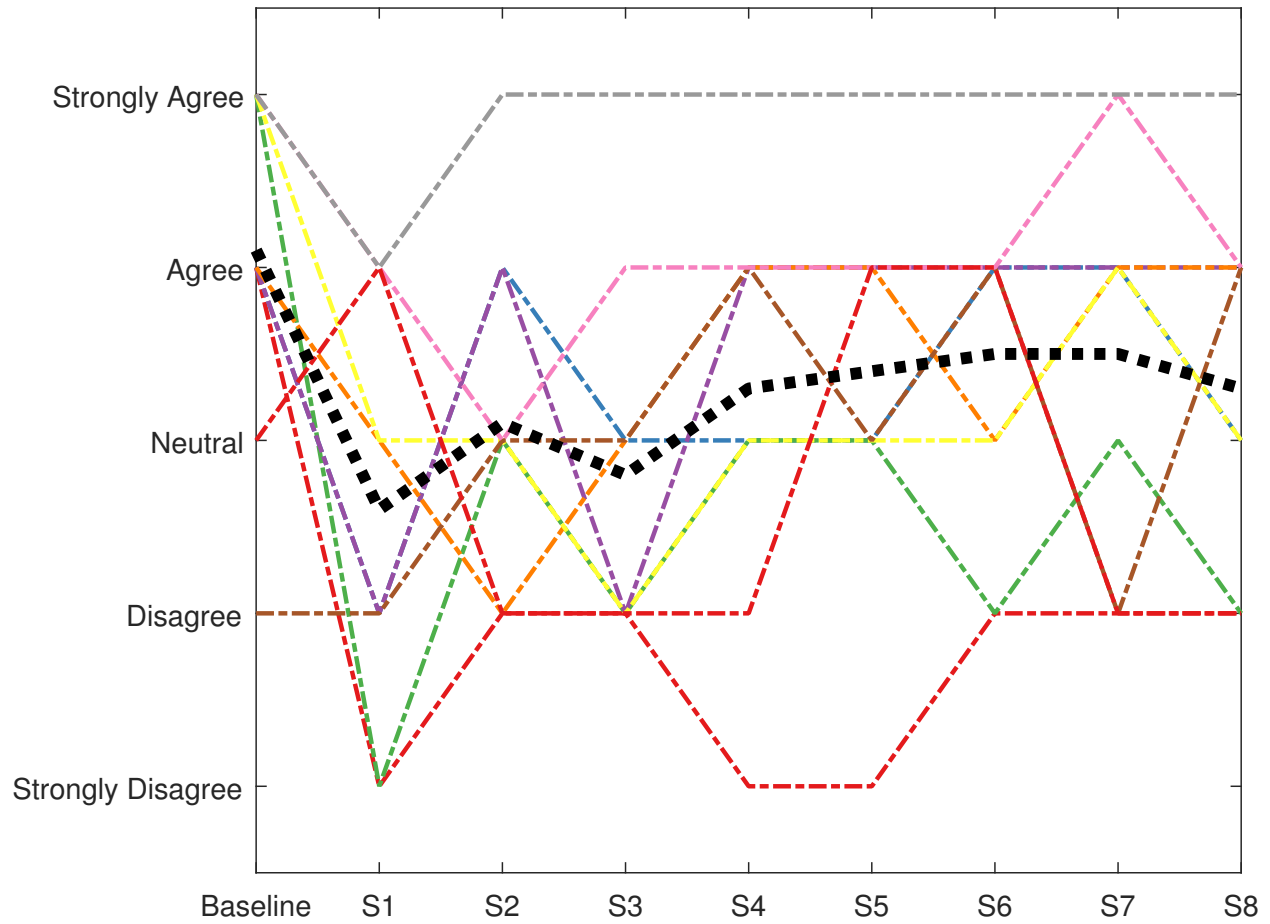


Figure 6.11: Acceptability of device's lag in response: The baseline response was 'Agree'. Once the authentication system kicked in, it dropped to slightly below 'Neutral', and improved and fell a couple time before settling at 'Neutral' by the end of the task sessions

6.5.2 Analysis

In the prototype interaction study, since each participant was wearing the wearable device while using the phone, all touches made by them should ideally have been validated. Therefore, we evaluate the performance of the prototype using True Acceptance Rate, the ratio of validated touch events over all touch events. Figure 6.7 shows the performance of the prototype in validating the touches made by the participants. While reviewing these results, it is important to keep in mind that this classifier is untrained on any of the users that were interacting with it in this study. On an average, 74% of touches made by the participants across all

sessions were accepted by the system. This can be attributed to the fact that when the participants started interacting with the system, they would not know how to get the system to respond to them. Figure 6.7 also shows the max accuracy obtained across all sessions by the participants. On an average, the maximum TAR across all participants was 88%, which confirmed our speculation that eventually the participants started to figure out how to get the system to respond to them.

We also analyzed the True Acceptance Rate of the system per session for each participant. The results can be seen in Figure 6.6. The aggregate performance for each session across all participants demonstrates the improvement in the TAR over the sessions. In the first session, the average performance across all participants was approx. 59%, while by the last session, the average performance improved to 82%. While this result is very encouraging, we would like to highlight three participants - P3, P7 and P10. Each of these participants, in one of the middle order sessions seem to have been able to improve their device responses. However, they did not seem to be able to replicate what worked for them, and eventually ended up staying almost at par with their performance in the first session.

