Enabling Visibility Into Building Energy Consumption Through Novel Metering Designs and Methods

by

Samuel Gordon DeBruin

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Computer Science and Engineering)
in the University of Michigan
2017

Doctoral Committee:

Associate Professor Prabal Dutta, Chair
Associate Professor Robert Dick
Professor Jerome Lynch
Assistant Professor Jenna Wiens

Samuel Gordon DeBruin

sdebruin@umich.edu

ORCID iD: 0000-0002-9294-1814

To my parents, David and Elizabeth, and to Holly. I could not have done it without you.

TABLE OF CONTENTS

LIST OF FIGURES

xv

LIST OF TABLES

xix

ABSTRACT

Energy consumption in buildings is an area of growing national concern, with almost 3,000 TWh going to residential and commercial buildings in the United States annually. This figure represents over 40% of total energy consumption, and 70% of total electricity consumption. These figures are representative of other countries as well. Improvements to devices and building practices reduce the energy use of household appliances and HVAC, but lower-power devices do not benefit from many of these advances. In particular miscellaneous electrical loads (MELs) – plug-load devices like televisions, coffee makers, and cell phone chargers – are a growing segment of consumption about which less is understood than heating, cooling, lighting, and other major loads. MELs are proliferating quickly, and although they individually draw moderately low power, in the aggregate they represent a significant fraction (30%-60%) of the total. Complicating the problem, existing tools for gaining insight into this segment are either too expensive, too difficult to deploy, too cumbersome to operate, or of too low accuracy to provide sufficient benefit.

In this dissertation, we propose the design of an energy metering architecture that is low power, inexpensive, and scalable to the needs of the long tail of miscellaneous electrical loads. This platform affords unprecedented visibility into the energy consumption of these devices, which can be used to guide energy management policies and regulatory action, as well as enable targeted energy management applications, to curb this growing segment. Realizing this energy metering architecture, which is composed of both circuit-level and plug-load energy meters, requires revisiting all aspects of the AC meter design. In particular we design and implement novel ways to power the meters including energy

harvesting from the AC, two different non-contact current measurement methods, a unique planar plug through form factor, and a new method for viewing and interacting with metering data on nearby smartphones. We deploy this metering platform alongside other sensors including light and occupancy, which allows us to generate unprecedented data on MEL energy consumption, illustrate how MELs can be identified from their electrical signatures, and identify faults in electrical appliances by detecting anomalies in their electrical fingerprints. Our findings show that a significant portion of total energy, as much as 50%, is spent on MELs, and also that electrical data contain detailed information about the identification, as well as the operational status, of AC devices.

CHAPTER 1

Introduction

Electricity consumption in buildings accounts for a significant portion of the total energy consumed in the United States, and around the world. The U.S. Energy Information Administration (EIA) finds in 2015 that total electricity consumption in residential and commercial buildings was 2,760 TWh [32]. This accounts for over 40% of domestic energy consumption and 70% of electricity consumption annually, and these breakdowns are mirrored in other countries around the world [98]. These figures have held relatively steady over the past several years, and present a startling reality of the state of consumption.

This strong dependence on electrical energy in buildings is at odds with U.S. Federal sustainability goals for achieving usable net-zero energy buildings by 2025 [19]. Although new building practices provide significant advances in energy goals, the timetable of the Federal target will require retrofits and modifications largely to the existing building set. This is corroborated by a recent study from the National Science and Technology Council, which finds that "for the foreseeable future, the greatest national energy saving potential lies with improvements to existing buildings" [105]. However, these improvements must be targeted to the realities of household consumption, and in particular to the recent proliferation of uncharacterized outlet loads.

Miscellaneous electrical loads (MELs) are becoming an increasingly dominant portion of total energy. According to the 2015 U.S. Department of Energy Quadrennial Technology Review, by 2020 MELs might make up 50% of the energy consumed in homes [31]. This results, in part, from programs that actively target other segments, whether through regulatory incentive or commercial action, like heating, cooling, lighting, and water heating. One of the primary barriers to addressing the problem of MELs is limited understanding of their energy use patterns. In this dissertation, we explore the potential of unprecedented visibility into this otherwise poorly understood segment. This level of understanding is improved through *submetering*, addressing the specific contributions that make up the reading on the utility meter.

(a) Energy Use	(b) Electricity Use
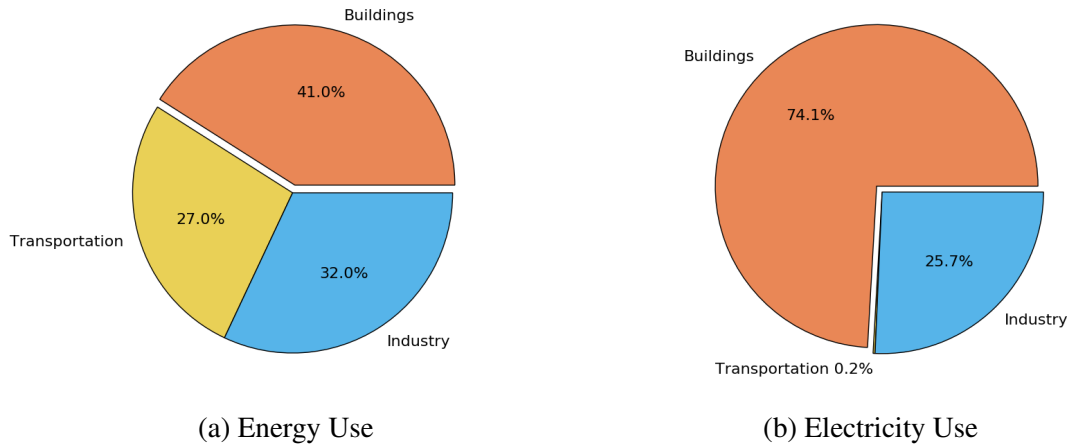
Figure 1.1: Total energy and electricity consumption broken down by category. Buildings account for 41% of total energy consumption and 74% of electricity use [31]. In addition to the burden on generation and infrastructure, this consumption is associated with significant greenhouse gas emissions

1.1 Building Energy: a Significant National Problem

Residential and commercial buildings consume 41% of total energy and 74% of electricity in the United States [31, 117]. As seen in Figure 1.1, the remainder is mainly consumed in industrial processes and transportation. Energy use in buildings is associated with large national expense for generation and distribution, approximately $430 billion annually [117]. Further, unchecked energy consumption in buildings violates all three of the U.S. Department of Energy's strategic objectives for the nation's energy system: energy security, economic competitiveness, and environmental responsibility [31]. Particularly with regard to the environment, building energy is associated with significant greenhouse gas emissions [98]. The Department of Energy Building and Technology Office (BTO) predicts that hitting building energy efficiency targets could reduce energy-related greenhouse gas emissions by 20% relative to 2010 figures [117]. Renewable sources like solar and wind provide improvements to energy security, economic competitiveness, and environmental responsibility, but they must be partnered with improvements to the way energy is consumed in buildings.

Improvements to the building energy problem can utilize advanced building practices for creating highly efficient new structures, but what is more difficult is finding a solution for the buildings that already exist. Efficiency programs like the LEED certification exist to make new construction highly energy efficient, but the National Science and Technology Council finds that the stock of existing buildings will continue to dominate, and that the greatest potential for energy savings lies with existing buildings [105]. This involves equipment

upgrades like replacing existing hardware with ENERGY STAR® and other high-efficiency appliances, but also utilizing the existing appliances and AC devices in a more efficient manner and with more efficient practices.

1.1.1 AC Devices Individualized and Diverse

In order to reduce the energy footprint in existing buildings, we must first understand where the energy is going. This is evidenced by calls for research and development in this area such as the Department of Energy wireless metering challenge and others [33, 106]. However, energy in buildings is not homogeneous with regard to the way power is drawn or the types of devices drawing power. This heterogeneity in devices makes a single all-encompassing metering solution difficult in practice. Devices can be largely resistive, meaning the current and voltage waveforms are sinusoidal and in-phase, but most devices have some distortion or displacement causing reactive components. This will be explored in later sections, but in practice it places a lower bound on the complexity of metering devices, particularly with regard to sampling quality.

In addition to diverse properties of their power consumption, devices have diverse fittings, plugs, and mechanical interfaces. Some high-power loads like HVAC are hard-wired directly into the electrical panel, and others like the oven or clothes washer in the United States use the large 14-20 plug given in the National Electrical Manufacturers Association (NEMA) specification [18]. The largest single category is plug-load outlet devices that either use the 1-15 or 1-15P 2-prong plugs or the 5-15P polarized 3-prong plug in the United States [18]. These are the televisions, phone chargers, refrigerators, and all other outlet devices that are familiar in households. The final category is built-in lighting, for which there are various bulb sizes and attachment fittings. Further, outside of the United States there are numerous other fittings and housings found in countries around the world [10]. Designing a single system to cover every global AC consumer may be impossible, and in practice a diverse metering system may be required.

1.1.2 Consumption Shifting to Miscellaneous Electrical Loads

The building energy problem is also not fixed in time; as technologies and consumer trends evolve, the energy problem evolves alongside. Specifically, the amount of energy consumed in buildings is shifting toward low power outlet devices. Programs have emerged that target the most significant energy consumers, including replacing appliances with ENERGY STAR® upgrades, adding demand response systems targeting HVAC and lighting, and more. As this has been happening, outlet devices have been largely ignored. Previously thought to

be too low power to have a significant impact, miscellaneous electrical loads (MELs) are increasingly becoming a significant contributor to the overall energy.

As various savings are realized, plug-loads may come to comprise 50% of the total [31]. This is because none of the programs and policies have been targeted toward this segment, and their growth has continued unchecked. According to the Building Technology Office, as "the building envelope moves toward the thermodynamic limit, the 'Other' end use category ... becomes the dominant contributor" [31]. Several methodologies are beginning to address this concern, but the common theme among them is increased understanding into this otherwise black box of energy consumption.

1.2   Impact Potential of Energy Data

Submetering, or breaking down the power draw of each AC device, provides the information needed for individuals to take action against their energy consumption and, if scaled to the city or national level, for regulatory agencies to find and develop effective means to reduce the overall contribution of MELs. Studies have shown that information and feedback on energy consumption can enable and empower users to reduce their energy consumption by 5%-20%. The effects have been shown to be long-lasting if applied correctly, and in particular in a way that is targeted to the specific usage patterns of the household in question. This requires not only a detailed understanding of how energy is used in that household, but also continued monitoring to provide responsive feedback over time.

Device energy data can also enable building energy applications beyond just feedback. Devices can be categorized based only on energy fingerprints, resulting in a human-understandable label for each energy stream. Studies have shown that device-level breakdown increases the effectiveness and permanence of energy feedback, and automatic classification reduces the human burden of device naming. Energy data can also be used to detect faults in devices. This is not uncommon in industrial facilities, but has untapped potential in residential situations. Detecting anomalies could be used to flag inefficient activities, occupant health and safety issues, and even the home router needing to be reset.

Energy data at the device level can also be combined with data from other sensors that measure occupancy and temperature to enable more detailed control and larger energy savings. This can involve lights turning on and off in response to user occupancy, which has been shown to have impact on total lighting energy, or preemptively activating the HVAC when heat-producing devices are turned on. By incorporating machine learning the system can also predict occupancy, which enables smart thermostats preparing the household for user arrival. These inexpensive sensors enable significant opportunities.

4

At the policy level, identifying the categories of devices that are the largest contributors to overall consumption allows regulators to target efficiency actions toward the greatest need. As MELs become more numerous and comprise a larger segment of energy consumption there may be opportunities for regulations to step in with an ENERGY STAR®-like certification for lower-power devices. Even a standard control interface - wireless radio and protocol - to enable a demand response system to actively control energy consumption in real time would require manufacturer buy-in that could be motivated by a large instrumented dataset. Submetering, and the data produced, provides an opportunity not only to inform those regulations in a way not currently enabled by existing technologies, but also to direct the future demand response investments to the areas and segments of greatest need.

Each of these applications or interventions requires device level energy data collected over time. From individual feedback to longitudinal energy evaluations and high-level applications, the platforms are enabled by consumption data for each AC device. There are multiple methods for collecting that data, both measuring at a single point and disambiguating and instrumenting each device individually, each with strengths and weaknesses.

1.3   Acquiring Dense Submetering Data

Providing fine-grained feedback on device AC energy consumption is difficult in residential and commercial environments for three reasons. First, there is a high density of devices. Homes and offices often have many more plug-load devices than occupants in the space, and these devices are often clustered on specific locations like surge strips. Second, the space is also densely occupied by users who care about the aesthetics of the space, and do not tolerate a high visibility large deployment. Third and finally, the devices to be measured are generally diverse and unique, with variations across plug fitting and in the nature of the power drawn between real, reactive, and apparent power.

Two options exist for acquiring this energy measurement. The first is to measure the aggregate power being consumed by all the devices in the space (a whole-home measurement), and disambiguate or disaggregate the individual contributions of particular devices. This type of system is very easy to deploy, it only requires one sensor and in some cases this can be the smart meter already deployed by the utility company. However the accuracy of these systems are limited to higher power household loads, and the state of the art is currently inadequate for measuring very low powered loads.

The alternative to disaggregation is to place meters on the devices themselves, or at the level of the circuit in the electrical panel. This has the benefit of accuracy, often these plug-load meters have accuracies of 99% or better. In addition, many of these devices offer
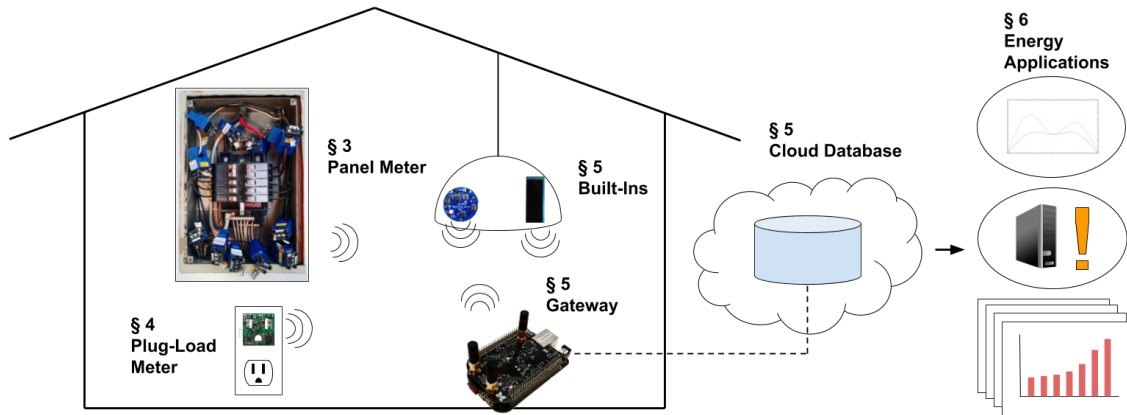
Figure 1.2: Energy metering system for measuring panel devices, plug-loads, and built-in lighting. Each sensor transmits data to one or more gateways, which upload data to a cloud database. From there, data can be used for a number of energy metering applications, including automatic device control, fault analysis, and breakdowns of device energy use.

control, which may be required to fully realize energy savings. However, the downside to this approach is the size, cost, and deployment time associated with these systems. These limitations often mean that deployments involving these sensors target specific subsets of the total device set, which effectively parallels the limitations of the disaggregation methods.

In particular regarding size, if the energy is metered at the point of the device itself there are strict size limitations imposed by the environment. In order to not require significant modifications to the wiring in a home or business, the meter must not block adjacent outlets either on the wall or on surge strips. If this were not the case then additional extension cords would be required, breaking the requirement on aesthetic. The plugs of outlet devices are defined by the NEMA standard for dimensions for wiring devices [18]. This means that adjacent outlets are about 1.1" apart on wall outlets, and will typically have similar spacing on surge strips. This imposes a size limit on the meter, at about 1" on a side.

In this dissertation we explore a system for measuring the energy delivered to household devices, particularly the underserved categories of plug-loads and built-in lighting. Figure 1.2 shows the architecture of an unobtrusive metering system, which requires exploring new designs that enable smaller and lower power energy meters. This has the potential to push the deployment narrative past existing implementations, to directly submetering one hundred percent of electrical energy consumption. Achieving this design point requires revisiting every aspect of the meter, in particular enabling miniaturization through dramatically driving down the energy cost to operate the meter.

1.4   Thesis Statement

Pervasive electrical plug-load metering, enabled by new energy meter designs that are inexpensive and nearly invisible, provides unprecedented visibility into plug-load usage that, combined with processes to produce human-understandable reports on consumption and wastage, yields energy management actions and enables applications to manage these loads and reduce their energy use intensity.

1.5   Contributions of this Dissertation

This dissertation presents an unobtrusive metering platform for collecting device level energy data in residential and commercial situations. In particular we target the size and usability of the metering devices to improve their applicability in high-density situations, as well as the price of the deployment to enable scale. This platform enables an energy dataset unprecedented in its coverage of electrical devices in a home, which can be used to inform future energy decisions.

First we present the design of a novel energy meter called Monjolo that addresses the problem of deployability, allowing for battery-free non-contact energy monitoring that, unlike other systems, has no idle power. Monjolo is built on the concept of harvesting energy from the side channel emissions of electricity transmission and use. The amount of energy harvested will correlate directly with AC current, and this relationship can be used to build an energy meter. This concept enables a design that is easy and safe to deploy even inside the electrical panel, presenting a new opportunity for residential and commercial sensing at the circuit level. This represents the first component of an ecosystem of devices for AC electricity metering, which we expand upon in subsequent sections. This technique was presented at SenSys'13 and was done in collaboration with Bradford Campbell and Prabal Dutta [64].

We continue to expand the coverage of the metering ecosystem by re-imagining the plug-load energy meter in an entirely planar form factor. The small size allows the meter to sit inconspicuously between the plug and the outlet, performing its function nearly invisible to the user. This enables deployments at a density unprecedented in other studies, both because of the physical overhead and the aesthetic concerns with the user. Achieving this form factor requires revisiting every aspect of the AC meter, including the power supply, voltage and current measurement, power calculation, and data communication. This new metering system appeared at SenSys'15 and was done in collaboration with Branden Ghena, Ye-Sheng Kuo, and Prabal Dutta [65].

7

Combining the systems from the previous chapters, as well as other sensors, we create an energy metering sensor system for collecting house-level energy data. We deploy this sensor system at eight residential locations and two commercial locations to evaluate the performance of the system and generate a dataset. We find that the unstructured Bluetooth Low Energy (BLE) network achieves sufficient coverage given transmission redundancy, and that in controlled experiments the metering system achieves full energy coverage on plug-loads and lighting built-ins with high metering accuracy.

In addition to evaluating the performance of the system, this dataset allows us to explore the increased visibility that the metering system provides. In particular we produce reports and breakdowns similar to what could be provided to the homeowner in the form of feedback, or to policy makers and regulators to make policy decisions. We augment the deployment system with a simple occupancy sensor to generate additional conclusions on savings potential and to enable applications that directly target the idle power in AC devices. We evaluate the dataset from the 10-location deployment for other building energy applications like automatic device identification and fault detection. Although the performance of each of these applications will be increased with larger deployments and more data, this provides the framework of the possibilities using PowerBlade data and allows for a comparison to existing datasets.

Finally we examine some future directions that could further improve the system, including technological improvements to the meter itself. In addition to increased safety and further decreased size in the meter, keeping time in the PowerBlade network could make the data more robust against packet loss and gateway downtime, and further increase the usability of the sensor platform.

CHAPTER 2

Background

Targeting energy efficiency in buildings is increasingly an area of national concern. The U.S. Department of Energy (DOE), Buildings and Technology Office (BTO), Energy Information Administration (EIA), and others publish annual and quadrennial reviews showing significant portions of the overall energy budget being consumed in buildings, with varying efficiency. This has lead to several targeted programs and solutions addressing certain aspects of the overall problem, from HVAC efficiency or building envelope to efficient appliances and water heaters. This continues to be an active area of development, with significant additional savings expected.

As energy savings have been realized in other sectors of the building environment, Miscellaneous Electrical Loads (MELs) have largely been ignored. These are the diverse outlet loads that are increasingly common in homes and businesses, and that are harder to classify and understand than the fixed high powered appliances and HVAC. Attention is turning to this previously ignored segment, but there remains significant opportunities for both understanding their consumption and limiting the future growth.

2.1 A Survey of Building Energy Consumption and MELs

Total residential and commercial energy consumption in the United States is currently on the order of three thousand terawatt hours (3,000 TWh) per year [32]. This represents, by many accounts, over 40% of total energy and 70% of electrical energy consumed in the U.S [31,98]. Although their energy footprint is often smaller, this energy breakdown is repeated in countries around the world [98]. Looking forward, a study by the National Science and Technology council predicts that "If current trends continue, by 2025 buildings worldwide will consume more energy than the transportation and industry sectors combined" [106].

This problem is receiving significant national attention, with increased prevalence of LEED certifications in buildings [12] and ENERGY STAR® rated appliances [6], but the

9

| (a) Residential Energy | (b) Commercial Energy |
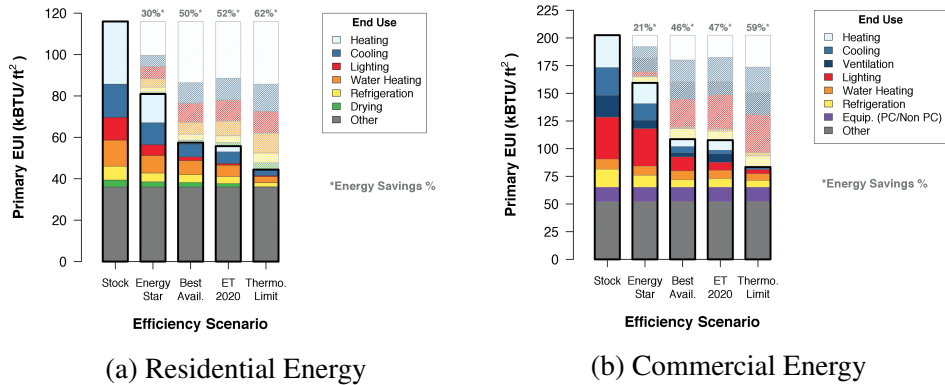
Figure 2.1: Energy use intensity (EUI) of residential (a) and commercial (b) buildings. This DOE report illustrates the EUI of MELs, "Other", under four scenarios of reduction in other sectors. If the 2020 goals are met MELs might make up over 50%, and at the thermal limit as much as 60%. If they are not directly addressed, MELs may soon dominate energy consumption [31].

bulk of the problem remains existing buildings. Another recent study from the NTSC finds that "for the foreseeable future, the greatest national energy saving potential lies with improvements to existing buildings." [105] Understanding these buildings, as well as their occupants and the way they consume energy, will be critical for targeting the demand response programs of the future.

As more attention is paid to the problem of energy consumption in buildings, the BTO predicts that the largest consumers of energy will gradually begin to show improvement. Figure 2.1 shows a projection from the DOE Quadrennial Technology Review 2015 [31]. Demand response systems will target HVAC and lighting, and high efficiency building materials will gradually reduce the importance of heating and lighting altogether [31]. However, as this is happening there is less attention paid to the diverse plug-load devices that are increasingly common in households and offices. These miscellaneous electrical loads (MELs) include televisions and set top boxes, computers, and cell phone chargers, and, if left unchecked, may eventually dominate overall energy consumption.

To understand what makes up MELs household energy consumption (HEC), Table 2.1 shows the residential breakdown of MELs devices today according to the EIA [34]. The installed base represents the number of each device present in households, and some households may have more than one of a certain device type. Televisions are the highest-energy category, due to a large installed base and medium to high power draw in each unit. Pool heaters and ceiling fans are traditionally non-outlet devices, and computers and set top boxes are also major contributors. Another study finds that, "televisions (23%), portable

10

and outdoor lighting (21%), and PCs (12%, including monitors and peripherals) are the largest contributors to average HEC and, together, represent just over half of average MEL HEC" [109]. Electronics, in particular televisions, are major factors in energy consumption.

The density of these devices, in particular consumer electronics, is increasing dramatically. MELs in the average commercial space accounted for 1 W/ft^2 in the late 1990s, today that figure has grown to 8 W/ft^2 [88]. One study finds that the "rapid market penetration of consumer electronics has expanded the MELs category significantly, however, the energy use and reduction strategies for MELs have so far received little attention" [88]. Even in cases where the devices themselves are more efficient than previous models, "in many MELs, decreasing energy consumption is offset in future years by growth in installed base" [34]. The dramatic increase in use of consumer electronics is having an effect on total energy.

In addition to being a significant amount of energy by absolute metrics, this accounts for a significant portion of the total energy spent in buildings. Currently MEL devices account for approximately 30% of total building energy [109], but in the future these devices may come to comprise upwards of 50% of total energy consumption [31]. Further, although 80% of this energy is active use, "idle, sleep, and off [account] for about 7 percent, 0.1 percent, and 13 percent of HEC, respectively" [109]. Even targeting just the idle, sleep, and off power for MEL devices can have a significant impact overall.

Device Type	Installed Base (millions)	Energy (TWh/yr)
Televisions	355	70
Pool Heaters	10.4	26
Desktop PCs	102	22
Set Top Boxes	176	22
Ceiling Fans	263	20
Audio Equipment	193	16
Monitors	130	13
Dehumidifiers	15.6	11
Laptop PCs	165	10
Portable Electric Spas	4.63	9
Modems & Routers	138	7
External Power Supplies	1,077	7
DVD Players	227	6
Non PC Rechargeable Electronics	1,200	4
Home Security Systems	28	1

Table 2.1: Installed base and annual energy consumption in residential buildings in the United States [34]. Televisions are the largest category for energy, due to both a large installed base and medium to high power draw. External power supplies and rechargeable electronics are the small but growing segment of smart home devices, smart phones, and other electronics.

Reining in this segment will require action on multiple fronts. Demand response, whether through regulatory incentives or consumer programs, has the potential for significant impact on MELs consumption. In particular studies have demonstrated the importance of increasing the regulatory or certified coverage over households [61, 98]. In addition, "47% of the total annual consumption of the top 20 residential and top 20 commercial MELs" can be reduced via upgrading to more efficient devices [96]. The overall scope of possibilities includes manufacturers upgrading products either voluntarily or through regulatory action, motivating customers to buy efficient products, and promoting behavioral initiatives that reduce consumption [96]. In the following section we explore these options more closely, including an analysis of where and how to collect data on the building energy problem.