We also asked the participants to mark their responses on a survey questionnaire at the end of each session. There were a total of nine sessions, of which the first one did not invoke the authentication system, and is considered the baseline for the responses. The participants were not informed about the differences in the sessions. The responses to the first survey question by each participant is shown in Figure 6.9. The average response to the question about their ability to get the device to respond to them in the baseline session seems to be 'Strongly Agree'. This can be attributed to the fact that the authentication system was not active in this session, and therefore all touches made by the participant were accepted by the system. In the second session, which is when the authentication kicked in, the participants started to face difficulties in getting the phone to respond to their touches. Corresponding to the second session, we observe a dip the average response from 'Strongly Agree' to 'Neutral'. Over the remaining sessions, this response improves and ends on 'Agree' by the ninth session.

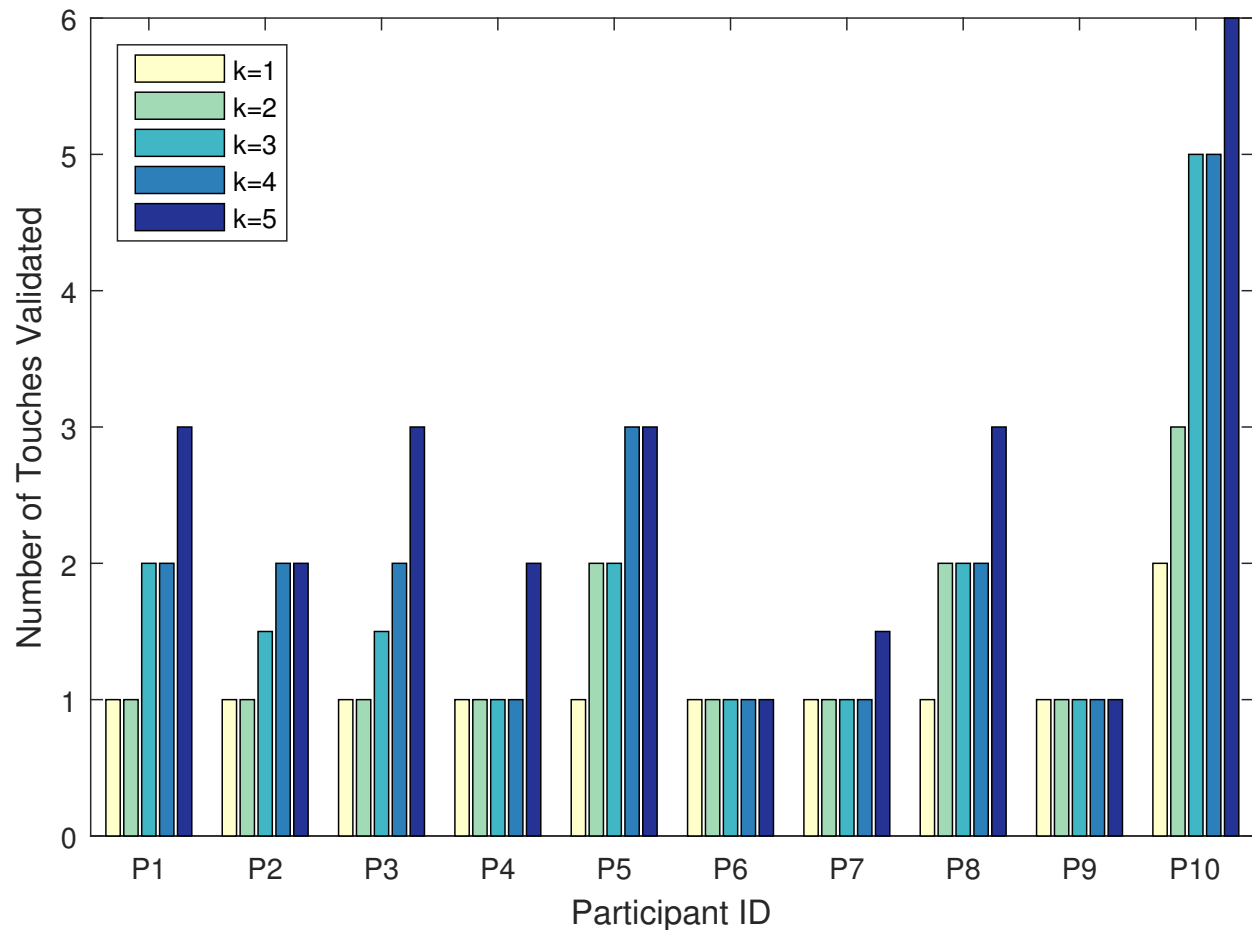


Figure 6.12: Median number of touches that were allowed, before the system displays a strong password screen upon detecting k number of consecutive invalid touches. The average number of touches across all participants for the values of k from 1 to 5 are 1.1, 1.4, 1.8, 2 and 2.55

Figure 6.10 shows the participants' responses to the second question of the survey. This question asked them to rate if they could notice a lag in the device's response to their touches. In the baseline session, the average response to this question is close to 'Disagree'. Its interesting to note that this response is not 'Strongly Disagree' for the baseline, which implies that the participants were able to notice the lag in Android's default response to touches. The biggest change in this response can be observed at the second session, which is when the authentication system starts working, and the average response reaches 'Agree'. However, this response starts dropping and ends at 'Neutral' by the ninth session. Similarly,

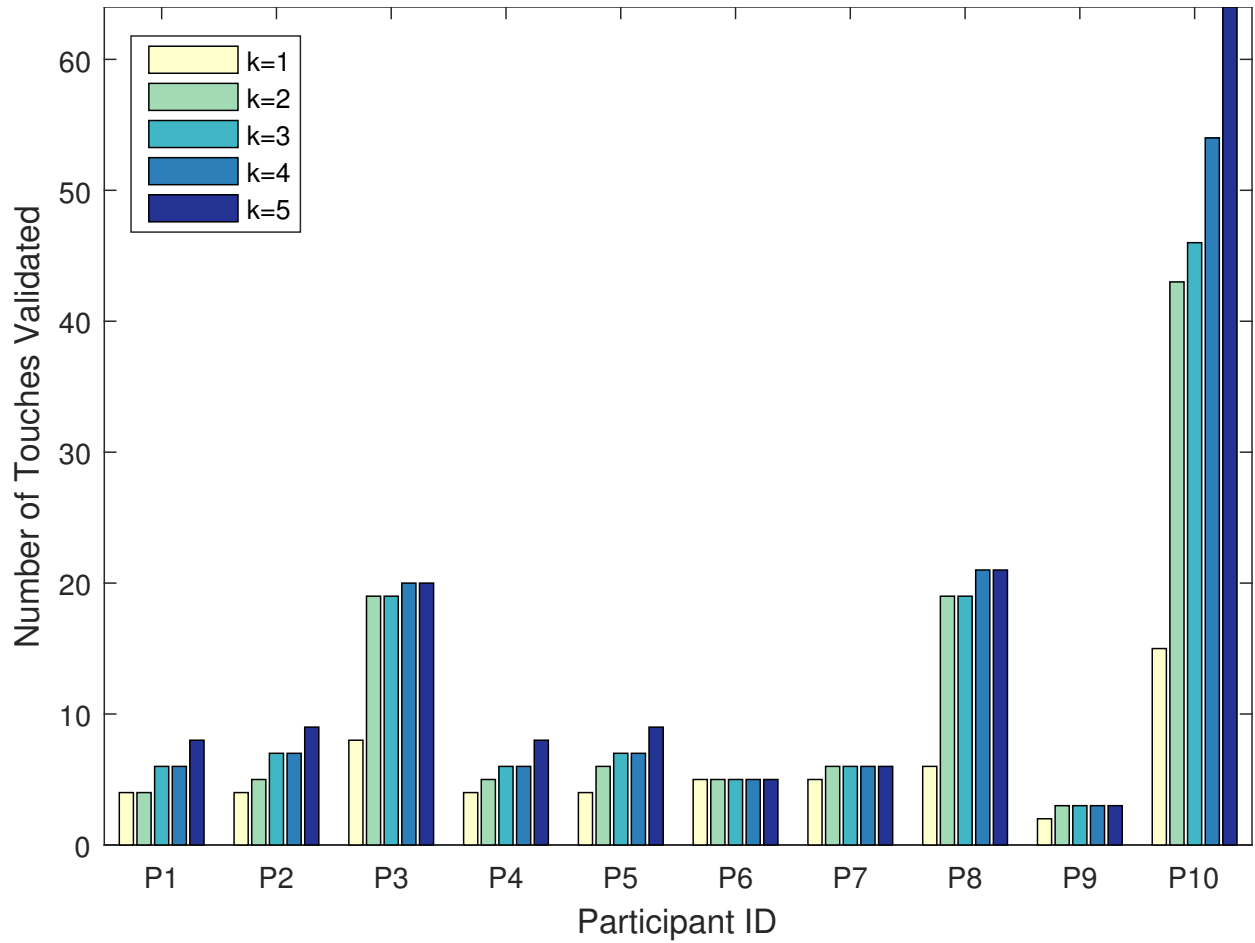


Figure 6.13: Maximum number of touches that were allowed, before the system displays a strong password screen upon detecting ‘k’ number of consecutive invalid touches. The average number of maximum touches the adversaries were able to achieve for the values of k from 1 to 5 are 5.7, 11.8, 12.4, 13.5 and 15.3

as can be seen in Figure 6.11, the acceptability of the device’s lag starts at ‘Agree’ at the baseline sessions (and not ‘Strongly agree’), and while it drops off at the second session to a point mid-way between ‘Neutral’ and ‘Disagree’, it settles back at ‘Neutral’ by the end of ninth session.

6.6 Adversarial Study

We earlier looked at an offline analysis of a false acceptance rate in the feasibility study, where we observed the ability of the classifier to differentiate between valid touch IMU features

and IMU features of other in-situ activities. We decided to put the system through a more rigorous test in active adversarial situations.

6.6.1 Procedure

The adversarial study was conducted with teams of two participants, where one participant acted as the adversary trying to gain access to the phone, while the other participant was the actual phone user, who was engaged in performing other activities. In this study, we recruited the participants who had previously participated in the Prototype Interaction study to act as adversaries, as we expected them to have gained some insight about how the authentication system might work.