2.2   Impact of Feedback on Residential Energy

Significant attention has been paid to the effect of feedback on residential energy consumption. This is, in effect, the first layer of application that can be applied to solve the building energy problem – inform the user of energy use. Studies have shown that consumers currently lack information required to make decisions on energy use in their households [89]. Further, direct feedback, along with goal setting, can be used to reduce users' electrical energy consumption [43, 49]. This provides the basic framework of the energy feedback application: empower the user to make decisions at the household level that can aggregate to a solution at the whole-problem level.

2.2.1   Basic Feedback Application

Numerous recent studies have examined the effect of feedback on household energy use. These studies vary in methodology and degree of feedback, and achieve different levels of result. Households in Japan after the 2011 tsunami achieved a voluntary reduction of 2%-4% that lasted over two years [74]. Other studies have achieved energy savings of 5% [70, 111], 10% [62, 67, 77], and even 20% [54]. The wide range demonstrates that the type of feedback, and how it is presented to the user, have significant impact on the effectiveness of the intervention. However, it should also be noted that participant willingness is also a contributing factor, and the permanence of the effect is an important metric.

Studies have shown that direct, targeted feedback is more effective than other types [62, 67]. Figure 2.2 shows the savings potential of two types of feedback, indirect and direct, with five total sub-types within [67]. The first is indirect feedback provided after the consumption occurs, like in the form of a bill or online report. The energy savings potential for indirect
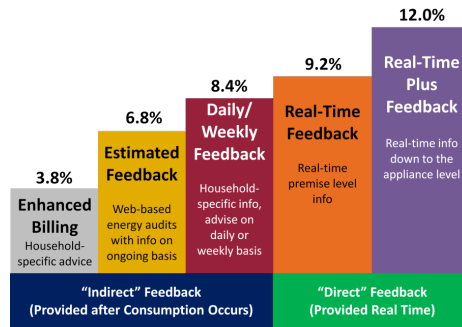
Figure 2.2: Potential energy savings due to residential energy feedback, found in [67]. Feedback can be classified between indirect feedback, which is provided after the point of consumption, and direct feedback, which is provided to the user as they are using energy. The savings potential increases between indirect feedback and direct feedback, and further if the application is able to provided analysis at the specific device level.

feedback ranges from about 4% for some basic targeted advice, to almost 9% for daily or weekly feedback. The second overall type is direct feedback, where information about consumption is available to the user while it is taking place. Direct feedback has the potential to save almost 10% for premise information, and up to 12% for information on specific appliances. This both demonstrates the overall potential of feedback, but also indicates that targeted and timely data are required for realizing that potential.

2.2.2   Formatting and Presentation are Linked to Performance

The formatting and presentation of energy feedback is directly linked to how effective it is at helping the user save energy. In particular a single audit is not effective at reducing household energy use, lasting results require numerous interventions [37]. The solution must be customized for each household [37, 47, 93], and should be broken down to the level of individual devices [72, 73, 83, 89]. Historical comparisons (comparison to the same household over time) are more useful than normative comparisons (comparison to other similar households) [37, 55]. Users need repeated feedback catered to their own situation and usage pattern.

The units of the data in the feedback mechanism are also related to the performance for energy savings. Users have difficulty relating to units of power and energy, and cannot visualize the benefit of savings. Converting to financial terms, or units of CO_2 emissions offset, has a larger effect [81, 83, 89]. In particular one study finds that the unit of "effective number of trees offset" is the unit that causes the user to save the most energy [83]. In fact, users presented with units of watts and kilowatt-hours actually increase energy use over time while the "tree" users reduce use. Users relate to the format of the feedback.

13

Regarding visual formatting on a display, either at the location or web-based, numerical outputs are preferred over pictorial representations [55, 83]. Color in the display does not lead to higher accuracy in understanding, but shortens the response time and increases user preference [55].

2.2.3 Permanence of the Effect of Feedback

Several studies have recognized a relapse effect, wherein a household will adopt unsustainable efficiency practices at a start of the study and end up using more energy at the end than before the study began [83, 93]. This is particularly noticeable in situations where the intervention is one-time, such as an energy audit, without follow up [72]. Even in situations where the intervention is a weekly email, engagement is high surrounding the email and drops off significantly on the days following [54, 83]. The effect is also unsustainable in situations where the users are presented only with units of energy and power, rather than numbers that are relatable like number of trees required to offset [83].

Permanence has been observed in certain categories of intervention more than others. Appliance upgrades result in a permanent effect, where the homeowner utilizes the feedback to identify a device in need of replacement [57, 111]. Similar, although easier to reverse, are set-point adjustments like to the fridge temperature [74]. Certain user actions have also been found to be more permanent than others, particularly watching less television and using lighting more efficiently [74, 111].

In addition, not all users are alike. For example the relapse effect can be observed in households with moderate to high energy use, but not in households with low energy use [93]. In addition methodologies that are customized to the usage type of the user, as opposed to applied generically, typically have a more lasting effect [77]. The interventions that last the longest are persistent, customized, and relatable.

In total, the work on feedback, particularly in residential energy use, presents a framework for what data must be collected, and how it must be presented to users, for maximum and longest-lasting effect. Direct feedback is not the only use of building energy data, but in many cases it serves as the basis for the more advanced applications. In the following sections we explore some of these other applications and their effect on the building energy problem.

2.3 Applications of Building Energy Data

The potential applications for building energy data go beyond feedback on consumption, verifying savings, and measuring persistence. In fact, specific information about each AC

device, potentially combined with other sensors measuring occupancy, temperature, and more, could be used as a platform for running multiple building management apps, and reusing these apps across multiple locations. Even at a location level, multiple applications could coexist on a sensor platform [99]. In addition, energy and building data could be part of a larger Internet of Things (IoT) platform infrastructure enabling city-wide or larger applications [113].

In order to engage users in a lasting way, the application must take a large volume and velocity of energy data and display in a way that is not complex to understand [122]. Targeted feedback can fit this criteria if the output is sufficiently simplified, but there are also opportunities for the system to take action on its own. One example from literature is a system that can detect when several devices that generate heat are turned on, like the television, lamps, and others, and preemptively engage the air conditioning to remain in a more efficient state [57]. In this section we explore a few of the other applications that have been analyzed in literature.

2.3.1   Device Classification and Identification

Although studies have shown that feedback is more effective if it is broken down to the level of appliance or device, this results in a naming problem. Even if a system were to be deployed at each household to collect individual device energy data, there would be significant user overhead in labeling which data stream corresponds to each of their many devices. In this case machine learning and neural networks can be applied to classify the device based on a large training set of seen devices.

There are multiple ways in which AC device classification can be solved, and the choice affects both the data and the processing required to classify. At the lowest level, it is possible to identify and classify devices based on the current waveform of the device (high frequency sampling) [46,68,75]. This leads to features in the classification vectors such as harmonics of the current waveform and the phase between the current and voltage waveforms. In addition, the classification problem has also been solved at the level of power measurements [45,108]. In this case the features include mean power, variance, and other features describing how power varies over time.

These classification applications typically solve either "seen" or "unseen" devices. The former is mainly applicable in situations where the loads are being disambiguated from an aggregate measurement, which will be explored later. However, it is also useful in detecting device movement and replacement. Many systems are able to achieve 90% or higher accuracy on the set of seen devices [68,78,108]. Alternatively, classification of unseen devices is useful

when a new house is instrumented, as the devices can be named and categorized without human intervention at all. Using high frequency sampling on the current waveform can result in 80%-90% accuracy on unseen devices [46, 75]. Applications that do not utilize high-frequency sampling, but instead utilize higher-level power measurements, classify unseen devices with accuracy of about 60%-65% [45]. Several studies also perform unsupervised classification, clustering only based on similar characteristics [82, 121]. Together the works in classification attempt to solve the naming problem, which leads to easier instrumentation in new households and more meaningful data representations.

2.3.2 Fault Detection and Diagnostics (FDD)

Fault detection and diagnostics (FDD) is an area of active research for industrial operations featuring heavy machinery, but it can also be applied to cyber security and even household user comfort. The basic concept behind fault detection is similar to that of device classification: construct a model that describes the performance and characteristics of the AC device. However, in classification the model contains many devices and the output is a certain device type. In fault detection, the model only needs to include one device over time. Certain features that do not typically vary with time, or that vary in predictable ways, can be used as markers for device health. Thresholds can also be applied, such that deviation can result in user notification or even automatic device shut off to prevent larger problems.

Much of the related work in this field can be broken down between malware on computing devices and faults in other devices or systems. This is an important distinction, as in one case there is a malicious action being taken whereas the other is typically an accident or simple wear-and-tear. At the system level, faults can even be generated when the output of multiple sensors deviates from a combined normal operation.

2.3.2.1 Malware on Computing Devices

Regarding malware, research has shown that data from a WattsUp energy meter can be used to detect malware anomalies in medical devices [27, 58]. This approach requires prior knowledge of the characteristics of the device, but this can also be accomplished with a training period. The untrusted device can be characterized, and then a random test application generated to test against expected performance [71]. Because the test application is random, this makes it more difficult for the attacker to mimic the correct operation of the device. Studies have also shown the effectiveness of using energy data to detect malware in SCADA systems that monitor and control critical infrastructure [86], and on software-defined radios (SDR) and cognitive radios (CR) [39]. These systems exploit low level energy data, often

16

considered a side-channel and a potential vulnerability, to ensure the correct operation of an untrusted device.

2.3.2.2 Device Faults and Failures

For accidental and wear-and-tear faults, it is common practice in industrial processes to monitor the state of machinery over time and flag a warning if anything changes from the ordinary [90]. This could include a user operating the machine incorrectly or the machine itself experiencing a fault. In addition, it is possible to detect faults in rooftop air conditioner units through electrical data [42]. Machine learning and a larger variety of IoT sensors can be used to detect faults in domestic appliances [112], aircraft engines [60], and car engines [114]. In all cases the system can be fully automated, simply looking for device behavior that is out of the ordinary.

One of the primary benefits of the fault detection application is that the user is not involved with the application until a fault occurs. Only when a fault actually occurs is a warning raised to the user, under normal conditions there does not need to be any output or user interaction.

2.3.2.3 System-Level Fault Detection

Previous applications of FDD presented here utilize the output of a single device, but it is also possible to generate faults at a system level. Across a household or office, the system could model typical energy use based on the occupants and their devices [56]. With this model of expected behavior, anomalies could be detected that represent inefficient energy use across the space (like a space heater being used while the air is on). For elderly users or those requiring other care, the system could model typical daily behavior and generate faults to off-site caregivers if, for example, the kettle is not used when it typically is [40]. This could enable a higher level of independence without an increased level of risk.

2.3.3 Occupancy and Other Applications

There are several other building data applications that do not fall into the categories described above. Many of these other applications use energy data alongside a simple occupancy sensor, which often adds significant benefit for minimal increased system cost. For lighting, studies have shown that occupancy-based lighting control can be used to save 24% on lighting energy [97]. A smart thermostat correlated with an occupancy sensor and predictive analytics can also save significant energy on building heating and cooling [69,102]. These applications can result in significant savings with the addition of only a simple occupancy sensor.

Also related to occupancy, the use of many AC devices is closely tied to use of the space. For example lights, televisions, blenders, and many other devices are only ever used when the room is occupied. Studies have shown the ability to use energy sensors to predict occupancy with 60%-90% accuracy [53, 94], and have explored the privacy implications of this dataset from the same perspective [104, 118]. Adding other sensors to create a building measurement platform can be further used to increase the accuracy of this system. Using illumination, sound, or CO_2 along with energy data can predict occupancy with 98%-99% accuracy [41]. These systems could be used to replace traditional occupancy sensors for the applications above, or to increase confidence in the simple occupancy output.

2.4 Estimation of Device Energy Data

Each of the applications in the previous section relies on individual device energy data, sometimes correlated with other sensors measuring occupancy and temperature. Direct, appliance-level feedback is the most effective source for meaningful and lasting behavior modification. Device identification by nature relies on individual device data, but in addition fault detection, occupancy prediction, and anomaly detection all rely on information at the level of the device, not the level of the building or the room.

Collecting device-level energy data can be done at the the device itself, but this requires a sensor for each device to be measured. Since residences and commercial buildings often have dozens or hundreds of devices, this can result in high sensor overhead. The alternative is measuring upstream of the device in the energy distribution tree, typically at the entry point for each location. This requires only one sensor per location, which eases the sensor burden and simplifies data collection. However, now the individual device components must be disaggregated from the total. This can be very accurate for high-power devices with distinctive energy signatures, but has lower accuracy on low-power and more generic devices. The choice of where and how to collect device level data has implications on the degree to which the application enabled serves the intended purpose.

2.4.1 Whole Home Metering with Disaggregation

The top-down approach to energy metering is called disaggregation or Non-Invasive Load Monitoring (NILM). The original NILM system was first proposed in [80], and focused on high-powered appliances. The approach examines aggregate power data to look for rapid edges which can be correlated with devices entering certain states. This applies to devices that can be described with fixed states, like a toaster or a clothes washer (fill, spin, drain).
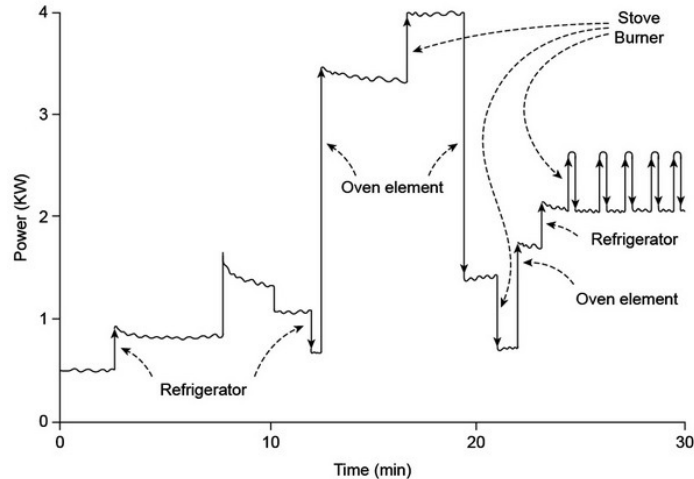
Figure 2.3: Basic NILM operating principle from the original work in disaggregation. The technique measures the energy consumption of the whole home at a single point, and disaggregates individual components to assign energy to particular devices. The original NILM technique, shown, searches for rapid edges in the energy signature which indicates devices are turning on or off [80].

Figure 2.3 shows the basic NILM concept in operation. The system disaggregates the signature of a particular appliance though the state transitions of that appliance. The approach in [80] has two limitations, first that it can only disambiguate loads above about 150 W and second that it only applies to loads that can be described with a finite number of consistent states. With these limitations, this system reports a disambiguation accuracy of 85%.

There have been numerous improvements since the original proposal of the NILM system. The original system is focused only on device state transitions, and accuracy can be improved by analyzing steady state components, turn-on transients, and even higher order harmonics of the 60 Hz signal [52, 76, 87]. Deep learning and recurrent neural networks (RNNs) on the device level disaggregation problem also improve accuracy and the number of devices disambiguated [91]. The system can meter at a single watch point without breaking the path, watching only for EMI generated by devices on the shared AC line [78, 107]. Similarly, EMI sensors placed throughout the location can detect changes of state of electrical appliances by the changes in the side-channel magnetic field given off by AC current. These sensors can be used in conjunction with a panel-level sensor to correlate changes in power with specific devices and appliances [110]. The system can even operate without a direct meter providing feedback, and simply approximate device energy breakdown from the monthly bill and by comparing to similar households [48]. Disambiguation accuracy rates are on the order of 90%-95%, and target on the order of 10-50 appliances and other devices in the home.

19

The state of the art in non-invasive load monitoring has been previously summarized in [51]. In this work the authors describe the benefits of individual device energy information, and that NILM has the greatest potential to address individual device metering in a scalable way. In addition, smart meters are increasingly being deployed by utility companies at no cost to the consumer, and provide sufficient data to enable certain NILM techniques. The authors find, however, that the state of the art in NILM is insufficient to solve the device metering problem [51]. Although NILM techniques may ultimately yield sufficient granularity, the state of the art is currently not a viable solution.

2.4.2 Individual Device Metering

Metering at the level of the device is accurate, plug-load meters are typically 95% accurate or better, but many units also have the ability to control power to the load. This introduces additional savings potential, with applications directly controlling rather than simply metering. Further, each device is directly measured, the contribution of the long tail of energy consumption can be directly analyzed. The downsides to this approach are the size and the cost of the individual meters themselves. In the following sections we describe the factors that determine the size and cost of an AC meter, as well as outline several deployments that have taken steps to mitigate these effects.

2.4.2.1 Systems and Designs

A plug-load energy meter must supply itself with power, measure current, measure voltage, calculate power from those measurements, and communicate the power readings to the user. The decisions for each of these five components will affect the functionality of the meter, as well as the size, cost, and other factors like internal power consumption. We explore each in tern in the context of existing meters, both from academia and from industry.

We analyze three AC power meters from academia - ACme-A [84], ACme-B [84], and PowerCube [66] - and three from industry - the Kill-A-Watt [11], WattsUp Pro [27], and Belkin Conserve Insight [4]. In addition to the decisions made by these meters, we include in this evaluation additional possible approaches, which are discussed more in subsequent chapters.

Power Supply. Converting the high voltage AC to a usable low voltage DC has a direct impact on the size of the system and the amount of power it has available. While the exact implementations vary greatly, there are two common components to each AC-DC supply: a rectifier and a component used to step down the voltage. Rectification can often be done with a few small components, but voltage step down requires a larger component.

The most common solution in wall wart supplies is an AC transformer. A transformer has the benefits of efficiency and isolation from the mains, but often introduces significant size overhead. The smallest AC transformer available from Coilcraft, for example, is approximately 0.31 in^3, and many are much larger. Of the meters we evaluate, only the ACme-B uses this method to step down the voltage.

More common among the meters we evaluate is to step down voltage using a high voltage capacitor. The capacitor is used to limit current from the AC, and the voltage can then be regulated by either a Zener diode or a voltage regulator. This method is often smaller than the transformer option, and has the benefit of volume scaling with the output current. The current will be inversely proportional to the capacitor impedance, which means it will scale directly with the capacitor value. A higher value capacitor can supply more current. However, these capacitors are often themselves quite large.

A third option is to harvest energy from the output of a current transformer attached to the load. This method is non-contact, thereby achieving equivalent isolation to that of power transformers, but current transformers can be found in packages much closer in size to the capacitors in other supplies. None of the meters that we evaluate in this section utilize this method, but it will be explored more in later chapters.

Current Measurement. All but one of the meters that we evaluate measures current by breaking the path of the AC and passing one leg through a current sense resistor. This method is inexpensive, accurate, and small, and only introduces a small power draw. The drawback to a current sense resistor is the requirement on breaking the path, any design that harvests energy through a current transformer, for example, does not have access to this option.

The only device that we evaluate that does not use a current sense resistor to measure current is ACme-B. Instead this device uses a Hall effect sensor, which measures the deflection of electrons in a conductor exposed to a magnetic field, like the one generated by a current.

A final option for measuring current is a current transformer, which outputs a signal proportional to the current flowing through the open core. This option is non-contact, and split core current transformers exist which would enable clip-on sensing.

Voltage Measurement. A true power meter requires real time voltage measurements, combined with current measurements, to generate real, reactive, and apparent power data. All of the meters evaluated in this section measure voltage directly through a voltage divider. While this is inexpensive, accurate, and occupies a small volume, direct voltage measurement requires that the measurement circuit be referenced directly to the AC line. In other words, selecting a transformer for the power supply element provides isolation, but then the voltage

measurement circuit would require isolation as well. This would add significant overhead to the design of the AC meters we evaluate, and is one explanation for why all use the capacitor fed supply.

Power Calculation. Converting raw measurements to metrics like volts and watts can happen in several ways, either locally on the meter or once the data have been transmitted. This is a case where local calculation can provide significant savings on the wireless transmission. Typical sampling rates for power meters are in the range of 500 Hz-5kHz, and for 8-bit ADCs on the voltage and current channels this represents 1 kBps-10 kBps of data. However, these samples can be succinctly described by industry-standard metrics like true power and power factor. Common reporting rates are 1 Hz-5 Hz, which means that calculating these values locally rather than transmitting raw samples can reduce the channel burden by a factor of 100 or more.

In the meters evaluated in this section, local calculation has taken one of two forms. Many commercial products like WattsUp, as well as the ACme-A, ACme-B, and PowerCube systems, use a power metering IC like the Analog Devices ADE7753 [1]. This is a device that measures both the line voltage and the output of a current sense device like a sense resistor. The ADE7753 has a serial output for reporting true power, apparent power, zero crossings, and other measurements to a microcontroller. Although this can be very accurate and simple to implement, the ADE7753 adds size and power overhead as well as cost to a design, in addition to still requiring a microcontroller.

A final option which is found in several commercial meters is to connect the voltage and current signals directly to ADCs on the microcontroller on board, and implement custom metering software. This has the potential to be very accurate, but at the cost of increased development time compared to packaged solutions like the ADE7753.

Data Communication. Once voltage and current have been measured, and the samples converted to power readings, that information must be conveyed to the user. This seemingly trivial step has deep implications on the design of the system, as it is often the component that requires the most physical space, energy, or both.

In all of the commercial meters that we evaluate (Kill-A-Watt, WattsUp, and the Conserve Insight), data communication to the user is done with a LCD screen mounted on the unit itself. This method has the least set up requirements: once the meter is installed the data is immediately available. However, this has the limitation of data only being available physically at the meter itself.

The alternative employed by all of the academic meters that we evaluate (ACme-A, ACme-B, and PowerCube) is to convey the measurements wirelessly from the point of the meter to another location before being delivered to the user. Belkin and others do offer

wireless options similar to the Conserve Insight with otherwise similar internals [9, 28]. This allows the unit to be much smaller, the radio itself takes much less space than an LCD screen, but also offers the important benefit of centralized data. This enables applications on the metering data that are otherwise impossible with visual readouts.

2.4.2.2 Experimental Results

Several studies generate useful power metering data with the bottom up metering approach. The meters are typically too large or expensive, or just physically impractical for achieving full coverage of every AC device. Instead, they select a statistically relevant cross section of devices and extrapolate, position sensors strategically upstream of multiple devices in the load tree, or some combination. In all cases the researchers construct a model that accounts for gaps in coverage by using the data that are collected.

Dawson-Haggerty et al. conducted a study using the ACme meter, deploying 455 meters and seven load balancing routers in a commercial office space [63]. The 455 devices they selected to meter were drawn from about 5,000 total, so this deployment achieved about 10% direct metering coverage. The devices selected were weighted toward devices that draw higher power (i.e. a computer monitor is more likely to be selected than the desktop speakers). The assumption is that direct metering data from this cross section of the total devices can be used to extrapolate and derive insights about the space as a whole.

Although this deployment is, in part, more of an evaluation of indoor sensing at scale than an explicit energy evaluation, three things prevent it from generalizing to the average metering case. First, selecting a statistically relevant cross section of devices is more appropriate in an office space, where a similar device footprint is repeated in each cubicle, than in a residential environment where consumption is individualized and diverse. Second, the ACme meter is large and blocks adjacent outlets, and 20% of the deployed devices were removed by the users during the study. Finally, at $70, the cost of these meters prevents larger deployment.

Another similar deployment involves the ACme sensor, as well as other sensors for measuring light and vibration [85]. This work combines these sources into a single dataset, using the light measurements to infer time in mode for overhead lighting and correlating with known power. Several aspects of this work can inform our own, in particular the multi-modal approach to sensing light as a side channel for the energy consumed by overhead lighting. However, this work still relies on a meter that is too large and expensive to deploy at scale. This leads to the requirement to measure upstream and infer the gaps, which does not necessarily apply in residential situations where each device is plugged in individually and most devices are unique.

The alternative to ACme that still measures real, reactive, and apparent power and transmits it wirelessly, but that also operates in a form factor suitable for dense deployments is PowerCube [66]. This design is based on ACme from a schematic perspective, but optimizes for space by dramatically reducing the power consumption, enabling shrinking in the power supply, and by removing the AC socket and implementing a custom socket instead. This enables a smaller meter, but at the expense of safety. The PowerCube design uses a thin insulation layer between the internal high voltage routing and its own low voltage components. This insulation layer is significantly thinner than the insulation provided by an AC socket, but also provides significantly less isolation. In some situations this thin insulation can even degrade to the point where the high voltage routing shorts to the low voltage components, and the meter is destroyed causing a dangerous situation. Safely shrinking to this size requires eliminating the need for internal high voltage routing.

2.5 Summary

AC energy consumption in buildings is an area of growing national concern, and one that can be addressed in part through submetering and feedback at the level of the individual AC device. Despite this opportunity, existing metering platforms and implementations do not scale to the requirements of dense metering, and are typically deployed with partial coverage of the AC devices. The primary limitation is not the quality of the data but instead the size and cost of the meter itself. In the following sections we detail the designs that result in an energy metering platform that is low cost, easy to deploy, and scalable to the diverse set of plug-loads and appliances.

CHAPTER 3

A Clean-Slate Energy Meter Design

AC energy metering at the circuit level can provide highly accurate data about the set of devices on that breaker, but metering in the electrical panel typically requires rewiring by an electrician. In this chapter we present the design and implementation of Monjolo, a wireless AC meter built on a unique principle in energy metering which is non-contact, thereby increasing the potential safety and deployability of the unit. In addition, the design is power proportional; unlike many systems the device draws less power when metering lower loads, and no power when the load is disconnected. Although aspects of the Monjolo design make it difficult to scale to the full requirements of plug-loads, it can inform multiple tracts of subsequent evaluation in addition to serving as a useful tool as a panel meter or in certain monitoring situations.

The Monjolo operating principle is simple: transmission or use of energy often results in a side channel emission of energy, which can be harvested to power a small low power wireless node. The activation rate of that node will increase with the harvested energy, and will be proportional to the amount of energy in the underlying transmission or use. The energy harvester becomes the sensor, with the amount of energy harvest-able as the metric.

In the case of AC energy metering, the Monjolo principle requires a single AC current transformer, which harvests energy from the magnetic field which is a side channel of the current being supplied to the load. This small amount of harvested energy is aggregated over time in Monjolo until the unit has stored enough energy to send an 802.15.4 packet. A listening receiver can measure the activation rate as the time difference between packets, and correlate this value with the underlying current. Because Monjolo operates solely on the output of the non-contact current transformer, there are no high-voltage components required to operate the basic principle and, if using a split-core transformer, the device can clip on.