In this study, the actual phone owners (wearing the wearable device), was asked to perform three different tasks - walk around the room, make notes on a white board and flip through the pages of a book. The total study was conducted for 15 minutes, and each task was performed for 5 minutes. Other than these broad instructions, the phone owners were not instructed on how they specifically executed each task. While the phone owner was engaged in performing these tasks, the adversary was given the smartphone, and asked to try to get the device to respond to their touches. They were given full freedom to move around as they liked, mimic the phone owners and do whatever they wanted to in order to get the phone to respond to them.

6.6.2 Analysis

The system's performance in terms of its resistance to the invalid touch events is measured by False Acceptance Rate, ratio of invalid touch events that are authenticated by the system to the total number of invalid touch events. The FAR for each participant as an adversary are reported in Figure 6.7 to contrast against the True Acceptance Rate of the participant in the Prototype Interaction Study. The average FAR across all participants is 13.8%. It is interesting to observe the participants (P6 and P9), who were able to get the device to

respond to them with the highest TAR in the prototype interaction study, were also the ones least able to hack the system, receiving the lowest FAR in the adversarial study. Opposite is true for P10, whose TAR in the previous study was the lowest, while the FAR in the adversarial study was the highest.

We further analyzed the output of the adversarial study to test if introducing a strong authentication mechanism, such as a password, after a certain number of consecutive invalid touch events would be helpful in limiting the amount of information that an adversary could have access to. We varied the number of consecutive invalid touches (k) required to trigger a password from 1 to 5. The result of this analysis can be seen in Figure ?? . Figure 6.12 shows the median number of touches that the adversary is able to get validated before the password is asked. The average number of touches across all participants for the values of k from 1 to 5 are 1.1, 1.4, 1.8, 2 and 2.55. While this number is encouraging, we also analyzed the worst case scenario - the maximum number of touches each of these adversaries were able to authenticate before the system would display the password screen in Figure 6.13. The average number of maximum touches the adversaries were able to achieve for the values of k from 1 to 5 are 5.7, 11.8, 12.4, 13.5 and 15.3. These numbers indicate that there is still room for improvement because in the worst case with 5-15 clicks the adversary can easily gain access to emails, and other personal data on the phone.

6.7 Limitations

While encouraging, our evaluation is limited in some ways. We recruited a total of 23 different users across the three different studies that we conducted. It would be interesting to evaluate our approach with a much larger user base. Also at this point, we demonstrated through offline analysis that our system was able to learn its user's pattern of interacting with a phone, and was able to achieve 95% accuracy after training on 84 samples across all participants. We also demonstrated that on an average the participants are able to learn how to use a

system that is not trained on them. The final system would incorporate a more symbiotic relationship between the phone and the user, where both would be learning at the same time, and therefore in such a scenario, we speculate the user should be able to get the device to respond accurately much faster. One of the main limitations that made it hard for us to build such a system was Android's security mechanisms that restrict apps from gaining access to all touch events in the phone. An ideal study would involve us actually building a lock screen app that would be monitoring all touches and validating them using an active learning classifier. Our system would clearly benefit from having an longitudinal study evaluating the insitu performance. Our proposed approach currently does not work in a particular hand position, when the wearable device is worn on the hand that is neither used to hold the phone or touch the phone. Our proposed approach also does not work in the scenario where a phone user is holding the phone in the hand wearing the wearable device, and someone else touches the screen. Future work in this area includes testing out approach in many other scenarios (such as walking, train), and even more rigorous adversarial attacks. With further efforts in this direction, we hope that the use of wearable device in implicitly and continuously validating every touch on smartphone would become a practical approach that can ensure security of a person's private data without being cumbersome to use.

6.8 Conclusion

We have proposed a new approach for implicitly and continuously authenticating a smartphone owner using a paired wearable device worn by the owner. Current security mechanisms are not only weak, in that once a device is authenticated, it remains such until the screen is switched off, but are also viewed as cumbersome by many smartphone users. The aim of this paper has been to discover a set of features that can be used to validate every touch made on the phone using data from the IMU sensors of the phone and the wearable device.

Using data from a 13-participant study, we evaluate the feasibility of our approach.

Classifiers such as Random Forest and Boosted JRip performed with a high True Acceptance Rate of 96% and a low False Acceptance Rate of 4%. We observed that training per participant gives a high average f-score of 96.4%, while leave-one-out method performs with an average f-score of 92.3%. We created a prototype of our approach to test how it works with users it has not been trained for. While the average TAR across all participants was 74%, the best TAR of the system across all sessions was 88%. We also conducted an adversarial study in which we evaluated the system's False Acceptance Rate when the phone's owner was engaged in other activities, while the adversary tried to interact with the phone. The average FAR across all participants was 13.8%. If we introduce a strong authentication mechanism upon detecting a certain number (\mathbf{k}) of consecutive invalid touches, the median number of valid touches an adversary would be able to make before being blocked by the system, for the values of \mathbf{k} from 1 to 5, are 1.1, 1.4, 1.8, 2 and 2.55.

Future work in this area includes validating these results in a longitudinal study, where the system is able to perform active learning based on the individual phone owner's touches. With further effort in this area, we hope that wearable devices will become a practical approach that users can rely on for providing security without having to engage in any explicit and cumbersome process.