In addition to safety and deployability, this design has the benefit of power proportionality. When the load is disconnected the current transformer does not harvest, and Monjolo draws

no power. As the current to the load is increased the harvested energy and wakeup frequency are increased, so Monjolo draws an amount of power proportional to, and only a small fraction of, the measured power.

In the following sections we detail the design and operation of a sensor built on this principle. For this evaluation we produce both a plug-load and a circuit-level meter, although in practice as a plug-load meter it does not meet the size or cost requirements.

3.1   Energy Harvester as an Energy Sensor

The basic Monjolo principle requires two subsystems: the Monjolo sensor itself and the listening receiver. Monjolo harvests from its current transformer until it can send a packet, which the receiver is always listening for. Although the packet can be empty, only the timing is important, we add three additional pieces of information to improve performance. First is a ID number to differentiate multiple Monjolo sensors. This allows multiple Monjolos in a single space to be differentiated from one another. The second is a wakeup counter to mitigate packet loss, and the third is a timer to prevent channel flooding.

In the basic principle, packet loss is misinterpreted as lower underlying current corresponding to a decreased wakeup rate. Since Monjolo employs no channel access protocol, and transmits whenever it is able, collisions and other forms of packet loss are likely. To mitigate the effect of dropped packets, we add a packet sequence number to each transmission. If the receiver observes gaps in that sequence it can still average over the time period for each lost transmission, thereby losing resolution but not overall accuracy.

Monjolo also transmits as soon as it is ready. This has the potential to flood the channel with packets if one or more Monjolo sensors are measuring high powered loads. To address this we add a non-volatile time keeping element on Monjolo. If the sensor is ready to transmit but the time threshold is not passed, it increments its count without transmitting, discharges its storage, and returns to charging. The discharging of storage is important so that the energy per wakeup remains fixed even without a wireless transmission. Each packet, therefore, contains a true packet sequence number used for wireless analysis, but also a wakeup count to quantify the number of suppressed wakeups between transmissions.

Figure 3.1 shows the operation of the Monjolo architecture. Multiple Monjolo units transmit to a single or a few fixed receiver(s). The Monjolo sensors mark their packets with an identification number as well as with the count of the individual wakeups between each packet. The receiver keeps track of each Monjolo in range and, upon receiving a new packet, calculates the wakeup rate for that Monjolo based on the differences in time and number of wakeups. Through calibration we generate a transfer function for each Monjolo, which the
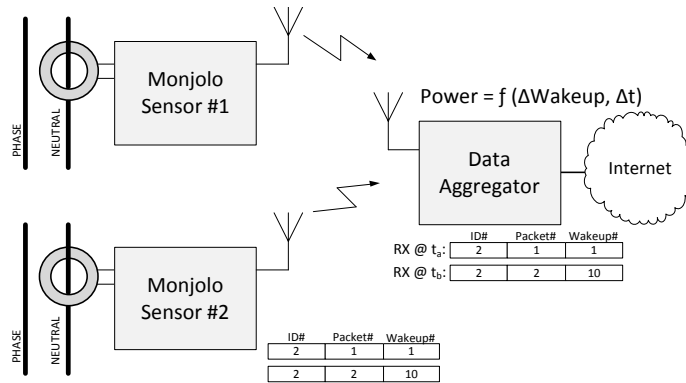
Figure 3.1: Monjolo system overview. Multiple Monjolo nodes transmit to a data aggregator, which estimates power based on the wakeup count contained in each packet. The data aggregator can use the data locally or transmit to the internet.
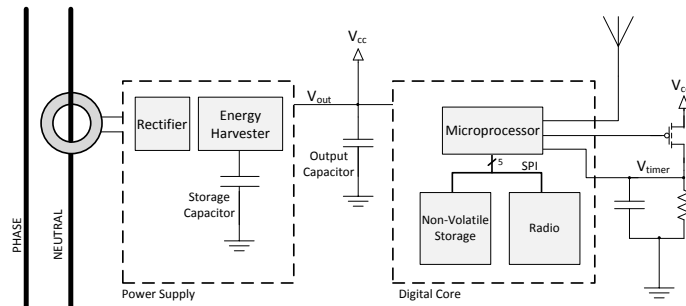
Figure 3.2: Monjolo sensor architecture. The power supply harvests from an AC current transformer, rectifies and boosts the output, and stores it on a bank of capacitors. When sufficient energy has been accumulated the power supply activates the digital core. The core boots, increments its non-volatile wakeup counter, and samples the time keeping circuit. If the voltage has decreased past the software-defined value, the core transmits and recharges the time keeping capacitor.

receiver can use to calculate watts from the measured wakeups per second. This information is now functionally similar to that from other energy meters, and can be forwarded from the receiver to the internet or used locally.

Figure 3.2 shows the block diagram of a Monjolo unit. There are three high-level components: the power supply, digital core, and analog time keeper. The power supply rectifies and harvests the AC signal from the current transformer. It stores energy on a bank of capacitors, and when it has harvested enough energy it connects this storage to the digital core. The digital core is a microcontroller and radio for generating and sending the packet, as well as non-volatile storage for storing the sequence and wakeup counts. The time keeping circuit is a non-volatile RC timer to give the node an approximate sense of how much time has passed since its most recent packet transmission.
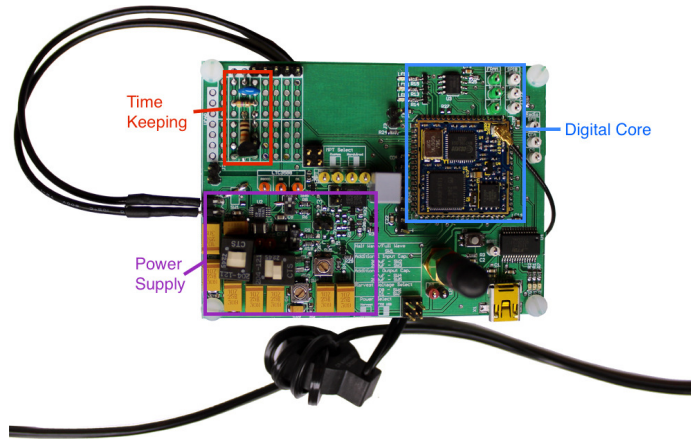
Figure 3.3: Monjolo test platform, containing the LTC3588 energy harvesting IC, storage capacitors, digital core, and time keeping circuit. Also shown is the current transformer, which is used to harvest energy from the AC current.

In this application the Monjolo sensor harvests energy from a current transformer, but any energy harvesting front end could be substituted and employ the same principle. A solar panel, for example, could be substituted to change Monjolo into a light sensor. In many cases this is useful as a relative activity monitor, increased packet frequency corresponding to increased energy harvesting without requiring calibration. This can apply to the power metering case as well, even an uncalibrated Monjolo could potentially be useful in a device health analysis. If the device has consumed energy in a consistent way, as indicated by a consistent or pattern-based wakeups per second, deviations from this steady-state might indicate a rotor or compressor needs maintenance. In all such cases this simplistic metering idea is built on a common back-end and applied to diverse applications.

3.2   Hardware Evaluation Platform

Figure 3.3 shows the test board that we use to evaluate the system. This contains the three high-level components directly mounted on the PCB and an off-board current transformer for harvesting power.

3.2.1   Energy Harvesting Power Supply

In the Monjolo prototype sensor the energy harvesting power supply is the Linear Technology LTC3588 energy harvesting integrated circuit [15]. The front end is a 1:3000 turns ratio current transformer from CR Magnetics, the CR2550. The LTC3588 has a minimum input voltage for harvesting which will correspond to a certain minimum wattage required to

meter. At the other end of the wattage range, a certain wattage will generate enough power to operate the system continuously. The wakeup functionality will no longer take place, and the operation of the sensor would be eliminated. This will correspond to the maximum wattage that can be metered. We find that six turns of the primary coil, for an overall effective turns ratio of 1:500 results in a system that can meter down to 17 W.

The LTC3588 contains an internal full wave rectifier, which takes two AC inputs and charges a capacitor. However, by shunting one of the AC inputs to ground we force the harvester to operate in half wave mode. This results in approximately half the power delivered to the device, but at the same voltage to maintain the same lower metering bound. Since many loads result in packet suppression, operating on half wave rectification reduces the power of the device without affecting accuracy or resolution.

The energy harvester has an internal boost converter, which we set to charge to 5.1 V before enabling the output. The digital core then draws power from the storage capacitor, which reduces its voltage. At 3.8 V the LTC3588 disconnects the output from the storage capacitor, and charging resumes.

We evaluate multiple options for the storage capacitor to determine the appropriate value for this application. The optimal value will be the minimum required to buffer sufficient energy to perform the operation of the node. Since a wakeup finishes only when the LTC3588 senses that the capacitor voltage has decreased to 3.8 V, increased capacitance reduces the resolution of the metering system without additional benefit. For our operating range we determine that a value of 500 μF is sufficient to update the counter, sample the timing capacitor, and transmit a packet. We use a tantalum capacitor due to their low leakage and high charge storage density.

Because the Monjolo operation relies on a fixed amount of energy to be harvested between each wakeup, the possibility exists that large variations in temperature will result in subsequent variations in Monjolo accuracy. Tantalum capacitors have similar temperature characteristics to aluminum electrolytic capacitors, and may result in some capacitance changes with regard to temperature. We explore this effect in future sections to determine if explicit temperature correction is required.

3.2.2 Packet Suppression Timer

The packet suppression timer on the Monjolo prototype is an RC circuit that can be recharged by the microcontroller. The timing element must be non-volatile as the node goes into hard shutdown between each wakeup, and has no on-board knowledge of duration in between. For this purpose we implement a 2.2 μF electrolytic capacitor which discharges through

its own leakage as well as through a 4.7 MΩ resistor. The microcontroller can charge the capacitor through a pull-up MOSFET.

Each time the microcontroller transmits a packet it also charges up the capacitor. On a subsequent wakeup it measures the voltage on the capacitor, if the voltage has decayed past a software-defined threshold the microcontroller transmits again. If it has not, the wakeup counter is incremented, storage is depleted, and the unit returns to shutdown without transmitting or charging the capacitor. We determine experimentally in the lab that, for a reasonable number of transmitting Monjolos, 5 s between packets is a reasonable update rate that does not flood the channel. The time constant of the RC circuit is about 10 s, but the delay value can be adjusted in software by adjusting the threshold voltage value.

3.2.3 Digital Core

The digital core in Monjolo is a Texas Instruments MSP430F1611 microcontroller [24], a CC2420 802.15.4 wireless radio [23], and a Ramtron FM25L04B FRAM module [21]. The FRAM is used for non-volatile storage of the sequence and wakeup counts, and is preferable to flash or EEPROM due to is low-energy writes and high-speed reads.

Each time the power supply charges the capacitor, it enables power to the digital core. Each time it depletes to the discharge value it cuts power. This back-and-forth causes the microcontroller to enter a cold boot each new activation, which both simplifies the software and corrects transient errors.

3.3 Quantization Resulting from Harvesting AC

After implementing the basic Monjolo design, we make two discoveries that impact the functionality of the system. The first is that allowing the system to continue supplying power to the storage capacitor while the core is running has an adverse affect on the power estimation algorithm. Second, certain aspects about harvesting from the AC result in quantization of wakeup rates, with multiple wattages eliciting the same Monjolo wakeup rate. Through experimentation, we demonstrate that both issues can be addressed with the same relatively simple solution, which is to disconnect energy harvesting while the digital core is powered.

3.3.1 Concurrent Charging and Discharging

The Monjolo principle has the constraint that a fixed amount of energy must be discharged each time the node wakes up, regardless of the wattage being measured. If this were not the
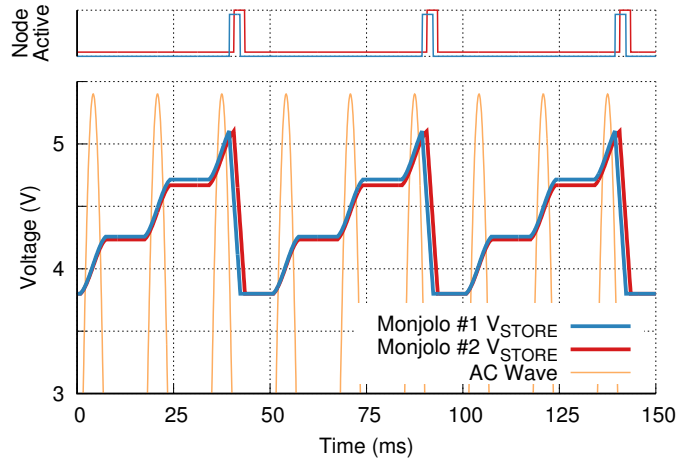
Figure 3.4: Charging cycles of two Monjolo sensors with half wave rectifiers. Monjolo #1 is attached to a slightly higher power load, and reaches the fully charged point slightly faster. However, both activations finish in the same inactive period of the charger, and wait until the subsequent active period to resume charging. This causes the quantization effect observed at high wattages.

case, if an increase in primary load caused an increase in Monjolo runtime as well as the expected decrease in recharge time, then the overall packet transmission rate, and therefore measured wattage, would be indistinguishable. The result might be a totally flat transfer function, the same rate regardless of wattage.

However, the normal operation of the LTC3588 continues to supply the storage capacitor with harvested energy while the output is enabled and the capacitor is discharging. The additional energy supplied will increase with increased wattage, thus breaking the constraint that a fixed energy be discharged under all conditions.

3.3.2 Quantization Issues from Harvesting AC

Another issue that we observe in the operation of Monjolo is the quantization of wakeup rates, especially at high wattages. This stems from the fact that in any AC-DC rectification there will be some period of the wave where the voltage is too low to activate the energy harvester. This is more pronounced with Monjolo's half wave rectifier, as this applies not only to part of the positive half wave but also to all of the negative half wave. This has the effect of quantizing wakeup rates from similar wattages; if two Monjolo devices finish their discharge at different times during the same negative half cycle, they will both wait to begin charging and their overall frequency will be the same.

We illustrate this point with a simple experiment involving two Monjolo devices measuring similar wattages. Figure 3.4 shows the voltage on the storage capacitor for each

31
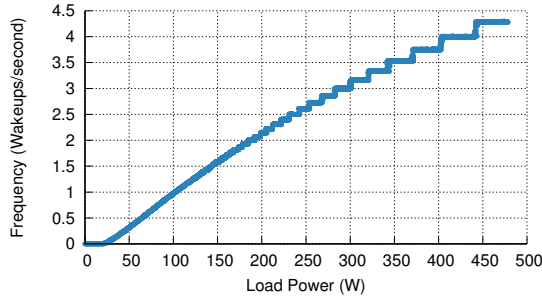
(a) Quantized Monjolo activation rate.   (b) Smoothed Monjolo activation rate.
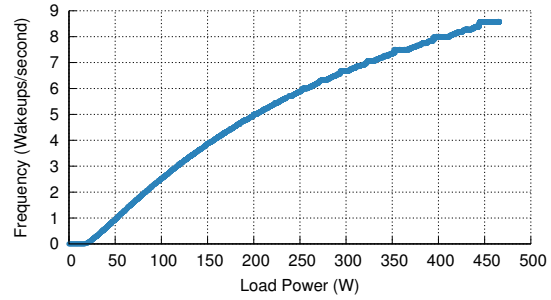
Figure 3.5: Quantized and smoothed Monjolo activation rates. In (a), multiple AC wattages alias to the same Monjolo wakeup rate, and cannot be disambiguated. Because the LTC3588 cannot harvest during the part of the AC wave, similar wattages whose cycle ends during the inactive period will all wait until the subsequent active period to restart the cycle. Their transmission intervals will be the same, and the wattages impossible to disambiguate. Although this effect is inherent in the system, it is exacerbated by allowing the LTC3588 to continue charging the capacitor during an activation. In (b), when the charger is disconnected during the activation period, the quantization effect is much smaller and occurs only at high wattages.

Monjolo, as well as the AC waveform and an activity marker for each Monjolo. Monjolo #1 is connected to a slightly higher power load, and reaches 5.1 V a few milliseconds before Monjolo #2. However, both units finish their discharge during the same negative half wave, and thus wait until the same point to begin charging again. Although the packet from #1 will arrive slightly before the packet from #2, their overall rate is indistinguishable. This is a separate effect from that in Section 3.3.2, and results in steps in the transfer function.

Figure 3.5a shows the wakeup frequency as a function of wattage, demonstrating both the quantization effect and that it becomes more pronounced at higher wattages. The increased quantization at higher wattages comes from fewer recharge cycles being required to fully charge the storage capacitor. In Figure 3.4 it only takes three AC cycles to charge the capacitor, many lower power loads require 200 or more AC cycles to charge. With a lower number of cycles required to charge, higher power devices are more heavily grouped into certain AC waves than lower power devices.

This effect is exacerbated by allowing the node to charge while the core is active. Figure 3.5b shows the wakeup rate as a function of wattage if the charger is disconnected during runtime, and the step size is reduced. It is only at high wattage that clear steps exist at all. This is because quantization does not happen if the node finishes discharging during a positive half cycle. In this case it is able to begin recharging immediately, actually benefiting from its fast recharge rate. If it continues to charge while discharging it will keep the node

alive long enough to enter the subsequent negative half cycle, exacerbating quantization. While there is no explicit solution to the quantization problem, this presents a workable fix.

3.3.3 Disconnecting While Charging

We disconnect the energy harvester from the storage capacitor during discharge, both because the operation of Monjolo relies on a consistent amount of energy being discharged each cycle and because disconnecting helps mitigate the effects of quantization of the wakeup rate. To disconnect the charger, we attach an N-channel MOSFET to the output of the current transformer and tie the gate to the output of the regulator. When Monjolo is powered, the output of the current transformer is shunted to ground and the device does not charge.

3.4 Evaluation of Monjolo Prototype

In this section we evaluate the Monjolo prototype for several metrics, including power metering performance and the usability and cost of the system. Although the lack of voltage measurement results in a weakness in the energy metering aspect, the design is accurate on resistive loads and draws less power than existing metering solutions. The power consumed by Monjolo is also proportional to the power in the load, and Monjolo draws no power when the load is not in use.

3.4.1 Monjolo as an Energy Meter

We evaluate Monjolo on both resistive loads and reactive loads, which together describe the total makeup of residential and commercial devices. Resistive loads are characterized by phase-synchronized sinusoidal voltage and current waveforms. True power is equal to, and the only component of, apparent power. In reactive loads, distortion and displacement factors in the current waveform create the reactive power component, wherein not all apparent power delivered to the load is consumed as true power. Any time current is flowing in the wire that Monjolo is sensing, it is able to harvest energy. This causes Monjolo to measure apparent power, the magnitude of the vector sum of real and reactive power, as opposed to real power itself [100].

The report from Monjolo is accurate if the desired quantity is apparent power. Even if not, the table lookup model that Monjolo uses is based on a unity power factor. Apparent power is always greater than or equal to real power, so Monjolo does not underestimate real power. Finally, power factor (PF) is defined as real power divided by apparent power, so

33
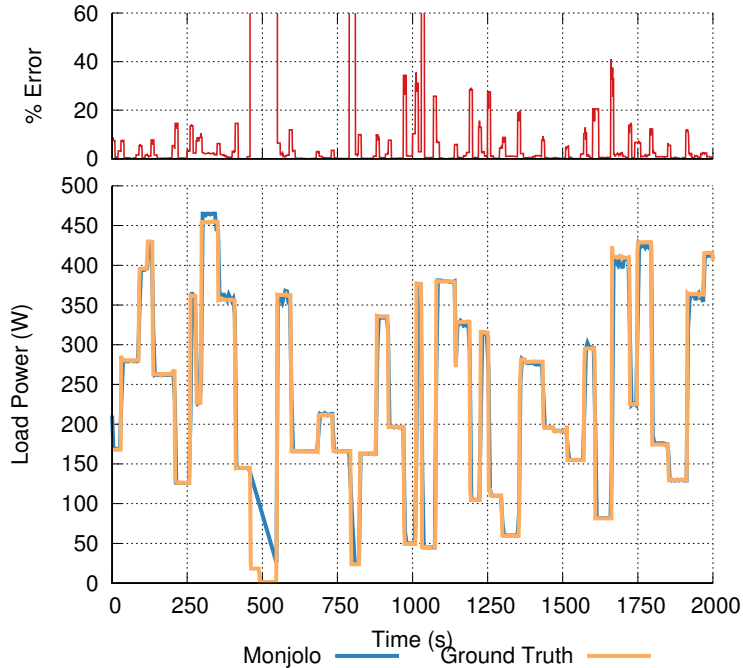
Figure 3.6: Accuracy of Monjolo's power measurements over a range of different loads. When the load is constant, Monjolo's error is very low at only 1%. When the load changes, Monjolo's relatively slow response time causes sharp instantaneous error while the data aggregator waits to receive a packet informing it that the load has changed. When the load is very small, as it is at t = 500 s, Monjolo is unable to transmit and the resulting error is quite large, although a less naive power estimation algorithm would be able to correct for this. There are three periods where low load power causes the Monjolo error to spike over 100%. If we ignore these periods as times when the Monjolo sensor is unable to function, the remaining periods show an average error of 3.7% across a randomly varying resistive load.

Monjolo will therefore be incorrect by a value of 1/PF. Mode transitions are still apparent so Monjolo's time-in-mode reporting remains accurate. Utilizing a single watch-point meter that measures the total power supplied to multiple AC loads being metered by multiple Monjolo nodes may allow the user to interpolate the value of the scaling factor.

We evaluate the differences between resistive and reactive loads through two experiments: first we test Monjolo on a programmable AC load for testing resistive loads, followed by some generic household and lab devices for testing reactive loads.

3.4.1.1 Resistive Loads

Using a software-programmable resistive load, we run an application that randomly selects power values for the AC load and switches after random intervals. This highlights Monjolo's

34

response to different load power levels and a range of power transitions. Figure 3.6 shows the results of a 30 minute run of this application. Monjolo tracks ground truth very closely when the load is constant with an error of only 1%.

During times that the load transitions, Monjolo cannot respond quickly enough. The load estimation error spikes at these transition points. These spikes typically last for a second or so, with error duration largely dependent on the new wattage. At points where the load transitions to a wattage below the harvesting threshold, Monjolo cannot transmit a packet and the error spikes. Because our simple algorithm does not use the sudden lack of packet receptions to make assumptions about what the load "must" be, transient errors can spike above 100%. This is quickly rectified when the next packet is received.

3.4.1.2 Reactive Loads

To examine the effects of reactive loads, we test Monjolo on four household and lab loads, as shown in Figure 3.7. These loads exhibit a variety of different real-world power profiles. Monjolo is able to track the actual power draw of the various loads but overestimates in cases where the power factor is low. This does not present a viable design for plug-load metering, but does allow us to evaluate the performance of Monjolo for future use.

Figure 3.7a shows Monjolo measuring a typical running desktop computer while the computer transitions through several operating modes. Between 250 s and 350 s, the CPU is fully active. At time 450 s, the computer is put into suspend mode until time 900 s. After it wakes up, it again runs at 100% and then after an idle period performs some normal activity. During the periods in which power factor is nearly (but not exactly) unity, (i) Monjolo is off by an average of roughly 6% and (ii) Monjolo exclusively overestimates the load. During the period in which power factor is near 0.3, the error increases dramatically but the transitions into different operating modes are clearly distinguishable, and Monjolo continues to exclusively overestimate.

Figure 3.7b through Figure 3.7d show Monjolo measuring an AC fan, heat gun for reflow soldering, and a soldering iron. These devices were selected specifically due to interesting properties, most household devices will appear similar to the desktop computer. The true power and power factor of the fan stay constant, and Monjolo overestimates by a constant amount. The heat gun shows high frequency oscillations in power, causing Monjolo to report an average value. Finally the soldering iron has a fluctuating power factor that causes Monjolo error to spike sharply, but it quickly recovers. Overall Monjolo shows higher error when measuring reactive devices than on resistive devices, but the error is generally a relatively constant factor. In the future this could be addressed by adding a voltage measurement into the Monjolo system to allow it to measure true power.

35

(a) Desktop Computer

(b) AC Fan

(c) Heat Gun

(d) Soldering Iron
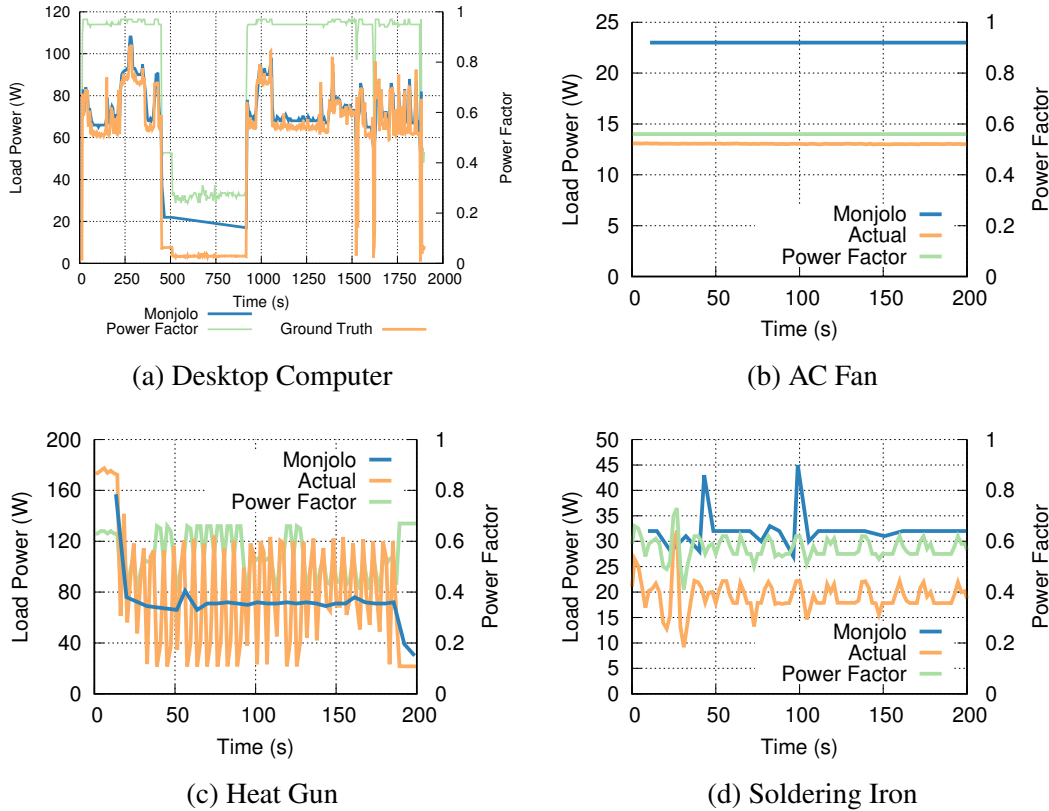
Figure 3.7: Monjolo performance on loads with non-unity power factors. Each figure shows the ground truth power and power factor measurement of the load, and Monjolo's estimate of the power. In Figure 3.7a Monjolo is able to track the power draw with a slight error in magnitude when the power factor is close to 1.0. When the power factor drops to 0.3, the load power also drops, preventing Monjolo from being able to charge. In Figure 3.7b the power factor remains constant, and Monjolo estimates the power with a constant offset. The heat gun in 3.7c is an interesting case when the high frequency oscillations cause Monjolo to estimate an average value. The fluctuating power factor in 3.7d causes Monjolo to sharply overestimate, but it quickly recovers. While the error is larger with non-unity power factor loads than a purely resistive one, Monjolo is off by a relatively constant factor. This still allows Monjolo to clearly identify the load's mode of operation over the course of the experiment. With some notion of the power factor, we expect Monjolo could more accurately estimate power in the future.