Chapter 7

Conclusion

7.1 Summary

The main focus of this dissertation was to investigate how different levels of data shared by the Wearable with the Smart Object can be used for performing implicit object user identification and authentication. Specifically, we looked at three different levels of data sharing and interaction. In the first level of interaction, when an object is used, it searches for the identity of the users closest to it. If only one user is detected, then the same user must have used the object. However, when there are multiple users detected, the first level of interaction is not sufficient. We looked at how heuristics based on past object usage can be used to make a learned guess about the most probable object user. If the object does not have the history of usage of every user it has to choose from, or if certain objects are used equally or randomly by individuals (e.g. lights), then the use of heuristics does not help in guessing who used the object. At this point, the object initiates the next level of interaction with the Wearable. At this level of interaction, the object queries the devices nearby to respond with how well their motion data matches the gesture required to interact with the object. However, this method is vulnerable to being spoofed, as devices may maliciously respond with a higher matching score. Therefore this method cannot be used for sensitive

and privacy critical applications. The third level of interaction is used for applications where it is important to validate that a user is the actual object user. In this stage, the identified user is queried for their raw sensor data, which is analyzed and correlated with the object's sensor data at a millisecond level accuracy.

7.2 Contributions

Some of the main contributions of this dissertation are:

1. Evaluated the use of room level granularity of location tracking in performing object user identification and compared it to more precise coordinate level tracking. Results show that user identification accuracy of heuristics augmented room level tracking granularity is comparable to coordinate level tracking.
2. Evaluated the use of potential users' wrist motion sensors to detect the unique fingerprint of the object used in three different studies - feasibility study which was conducted in controlled experiments, real world task emulation study with pairs of participants, as well as a simulated study in which in-situ hand gesture data was used to emulate multiple potential object users.
3. Evaluated a set of features that can be used to validate touches made on a smart phone using data from the IMU sensors of the phone and the wearable device. Offline analyses of our approach using a 13 participant feasibility study, show promising results. We also developed and tested a prototype with 10 other participants to validate the true acceptance rate of the system. We also conducted a 10 participant adversarial study to determine the false acceptance rate of the system.

7.3 Limitations

While the results of our work are promising, there are some known limitations in our evaluation:

1. We evaluate the first level of interaction, which uses location information, only with objects in a home environment with pairs of residents. A more generic evaluation would look at the accuracy of this approach in different environments, and larger number of potential object users.
2. We evaluate the next level of interaction, which uses a matching score to detect which potential object user's hand motion best matches the gesture required to interact with the object in a constrained conditions: there were only sets of two potential object users to select from, and the objects had a fixed location and position in the home environment. A more robust evaluation would involve evaluating this approach with larger sets of potential object users, as well as with mobile objects.
3. Our user authentication work, which was the third level of interaction, was evaluated in a Smartphone authentication use case. To evaluate how generalizable the approach is, we need to evaluate this approach with other objects as well.

7.4 Future Research Directions

The limitations of our work set the tone for future work in this area. We need to be able to make our approach more generic, require less training, and be robust to a large set of potential object users. This is a much harder problem, because currently the features used for gesture recognition rely on a database of training samples from every potential object user. Using features that are independent across users will help reduce the training requirement of this approach. Its also interesting to explore the authentication work in domains of other objects, specially objects that are fixed, such as a whiteboard. We want to explore fixed

objects, because their motion data would be much more weakly co-related with the wearable device, than a mobile object, such as a smartphone.

Bibliography

- [1] Juhi Ranjan and Kamin Whitehouse. Rethinking the fusion of technology and clinical practices in functional behavior analysis for the elderly. In *Human Behavior Understanding*, pages 52–65. Springer, 2015.
- [2] WBCSD. Transforming the market: Energy efficiency in buildings. Survey report, 2009.
- [3] Dominik Schmidt, Fadi Chehimi, Enrico Rukzio, and Hans Gellersen. Phonetouch: a technique for direct phone interaction on surfaces. In *Proceedings of the 23rd annual ACM symposium on User interface software and technology*, pages 13–16. ACM, 2010.
- [4] Stefan Poslad. Ubiquitous computing smart devices, smart environments and smart interaction. *Ubiquitous computing smart devices, smart environments and smart interaction*, 2009.
- [5] Gerd Kortuem, Fahim Kawsar, Vasughi Sundramoorthy, and Daniel Fitton. Smart objects as building blocks for the internet of things. *IEEE Internet Computing*, 14(1):44–51, 2010.
- [6] Samsung Smart TV. <http://www.samsung.com/us/experience/smart-tv/>.
- [7] Ecobee: Smart WiFi Thermostat. <https://shop.ecobee.com/products/ecobee3>.
- [8] Nest: Smart Learning Thermostat. <https://nest.com/thermostat/meet-nest-thermostat/>.
- [9] Whirlpool Energy Smart Electric Water Heater. <http://www.whirlpoolwaterheaters.com/products/electric-water-heaters/es50r9-55/>.
- [10] Amazon Echo. <https://www.amazon.com/Amazon-Echo-Bluetooth-Speaker-with-WiFi-Alexa/dp/B00X4WHP5E>.
- [11] Computer World: Innovative Tech Features in Vehicles. <http://www.computerworld.com/article/2910256/10-amazing-innovative-tech-features-in-the-latest-vehicles.html>.
- [12] BankRate: Advances in Auto Technology. <http://www.bankrate.com/finance/money-guides/8-great-new-advances-in-auto-technology-1.aspx>.