3.4.1.3 Rate-Limiting Transmissions

Timely and successful packet transmissions are essential to Monjolo operation. To enable more than one Monjolo to co-exist in a single collision domain, each node must not saturate the shared wireless channel. To verify that our RC timer successfully limits packet transmission rate, we measure the rate over a range of primary load power. The results are shown in Figure 3.8. As expected, low wattages cause Monjolo to transmit infrequently. As the load
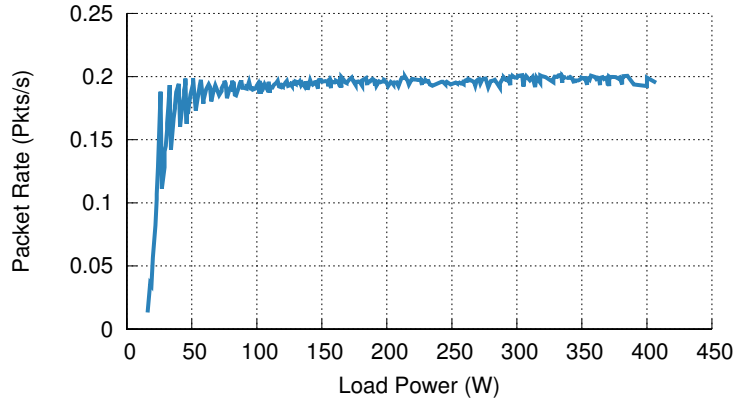
Figure 3.8: Monjolo packet transmission rate over a range of load power. The packet rate increases until it saturates at approximately one packet every five seconds. The oscillations occur due to the increasing wakeup frequency interacting with the static 5 s rate limiter. This limit on the transmission rate ensures that a single Monjolo sensor will not saturate the wireless channel. While we use a static rate limit, other incarnations could employ feedback from the data aggregator to dynamically adjust the maximum rate by adjusting the Monjolo threshold voltage.

power increases, the packet rate also increases until it is capped at one packet every five seconds. Oscillations around 50 W are caused by the increasing wakeup frequency interacting with a static 5 s timer. This shows that Monjolo can operate in dense environments, and the exact transmission rate can be adjusted in software by adjusting the voltage threshold.

3.4.2 Environmental Effects on System Accuracy

In prior sections, we evaluate Monjolo's maximum attainable accuracy due to the physical properties of the system and the interactive nature of real, reactive, and apparent power. In this section, we relax the ideal environmental conditions and evaluate how Monjolo performs under more realistic and varied situations.

3.4.2.1 Temperature Changes

Monjolo depends on a consistent relationship between activation rate and load power to accurately estimate one from the other. Capacitors tend to slightly change their leakage and capacitance values with temperature. If significant, these physical changes would affect Monjolo's activation rate, requiring a temperature-dependent calibration procedure to ensure Monjolo remains accurate across a range of field conditions.

Over a temperature range spanning 21C (room temperature) to almost 60C, there is a slightly downward trend in the power estimate, so Monjolo slightly underestimates power at

37
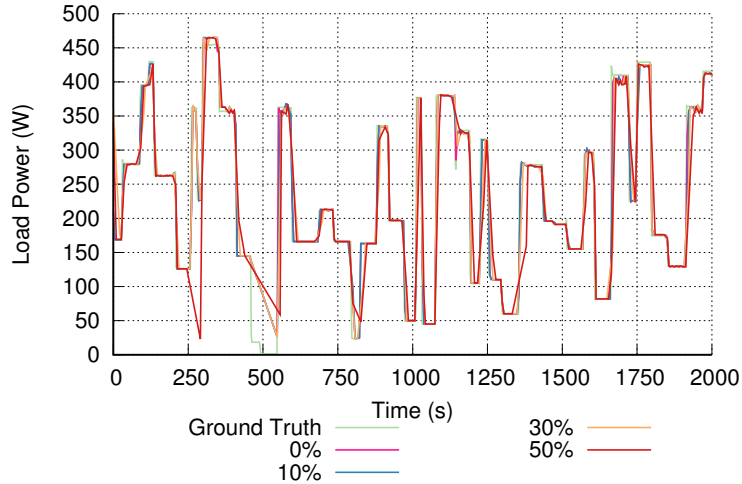
Figure 3.9: Effect of packet loss on Monjolo's accuracy. To test how Monjolo performs in a lossy and noisy RF environment, we simulate the effect of varying degrees of packet loss. By manually removing packets to simulate 10%, 30% and 50% packet loss, and then running the remaining data through our model, we see that Monjolo is still able to track the actual power draw of the load. Even at 50% packet loss, most transitions are measured with only a few high frequency power changes going undetected. While such high packet loss rates are not expected, this demonstrates that Monjolo is robust to communication failures.

high temperatures. This estimate, however, accounts for a maximum error of 2%. Given the low error, and the power, time, and cost overhead required to add and sample a temperature sensor, we decide that the temperature effects are not significant enough to warrant adding temperature compensation circuitry or software to Monjolo.

3.4.2.2 Packet Loss

Monjolo has the potential to operate in a noisy RF environment. Because Monjolo's fundamental ability to sense, and its metering resolution, are predicated upon the data aggregator successfully receiving packets in a timely manner, we evaluate the effect that packet loss has on Monjolo's measurement accuracy.

Figure 3.9 shows Monjolo's response to packet loss under conditions of 0%, 10%, 30%, and 50% loss rates, as well as the ground truth AC power. The graph shows that while the most accurate Monjolo report comes from the dataset with 0% packet loss, even 50% packet loss allows the system to track the load very well. The most significant errors occur when the load pulses rapidly and the dropped packet(s) occur during that pulse. During periods of constant load, packet loss has no effect on accuracy as it does not affect average reported power. When packets are dropped as the load changes, temporal granularity is lost but overall average accuracy is maintained.

Power Meter	Overhead (mW)	
	0 W Load	60 W Load
Belkin Conserve Insight*	440	560
Ensupra PM0001*	280	360
Kill A Watt PS-10*	180	690
Kill A Watt Wireless	670	720
SmartHome iMeter Solo	810	820
UPM EM100*	240	280
Watts Up? .net	1510	1590
Monjolo	0	≈ 4

Table 3.1: Power draw of several commercial power meters and Monjolo with no load and with a 60 W load. Meters marked with an asterisk have no ability to transmit their readings, whether wired or wirelessly. Due to the active circuitry and displays common in commercial plug-load meters, these devices consume nonzero power with no load. Monjolo, however, remains completely off when no load is attached. At 60 W the power draw of Monjolo is less than 10 mW (granularity of our measurement system) but draws approximately 4 mW based on the theoretical model.

3.4.3 Energy Cost of Metering

While no system can monitor power draw with zero overhead, Monjolo adds a much smaller power overhead than today's commercial power meters. In particular, when a load is completely off Monjolo draws no power due to its load current based energy harvesting technique. In contrast, commercial meters draw between 180 mW and 1.51 W with no load, as shown in Table 3.1.

To estimate the power overhead of Monjolo with a load attached, we model Monjolo's overhead at a given load power as:

$$P_{overhead} = f_{wakeup} \cdot P_{wakeup} \cdot t_{wakeup} + P_{loss}$$

P_{wakeup} is constant at approximately 66 mW (20 mA at 3.3 V). t_{wakeup}, the time the output voltage regulator is active and the microcontroller is powered on, is also constant (as required by the system) at 58 ms. This is itself a function of P_{wakeup} and the capacitance of C_{store}. P_{loss} is the efficiency power losses of the system, including the overhead of the coil, rectifier, and regulators.

The power overhead of Monjolo scales with primary load power as the activation rate increases. This makes Monjolo's operation power-proportional – certainly a desirable trait.
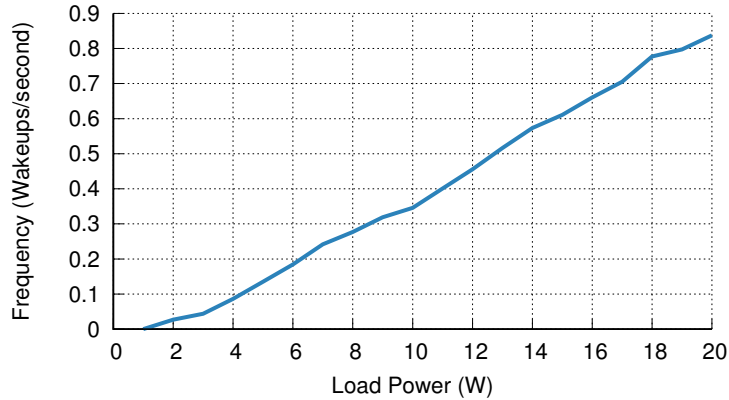
Figure 3.10: Monjolo activations over a range of low power loads with a 1:50 current transformer turns ratio. By changing the current transformer front-end, Monjolo can operate with primary loads as low as 1.15 W. This demonstrates that the minimum load for operation is not a fixed limitation but rather a tunable parameter of the Monjolo design.

3.4.4 Parameter Tuning

Many factors affect Monjolo's accuracy, measurement range, and overhead. The current transformer's turns ratio, rectification scheme, and capacitance values can all be tuned to optimize for different desired characteristics.

An obvious limitation of the Monjolo design is the minimum load requirement. Many household loads are below the 17 W minimum described in Section 3.2.1. This minimum is due to the LTC3588 minimum voltage threshold. In order for the input boost converter to operate, the rectified wave must meet a minimum voltage. The minimum primary load is therefore fully configurable in turns ratio. By selecting a higher turns ratio, we can cause low wattages to generate higher voltages than at the 1:500 ratio that we use to evaluate Monjolo.

Figure 3.10 shows the range of wattages from 1.15 W to 20 W generating wakeups by means of a 1:50 turns ratio. This 1.15 W minimum could be decreased further but already covers most, if not all, household devices. There are two reasons why a user might desire a higher turns ratio like the one on which Monjolo is evaluated. First, a lower turns ratio will generate fewer wakeups and therefore cause a lower power overhead. Second, Figure 3.5 shows the quantization error increasing with increasing wattage. Shifting the transfer function to lower wattages will be at the expense of relatively higher quantization error at high wattages. This inaccuracy will be counteracted in some regards by the fact that a higher turns ratio will generate more wakeups and therefore provide higher granularity.

Turns ratio is not the only way to adjust the nominal activation rate or power overhead. Switching from half-wave to full-wave rectification provides higher granularity with higher power overhead, but does not affect runtime. Increasing the storage capacitance decreases

40

(a) Plug-load version of Monjolo

(b) Split core version of Monjolo
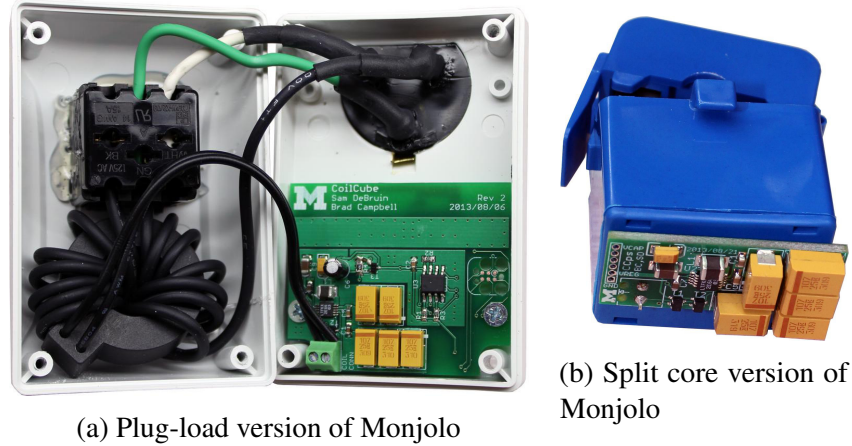
Figure 3.11: Two versions of the Monjolo sensor. In (a), a plug-load version with the case open, so the female socket is visible on the left and the male plug behind the case on the right. The power supply is visible on the front of the PCB, and the digital core is on the back. In (b), a split core version suitable for panel metering and other non-contact applications. Because Monjolo does not require additional power or wiring and can clip on without breaking the path or rewiring the circuit, it dramatically changes the landscape of panel-level energy metering.

wakeup frequency but increases runtime without affecting power overhead. Tuning of these parameters to the application allows the Monjolo design to be versatile in diverse situations.

Monjolo's sample rate as observed by the data aggregator is tunable by adjusting R_{timer} and C_{timer}. We set this rate at 0.2 Hz, but it could be adjusted upward for higher fidelity data collection. This parameter is also software tunable as each sensor can tune the V_{timer} level at which it transmits a packet. For comparison, many of the commercial meters we evaluate sample at 1 Hz, although some sample slower and the Kill-A-Watt Wireless samples once every two minutes.

All of these parameters are user-selectable. Whether optimizing for accuracy, measurable load range, or power overhead, Monjolo is configurable for target load power.

3.4.5 Monjolo Prototypes

Figure 3.11 shows two working versions of the Monjolo prototype, a plug-load variant for measuring outlet loads and a split core variant for measuring in the AC panel. As a clip-on panel meter, the unit is inexpensive to deploy, and deployment can be done safely without disconnecting any wires. The relatively high cost and size of the current transformers and storage capacitors has less overall impact than in plug-load metering, which requires many more sensors than in the panel.

This positioning as a panel meter offers two opportunities for further development. First, even without calibrating the sensor Monjolo can operate as a fault detector or device health monitor. Many industrial devices are directly wired without a plug, so a clip-on sensor is the only possibility that does not require rewiring. In such cases the machinery will typically draw power in a consistent or pattern-based way, which will manifest as patterns in the wakeup rate of Monjolo. Even without having calibrated the Monjolo, changes from this pattern could indicate to a listening receiver that the machinery requires servicing, which could potentially avert larger issues. This is an open area of investigation, and one that we explore more in later sections.

In addition, as a panel meter Monjolo could perform the function of metering (true power) as well as monitoring (relative changes) by adding a measurement of voltage. Although studies have shown that voltage measurement can be virtualized, i.e. measured off-site and transmitted wirelessly to the node, the limited wireless capability of Monjolo would make such synchronization difficult, although not impossible [50]. Instead Monjolo could measure a capacitively-coupled induced voltage, thereby both measuring voltage and maintaining zero AC contact [101]. Although this offers potential to improve Monjolo, that subsequent evaluation is not the focus of this work.

3.5 Summary

Although the panel variant of Monjolo offers potential for future research, as a plug-load meter the design is too expensive to scale to the density required for outlet devices. However, a whole-home solution could incorporate the panel variant of Monjolo alongside a miniature device for plug-load measurement. In the next chapter we leverage the learnings from both Monjolo and our earlier PowerCube work to scale the plug-load meter to an unobtrusive form factor and inexpensive design point. This has the potential to provide much greater visibility for plug-loads than previously possible.

CHAPTER 4

Non-Intercepting Sensing Enables a Nearly-Invisible AC Energy Meter

To scale the wireless AC energy meter past the design points of ACme, PowerCube, and Monjolo to a size and price point that is viable to deploy at scale, we must revisit all aspects of the design. In this chapter we present PowerBlade, a 1" × 1" × 1/16" plug-load energy meter that measures real, reactive, and apparent power and transmits this information over a standard Bluetooth Low Energy (BLE) link. Devices to be measured literally plug through PowerBlade before plugging into the wall. Despite its small size, PowerBlade measures real, reactive and apparent power, and unlike Monjolo it measures voltage directly. In large quantities of 10,000 or more it is available for $11 or less. PowerBlade is the smallest, lowest cost wireless true power AC meter, and it enables deployments of individual submetering on AC devices.

Like Monjolo, PowerBlade measures current unobtrusively by detecting the magnetic field generated by the current passing through a wire. However, while Monjolo utilizes a large, expensive current transformer, PowerBlade measures the signal induced in a standard horizontally wirewound inductor. This small, inexpensive component can be placed near the plug as a planar current sensor. Although the transfer function changes slightly, this is effectively the same as a split core CT where most of the core encircling the plug is air and only a small amount of the core is the ferrous material contained in the inductor. This results in less output, no longer enough to harvest power, but at a size point viable for PowerBlade.

The planar form factor of PowerBlade requires revisiting all aspects of the plug load power meter beyond just the current measurement. Although current measurement, and the high voltage routing commonly associated, is often a driving factor of size, the AC-DC power supply also commonly requires large volume capacitors, inductors, or transformers. Where and how to calculate power as well as the communication topology are two components that require little volume, but higher current draws will typically require a larger volume power supply. Finally voltage measurement can be a simple voltage divider, but how to
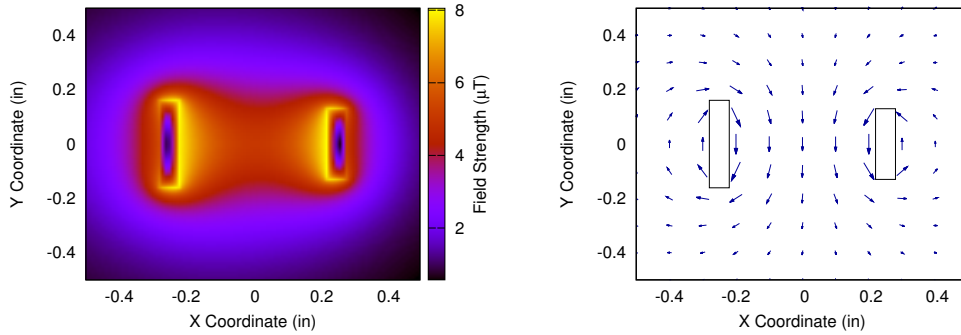
Figure 4.1: Relative strength and direction of the magnetic field on the surface of PowerBlade. The current passing in either direction through each plug will generate magnetic fields of opposite polarity. These fields will add constructively between the plugs and destructively outside. This figure can be generated with the Biot-Savart saw, and is modeled with a current of 1 A. This model allows us to select a measurement technique and position it in this field.

make physical contact with the AC, which is also required to power the sensor, can have an impact on size. In the following sections we describe the design and implementation of the PowerBlade sensor.

4.1   System Design

PowerBlade is designed to be the smallest, lowest power wireless true power meter. The size limitations of plugging through the device impact all aspects of the meter and influence each implementation decision. The specific bounding box of volume is imposed by the placements of outlets on walls and surge strips and by the pairing of the length of the plug and the depth of the receptacle that will decide the maximum thickness of the system. NEMA outlet spacing is nominally 1.1" [18], and although there is variability in surge strips they typically have similar spacing. Like in the original PowerCube design, therefore, PowerBlade must be no more than 1" on a side. The NEMA specification does not list tolerances that might indicate how thick PowerBlade could be, but experimentation indicates that about 3 mm is the maximum thickness to maintain a good connection.

4.1.1   Planar Current Measurement

The PowerBlade form factor requires a planar current measurement method, and one that does not rely on breaking the AC path. This means a non-contact sensor like a current transformer, but also one that, like a current sense resistor, exists almost exclusively in the

plane of the PCB. For this application we leverage the same property as current transformers, which measure via the magnetic field generated by the AC current flowing through the wire.

On the surface of PowerBlade, the fields generated from either plug will add constructively between and destructively outside. The magnitude and orientation of the resulting combined field can be modeled using the Biot-Savart law and is shown in Figure 4.1. The model is generated using a current of 1 A, effectively a 120 W load for this application. This can be used to optimally select a sensor for, and position it in, this field.

In addition to guiding the optimal orientation of sensors, this model provides two significant data points. First, the overall magnitude of the field is on the order of 10-100 nT. This establishes a lower bound on the sensitivity of the transducer. Second, although intuition might indicate that the maximum field would exist in the center where the fields add most constructively, rapid falloff of signal strength means the optimal location is as close to either plug as possible.

Existing Hall effect sensors or other vector magnetometers fail to meet the requirements of this sensing application. Although many have sufficient sensitivity to measure the field, none meet the requirements from a volume or power perspective. Instead, we note that a horizontally wirewound inductor - where the winding of the coil is along the major axis of the device and not vertical - can act as a magnetic sensor. Located close to the plug and oriented such that the field lines from Figure 4.1 are passing through the long axis, the inductor acts as a type of magnetometer called a search coil [116].

$$V = -\mu N A \frac{dH}{dt} \tag{4.1}$$

The signal in the coil is governed by Faraday's law of induction, and can be described using Equation 4.1. μ is the magnetic permeability; N and A are the number of turns and the cross sectional area of the coil, respectively. Although the unit will require calibration to determine the scaling value, Equation 4.1 does illustrate an important relationship. Namely, the signal is proportional to the time derivative of the magnetic field, and will require integration before it can be used to model the underlying current.

In addition, because this inductor is exposed in the system, there is the possibility for interference from other sources than the magnetic field generated by the current in the wire. In particular on the surface of an AC outlet there is the possibility for interference with other 60 Hz signals. This may require shielding, or in some cases it may actually require spacing from the outlet to reduce the effect. The latter would decrease the effectiveness of the planar measurement mechanism, and will be explored in later sections.
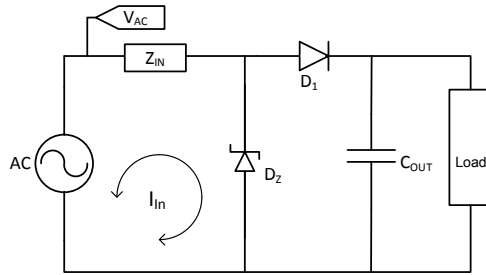
Figure 4.2: Zener regulated, half wave rectified power supply. Z_{IN}, whether a resistor or capacitor, limits the current supplied from the AC. The Zener regulator then provides a constant voltage for the load, and the capacitor C_{OUT} is charged on every positive half cycle of the AC wave. This supply requires only four components, and only Z_{IN} must be rated for high voltage.

4.1.2   AC/DC Power Supply Scaling

AC-DC power supplies often rely on large components like transformers and high voltage capacitors that do not scale to this volume. As an example, consider the case of a capacitor that is commonly used to limit input current to a rectifier [59]. The capacitance is inversely proportional to the impedance, so the current supplied will scale directly with capacitor value. For a line voltage of 120 V, an AC-DC supply operating on this principle would require approximately $\frac{C}{V} = \frac{1}{2*\pi*120V*60Hz} = 22$ nF per output milliamp. The Supertex SR-10, which operates by this principle, requires a minimum of 220 nF for a 5.7 mA output [22]. Further, safety concerns mandate a capacitor with a rating of either X1 or X2. A 220 nF capacitor with a X rating is only available in film packaging, which introduces significant size overhead. This really makes this design inappropriate for the application. However, if the system is designed around a significantly lower power budget, the capacitor can be sized lower and significantly smaller ceramic capacitors become viable.

Figure 4.2 shows the diagram of a passive series shunt supply, where a series impedance Z_{IN} limits current and a Zener diode regulates voltage for the load. This design does not require an IC, and only has one high voltage component; the three low voltage components can all be 0402 package or a similar small surface mount size. The design has drawbacks, namely static power dissipated by the Zener (as well as by the series element, if a resistor) and it does not provide isolation from the mains. However, if Z_{IN} can be effectively reduced the size of the supply is small and the thickness is single millimeters.

To determine the design space of Z_{IN} we examine resistors and capacitors across six families and chemistries. For each component we record the listed volume, and calculate the

46
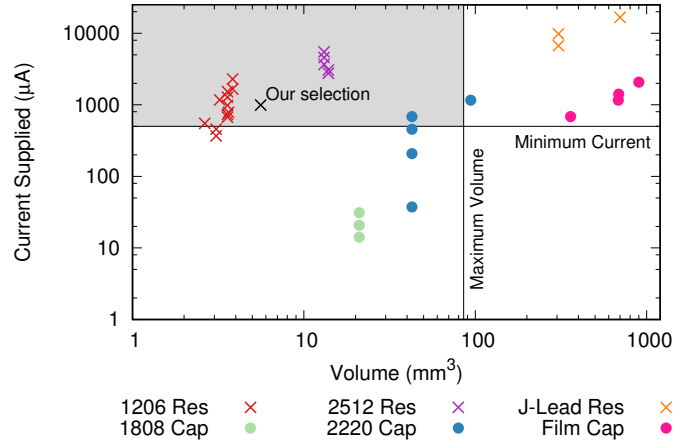
Figure 4.3: Design space for Z_{IN}. If Z_{IN} is a capacitor the current increases with capacitor value, and for high voltage capacitors volume increases with capacitance as well. If Z_{IN} is a resistor, a lower resistance will deliver more current, but requires a larger package to dissipate the additional power. There is a maximum volume that the system can occupy, as well as a minimum current required to operate. The upper left quadrant is the design space for Z_{IN}, and contains Pareto optimal points for current and volume. There is an additional Pareto optimal point for idle real power, which is zero for capacitors and non-zero for resistors. The selection for PowerBlade is also marked, selected back from the frontier to allow tolerance in the design.

current that it can deliver to the load. Current is calculated according to Equation 4.2 [59]. This is the design space for Z_{IN}.

$$I_{MAX} = \frac{V_{IN} - V_Z}{2 * Z_{IN}} \tag{4.2}$$

For a system with a certain minimum current required to operate and maximum allowable volume, Figure 4.3 represents this design space. These limits are determined iteratively with the rest of the design, and the approximate limits for the final PowerBlade system are shown. Any component in the upper left quadrant is viable for the design, but there are Pareto optimal points for both volume and current. A third dimension, idle real power, would be zero for capacitors (only reactive) and nonzero for resistors, and this introduces a third Pareto optimal point. The value we select for PowerBlade is also shown, a 1210 resistor selected back from the frontier to allow tolerance in the design.

4.1.3   Socket-Free AC Voltage Acquisition

In order to power itself from the AC, PowerBlade contacts the AC plugs passed through it. Unlike Monjolo, PowerBlade also measures voltage directly, which allows it to measure true

47
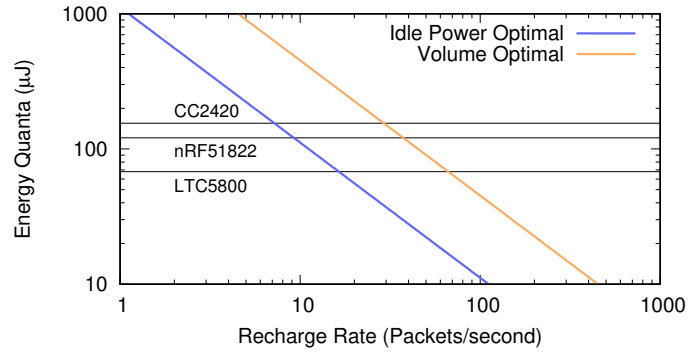
Figure 4.4: Wireless design space in PowerBlade. For two of the optimal power supply options for Z_{IN}, this figure demonstrates the relationship between energy quanta and recharge rate as a function of capacitance. Increasing C_{OUT} results in a higher energy quanta at a slower recharge rate, i.e. moving up and to the left on a given line. For a particular option for Z_{IN} and a particular required data rate, this helps motivate the selection of a certain radio topology.

power in real time. Although voltage measurement on a microcontroller can be done with a voltage divider, this, too, requires contact with the AC. The difficulty in making contact in this application arises from the form factor: a typical AC socket is many factors too large. PowerBlade requires contact either via the PCB itself or with a method recessed into the plane of the PCB.

4.1.4 Extracting AC Waveform Parameters

As noted in Chapter 2, power metering is a case where local calculation can offer significant savings over transmitting raw samples. PowerBlade does not have the size or power budget to accommodate an energy metering IC like the ADE7753 used in PowerCube and other designs, so the only remaining option is to implement a custom measurement algorithm on the on-board microcontroller.

4.1.5 Low-Energy Wireless Communications

Due to the significant volume cost of power in this application, and to the fact that a significant portion of the overall power will be directed to the radio, careful consideration must be made in selecting the wireless topology and implementation. This is typically the only component whose power draw is higher than the power supply can provide, and the system must buffer sufficient energy to power the radio and send a transmission. In most radios there is a fixed minimum action, typically initializing and sending a packet, that must be performed as a continuous action. This energy which must be buffered by the system

is referred to here as *energy quanta*, and the amount of energy available is a factor of the selections in the power supply.

For the supply in Figure 4.2, Z_{IN}, D_Z, and C_{OUT} determine the maximum energy quanta, as well as the recharge rate of the system. The energy required is stored on C_{OUT}; increasing this value results in more energy available but at the expense of a slower recharge rate. Decreasing Z_{IN} results in more current to the system overall, which both decreases recharge time and increases energy quanta as some additional current can be supplied during discharge. However this also results in a higher idle power, and typically a larger component. For D_Z, increasing the Zener voltage V_Z will increase both energy quanta and recharge time, but will increase energy quanta more due to the quadratic factor in energy calculation. Higher values for V_Z are therefore more optimal, but there is an effective limit imposed by capacitor and downstream regulator voltage limits.

These three components make up the wireless design space for PowerBlade. Figure 4.4 shows this tradeoff for the volume optimal and idle power optimal selections for Z_{IN}. This represents the capacitor and resistor implementations of a low-power system, the current optimal option introduces significant idle power. V_Z is fixed at 10 V, which is roughly the maximum voltage for which the smallest output capacitors and regulators are viable. On the curve for a given Z_{IN}, increasing capacitance results in moving up and to the left.

Also shown in Figure 4.4 is the minimum energy required to wake up and send a single packet for three low power radios, the CC2420 [23], LTC5800 [14], and nRF51822 [20]. The measurement of the CC2420 was performed by Yerva et al. [119], the LTC5800 value is available from its datasheet, and we measure the energy for the nRF51822. The values for energy quanta can be approximate, their purpose is to drive the selection of a certain radio. This calculation allows us to select the radio based on maximum recharge rate, which allows us to calculate a maximum date rate as well as a minimum required power supply capacitance.

4.2 Implementation Details

Figure 4.5 shows four generations of the PowerBlade sensor, each representing several iterations of a certain concept. Version 1 is the only design to exceed 1" on a side, the design is 1.1" square. Version 1 has footprints for both a resistor and capacitor for Z_{IN}, allowing flexibility in the design. Version 2 optimizes for space by selecting only a resistor, and both versions 1 and 2 make contact with the AC plugs with flexible tabs built into the PCB itself. In version 3 we remove the tabs, and make contact instead with spring loaded pins mounted on the PCB. In version 4 we add to this idea by plating the interior of the AC prong cutout.
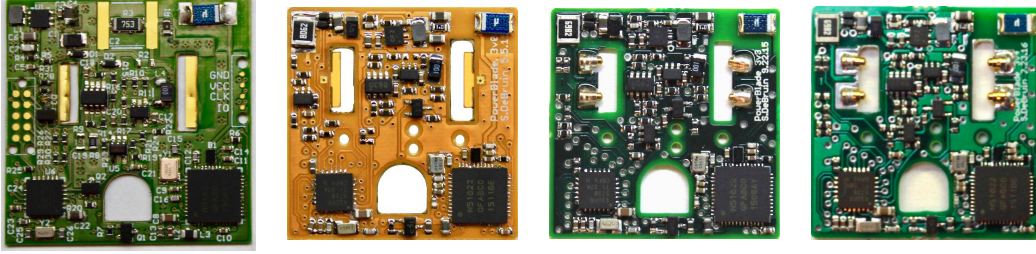
Figure 4.5: Four generations of the PowerBlade sensor. Versions 1 and 2 make contact with the AC plugs via flexible tabs built into the PCB. Version 1 has footprints for both a resistor and capacitor, and it is the only version larger than 1" on a side (1.1"). Version 2 optimizes for space by including only a footprint for a resistor. Versions 3 and 4 make contact with spring loaded pins mounted in the plane of the PCB, and in addition Version 4 has plating all around the AC cutout to increase the potential contact surface area.

Z_{IN} (Z)	Expected	Actual	Volume	Cost	Idle Power
22 nF (120 kΩ)	456 µA	451 µA	43.0 mm^3	\$1.24	0 mW
33 nF (80 kΩ)	705 µA	740 µA	43.0 mm^3	\$1.63	0 mW
56 nF (47 kΩ)	1161 µA	1137 µA	95.0 mm^3	\$1.34	0 mW
100 kΩ	550 µA	538 µA	2.6 mm^3	\$0.03	140 mW
80.6 kΩ	687 µA	658 µA	3.0 mm^3	\$0.03	170 mW
75 kΩ	733 µA	716 µA	3.0 mm^3	\$0.03	190 mW

Table 4.1: Possibilities for Z_{IN} evaluated for use in PowerBlade. We select the most promising options from the power supply design space, and implement a power supply on a test board to evaluate the maximum power, size, and idle power introduced. We find that for a similar impedance the capacitor and resistor options perform similarly. The resistors introduce idle real power, but their significantly lower price point outweighs the potential energy savings. For our design we select a 80.6 kΩ resistor.

Only versions 3 and 4 have been deployed at scale, in the following sections we detail the design decisions that result in a deployable system.

4.2.1 Selecting the Series Impedance Z_{IN}

For the supply circuit shown in Figure 4.2, the choices for Z_{IN} and D_Z guide the amount of power available to the load and the idle power in the device, as well as the size and price point of the system. To select the value for Z_{IN} we implement and evaluate several components for their power output, cost, volume occupied, and idle power. What we find is that for a similar impedance, the capacitor and resistor options output similar amounts of current. The capacitors do so without adding a power draw, but at a volume and price point significantly higher than that of the resistors.

PowerBlade Version 1 has space for both a 2220 package capacitor and a 1210 package resistor, and in this design we utilize either a 33 nF capacitor or a 80.6 kΩ resistor. In subsequent designs we optimize for space by only providing a footprint for a 80.6 kΩ resistor. In addition to the significantly smaller size, the low price point of the resistor outweighs the idle power it introduces. At $0.12 per kilowatt-hour, for example, it would take almost nine years of continuous operation for the 170 mW to overcome the additional $1.60 in price for a 33 nF capacitor.

The remainder of the components from Figure 4.2 are a 10 V Zener and 250 mA rectifier diode, both small SC-79 packages, and a combined 55 µF for C_{OUT}. The load will nominally experience 8.9 V: the Zener voltage minus the rectifier forward voltage of 1.1 V. To regulate the output to 3.3 V we use a buck regulator, the TPS62122 from Texas Instruments [26]. For an input of 8.9 V this regulator has an efficiency of 85-90%.

4.2.2 Flexible Tabs vs. Spring Loaded Pins

The four PowerBlade designs shown in Figure 4.5 use two different options for contacting the AC plug. Versions 1 and 2 make contact via flexible tabs built into the PCB and that extend into the opening to touch the plug. This requires a specialized manufacturing process, and has other drawbacks regarding lifetime, but requires no additional components to implement. Versions 3 and 4 make contact with spring loaded pins mounted in the plane of the PCB. This lasts longer, but adds overall cost and requires more space.

4.2.2.1 Flexible Tabs

Early versions of PowerBlade make contact with the AC via flexible tabs. The manufacturing process of the PCB itself produces the contact method, there are no additional parts or assembly required. However, the manufacturing process for flexible PCBs is not designed as a mechanical interface, and over time the tabs do not provide a strong contact to the AC.

The basic concept is shown in Figure 4.6, which, in Figure 4.6a, shows a cross section of the Version 1 PCB with the flexible tab extending out. The PCB is four copper layers, the inner two layers on a flexible polyamide core and the outer layers on rigid FR4. Through most of the area of the board, the stackup is rigidized by the FR4 and functions as a standard four layer PCB. The inner layers are used for Power and Ground planes as in a typical system. However, in the area of the tab, the inner flexible polyamide and the two inner copper layers extend unsupported. This results in the flexible tab, conductive on either side. This can be electrically isolated from the Power and Ground planes inside the system, and routed to the power supply and voltage sensing circuits.

(a) Version 1 layer stackup.   (b) Version 2 layer stackup.
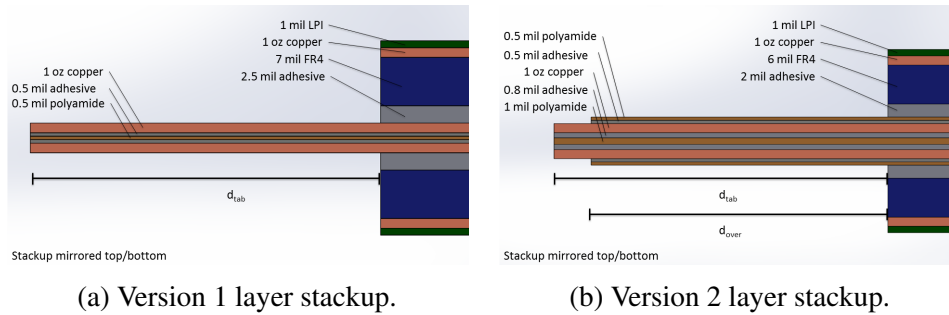
Figure 4.6: Flexible tab layer stackups from PowerBlade Versions 1 and 2. In Version 1 the tab is just a small flexible core with conductive copper covering. The FR4 rigidizes the PCB except in the tab, and the tab bends to make contact as the plug passes through. The Version 1 tab deforms after a few uses, so in Version 2 we reinforce the tab by thickening and adding an additional polyamide overlay. We adjust the other layers to maintain the same overall PCB thickness, and we evaluate multiple options for the length of the tab, d_{tab}, and the length of the overlay, d_{over}.

Although the Version 1 stackup shown in Figure 4.6a makes contact with the AC plug, the tab has very little elasticity and does not return to original shape when the plug is removed. In Version 2 of PowerBlade we improve on this design by making two modifications to the tab construction. First is to increase the tab thickness by 3 mil overall and to strengthen it with additional overlays, which both increase elasticity and reduce the break risk. Second is to significantly reduce the length of the tab, decreasing the total deflection while the plug is connected and increasing the lifetime of the tab.

The augmented layer stackup for Version 2 is shown in Figure 4.6b. In this design the inner polyamide layer is doubled from 0.5 mil to 1 mil. In addition, on either side of the tab we add a 0.5 mil polyamide overlay with 0.5 mil adhesive that extends through most of the length of the tab leaving the conductive tip exposed. This overlay, combined with the increased core thickness, results in a stronger and more elastic tab. By reducing the thickness of other layers we keep the overall PCB thickness unchanged, which maintains the same available component volume.

In addition to modifying the PCB construction, in Version 2 we significantly reduce the tab length to reduce bend. The total cutout width for each prong is about 0.1", and the prong itself occupies nominally 0.06" [18]. The cutout is not centered on the plug, the free space is distributed about 0.01" away from the tab and 0.03" toward the tab but exact plug dimensions vary widely. A tab length of 0.03", therefore, will only just make contact with the nominal plug width.

To determine the appropriate lengths of the tab and the overlay, we implement four combinations and evaluate the performance. The four options are shown in Table 4.2. Tab

Option #	d_{tab}	$d_{over,top}$	$d_{over,bot}$
1	0.035"	0.012"	0.012"
2	0.035"	0.030"	0.030"
3	0.045"	0.040"	0.040"
4	0.045"	0.045"	0.012"

Table 4.2: Tab length and overlay length options evaluated for use in PowerBlade. To determine the configuration that results in maximum elasticity and lifetime we implement four options of both tab length, d_{tab}, and overlay length, d_{over}. We control top overlay and bottom overlay separately. We find that tab option 4, which requires plugs to pass from bottom to top, provides the longest lifetime, but potentially still not sufficient to deploy at scale.

option 1 is 0.035" long, with minimal bend as only 0.005" of the tab overlaps the plug. This option has only 0.012" of overlay on either side, just enough to strengthen the vertex. Tab option 2 is the same overall length, but with a longer overlay almost to the edge to determine the potential to improve elasticity. Tab option 3 is longer with more bend, but also a long overlay. Finally tab option 4 is a special case - a longer tab with the top side fully covered and therefore non-conductive. The plug must therefore be inserted from the back of the unit, but this option evaluates the maximum potential support provided by a single overlay.

By implementing these four designs we arrive at several conclusions. First, options 1 and 2 with 0.035" overall length make contact with some plugs but not others, the tight tolerance does not allow for the large diversity in plug widths. Second, any tab area unsupported by overlay bends significantly more than the areas that are supported. This means that options 1, 2, and 3 all have second vertex areas at the juncture of the overlay and the exposed conductive tip. Essentially the area with the overlay returns to form, but the tip is permanently deflected. Finally, option 4 has the drawback of single-sided operation, but with the fully covering overlay it is able to maintain a strong connection for 20 or more plug-unplug cycles. Since PowerBlade is designed to be left on the device semi-permanently, this may be sufficient for deployability. However, in the next section we evaluate another method for contact that requires more cost and labor to manufacture, but results in a more long term connection.

4.2.2.2 Spring Loaded Pins

An alternative to the flexible tabs is spring loaded pins mounted in the plane of the PCB [16]. This has several drawbacks, the pins themselves add significant size overhead and cutouts must be provided in the PCB for them to attach. In addition the price of the pins is higher than the increased cost associated with the flexible PCB, and the manual assembly process
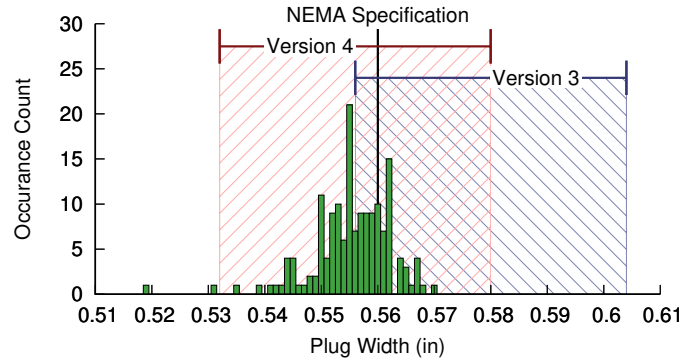
Figure 4.7: Plug widths for 162 NEMA plugs, both 1-15(P) two pronged devices and 5-15 three pronged devices. PowerBlade sits at the base of the plug, and makes contact with the outside edges, so this distance is measured from the outside edge of each plug. The NEMA specification indicates that this should be 0.056", and PowerBlade Version 3 is designed to this value. However we find that a significant number of devices are narrower, and Version 3 only makes contact with 73 of the 162. By moving the pins 0.024" closer together we are able to make contact with 160 out of 162.

adds cost beyond that. However the spring loaded pins do afford the benefit of repeatability, they are designed to be used for a similar purpose and provide a significantly longer lifetime than tabs.

PowerBlade Versions 3 and 4 both utilize this method for contacting the AC. In Version 3 the plug slot is augmented with trapezoidal cutouts, which are plated on the interior to allow the pin to connect. The pin can then be attached by hand as a thru-hole part after the remainder of the board has been reflowed. In Version 3 we mount the pin according to the NEMA specification, but with some overlap to accommodate variation in plugs. Since the pin is rounded only at the tip, excessive overlap could prevent the plug from being able to compress the pin. This design is viable for deployment, and many were deployed for evaluation.

In deploying PowerBlade Version 3, however, we discover that significant variation in plug widths, combined with the relatively short compression length of the pins, means that many plugs do not make contact with PowerBlade. Figure 4.7 shows a histogram of the plug widths of 162 devices from our lab and various homes and apartments. Also shown on the figure is the nominal NEMA width, the span of the pins on Version 3, and the modified location used in the final Version 4. Of the 162 devices, only 73 make contact with the pins on Version 3. By moving the pins 0.024" closer together for Version 4, we cover 160.

To cover the final two devices we introduce an additional change in Version 4: plating the entire AC cutout. In Version 3 only the area immediately around the spring loaded pin is plated, with metal covering the inside wall of the cutout. In Version 4, in addition to
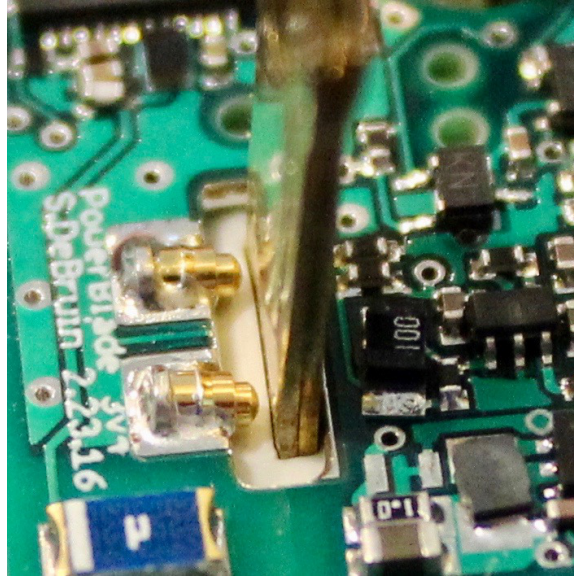
Figure 4.8: A narrow plug does not make contact with the spring loaded pins in PowerBlade. Although moving the pins 0.024" closer together results in PowerBlade making contact with 160 of the 162 device we evaluate, 2 devices are too narrow to make contact with the pin. In this case the other side of the cutout is also plated, visible in the picture. For those plugs, the inner side of the plug contacts the board directly.

moving the pins inward we also plate the rest of that AC cutout so that if the plug makes any contact with PowerBlade it will power the system. This only applies for plugs whose pins are much smaller than the NEMA specification, which are also the few plugs too small to make contact with the spring loaded pins. Figure 4.8 illustrates this effect. With these changes, PowerBlade Version 4 makes contact with all 162 of 162 devices evaluated. While plug variation, along with this relatively small sample size, will likely mean some plugs do not make contact, this represents a positive result for the general case.

4.2.2.3 Leaf Springs and Battery Contacts

A third contact method that we do not evaluate is leaf springs – a mechanical mechanism found in various applications from vehicle suspensions to battery contacts. Similar to the concept of flexible tabs, the concept of leaf springs is to deflect a small material to apply pressure to the AC prong [92]. The tab to be deflected can cover the entire prong opening, eliminating the problems of the short throw on the spring loaded pins not making contact with every plug. Further, since the leaf spring, similar to the spring loaded pin, is designed for a mechanical application it will not degrade as quickly as the flexible tabs. Leaf springs are typically stainless steel plated with a conductive metal [92], and are significantly more elastic than the polyamide-copper construction of the flexible tabs.

(a) Current Measurement

(b) Voltage Measurement
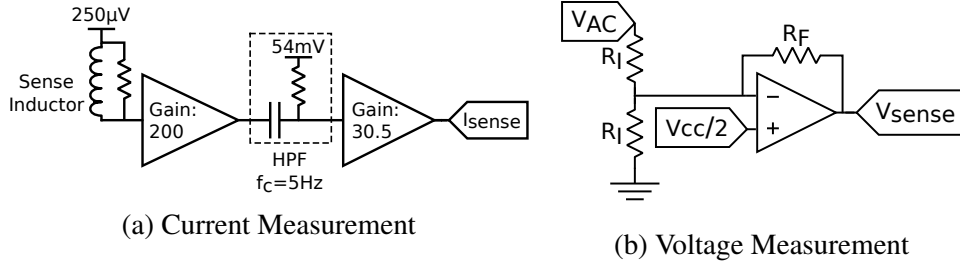
Figure 4.9: Current and voltage measurement circuits for PowerBlade. Current is measured with the sense inductor referenced to 250 μV, and the signal is first amplified by the OPA333 with a gain of 200. Low frequency noise is removed with a high pass filter, followed by a second stage amplifier with a gain of about 30. Voltage is measured through a voltage divider with a $V_{CC}/2$ offset, so both the voltage and current signals are referenced to $V_{CC}/2$.

There is another potential benefit of leaf springs over spring loaded pins. In some rare situations, the spring loaded pins can be caught in the small hole on some AC prongs. In most situations the two spring loaded pins on either side of PowerBlade span the hole. However, if the AC prongs are narrow then the PowerBlade unit is able to torque slightly and the pin can sometimes slip in. A leaf spring would not have this problem.

We do not directly evaluate leaf springs for this application due to the lack of a suitable part for purchase, but future effort could work with the manufacturer to produce a pin suitable for the PowerBlade design. This would be longer lasting than the flexible tabs as well as being less expensive and more versatile than spring loaded pins.

4.2.3 Voltage and Current Measurement

PowerBlade measures current with the inductor method described in Section 4.1.1, and makes contact to measure voltage via the contact method from Section 4.1.3. To amplify the signals before they are connected to the ADCs on the microcontroller, we use the circuits shown in Figure 4.9. The voltage measurement circuit is designed to operate on either 120 V_{RMS} (United States) or 220 V_{RMS} (much of Europe and China). It is a voltage divider with a $V_{CC}/2$ offset and with the resistor values $RF = 4.99$ kΩ and $RI1 = RI2 = 953$ kΩ. The equation for V_{SENSE} can be described by Equation 4.3.

$$V_{SENSE} \approx \frac{V_{CC}}{2} - 5.24 \times 10^{-3} VAC \tag{4.3}$$

With this configuration the peak to peak voltage will be approximately 1.79 V in the United States and 3.28 V in much of Europe and China. The rail voltage in PowerBlade is 3.3 V, so the same circuit operates in either condition.

The current measurement inductor is referenced to 250 µV and amplified in two stages before digitization in the ADCs. Between the first and second stages low frequency noise is removed with a high pass filter with a voltage offset so the final signal is centered around $V_{CC}/2$. The first stage amplifier is specialized for this application due to the low voltage output from the inductor, and we evaluate multiple amplifiers and gain options to determine the best design.

Device	I_{BIAS}	V_{OFFSET}	GBP	I_{SUPPLY}
MAX9910	1 pA	200 µV	200 kHz	4 µA
MCP6V3	5 pA	8 µV	300 kHz	23 µA
OPA333	70 pA	2 µV	350 kHz	17 µA
MAX4238	1 pA	0.1 µV	1000 kHz	600 µA

Table 4.3: Amplifiers evaluated for the PowerBlade design. To determine the correct amplifier to use in the first stage current measurement we implement four different options and compare the results. We find that the MAX9910 and MCP6V3 are insufficient due to their high input offset voltage. Of the two that remain, the OPA333 and MAX4238, performance is similar indicating the system has encountered another threshold elsewhere. For its lower supply power we select the OPA333 for this design.

Table 4.3 shows the four amplifiers tested for PowerBlade, as well as key properties of those amplifiers. The range of amplifiers selected have either a low current offset, low voltage offset, or both, but with a higher overall power draw. For this purpose we utilize a custom test rig, shown in Figure 4.10. This PCB has the inductor front end as well as footprints for each amplifier to test. This board does not contact the AC, we use a separate power supply to allow us to test amplifiers without power and safety limitations. We implement a different copy of the board for each amplifier, connect the signal into an MSP430 as in PowerBlade, and compare the results.

We find that three of the four amplifiers tested are capable of measuring the inductor. The last, the MAX9910, has a voltage offset of 200 µV that is not met by the signal. Under most AC loads there is no output. Of the others, the MCP6V3 has a voltage offset of 8 µV and is unable to measure low power AC loads below 200 W. The performance of the OPA333 and MAX4238 are similar, and are the best among the amplifiers we evaluate. This indicates that the issue is not one of current offset, as the relatively high current offset in the OPA333 seems to have little effect. Further, the wattage measurement range of the MAX4238 is not significantly lower than that of the OPA333 despite a lower voltage offset. Although the low voltage offset of the OPA333 is critical, further decreasing by a factor of 10 does not provide additional benefit. This indicates that the system has encountered an additional threshold and the amplifier is no longer the limiting factor. Because of its good performance and
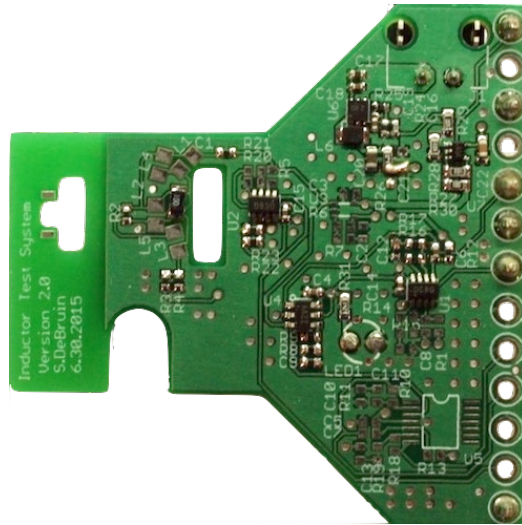
Figure 4.10: PowerBlade amplifier test platform. This system contains the sense inductor front end, as well as the footprint for the amplifiers evaluated for use in PowerBlade. We create a copy of this board for each amplifier considered for the design, and evaluate the performance.

low system current, we use the OPA333 as the first stage amplifier in PowerBlade. Further evaluation of the additional limiting factor and other improvements are left for future work.