- [13] VSPGlobal: Level, tech infused frames. <http://newsroom.vspglobal.com/vsp-global/news-releases/vsp-global-and-usc-launching-major-wearables-study-with-tech-infused-frames>.
- [14] Recon JET: Smart Eyewear. <http://store.reconinstruments.com/Recon-Jet>.
- [15] Jins Meme. <https://jins-meme.com/en/>.
- [16] Hexoskin. <http://www.hexoskin.com/>.
- [17] Athos. <https://www.liveathos.com/>.
- [18] Sensoria. <http://www.sensoriafitness.com/>.
- [19] Owlet. <http://www.owletcare.com/>.
- [20] Kwikset Kevo Plus Smart Lock. <http://www.kwikset.com/kevo/plus>.
- [21] Schlage Sense Smart Deadbolt. <http://www.schlage.com/en/home/keyless-deadbolt-locks/sense.html>.
- [22] August Smart Lock. <http://august.com/products/august-smart-lock/>.
- [23] IHome Smart Plug. <http://www.ihomeaudio.com/isp5wc/>.
- [24] mFi In-Wall Manageable Devices. <https://www.ubnt.com/mfi/inwall/>.
- [25] Belkin WiMo Switch. <http://www.belkin.com/us/F7C029-Belkin/p/P-F7C029/>.
- [26] Phillips Hue. <http://www2.meethue.com/en-us/>.
- [27] Samsung SmartThings. <https://www.smarthings.com/>.
- [28] Chen-Rong Yu, Chao-Lin Wu, Ching-Hu Lu, and Li-Chen Fu. Human localization via multi-cameras and floor sensors in smart home. In *Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference on*, volume 5, pages 3822–3827. IEEE, 2006.
- [29] David Lo. Multimodal human localization using bayesian network sensor fusion.
- [30] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury. No need to war-drive: unsupervised indoor localization. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*, pages 197–210. ACM, 2012.
- [31] Dimitrios Lymberopoulos, Jie Liu, Xue Yang, Romit Roy Choudhury, Vlado Handziski, and Souvik Sen. A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, pages 178–189. ACM, 2015.

- [32] Robert J Orr and Gregory D Abowd. The smart floor: a mechanism for natural user identification and tracking. In *CHI'00 extended abstracts on Human factors in computing systems*, pages 275–276. ACM, 2000.
- [33] Timothy W Hnat, Erin Griffiths, Ray Dawson, and Kamin Whitehouse. Doorjamb: unobtrusive room-level tracking of people in homes using doorway sensors. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, pages 309–322. ACM, 2012.
- [34] Magnus Bang, Carin Torstensson, and Cecilia Katzeff. The powerhouse: A persuasive computer game designed to raise awareness of domestic energy consumption. In *International Conference on Persuasive Technology*, pages 123–132. Springer, 2006.
- [35] Jie Chi Yang, Kun Huang Chien, and Tzu Chien Liu. A digital game-based learning system for energy education: An energy conservation pet. *TOJET: The Turkish Online Journal of Educational Technology*, 11(2), 2012.
- [36] Maarit Haakana, Liisa Sillanpää, and Marjatta Talsi. The effect of feedback and focused advice on household energy consumption. 1997.
- [37] Anton Gustafsson and Magnus Gyllenswärd. The power-aware cord: energy awareness through ambient information display. In *CHI'05 extended abstracts on Human factors in computing systems*, pages 1423–1426. ACM, 2005.
- [38] Dane Petersen, Jay Steele, and Joe Wilkerson. Wattbot: a residential electricity monitoring and feedback system. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*, pages 2847–2852. ACM, 2009.
- [39] Karin Kappel and Thomas Grechenig. Show-me: water consumption at a glance to promote water conservation in the shower. In *Proceedings of the 4th international conference on persuasive technology*, page 26. ACM, 2009.
- [40] Stacey Kuznetsov and Eric Paulos. Upstream: motivating water conservation with low-cost water flow sensing and persuasive displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1851–1860. ACM, 2010.
- [41] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 32–43. ACM, 2000.
- [42] Andy Ward, Alan Jones, and Andy Hopper. A new location technique for the active office. *IEEE Personal Communications*, 4(5):42–47, 1997.
- [43] Yasuhiro Fukuju, Masateru Minami, Hiroyuki Morikawa, and Tomonori Aoyama. Dolphin: An autonomous indoor positioning system in ubiquitous computing environment. In *WSTFEUS*, pages 53–56, 2003.

- [44] Thomas King, Stephan Kopf, Thomas Haenselmann, Christian Lubberger, and Wolfgang Effelsberg. Compass: A probabilistic indoor positioning system based on 802.11 and digital compasses. In *Proceedings of the 1st international workshop on Wireless network testbeds, experimental evaluation & characterization*, pages 34–40. ACM, 2006.
- [45] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. Spotfi: Decimeter level localization using wifi. In *ACM SIGCOMM Computer Communication Review*, volume 45, pages 269–282. ACM, 2015.
- [46] Moustafa Youssef and Ashok Agrawala. The horus wlan location determination system. In *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, pages 205–218. ACM, 2005.
- [47] Kaishun Wu, Jiang Xiao, Youwen Yi, Min Gao, and Lionel M Ni. Fila: Fine-grained indoor localization. In *INFOCOM, 2012 Proceedings IEEE*, pages 2210–2218. IEEE, 2012.
- [48] Brian Ferris, Dieter Fox, and Neil D Lawrence. Wifi-slam using gaussian process latent variable models. In *IJCAI*, volume 7, pages 2480–2485, 2007.
- [49] Souvik Sen, Jeongkeun Lee, Kyu-Han Kim, and Paul Congdon. Avoiding multipath to revive inbuilding wifi localization. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, pages 249–262. ACM, 2013.
- [50] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N Padmanabhan. Indoor localization without the pain. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, pages 173–184. ACM, 2010.
- [51] Anshul Rai, Krishna Kant Chintalapudi, Venkata N Padmanabhan, and Rijurekha Sen. Zee: zero-effort crowdsourcing for indoor localization. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 293–304. ACM, 2012.
- [52] Zheng Yang, Chenshu Wu, and Yunhao Liu. Locating in fingerprint space: wireless indoor localization with little human intervention. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 269–280. ACM, 2012.
- [53] Fan Li, Chunshui Zhao, Guanzhong Ding, Jian Gong, Chenxing Liu, and Feng Zhao. A reliable and accurate indoor localization method using phone inertial sensors. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 421–430. ACM, 2012.
- [54] Nishkam Ravi, Pravin Shankar, Andrew Frankel, Ahmed Elgammal, and Liviu Iftode. Indoor localization using camera phones. In *Seventh IEEE Workshop on Mobile Computing Systems & Applications (WMCSA'06 Supplement)*, pages 49–49. IEEE, 2006.

- [55] Stephen P Tarzia, Peter A Dinda, Robert P Dick, and Gokhan Memik. Indoor localization without infrastructure using the acoustic background spectrum. In *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pages 155–168. ACM, 2011.
- [56] Patrick Lazik and Anthony Rowe. Indoor pseudo-ranging of mobile devices using ultrasonic chirps. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, pages 99–112. ACM, 2012.
- [57] Simon Hay and Andrew Rice. The case for apportionment. In *Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, BuildSys '09, pages 13–18. ACM, 2009.
- [58] Simon Hay, Andrew Rice, and Andy Hopper. A global personal energy meter. In *Adjunct Proceedings of the 7th International Conference on Pervasive Computing*, 2009.
- [59] Jeff Hsu, Prashanth Mohan, Xiaofan Jiang, Jorge Ortiz, Sushant Shankar, Stephen Dawson-Haggerty, and David Culler. Hbci: Human-building-computer interaction. In *Proceedings of the 2Nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building*, BuildSys '10, 2010.
- [60] Yun Cheng, Kaifei Chen, Ben Zhang, Chieh-Jan Mike Liang, Xiaofan Jiang, and Feng Zhao. Accurate real-time occupant energy-footprinting in commercial buildings. In *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, BuildSys '12.
- [61] Seungwoo Lee, Daye Ahn, Sukjun Lee, Rhan Ha, and Hojung Cha. Personalized energy auditor: Estimating personal electricity usage. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*.
- [62] Cheng Bo, Lan Zhang, Xiang-Yang Li, Qiuyuan Huang, and Yu Wang. Silentsense: silent user identification via touch and movement behavioral biometrics. In *Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 187–190. ACM, 2013.
- [63] Christian Holz and Patrick Baudisch. Fiberio: a touchscreen that senses fingerprints. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 41–50. ACM, 2013.
- [64] Chris Harrison, Munehiko Sato, and Ivan Poupyrev. Capacitive fingerprinting: exploring user differentiation by sensing electrical properties of the human body. In *Proceedings of the 25th annual ACM symposium on User interface software and technology*, pages 537–544. ACM, 2012.
- [65] Nicolai Marquardt, Johannes Kiemer, David Ledo, Sebastian Boring, and Saul Greenberg. Designing user-, hand-, and handpart-aware tabletop interactions with the touchid toolkit. In *Proceedings of the ACM international conference on interactive tabletops and surfaces*, pages 21–30. ACM, 2011.