4.2.4 Efficiently Computing Power

PowerBlade digitizes voltage and current, calculates power, and transmits the data wirelessly. Our implementation is the MSP430FR5738 microcontroller from Texas Instruments [25] and the Nordic nRF51822 BLE transceiver [20]. The MSP430FR5738 has internal non-volatile FRAM, which is important for maintaining state as the device is unpowered or moved around. The nRF51822 is selected both because of its suitability from the energy quanta analysis in Section 4.1.5, and because the data format, BLE, is readily readable by smartphones. This means that the PowerBlade system can either operate through a fixed gateway or by relaying the data through listening smartphones.

4.2.4.1 Calculation and Calibration

Power calculation begins with the MSP430 taking voltage and current measurements with internal ADCs at 2,520 Hz. This frequency is both an even divisor of the sampling frequency, 32,768 Hz, and as an even multiple of the 60 Hz signal being measured. Next, the MSP430 integrates the current signal in software to correct for the derivative factor in Equation 4.1. This results in a virtual representation of the true current waveform. True power is

the instantaneous product of voltage and current. This is averaged for 2,520 samples and reported at 1 Hz. Apparent power is the product of V_{RMS} and I_{RMS}, which are calculated at 60 Hz. This is then also averaged and reported at 1 Hz. Finally, the MSP430 calculates watt hours, the integral over time of true power.

A significant benefit of the MSP430's on-board FRAM is storing watt hours as a non-volatile number. PowerBlade travels with the plug it is measuring, which means that it will have full knowledge of that device's energy consumption over time. This is only valuable if that accumulating value is maintained as devices like computer chargers and kitchen appliances are moved around. Initial versions of PowerBlade provide a wireless interface to reset this accumulator, but later versions leave this feature to the application level.

Calibrating a PowerBlade requires four measurements, each performed by the PowerBlade unit itself in a known environment. Each of the voltage and current measurements have an offset in hardware slightly different than $V_{CC}/2$, and even after removing this, the software integrator will generate an additional integration offset. Finally, the power measurement must be scaled to accommodate the unknown transfer function from inductor measurements to current. We perform all four calibrations in one step, by connecting PowerBlade to a known load with a zero-biased current signal. The MSP430 can then calculate its own offsets, which it then subtracts from subsequent samples. It also calculates its own scaling value, but to reduce the burden on PowerBlade it transmits this value with subsequent transmissions and the receiver performs the multiplication.

4.2.5   BLE Communications

PowerBlade sends metering data in BLE advertisement packets. The choice of BLE means the data is available not only on fixed gateways but also on modern smartphones for real-time visualization [120]. Further, including the data in the advertisement, rather than requiring the phone or gateway to establish a connection, means that the data is available in real time on multiple devices.

Once per second the MSP430 transmits new values to the nRF51822 over UART. To increase the likelihood of reception the advertisement rate is 5 Hz, five repeated messages are transmitted from the nRF51822 over BLE for each unique UART message. A sequence number transmitted with the data distinguishes repeat transmissions in the receiver.

PowerBlade is robust against packet loss. The value of watt hours transmitted with each packet is the integral over time of power. If a packet is lost, the overall watt hours number retains a valid measure of total energy. The resolution provided by that packet is lost, but overall accuracy is maintained.
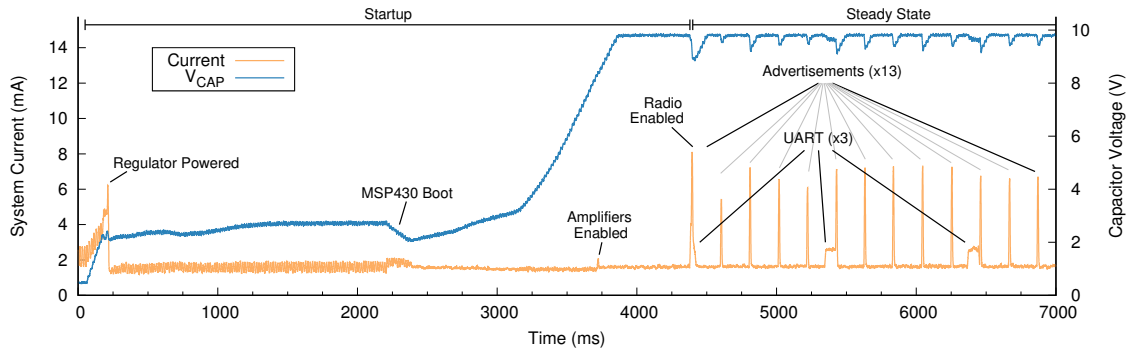
Figure 4.11: PowerBlade timing: storage capacitor voltage and 3.3 V regulator output current for the first 7 s of operation. At startup the measurement circuits and the nRF51822 are automatically disabled. For approximately 2 s the capacitors charge gradually, and around 2.25 s the MSP430 boots. When the MSP430 detects the capacitor has charged to the nominal voltage of 8.9 V it enables the voltage and current measurement circuits. For the first 1 s of the amplifiers being powered there is no data to transmit, so the nRF51822 is kept disabled. After 1 s of measurement the MSP430 enables the nRF51822 and sends data over UART. At this point the nRF51822 begins advertising data at 5 Hz for the remainder of operation and the MSP430 updates with new information over UART at 1 Hz. This results in up to five identical packets transmitted, which increases the likelihood of reception, and a sequence number transmitted with each packet denotes the duplicate.

4.2.6   System Operation

Figure 4.11 shows the startup phase of PowerBlade and 7 s of steady-state operation. When the system first starts, there is only power to boot the MSP430; if more components are drawing power, the 3.3 V power rail will never enable and the system will lock up. Instead, MOSFETs separately power gate the sensing circuits and BLE radio. With only the MSP430 running, the capacitor charges, enables the 3.3 V rail, and eventually reaches a nominal 8.9 V.

When the MSP430 detects that the capacitor has charged to the nominal voltage it enables the sensing circuits, and these remain powered for as long as PowerBlade is powered. The MSP430 spends 1 s collecting measurements before enabling the nRF51822, which also remains powered for the duration of the trace. At this point the device has entered steady-state operation.

4.3   Evaluation

We evaluate PowerBlade's accuracy in two parts. First, we explore bench top accuracy, where we use PowerBlade to measure the power draw of a programmable AC load—the

60
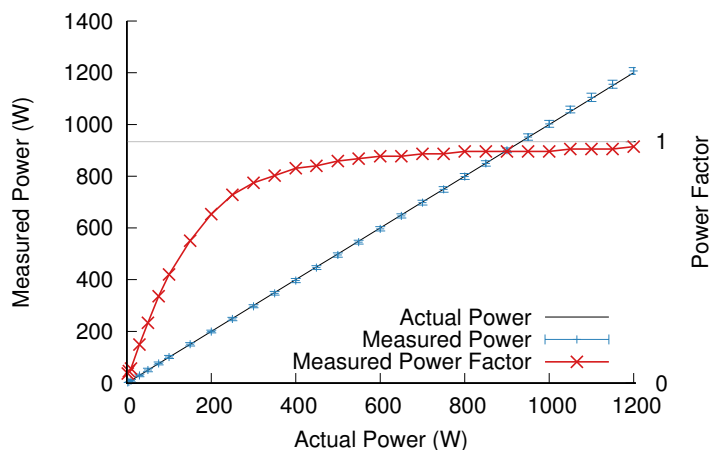
Figure 4.12: Metering accuracy for a variable resistive (unity power factor) load: measured power vs actual power as well as measured power factor. Also shown are the 95% confidence intervals for real power. The minimum AC load for accurate metering is about 2 W and over the range from 2 W to 1200 W the average error in real power is 1.13%. PowerBlade's metering and reporting system is accurate over a range of resistive loads.

APS 3B012-12 [2]—set to unity power factor. This allows us to measure a large part of the metering range (up to 1200 W) in defined increments and in a controlled setting. Second, we use PowerBlade to measure the power draw of several household items. Although not an exhaustive list, this is representative of PowerBlade's target usage. We also present benchmarks that affect the usability of the system, including PowerBlade's volume, cost, wireless performance, and safety of using the system.

For the bench top accuracy, ground truth is provided by the 3B012-12 programmable AC load itself via its serial interface. For the household tests, ground truth is taken from two sources. On the low range, we use a professionally calibrated Power Line Meter (PLM) [5]. This device is limited to 480 W, however, so to measure larger loads we take ground truth from a Watts Up [27]. We report the Watts Up measurements over the low range as well, and it is clear that it is less accurate than the PLM (typical error of about 1.72%). The need to use two different meters for ground truth demonstrates the difficulty in creating a highly accurate whole-range metering solution. For each test we take 30 PowerBlade measurements, 30 ground truth measurements, and report the mean and 95% confidence interval of both.

4.3.1 Metering Accuracy on a Resistive Load

Resistive loads with a unity power factor, which include incandescent lights and power-factor-corrected devices, exhibit a sinusoidal current waveform in-phase with voltage. To measure PowerBlade's accuracy in this simple but common case, we use an APS 3B012-12

Device	PF	Power	Watts Up Error		PowerBlade Error	
150 W Bulb	1.00	162.17 W	1.22 W	(0.75%)	-0.99 W	(0.61%)
Fridge	1.00	108.22 W	–		-5.30 W	(4.90%)
Drill (Max)	0.99	235.21 W	1.69 W	(0.72%)	2.96 W	(1.26%)
Toaster	0.99	827.87 W	–		-22.11 W	(2.67%)
Vacuum	0.98	1246.96 W	–		15.24 W	(1.22%)
Microwave	0.92	1729.73 W	–		16.01 W	(0.93%)
Hot Air	0.83	305.54 W	0.88 W	(0.29%)	-1.93 W	(0.63%)
TV (Normal)	0.62	196.23 W	0.86 W	(0.44%)	-9.03 W	(4.60%)
50 W CFL	0.61	48.57 W	-1.08 W	(2.22%)	-9.51 W	(19.58%)
TV (Static Image)	0.61	129.51 W	-0.04 W	(0.03%)	-4.00 W	(3.09%)
Xbox	0.57	50.44 W	0.64 W	(1.27%)	-0.83 W	(1.65%)
MacBook	0.51	52.68 W	-0.41 W	(0.78%)	-4.49 W	(8.52%)
Blender	0.49	106.63 W	1.95 W	(1.83%)	36.97 W	(34.67%)
Router	0.46	9.11 W	0.22 W	(2.41%)	-0.62 W	(6.81%)
Drill (Low)	0.30	51.10 W	4.18 W	(8.18%)	20.40 W	(39.92%)

Table 4.4: Metering accuracy for a cross section of household devices. On this selection, the average percent error in real power is 6.5%. Although these devices produce more complex waveforms than a fully resistive load, PowerBlade remains acceptably accurate.

programmable AC load set to a fixed unity power factor. Figure 4.12 shows the end-to-end accuracy for PowerBlade metering this resistive load. Displayed are the reported real power and power factor from PowerBlade, as well as the ground truth power up to the programmable load's maximum of 1200 W. Note that the true power factor is equal to one throughout the test.

We measure 29 wattages from 2.2 W to 1200 W: 50 W to 1200 W in increments of 50 as well as 2.2 W, 5 W, 10 W, and 75 W. For these measurements the average error is 2.3 W and the average percent error is 1.13%. At 2.2 W the error is 0.21 W (9.5%) and at 1200 W the error is 7.01 W (0.6%). PowerBlade is very accurate in the benchtop experiment on a resistive load.

4.3.2 Accuracy on Household Loads

Resistive loads constitute a large fraction of household devices, but not all loads have sinusoidal current waveforms. Table 4.4 shows PowerBlade's accuracy for several devices found in a common household. For these devices, we simultaneously take a PowerBlade measurement, a Watts Up measurement and, if the load is below 480 W, a PLM measurement. If available, the PLM is used as ground truth. If not, the Watts Up is used. Although the fridge draws about 100 W, its start-up power tripped the 4 A fuse in the PLM, preventing PLM

measurements for that load. For each device, we report its power and power factor, as well as PowerBlade's error and, if not used as ground truth, the Watts Up error for comparison.

This set of devices has a range of power from 9 W to 1730 W and a range of power factor from 0.30 to 1. The average absolute error for PowerBlade measurements of these devices is 10 W (4.3x higher than the resistive load), and the average percent error is 6.5% (5.8x higher). This error is dominated by two devices with highly inductive (low power factor) draws: the blender (PF=0.49) and the drill set to low power (PF=0.30). Each of these devices has an error of over 20 W and a percent error of over 30%, and excluding these two, the average absolute error drops to 7.2 W (3.1x the resistive load) and the average percent error drops to 4.3% (3.8x the resistive load). Correctly metering such devices is an area of future development.

The difficulty in measuring highly inductive loads, and further the difference in accuracy between the programmable load and the household devices, can be partially explained by
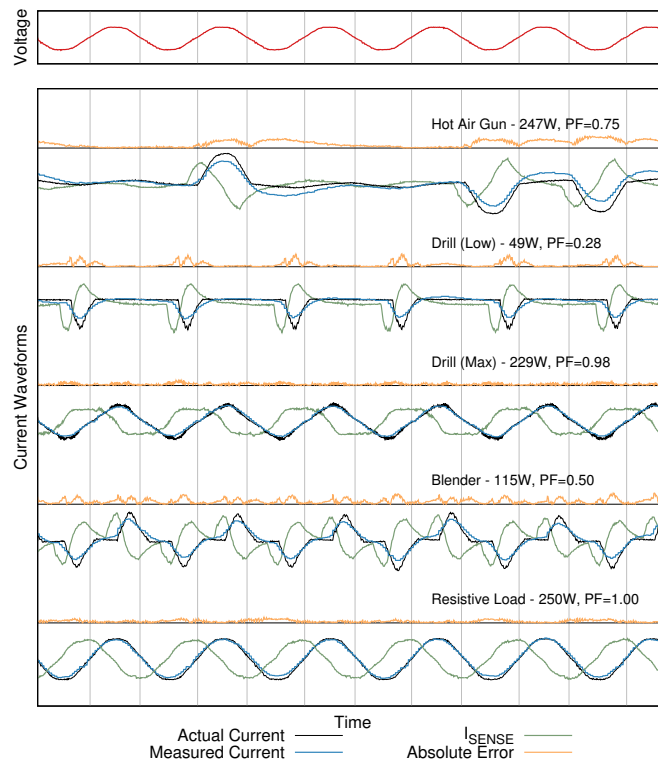
Figure 4.13: Current sensing fidelity. Output of current sense amplifier, I_{SENSE}, and the internal representation of current are shown for several household devices. Also shown is the true current waveform as measured by a commercial current transformer [13]. Visible is the derivative relationship between actual current and I_{SENSE}, as well as the distortions introduced by this sensing and integrating technique. PowerBlade's current sensing method reasonably captures the current waveform.

examining the current waveforms. Figure 4.13 visualizes the current measurement process in PowerBlade for a few loads from Table 4.4. The known current signal, measured by a commercial current transformer [13], is shown along with I_{SENSE}, the signal output from PowerBlade's current amplifier. As described in Section 4.1.1, this represents the derivative of the current waveform. Also shown is the post-integration representation of current. Voltage for each load is synchronized, and zero crossings of the common voltage are denoted by vertical lines.

Visible on the figure is the integral/derivative relationship between I_{SENSE} and known current, as well as the fidelity of the integrated signal to that known current. For devices with sinusoidal or otherwise smooth-transitioning current waveforms, the integrated signal tracks well with known current. For other devices, however, high frequency components in the current waveform are suppressed by the integration process, resulting in increased error. Although this is lower quality than the signal from a current sense resistor, it is still usable in most situations.

4.3.2.1   Watt-hour Accuracy

PowerBlade is designed to measure and report both instantaneous power and watt-hours, the sum over time that will be used by the utility company to levy charges. The figure of watt-hours also accounts for the possibility that one or multiple packets are not received: resolution is lost but watt-hours remains an accurate long-term measure. To evaluate PowerBlade's accuracy in reporting watt-hours, we take simultaneous measurements from the PLM, Watts Up, and PowerBlade for a television in normal viewing use. Figure 4.14 shows the measurements over time from Watts Up and PowerBlade as compared to the PLM.

After 15 minutes of normal television use, the PLM reported 49.07 Wh, the Watts Up reported 49.28 Wh (0.42% error), and PowerBlade reported 46.80 Wh (4.62% error). In instantaneous measurement trials the PowerBlade measurements for the television in full use were off by an average of 4.60%, the watt-hours figure of 4.62% error is consistent with the instantaneous readings.

4.3.3   Deployment Configurations and Shielding

Originally we calibrate PowerBlades and run accuracy experiments on a custom rig created using a GFCI outlet. We find that when these PowerBlades are deployed on outlets in homes the error is higher than on the test rig. To determine the source of error, we consider the configuration in which the PowerBlade is calibrated and the configuration in which it is deployed.
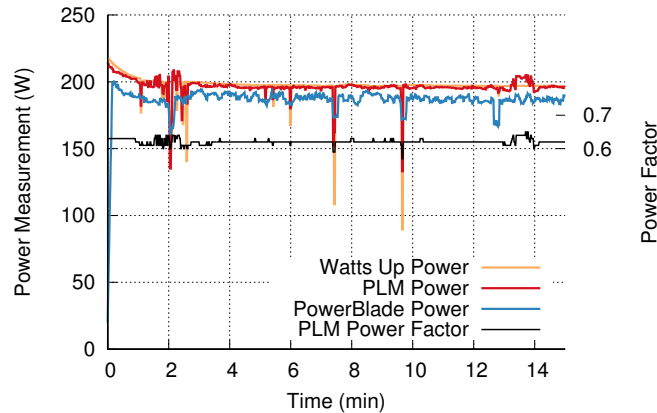
Figure 4.14: Metering accuracy over time for a television in use. Shown on the figure are reported power from the PLM (ground truth), Watts Up, and PowerBlade, as well as power factor reported from the PLM. At the end of 15 minutes the PLM reported 49.07 Wh, the WattsUp 49.28 Wh (0.42% error), and PowerBlade 46.80 Wh (4.62% error). This is consistent with PowerBlade's instantaneous error of 4.60% for the television in full use.

We do these experiments on a programmable AC load, which allows us to test a wide range of power levels, current waveforms, and power factors to mimic household conditions. We have four different calibration and deployment configurations: GFCI outlet, normal outlet, surge protector, and a short 1' extension cord that we call a jumper. We calibrate PowerBlades for each calibration configuration, and measure in every deployment configuration (sixteen calibration/deployment pairs total).

We calibrate all of our PowerBlades on the programmable load set to 100 W with a unity power factor. We take measurements at two apparent power levels and at five different waveform settings that have approximate power factors of 1, 0.84, 0.68, 0.57, and 0.48. For each of the sixteen calibration/deployment pairs we take 10 sets of measurements, one for each power level, and average together the results.

Figure 4.15 shows the results of the configurations experiment. The nomenclature for the x-axis is calibration.deployment, i.e. calibrating on a GFCI and deploying on an outlet results in greater than 100% average error. This agrees with our observations in the field. Calibrating and deploying on a GFCI is much better, approximately 12% error which also agrees with our above experiments. An outlet is the highest overall error, and calibrating and deploying on a jumper or a surge strip are the lowest.

We consider two possibilities for the cause of this phenomenon. First, that the outlet vs. GFCI vs. surge/jumper apply pressure on the prongs in a different way and that this affects the distance between the prong and the inductor. However, this seems somewhat unlikely. Plugs and outlets that are different ages, different prong spacing, etc, all exhibit
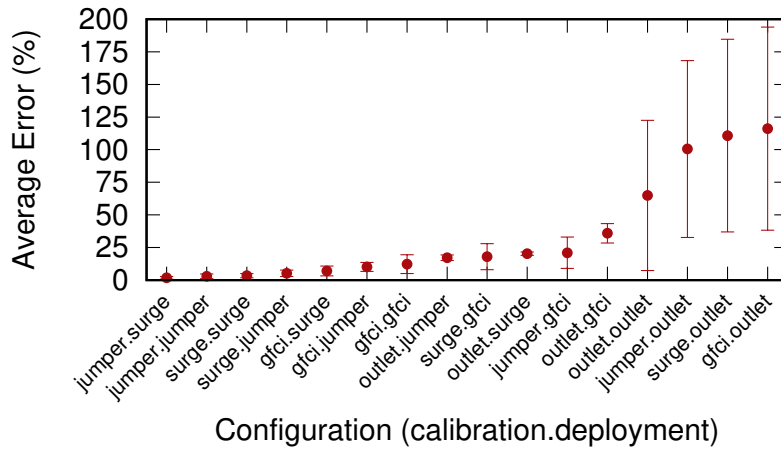
Figure 4.15: Measurement error in each calibration and deployment configuration. PowerBlades are calibrated in four configurations, GFCI outlet, normal outlet, surge protector, and short extension cable (jumper). Each is cross-validated with the four configurations. The deployment configuration with the highest error is outlets, so we do not deploy PowerBlades directly on GFCI or normal outlets. We calibrate on surge strips, and deploy with surge strips or jumpers for an expected deployment error of around 4%.

very consistent results for the experiment shown in Figure 4.15. The second option is that outlets and GFCI are the source of magnetic noise that affects the measurement. This is corroborated by the fact that certain inductive loads, in particular an electric knife sharpener, cause significant noise when operated near a PowerBlade. We evaluate against this second hypothesis, and attempt to solve the issue through shielding.

To determine if the source of the configuration accuracy issue is EMI, we evaluate four types of shielding to build in to the PowerBlade sensor. Because the exact frequency and intensity of the EMI source are unknown we evaluate multiple thicknesses and materials. Table 4.5 shows the four shielding options. We evaluate a thin piece of MuMetal, two different thicknesses of ferrite powder sheets, and a woven mesh shield. The ferrite powder sheets are non-conductive, and can be applied directly to PowerBlade, but the other two require electrical isolation.

In order to determine the effectiveness of each shielding type, we first build a large box around the PowerBlade and the plug and cover the box with the shield material. We then use the knife sharpener placed outside the box to determine if the shield has any effect before actually building the shield into the PowerBlade sensor itself. Of the four, only the MuMetal has any efficacy in shielding the PowerBlade signal from the knife sharpener's interference. This allows us to confirm that the MuMetal is effective to EMI of the frequency and type that PowerBlade is susceptible to.

66

Type	Thickness (in)	Percent Increase	Material	Conductive?
MuMetal	0.002	3.2%	MuMetal (nickel-iron)	Yes
Mesh	0.004	6.4%	Nickel-Copper Polyester	Yes
Ferrite Powder	0.024	38.4%	Ferrite Powder	No
Ferrite Powder	0.039	62.4%	Ferrite Powder	No

Table 4.5: Shielding materials evaluated for use in PowerBlade. Because the exact source and frequency of the interference is unknown, we evaluate multiple types and thicknesses to determine if any has any effect on the outlet accuracy issue. We evaluate 2 mil MuMetal foil, 4 mil nickel-copper mesh, and two different thicknesses of ferrite powder. We also calculate the percent increase that this would cause in the total PowerBlade thickness. Of these four, MuMetal is very effective at shielding from EMI and only adds 3.2% to the current system thickness.

Next, we cut a small piece of MuMetal to build a shield to place between PowerBlade and an outlet. This allows us to not only confirm that EMI is the source of the error on outlets, but also that the MuMetal is an effective solution. For this experiment we take six total sets of measurements: on both a surge and an outlet we measure a bare PowerBlade with no shield, a shielded PowerBlade that was calibrated without the shield, and a shielded PowerBlade that was calibrated with the shield. Figure 4.16 shows the results of this experiment, which confirms that MuMetal is an effective solution.

Specifically, in the bare configuration we see what we expect: that on the surge the PowerBlade measures very accurately but that the error is higher on the outlet. The disagreement between the measurements, calculated as the percent error of the outlet measurement relative to the surge, is almost 18%. After adding the shield the error of both configurations is higher, but the disagreement is much lower (3%). This implies that the presence of the shield affects the calibration, but that it affects both configurations similarly. Finally by re-calibrating the PowerBlade on the surge strip but *with* the shield, we see that the error of both measurements, as well as the disagreement, are all low. This demonstrates that if shielding is included in the PowerBlade, the accuracy issues are mitigated.

4.3.4 Usability Benchmarks

PowerBlade's accuracy makes it comparable to other power metering systems, but it is the usability of the system, and in particular the size, cost, and wireless communications, that most distinguish it. In this section we also evaluate the safety of using the system.

	Bare	Shield	Re-cal
Surge Error	1.52%	19.84%	1.46%
Outlet Error	16.10%	23.91%	2.42%
Disagreement	17.90%	3.39%	3.94%
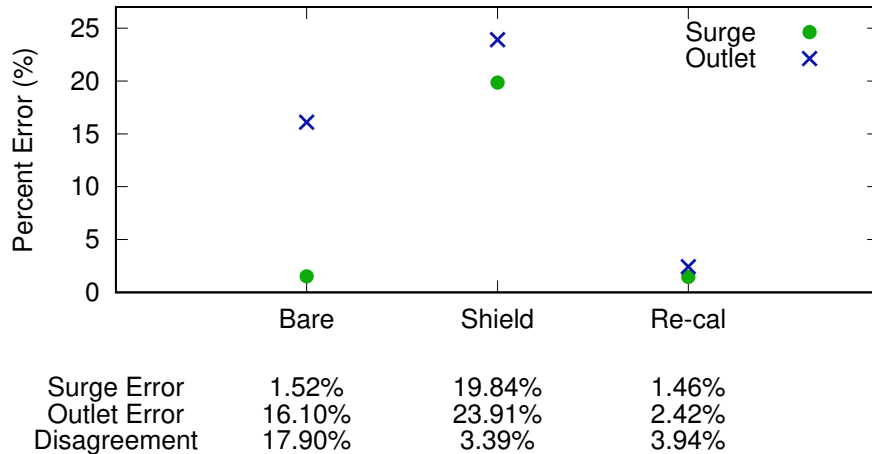
Figure 4.16: Accuracy improvement using a MuMetal shield. We take measurements with a PowerBlade without any shield, with a MuMetal shield but without recalibrating, and with the shield and recalibrating. There is significant error in the bare outlet configuration, and significant disagreement between the measurements. With the shield there is error in both measurements, but the disagreement is lower. This indicates that the shield affects the calibration similarly for both configurations. Finally if we recalibrate the PowerBlade with the shield both errors and the disagreement are all low. This shows that the source of the error is EMI from the outlet, and that the MuMetal is an effective shielding solution.