- [66] Ulrich von Zadow, Patrick Reipschläger, Daniel Bösel, Anita Sellent, and Raimund Dachzelt. Youtouch! – low-cost user identification at an interactive display wall. In *Proceedings of the 14th International Working Conference on Advanced Visual Interfaces 2016*, New York, NY, USA, 6 2016. ACM.
- [67] Muhammad Shahzad, Saira Zahid, and Muddassar Farooq. A hybrid ga-pso fuzzy system for user identification on smart phones. In *Proceedings of the 11th Annual conference on Genetic and evolutionary computation*, pages 1617–1624. ACM, 2009.
- [68] Mauro Conti, Irina Zachia-Zlatea, and Bruno Crispo. Mind how you answer me!: transparently authenticating the user of a smartphone when answering or placing a call. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, pages 249–259. ACM, 2011.
- [69] Davrondzhon Gafurov, Kirsi Helkala, and Torkjel Søndrol. Biometric gait authentication using accelerometer sensor. *Journal of computers*, 1(7):51–59, 2006.
- [70] Jani Mantyjarvi, Mikko Lindholm, Elena Vildjiounaite, S-M Makela, and HA Ailisto. Identifying users of portable devices from gait pattern with accelerometers. In *Proceedings.(ICASSP’05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 2, pages ii–973. IEEE, 2005.
- [71] Mark R Hodges and Martha E Pollack. An object-use fingerprint: The use of electronic sensors for human identification. In *UbiComp 2007: Ubiquitous Computing*, pages 289–303. Springer, 2007.
- [72] Vijay Srinivasan, John Stankovic, and Kamin Whitehouse. Fixturefinder: Discovering the existence of electrical and water fixtures. In *Proceedings of the 12th international conference on Information processing in sensor networks*, pages 115–128. ACM, 2013.
- [73] Sidhant Gupta, Matthew S Reynolds, and Shwetak N Patel. Electrisense: single-point sensing using emi for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 139–148. ACM, 2010.
- [74] Jon E Froehlich, Eric Larson, Tim Campbell, Conor Haggerty, James Fogarty, and Shwetak N Patel. Hydrosense: infrastructure-mediated single-point sensing of whole-home water activity. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 235–244, 2009.
- [75] Younghun Kim, Thomas Schmid, Zainul M Charbiwala, and Mani B Srivastava. Viridiscope: design and implementation of a fine grained power monitoring system for homes. In *Proceedings of the 11th international conference on Ubiquitous computing*, pages 245–254. ACM, 2009.
- [76] Younghun Kim, Thomas Schmid, Zainul M Charbiwala, Jonathan Friedman, and Mani B Srivastava. Nawms: nonintrusive autonomous water monitoring system. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, 2008.

- [77] Matthai Philipose, Kenneth P Fishkin, Mike Perkowitz, Donald J Patterson, Dieter Fox, Henry Kautz, and Dirk Hahnel. Inferring activities from interactions with objects. *Pervasive Computing*, 2004.
- [78] Vijay Srinivasan, John Stankovic, and Kamin Whitehouse. Using height sensors for biometric identification in multi-resident homes. In *Pervasive Computing*, pages 337–354. 2010.
- [79] Juhi Ranjan, Yu Yao, and Kamin Whitehouse. An RF Doormat for Tracking People’s Room Locations. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*, pages 797–800. ACM, 2013.
- [80] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Kdd*, volume 96, pages 226–231, 1996.
- [81] UX90-002M: HOBO UX90 Light On/Off Logger with Extended Memory. <http://www.onsetcomp.com/products/data-loggers/ux90-002m>.
- [82] UX90-001: HOBO UX90 State Logger. <http://www.onsetcomp.com/products/data-loggers/ux90-001>.
- [83] CSV-A8: AC Current Switch. <http://www.onsetcomp.com/products/sensors/csv-a8>.
- [84] TED: The Energy Detective. <http://www.theenergydetective.com/home>.
- [85] FTB4700: Low Flow Liquid Flowmeters. <http://www.omega.com/pptst/FTB4700FTB4800.html>.
- [86] UX120-017M: HOBO 4-Channel Pulse Data Logger. <http://www.onsetcomp.com/products/data-loggers/ux90-001>.
- [87] People Do Not Secure Phones, 2011. <http://www.bullguard.com/news/latest-press-releases/press-release-archive/2011-06-21.aspx>.
- [88] 10 Ways to Secure Your Smartphone, 2014. <http://www.androidcentral.com/10-best-ways-secure-your-smartphone>.
- [89] Fixed Screen Photo Sharing, 2016. <https://play.google.com/store/apps/details?id=com.risesoftware.focus.free>.
- [90] Juhi Ranjan and Kamin Whitehouse. Demo: Automatic authentication of smartphone touch interactions using smartwatch. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 361–364. ACM, 2016.
- [91] Android Smart Lock, 2014. <https://support.google.com/nexus/answer/6093922?hl=en>.

- [92] Muhammad Shahzad, Alex X Liu, and Arjmand Samuel. Secure unlocking of mobile touch screen devices by simple gestures: you can see it but you can not do it. In *Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 39–50. ACM, 2013.
- [93] Fabian Monrose, Michael K Reiter, and Susanne Wetzel. Password hardening based on keystroke dynamics. *International Journal of Information Security*, 1(2):69–83, 2002.
- [94] Saira Zahid, Muhammad Shahzad, Syed Ali Khayam, and Muddassar Farooq. Keystroke-based user identification on smart phones. In *International Workshop on Recent Advances in Intrusion Detection*, pages 224–243. Springer, 2009.
- [95] Nan Zheng, Kun Bai, Hai Huang, and Haining Wang. You are how you touch: User verification on smartphones via tapping behaviors. In *2014 IEEE 22nd International Conference on Network Protocols*, pages 221–232. IEEE, 2014.
- [96] Hassan Khan, Urs Hengartner, and Daniel Vogel. Targeted mimicry attacks on touch input based implicit authentication schemes. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, pages 387–398. ACM, 2016.
- [97] Ahmed Kharrufa, James Nicholson, Paul Dunphy, Steve Hodges, Pam Briggs, and Patrick Olivier. Using imus to identify supervisors on touch devices. In *Human-Computer Interaction*, pages 565–583. Springer, 2015.
- [98] Wearable Sensing Technology, 2016. <http://www.shimmersensing.com/>.
- [99] Supervised random subsampling of a dataset, 2016. <http://weka.sourceforge.net/doc.stable/weka/filters/supervised/instance/SpreadSubsample.html>.