4.3.4.1 Wireless Range

We test the effectiveness of PowerBlade communications by measuring packet reception rates in three configurations of PowerBlade units. First we deploy and evaluate a single PowerBlade as a baseline. Next we place three PowerBlade units throughout a room as a more typical deployment case. Finally, we place three units on a single power strip and activate them simultaneously to test for possible packet collisions on worst-case synchronized nodes. We record both unique and total packets received per second. PowerBlade updates data at 1 Hz, and BLE packets are sent at 5 Hz, so nominally we should receive 1 unique and 5 total packets per second.

For each configuration we evaluate packet reception rate at three transmission distances, and perform the entire experiment both in an apartment and in a campus lab. The apartment consists of three rooms in a row, with measurements taken in the same room, adjacent room, and two rooms away. Only a single other BLE device is active in the apartment. The lab consists of one room and adjacent hallways, and measurements are taken in the same room, immediately outside of the room in the hallway, and 20 m down the hallway. The lab environment includes 16 non-PowerBlade BLE devices as well as numerous other devices active in the 2.4 GHz band.

Component	Cost	Component	Cost
MSP430FR5738	$1.71	Sense inductor	$0.30
nRF51822	$1.66	Amplifiers	$2.51
Antenna, balun, & crystals	$1.70	MOSFETs	$0.12
Buck converter	$0.65	Other passives	$0.93
		PCB	$1.47
		Total	$11.05

Table 4.6: Cost of the PowerBlade sensor system. The unit is roughly $11 in quantities of 1,000. We believe this represents an acceptable price point for effective plug-load metering deployments.

In all cases, the unique reception rate is at or close to the nominal 1 Hz when in the same or adjacent rooms, but the total reception rate decreases in the adjacent room. Further, the total reception rate is higher in the apartment than in the lab for both distances. Taken together these three results suggest that range and interference do affect BLE transmissions, but the redundancy in PowerBlade helps ensure reliable data communication.

The distant measurements show continued decline in total packets, but also a decrease in unique packets: 20% to 50% of the unique packets are not received at all (all five redundant packets were all dropped). This indicates that this distance—whether two rooms separated in a residential setting or 20 m down a hallway in a university building—exceeds PowerBlade's usable range. We note that RF designs require some degree of lumped-parameter tuning to achieve maximum performance but PowerBlade's RF circuitry has not be tuned yet, so that may help explain these results.

4.3.4.2 Volume and Cost

The defining characteristic of PowerBlade is its volume: the entire system is a single PCB. This circuit board is 1.0" on a side, and the PCB itself is 0.023" thick. The thickest component on the surface is the antenna at 0.043", so the combined total thickness of the system is 0.066". This is the same thickness as the pass-through section of the FlipIt charger [7], which is a certified commercial product.

The component breakdown in PowerBlade with costs is listed in Table 4.6. Prior to consumer use, PowerBlade needs an enclosure, but the system could be largely assembled for $10-$15 per unit. Although it is important to note the distinction between the *cost* of PowerBlade and the *price* of other systems, this is slightly less than the price of Kill-A-Watt ($23.99) and significantly less than the price of Watts Up ($130.95). The cost of $10-$15 for PowerBlade is also an un-optimized reporting of DigiKey pricing; the minimum viable cost would likely be much lower.
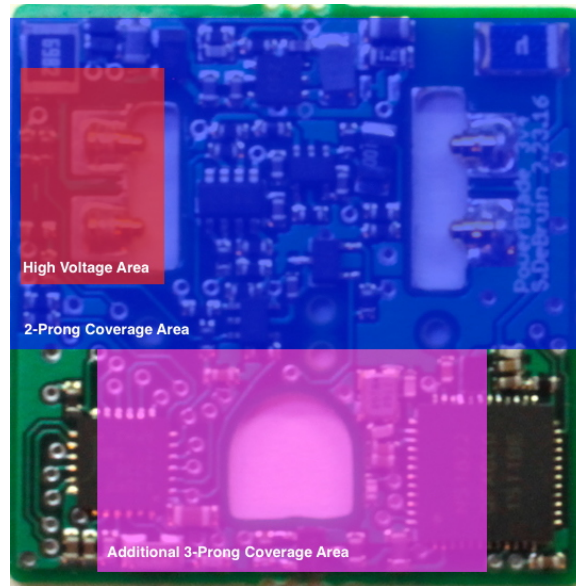
Figure 4.17: The high voltage region on PowerBlade covered by both 2-prong and 3-prong plugs. The small area of PowerBlade that makes contact with high voltage AC is shown, as are the regions covered by a 2-prong plug and the additional region covered by a 3-prong plug. Although not sufficient for UL certification, this effectively shields the user from any high-voltage exposure.

4.3.4.3 Safety

PowerBlade makes contact with high-voltage AC, and because there is no protective enclosure there is the possibility for user exposure. However, we take steps in the design to address this issue, and in addition the system could receive a protective overmolding.

At the level of the PowerBlade design, the power supply circuit shown in Figure 4.2 is agnostic to orientation on the AC lines. The system functions similarly whether V_{AC} is the Phase AC path or the Neutral AC path. However, whichever path is not selected as V_{AC} is used as PowerBlade system ground, and will be pervasive throughout the PCB. If Phase is selected as ground, most of the area of the board is energized at nearly 120 V relative to earth ground. This could present a safety issue as contact with nearly anywhere on the board could result in shock.

Instead, we use the Phase AC path as V_{AC}, electing to use Neutral as the system ground. Although Neutral is not identical to earth ground, it has a much lower potential than Phase relative to ground. Now, touching the system in most locations is imperceptible, similar to touching both terminals on a 9 V battery. There is still a small 120 V area, shown in red in Figure 4.17, but this region is fully covered by both 2-prong and 3-prong AC cords. Although not perfect, this presents a viable design point for deployability without fear of injury, and in the future the PowerBlade system can be fully encased in plastic overmolding.

4.4 Thickness Reduction and Overmolding

There are three primary areas where the design of PowerBlade could be improved. First, although the shielding experiment in Section 4.3.3 indicates the effectiveness of the MuMetal shield, integrating this into the manufacturing of the system is a challenge. Second, the thickness of the system can be further reduced. Third and finally, protective overmolding on the design increases safety and facilitates both adding the shielding and reducing the thickness. All three of these changes require significant manufacturing volume to make cost-efficient, but we describe them here for future applications.

Regarding shielding, although the shielding experiment indicates that a 0.002" MuMetal shield is effective, we are currently unable to directly incorporate this result into our measurement system at scale. We cut the shields by hand and take great care to insulate and position them to prevent shorts. Although the shield could be stamped and the unit costs likely very low, the issue remains integrating it into PowerBlade in a way that does not short to components or AC prongs. If the PowerBlade itself incorporated electrical isolation, as in plastic overmolding, this issue would be simplified.

The overall thickness of the system is determined by four things. The first is the PCB itself, at 0.029" this is approximately half of the total thickness. The other half is components, and in particular the antenna, sense inductor, and pogo pins. The antenna and sense inductor provide value in proportion to their height – changing either component will have implications on the wireless radio and current measurement subsystems, respectively. Reducing the heights of the pogo pins and PCB can have dramatic effects on thickness without affecting performance, but require larger deployments to justify the larger up-front expenses.

We have worked with the manufacturer of the pogo pins to identify the pathway to, and costs associated with, a custom pin solution. This requires a large up-front investment in non-recurring engineering (NRE) costs, but results in comparable unit cost at scale. Further, we believe that with the correct design we can move from two pins per side to one pin per side, effectively cutting the pin cost in half.

Regarding the PCB, in Figure 4.6 rigid FR4 is a sizable portion of the total thickness. Removing the FR4 regions saves 0.012". Optimizing for adhesive and polyamide thicknesses saves another 0.004". Together this reduces the PCB thickness by 55%, and reduces the total thickness by 26%. This results in room available for overmolding and shielding, but at the expense of board structure. The original PowerBlade flexible design was actually made with this construction, but without an enclosure the board would flex and components break off. This indicates that this step should be performed along with a change to an overmolded design, which can provide the rigidity separately from the PCB.

71
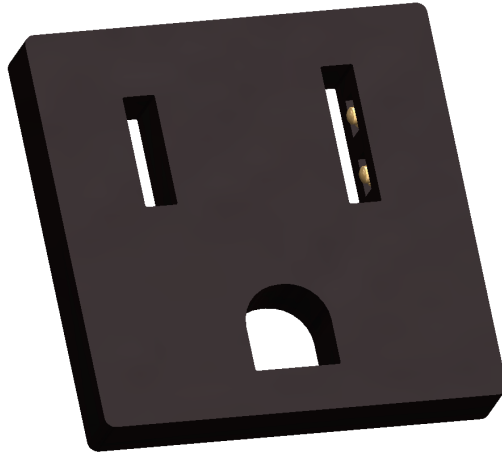
Figure 4.18: Three dimensional rendering of a PowerBlade unit with overmolding installed. Only the pogo pins are exposed for making contact with the AC prongs.

Issues regarding safety, shielding, and thickness can all be mitigated by adding over-molding, shown as a model in Figure 4.18. Safety is increased as the high-voltage and low-voltage components are all hidden inside a resistive coating. The prospects for shielding are also improved as the shield could be incorporated into the overmolding process. Finally the thickness issue is mitigated by changing to a flexible board, which requires the rigidity of overmolding for strength. Similarly to custom pogo pins, overmolding requires significant NRE that makes it inappropriate for small deployments, but future larger scale deployments could take advantage of this solution.

This discussion also has implications on the safety of the system. Until the shield can be directly incorporated into the design, PowerBlade is deployed with jumpers as described in Section 4.3.3. This provides sufficient spacing between the PowerBlade and the outlet to eliminate the accuracy issue, but presents a different safety issue. In some cases in kitchens or bathrooms the outlet is less than 1' above the counter, so the jumper-PowerBlade-plug junction actually rests on the countertop itself. If water ever pools on this countertop, this leads to the possibility of a short through the PowerBlade system. This issue can be mitigated by incorporating shielding directly into the PowerBlade design. If the jumpers are not required then PowerBlade will no longer rest down on the countertop and any possible short due to water pooling is prevented.

4.5   Summary

In this section we reexamine each aspect of the wireless AC power meter, allowing us to arrive at a form factor viable for unobtrusive AC power metering. PowerBlade is 1" x 1" x

1/16", devices plug through it before plugging into the wall. It has comparable accuracy to larger and more obtrusive energy meters, and enables broader and more long lasting energy metering studies.

Two questions remain regarding the deployability of the device. First, in Section 4.1.3 we evaluate the ability of PowerBlade to contact a diverse set of AC plugs, but the plug must still make contact with the outlet with PowerBlade in between. Because this is a function of both plug width/length and properties of the outlet, this can only be evaluated by deploying in several households. Second, the fire-and-forget communication method may not be sufficient for households, or may require so many gateway devices that the unobtrusive meter results in a very obtrusive system.

In the following section we present a deployment system involving PowerBlade and other sensors, a gateway for forwarding data, and a cloud database for storage and analysis. We deploy this system in eight residential households and two other locations, which allows us explore the two open questions in the PowerBlade design. A small deployment of PowerBlade also provides an opportunity to generate a dataset of metering data. The size and price point of the device affords the ability to deploy at a scale not previously possible, and the resulting dataset can motivate the location of future deployments and begin to formulate regulatory recommendations. In the next chapter we present a small deployment involving PowerBlade and other sensors, and the dataset that results.

CHAPTER 5

High-Coverage Residential and Commercial Energy Metering System

In this section we present a residential energy metering system involving PowerBlade and other sensors. The PowerBlade sensor is used to measure plug-load devices, and a combination of two different light sensors measure the energy use of built-in lighting fixtures. Monjolo units can also be used in the AC panel to measure high-power devices like the oven and HVAC. All sensors relay their data through a gateway, typically one or two per location, to a cloud database for storage and analysis. This is an unobtrusive metering system that, once installed in a household or other location, does not affect the aesthetics of the space.

To evaluate the performance of this system and generate a dataset to motivate future work, we deploy this system at eight residential households and two other locations. We install an average of 41 sensors at each location, and leave them for an average of 62 days. This provides an opportunity to analyze the metering accuracy and wireless performance of this system in real situations. We find that the aggregate error of the measurement system is generally less than 5%, and that the average measurement latency is 1.16 s. In following chapters we explore the energy analysis of this dataset, but these accuracy and wireless results indicate that the system is usable in the residential environment.

5.1   Goals and Requirements

We describe two high-level requirements for a residential and commercial metering system. The first is full coverage over the devices to be measured, whether plug-loads, plug-loads and lighting, or every AC consumer. Not only is this required for determining the fraction each device contributes to the total, it enables realizations that certain devices are drawing more than expected. The second requirement is data availability; the sensor system must be

74

able to relay the large amount of data to a central aggregator. In this section we explore the effects of each of these requirements on the design of the sensor system.

5.1.1   Full Residential Coverage

Although central to this evaluation is the energy meter PowerBlade, not all energy consumers are outlet devices. Electrical devices can be grouped into four categories: outlet devices, lighting, fixed high powered loads like HVAC, oven, and clothes dryer, and fixed low powered loads like the bathroom fan. Outlets loads are covered by PowerBlade, and each of the other three can be addressed with a unique solution. Differing sources and types of data can be aggregated into a common dataset.

In addition to plug-loads, the primary segment of miscellaneous electrical loads is lighting. In this case an inline sensor is inappropriate, both because light sconces and other fixtures are often closely formed around the bulb and because of the large diversity in shapes and fittings of indoor and outdoor lighting. Instead, we note that unlike many electrical loads, light bulbs are labeled with their power draw in watts. To measure the energy consumed by lighting, we use a light sensor to determine whether that bank of lights (controlled by a single switch) is on, and correlate with the known power. This sensor could be the solar panel version of Monjolo: the presence of packets from the Monjolo indicates the light is on.

As noted in Chapter 2, fixed high powered loads are already well addressed by NILM/disaggregation and are therefore not the focus of this study. Armel et al. predict that smart meters deployed by the utility are sufficient for accurately metering these high powered loads, and these are increasingly common and provided at no cost to the consumer [51]. A submetering deployment can reasonably rely, in the near future, on data on high powered devices already being available. However, this coverage analysis requires ground truth to determine if all watt hours are being accounted for. Whether by actually implementing a NILM algorithm or through some other method, these high powered loads must be accounted for in the aggregate as they will have a significant effect on ground truth.

Finally, a few loads are on the same circuits as outlets and lighting but are too low power to be accurately measured by NILM techniques. This includes the bathroom fans and the oven fan in the kitchen. The multi-modal sensing technique from [85] using vibration or airflow sensors could be used to target this group of devices. However, since those loads are very low power and infrequently used, for the purposes of this study they are ignored.

Figure 5.1 shows the architecture of this deployment. PowerBlades on outlets and light sensors on lights transmit to a few always-on gateways, which forward the information to a cloud database. In parallel, a smart meter measures the aggregate current supplied to the
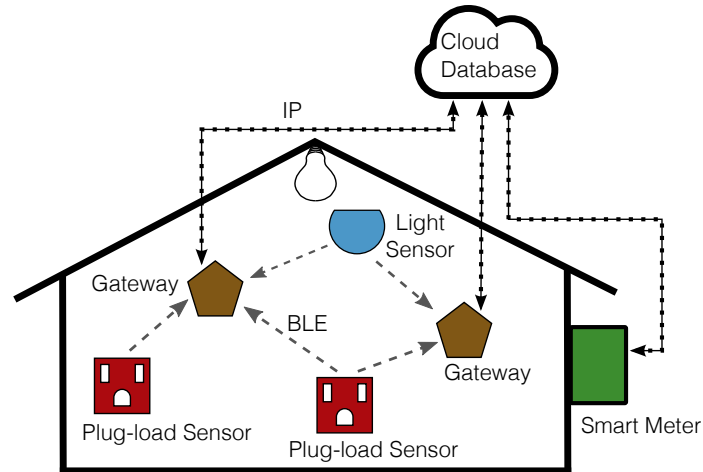
Figure 5.1: Deployment architecture. Energy meters are connected to plug-loads. Light sensors are placed near lights to determine state. Gateways collect data from sensors and log to a cloud database. A smart meter attached to the apartment monitors ground truth power.

apartment, which is used for ground truth. Not pictured in this diagram are Monjolo sensors, which reside in the panel and also transmit to the gateways.

5.1.2 Gateway and Other Wireless Provisioning Requirements

In addition to testing full coverage residential submetering, this deployment is an opportunity to test the PowerBlade unit at scale. In this section we develop a theoretical expectation for how the wireless aspects of the system perform. This can provide an initial guide for system provisioning, i.e. how many gateways are required to effectively cover a space, but this will also be refined experimentally. The performance of the network will be affected by two factors: transmission range/power and wireless interference. We explore both to determine the expected effect and expected performance of our deployment system.

5.1.2.1 Transmission Range

The BLE specification suggests that the range of a transmission is up to 100 m, but in practice the transmission distance will be lower. In particular the location of PowerBlade, between the plug and the outlet and under appliances and furniture, has the potential to strongly attenuate the wireless signal. A deployment allows us to generate a dataset to approximate the range of a PowerBlade transmission, both in distance and in maximum number of obstructing walls. This results in an approximate number of gateways required based on square footage or number of rooms. In general we expect about one gateway per room, but gateways will receive packets from multiple adjacent rooms.

5.1.2.2 BLE Broadcast Contention

BLE advertisements are sent naively without performing any channel sense prior to transmission. Packet collisions can occur when transmissions from two devices overlap in time and space. The introduction of a uniform random delay in the BLE specification ensures that devices that collide on one transmission will drift apart and be unlikely to collide on subsequent transmissions, but the same drift can cause devices to collide which previously did not.

The uniform random delay causes transmission times to evenly distribute over the transmission period. Even if all BLE devices start simultaneously, given enough time their transmissions will distribute over their broadcast interval. Knowing that transmissions themselves follow a uniform distribution, the probability that k transmissions occur simultaneously follows a Poisson distribution.

We find that BLE advertisements are well modeled in the literature by Pure Aloha [38]. The probability of a successful transmission follows Equation 5.1, where G is the mean number of transmissions during the time to transmit a single packet [115]. In BLE, G is one less than the number of devices multiplied by the ratio of the time to transmit an advertisement to the advertising period.

$$Pr[success] = e^{-2G} = e^{-2(N-1)\frac{T_{tx}}{T_{interval}}} \tag{5.1}$$

With a T_{tx} of about 400 μs and an advertisement period of 200 ms, the likelihood of a BLE collision is very low until the number of devices reaches the hundreds, matching the findings from Beacon Trains [95]. Since the range of BLE is limited to a few rooms, this effectively means that this technique scales to 50-100 devices per room.

5.1.2.3 RF Interference

BLE operates in the 2.4 GHz spectrum, and it is susceptible to interference from other RF sources such as WiFi. Each advertisement transmission actually consists of three sequential advertisements, one on each of three different advertisement channels. A receiver can only scan one channel at a time and thus is likely to receive only one packet per advertisement event. Over time, however, this process allows at least some packets to be received, even with heavy interference on other channels. These advertisement channels were opportunistically selected to avoid the primary WiFi channels (1, 6, and 11), but some interference from WiFi will still occur. Other sources of acute interference, such as microwave ovens, will cause some data loss as well.
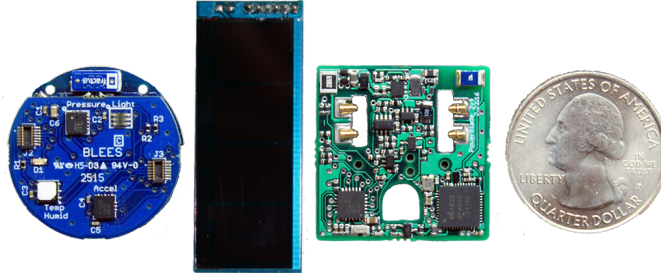
Figure 5.2: Sensors used in the residential energy metering deployment. Shown left to right are the the lighting sensors, BLEES and Ligeiro, and the plug-load sensor, PowerBlade. BLEES and Ligeiro are light sensors that can be used to determine time-in-mode of overhead lights, each with strengths and weaknesses. PowerBlade is a single circuit board, and measures the power draw of loads plugged through it.

5.2   Deployment Methodology

In this section we present a system that meets the goals and requirements for residential and commercial electricity metering. This involves the small plug-load sensor PowerBlade, two different light sensors covering built-ins, a gateway to relay data to the cloud, and a cloud database to store and process the data. In addition we describe the ground truth methodology that we utilize to evaluate the metering accuracy of the system. Finally in this section we describe the real-world testbed that we instrument to evaluate this system.

Figure 5.2 shows the sensors used in this deployment: the plug-load meter PowerBlade and the light sensors BLEES and Ligeiro. Both are small and low power. BLEES operates on a similar principle to PowerBlade, and Ligeiro is similar to Monjolo but with a solar panel. In the following sections we describe these other sensors in more detail.

5.2.1   Plug-Loads

Plug-loads in this deployment are monitored by PowerBlade. The ground truth, discussed below, only reports true power, so for this application we do not use the apparent power reported from PowerBlade. For each PowerBlade packet we store the true power, RMS voltage, watt hours, and metadata about the packet like sequence number, the PowerBlade MAC address, and the gateway MAC address. Although true power can be reconstructed by taking the derivative of watt hours, we retain both values for ease of computation later.

5.2.2   Built-Ins

To measure overhead lighting we use a combination of two different sensors. Although it would be difficult to unobtrusively measure the power of the light with an inline sensor,

the wattage is known and is often written on the bulb. A light sensor next to each overhead light can determine when the light is switched on or off, and these activity periods can be correlated with the known wattage of the bulb. The device is not an energy sensor like PowerBlade is, this would require an over-time aggregation of the correlated power, but the data can be post-processed to make up for this. Future systems could also incorporate this functionality into the sensors directly.

The two sensors that we use in this system both perform this light-sensing function, but have different strengths and weaknesses. The primary light sensor is an energy harvesting device called Ligeiro. Ligeiro operates according to the Monjolo principle [64]; it harvests solar energy from the light until it can send a packet. The presence of the packet acts as the sensor measurement indicating that the light is on. Like PowerBlade, this has the benefit of never needing a battery replacement. However, for certain rarely-used lights the device never activates, and it is difficult to distinguish this from sensor failure.

For rarely-used and far-away devices we use a battery-powered light sensor called the BLE Environmental Sensor (BLEES). BLEES functions similarly to PowerBlade in that data are broadcast at 1 Hz in BLE advertisement packets. The benefit of this sensor is that even if the light is rarely used, we get an indication that the sensor is active and working. However, the significant downside is the need to replace the battery, on average every 3-4 months. In addition, BLEES data are harder to process. For Ligeiro the mere presence of a measurement has meaning, for BLEES the contents of that measurement have to be inspected to find meaning. For these reasons most of the light sensors deployed are Ligeiros.

5.2.3   Other Devices

The only devices not covered by PowerBlade or BLEES are the fixed high powered loads light HVAC and oven, and the fixed low powered loads like the bathroom fan. Neither is the target of this study, high powered loads are already addressed by NILM techniques and low powered devices do not contribute significantly enough to warrant direct measurement. For controlled experiments comparing our measured total to the ground truth value we must identify periods where these devices are inactive. To remove the high power loads from our sense of ground truth, we use Monjolo sensors installed in the electrical panel in the apartment. This is easy and safe to deploy, connects with our existing infrastructure, and is reasonably accurate on devices like the oven and HVAC. Although for the purposes of this study we only use Monjolo to isolate periods from the ground truth, in the future we hope to install a panel meter that measures true power.

5.2.4 Gateway

To collect data on a continuous basis, we deploy a gateway node with radios and an Ethernet interface. This gateway system is built on a BeagleBone Black [3, 8], and data are moved through the gateway using MQTT [17]. When the gateway receives a packet that it can decode, it relays that measurement to a cloud database for processing.

In addition to transparently relaying data the gateway also performs some basic processing. To reduce processing load on PowerBlade each device transmits raw measurements and unit-specific scaling factor together, and the gateway multiplies these values to produce the final power reading. BLEES and Ligeiro do not require an additional processing step.

5.2.5 Ground Truth

Ground truth for this study is provided by the locations' utility companies in various forms. Due to deployment of smart meters for monitoring aggregated energy use, many locations have access to finer grained data than just a monthly bill. This is often available in an app or other electronic form, and sometimes only in paper form.

5.2.6 Data Storage and Processing

All deployment data are stored in a MySQL cloud database hosted by Amazon Web Services. The database stores the raw data from each sensor type (PowerBlade, BLEES, and Ligeiro) as well as metadata for the deployed sensors.

5.2.6.1 Database Structure

Each of the three raw data tables, *dat_powerblade*, *dat_ligeiro*, and *dat_blees* receives data directly from the gateways. There is a row for every measurement sent by a sensor and received by a gateway. Some measurements can appear multiple times if they are received by multiple gateways, but mostly each row in a raw data table represents a unique measurement. Each contains generic information common to all sensors, the timestamp of the packet and the MAC addresses of the device that sent and the gateway that received it, as well as sensor-specific information. For PowerBlade this is the instantaneous power and total accrued energy, for BLEES it is the instantaneous lux measurement, and for Ligeiro it is simply the count value indicating how many wakeups have occurred.

In addition to the raw data table, each sensor type also has a lookup table describing the metadata for those sensors. Again this includes some generic information for all sensors like when the sensor was deployed and where, as well as some sensor-specific information. For
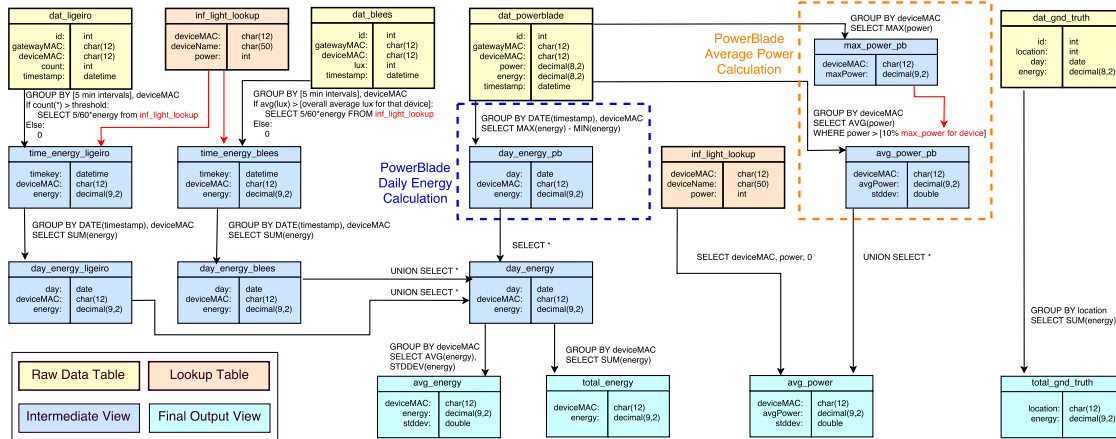
Figure 5.3: Energy and power calculation using database views. We calculate daily energy for each PowerBlade, BLEES, and Ligeiro device, and in addition calculate the average power for PowerBlade sensors. The four final views can be easily queried at any time for energy information.

PowerBlade this is the type of device (television, refrigerator, etc) for grouping analysis, and for BLEES and Ligeiro this is the power of the light being metered.

The raw data and lookup tables are joined on the MAC address of the sensor. However, since a device can be redeployed and its metadata updated, the join between these tables is actually on the combination of the device MAC address *and* specific time period(s). This is computationally more difficult, and in future deployments a unique key could be used to link these tables on the combination of device MAC address and time.

The final component of the database structure is the ground truth data table. This data is not sourced from the gateways but instead from the utility company. This information is recorded manually from the app or other source, tagged with the appropriate location information, and stored in the ground truth table for analysis.

5.2.6.2 Energy and Power Calculation

The complete energy and power calculation at the location level or whole-deployment level can be broken into three sub-processes. First is calculating energy and power for each PowerBlade. Second is calculating the daily energy for BLEES and Ligeiro. There is no need to calculate the power for BLEES and Ligeiro as the bulb wattages are known. These calculations apply uniquely to each device in *each* of the locations where it was deployed. Finally the third step is calculating the average daily ground truth for every location to compare to the measured totals.

Figure 5.3 shows the set of queries to select what devices are active at what locations at what times, and calculate energy and power. To query for energy, we take the maximum energy for each device on each day and subtract the maximum from the previous day. This requires a multi-column MySQL index on (deviceMAC, timestamp, energy). The first index groups by device, then into specific timestamps that can be grouped into days, and then the maximum energy is a direct indexed lookup.

To query for power, for each device we take the average of the power measurements that are greater than 10% of the maximum for that device. This is used to eliminate zero-measurements and idle power from the average (although these values are still integrated within PowerBlade to get energy). This query requires a multi-column index on the columns (deviceMAC, timestamp, power). Again we include timestamp to handle querying certain devices being active at certain times.

Calculating energy for BLEES and Ligeiro requires similar complexity to select which devices are active at what times, but the energy calculation is very different. PowerBlade has an accumulated energy metric, allowing us to subtract each day's final from the previous. At best BLEES and Ligeiro are power sensors, and even then only when correlated with the known power, so energy must be accumulated over time. This is done at the database level, but could in future deployments be done at the gateway or even sensor level.

For Ligeiro, the deployment is broken into 1-minute periods, and the light is assumed to be switched on if the number of Ligeiro counts in that period is above a threshold. If it is, that Ligeiro for that minute is assigned an energy value of $power/60$ to convert power in watts into energy watt-hours. Energy can then be added over the course of a day. The calculation for BLEES is similar, but rather than a simple count during each minute, the count is the number of lux measurements above that device's average during that minute. Each of these requires an index on (deviceMAC, timestamp, [lux or count]) but since there is significantly less data the indexing is less critical.

5.2.6.3 Daily Database Alerting

The database is also used for monitoring the status of the system itself. We find that sensors and gateways deployed in real-world situations are subject to problems, both technology-based and human-based. Gateways can sometimes go down, either temporarily or requiring a reset. Although we are able to improve gateway reliability over time by identifying issues as they arise, gateway downtime can be a daily occurrence. On the human side the types of problems are much more varied from devices being removed or replaced, gateways being moved or unplugged, and other issues. To ensure the integrity of the data these issues have to be caught nearly in real time.

Figure 5.4: Daily hitrate calculation for each device and each gateway. We calculate the number of packets for each sensor in each 15-minute interval in the previous day, which can be used to detect periods when the sensor or gateway drops out. The processing is split between the database and a python script for sending an email. This process allows the research team to quickly identify problems without human interaction.

Figure 5.4 shows the process of detecting faults in the deployment system, split between the database and a python script. We determine a daily 'hitrate' for each device and gateway in the deployment. We use a MySQL view that divides the previous day into 15 minute blocks. For each PowerBlade, BLEES, and Ligeiro we take the number of unique packets (removing the repeat packets received by multiple gateways) in each block. For the gateway hitrate we join the sensor hitrate tables and average across each gateway, which normalizes each gateway hitrate against the number of sensors in range.

At 5 am daily, a python script queries this hitrate view for all the devices that should be active (as found in the metadata tables). Any device with a hitrate below a threshold is marked with a warning. Any device with a hitrate of zero (no packets received in the previous day) is marked with an error. All devices with a warning or error, as well as a summary of OK devices for each location, are sent in a daily email to the research team. This allows us to monitor the deployment and quickly address issues without manually sifting through raw data.

Figure 5.5: Floor plan of the sample apartment. This is one of eight total residential locations (ten locations overall), and is over-instrumented with gateways to allow network evaluation. In this diagram PowerBlade sensors are represented by red squares, BLEES/Ligeiro by blue circles, and gateways by gold pentagon-shaped icons. This location has a total of 74 devices: 56 PowerBlades, 11 light sensors, and 3 gateways. Not pictured are the Monjolo sensors in the panel, of which there are four.

5.2.7 Deployment Statistics

We install this measurement system at eight residential households and two other locations. The average deployment duration is 62 days, the longest is 168 days and the shortest is 16 days. These locations cover a range of number of occupants, age, and demographics.

In addition we select one location to examine more closely. The sample apartment is 1,100 square feet, with one bedroom on either side of a common living/dining room with an open kitchen. There are plug-loads and lighting areas in each of the three rooms and two bathrooms, and lighting areas in both bedroom closets, as well as in the outdoor porch area. See Figure 5.5 for an image of the sample apartment layout. The other locations include apartments and standalone homes, most with 1–3 bedrooms, a kitchen, living room, and various closets and bathrooms.

Every location in this study has, at least, PowerBlade sensors and gateways. This results in 338 PowerBlades deployed and 16 gateways. The average number of PowerBlades in each residential location is 34; the maximum is 59 and the minimum is 22. In addition, six of the eight locations have BLEES and/or Ligeiro sensors to measure built-ins. This results in a total of 43 sensors: 7 BLEES and 36 Ligeiros. Although BLEES works better in situations where the source is rarely used, the Ligeiro have the benefit of never needing battery replacement. All eight locations have ground truth data provided by the utility company. The sample apartment has second-level power and energy data. Four more have daily energy data, and two have monthly energy data in the bill. For this last category we

align our deployment with the billing period at each location, and otherwise all data are treated in terms of daily average for the purposes of calculation and analysis.

In total, the PowerBlade table accumulated 3.51 billion rows, for BLEES 30.82 million, and for Ligeiro 6.65 million. Most locations have only one gateway, so each row is a unique measurement collected in real time.

For these deployments we utilize jumpers and surge strips rather than the shielding described in Section 4.3.3. This is due to difficulty in manufacturing the shielding required, but the jumper/surge strip provides adequate accuracy for deployments. For each PowerBlade device that otherwise would plug directly into an outlet, we deploy a jumper or surge strip alongside the PowerBlade unit.

5.3   Deployment System Evaluation

We evaluate this deployment system from the perspectives of measurement quality, wireless performance, and deployment cost. The first two, measurement quality and wireless performance, directly describe the usability of the system and mirror, in part, the two open questions from PowerBlade in Chapter 4. Deployment cost is an additional useful metric for assessing the viability of the deployment system, but will be reduced in the future as prototype quantities are increased.

5.3.1   Measurement Quality

In this section we determine the aggregate error of multiple PowerBlades, as well as any error in the Ligeiro/BLEES light sensor methodology. For this purpose we utilize the sample apartment that is equipped with a utility company-supplied unit that gives us access to second-resolution data. Overall measurement error is difficult to disambiguate from incomplete coverage, but we attempt to evaluate these separately by performing two steps.

First, we evaluate the accuracy of the system in a simple experiment by selecting five dynamic loads, and turning off or unplugging any other load that draws power. We then measure a baseline of the remaining static devices to subtract them from the ground truth. Finally, we manipulate the five selected devices and evaluate their sum against the adjusted ground truth. We select a cross section of devices that represent a range of both power and power factor (which may effect real power accuracy), as well as two overhead lights. The three plug-load devices are a floor lamp, laptop charger, and refrigerator.

At the beginning of the test, the fridge and floor lamp are powered, and the additional baseline is 206 W. We subtract this baseline from the remainder of this test. During the test
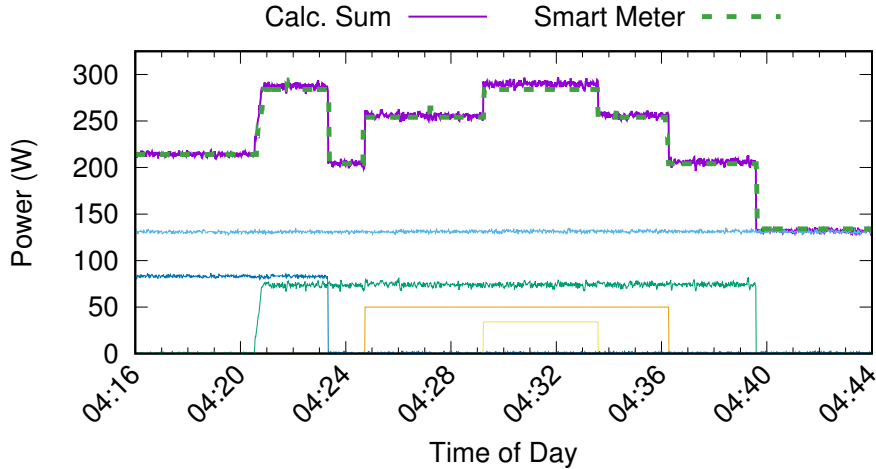
Figure 5.6: Simple aggregate accuracy test. We select five dynamic loads and subtract from the smart meter the baseline of the remaining static loads. We can then manipulate the five loads, selected across a range of average power and power factor, to determine any error introduced by each type of device. For this simple test we see that the aggregate is very accurate, with an average error of 3.44±0.25 W (1.51±0.1%).

the fridge finishes a cycle, the laptop is plugged and unplugged, and the overhead lights are each turned on and off. Figure 5.6 shows the results, including all five devices, their sum, and the baseline-adjusted smart meter value. The average error between measurements of our calculated sum and the smart meter is 3.44±0.25 W (1.51±0.1% error). This indicates that the methodology for aggregating Ligeiro/BLEES and PowerBlade measurements is accurate for a controlled set of devices.

Second, to test coverage and accuracy together we compare the sum of *all* of the Ligeiro, BLEES, and PowerBlade sensors in one location to the ground truth for a given hour during the deployment. During this time period we do not use the oven, HVAC, dishwasher, or clothes washer/dryer but otherwise do not have any explicit controls. Figure 5.7 shows the results of this experiment, including all the active loads, our computed sum, and the ground truth total. During this experiment the average error is 7.8±10 W (1.19±1.57%). Again this indicates that the system is accurate as a whole-home metering system.

Finally, to test accuracy over a longer period use we use a panel-level sensor to detect times when the oven, HVAC, dishwasher, and clothes washer/dryer are all inactive. We can take the sum of our measurements over these time periods as well as the sum of the energy bridge, and compare the results. We select 24 hour-long periods, during which time the smart meter reports a total of 10.607 kWh compared to our aggregated sum of 10.469 kWh (1.30% overall error). However, the average instantaneous second-by-second error is 27.8±0.27 W (7.28±0.08% in this sample).
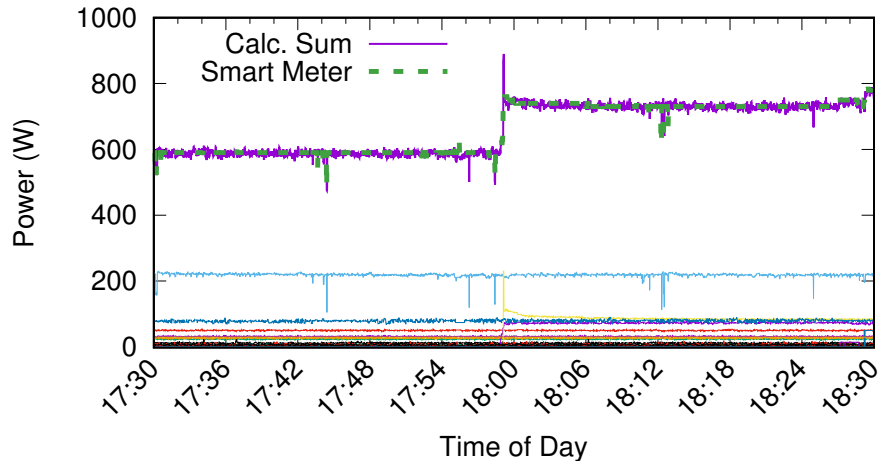
Figure 5.7: Full accuracy and coverage test. Power measurements and the reported value from the smart meter over an hour window are shown. While noise increases when aggregating over more PowerBlades, accuracy remains high (average error of 7.8±10 W, 1.19±1.57%).

Error in this application comes from multiple sources including PowerBlade error, which tends to overestimate on low-power loads and underestimate on very reactive loads, and Ligeiro/BLEES error, which occasionally overestimates due to false positives. Error is also mitigated by many sensors reporting quite accurate results. That the total overall agreement is significantly better than what might be expected from the instantaneous error indicates that these individual errors are approximately evenly distributed above and below the nominal value. Although not ideal, the combination of Ligeiro, BLEES, and PowerBlade measurements provide reasonably high accuracy measurements on plug-loads and lighting.

5.3.2 Wireless Performance

In this section we evaluate the wireless performance of the sensor network itself, in three steps. First, for every PowerBlade and BLEES device in the study we calculate the effective data resolution, given packet loss. Second, most locations require just one gateway, but in the sample location we install one per room. This allows us evaluate the performance of the system as a function of number of gateways. Finally, because Ligeiro has no expected packet rate it is difficult to calculate network performance, but we can qualitatively evaluate Ligeiro in the sample location as well.

5.3.2.1 Measurement Latency

To determine the effective resolution of our sensor system we look at PowerBlade, both because there is an expected rate and because each packet contains a sequence number that
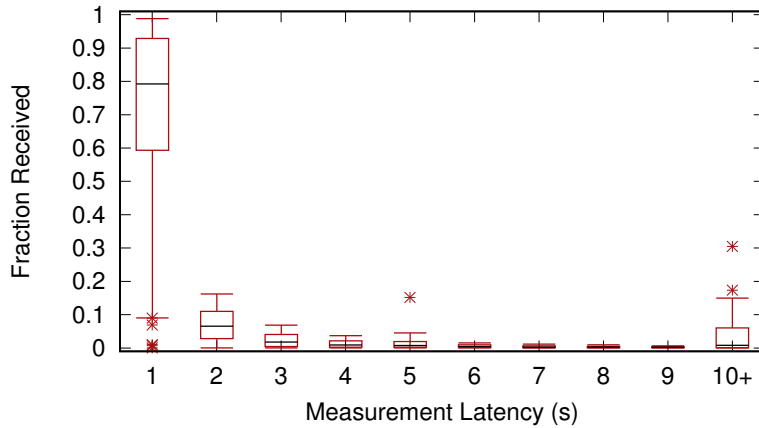
Figure 5.8: PowerBlade Measurement Latency. Across all the devices in the study, most measurements are received at the nominal rate of 1/s. The average latency is 1.16 s. Of the 338 PowerBlades deployed only 4 (1%) have an average latency of greater than 10 s.

allows us to calculate total loss. Figure 5.8 shows, for every PowerBlade at all ten locations in our study, the fraction of packets received with latency of 1-9 s, as well as 10+ s. The strong concentration of counts around 1 and 2 s shows that most measurements are received within 1 or 2 s of the previous for that sensor. The average measurement latency across all PowerBlades is 1.16 s. The worst-performing sensor has an average latency of 600 s, but 98.8% PowerBlades in the study have an average latency of less than 10 s. Although there are very few BLEES, a similar evaluation yields an average latency of 8.3 s with the difference being attributable to the redundancy in PowerBlade transmissions.

5.3.2.2 Gateway Coverage

Although at most locations a single well-placed gateway is in range of all deployed devices, at the sample location we deploy one gateway per room. This allows us to determine effective coverage requirements as well as the benefits of redundancy. For this evaluation we take the total fraction of measurements received for each device, cumulatively as well as on each of the three gateways. Figure 5.9 shows the results of this experiment, with the devices arranged on the x-axis in order of reception fraction. This shows that each of the three gateways receives about 60% of all measurements, but the addition of redundancy increases this to about 85%. Further, Gateway 2 is centrally located, with Gateways 1 and 3 in the bedrooms. Gateway 2 visibly performs better in the mid-range of devices, where most devices are centrally located, but Gateway 1 performs significantly better on the periphery devices (i.e. its minimum slope point is not as flat). This indicates a superior location, as incremental increases on the periphery devices are worth more than on central devices.
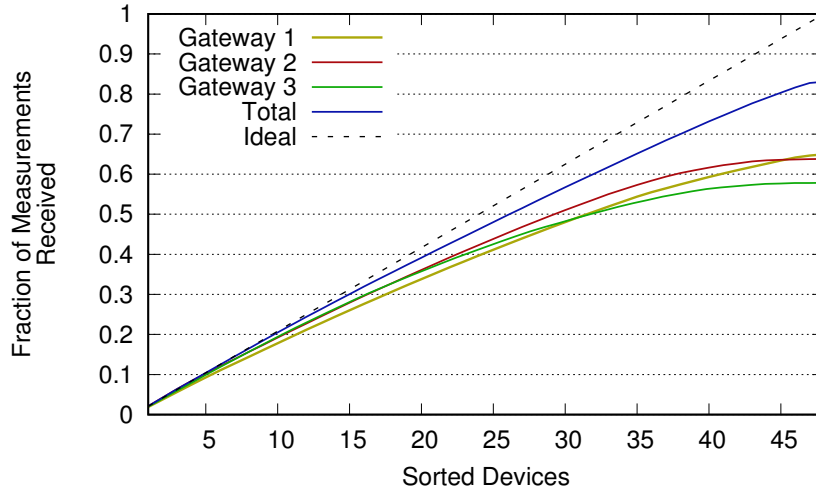
Figure 5.9: CDF of Measurement Reception. Shown are the CDF of measurement reception fraction for each gateway, as well as the combination of the three gateways and the ideal. Each gateway receives about 60% of packets, and the addition of redundancy increases this to about 85%.

Although the additional 20% of receptions due to redundancy is a positive result, we also want to determine a minimum number of gateways required. In this experiment both Gateways 1 and 2 are in range of all devices in the sample deployment, and either would be a good candidate for a single location. This agrees with the latency experiment, which is pairwise on (gateway, device) to ignore the effects of redundancy. In both cases we find that a single gateway has good coverage of almost every device in three rooms.

5.3.2.3 Ligeiro Performance

Although it is difficult to determine an expected packet rate for Ligeiro, we can evaluate how many Ligeiro packets are received by each gateway over the course of an average day. In the sample location we have Ligeiros in each of the three rooms as well as both closets and both bathrooms. All seven of the Ligeiros are in range of all three of the gateways, and for a given Ligeiro the receive counts on the three gateways are similar. This indicates that distance-related packet loss is minimal over these distances.

Although receive count between the gateways is similar, there are differences between Ligeiros based on expected usage level. The kitchen sensor is the most-activated device, at 8.6k±2.8k packets per day on a bedroom gateway to 14.7k±4.4k packets per day on the living room gateway. The dining room and stove light are approximately tied for second, with averages of 2.8k±2.2k and 2.8k±1.7k packets per day. The two bathrooms and master closet are generally tied for fourth, ranging from 242±209 to 313±236, and finally the

Component	Average Qty	Unit Cost	Total Cost
Gateways	1.25	$133.00	$166.25
PowerBlades	34	$26.00	$884.00
BLEES	1	$36.00	$36.00
Ligeiro	6	$68.00	$408.00
		Total	**$1494.25**
Jumpers	21	$1.79	$37.59
Surge strips	1.25	$4.52	$5.65
Total w/ jumpers & surge			**$1537.49**

Table 5.1: Average deployment cost. The average location has 1.25 gateways, 34 PowerBlades, 1 BLEES, and 6 Ligeiros. The total cost for this equipment at the current production quantities is $1494.25, and with the jumpers and surge strips $1537.49.

rarely-used other closet is 111 ± 76. This indicates that Ligeiro has sufficient range to cover a basic apartment, and this is backed up by the other Ligeiro sensors in the deployment.

5.3.3 Installation Costs

In this section we evaluate the cost of the deployment for the average location. For this we evaluate the eight residential locations for an indication of the average residential cost. All but one of these locations has one gateway, so we will use a conservative average of 1.25 gateways per location. There are an average of 34 PowerBlades, 1 BLEES, and 6 Ligeiros. Table 5.1 shows the cost breakdown for the average location. The average deployment cost is $1494.25, and including the jumpers and surge strips that we deploy alongside the total becomes $1537.49.

Although in most deployments we are able to identify at least one opportunity for significant energy savings, at almost every location the payback period on the average deployment cost of $1494.25 would be significant. However, the sensors are produced in very small quantities, increasing the quantity will drive down the cost. In addition, we can reduce the cost of PowerBlade by 15% by combining the processor and radio, which will also save space. We have identified an overseas manufacturer that will reduce the PowerBlade cost by an additional 18%. On BLEES we can eliminate the other environmental sensors besides the light sensor, reducing the cost by 49%. On the gateway we can remove the 802.15.4 radio. We estimate that with these modifications the average deployment cost will be reduced to $834.

5.3.4 Comparison of Residential and Commercial

Due to the relatively few commercial locations in our deployment it is difficult to draw conclusions about commercial energy statistics. In analyses in later sections we specify where data from the commercial locations are used, but the comparison between the deployment *experiences* of residential and commercial locations is also significant. Most notably, non-residential locations present a more difficult scheduling problem, as devices are often plugged in and running, and providing value. Capturing a time when all AC devices are down for maintenance, or are simply not in use, is difficult, and in practice requires multiple partial installations. By comparison, at each residential location the homeowner can turn off each device, install the PowerBlade, and turn it on without dependencies on other users.

We test our system at two non-residential locations: a working office that we instrument with full coverage (all plug loads and lighting) and a university datacenter that we instrument with partial coverage (only a few devices). At these locations the device breakdown is shifted toward computing. For the larger commercial location that we instrument, there are $6.3\times$ as many computers as the average residential location in our study as well as $9.3\times$ the number of computer monitors and $2.6\times$ the number of televisions. The average residential location has 3.8 lamps, whereas the commercial locations are all built-in lighting.

There are also other practical differences. Installation at all eight residential locations can be done without a ladder, but the light fixtures at one commercial location were sufficiently high that a ladder is required to install BLEES and Ligeiro. The residential locations also experience less device turnover with sensors being moved or simply lost, presumably because employees are not as informed about the purposes of the study as residential participants.

5.4 Summary

In this section we combine PowerBlade with other sensors, a gateway, and cloud storage to create an unobtrusive energy metering system. The PowerBlade sensor measures plug-load devices, two different light sensors monitor built-ins, and the gateways forwards data to the cloud database. The Monjolo sensor can also be used in the panel to measure high-power devices like the oven and HVAC. This deployment system is made up entirely of unobtrusive sensors, and can be deployed in residences and commercial spaces without affecting the aesthetics of the space.

To evaluate the performance of this system we deploy it in eight residential locations and two other locations. This allows us to evaluate the open questions in the PowerBlade design, as well as generate a dataset that can be used to motivate future deployments. We find that

the metering system is able to cover all plug-load and built-in devices, and that the aggregate measurement is an accurate representation of the total. In addition, although increased gateways results in slightly higher data reliability, most devices successfully transmit data at about 1 Hz.

In this following section we explore the possibilities of this dataset from an energy perspective, in particular certain applications that can be used to save energy, save time, or increase user comfort. In addition, we add an occupancy sensor to the deployment system, which enables additional analysis from new perspectives.

CHAPTER 6

Energy Analyses and Applications

In previous chapters we describe the components that can be combined to produce an unobtrusive energy metering system, and in this chapter we explore the possibilities of the high-coverage data that it provides. First, we examine the data directly, in the format that it might be presented in energy feedback to homeowners or reports to policy makers. Through the fine-grained dataset we find that MELs vary from household to household, but that their overall contribution, including the contribution of the long tail of energy devices, is significant. Next, we recognize that building energy is subject in part to the human workload, and will be affected by how occupants use the space. Taking occupancy as a marker for utility, how do occupancy and utility vary between locations? Could devices become more efficient using occupancy data? After that, we recognize that meaningful scale-up of this system requires automatic device identification to reduce the burden of naming during deployment. We evaluate ideas from literature to determine their quality in the PowerBlade dataset, and find that this process can be used to simplify installation without significantly affecting energy results. Finally, when equipment is degrading, or software processes get stuck utilizing significant CPU, the electrical fingerprint may encode a marker to this condition. We evaluate a few cases where this might be possible in residential and commercial environments. These four examples mirror a few of the applications from Chapter 2, and other applications and analytics have been evaluated in literature. We use the energy dataset collected in Chapter 5, which includes eight residential households and two other locations, an average deployment duration of 62 days per location, 338 PowerBlade sensors on plug-load devices, and 43 light sensors on built-ins (36 Ligeiro and 7 BLEES) [1].

---

[1]These experiments and analyses are approved by the Institutional Review Board (IRB) of the University of Michigan.
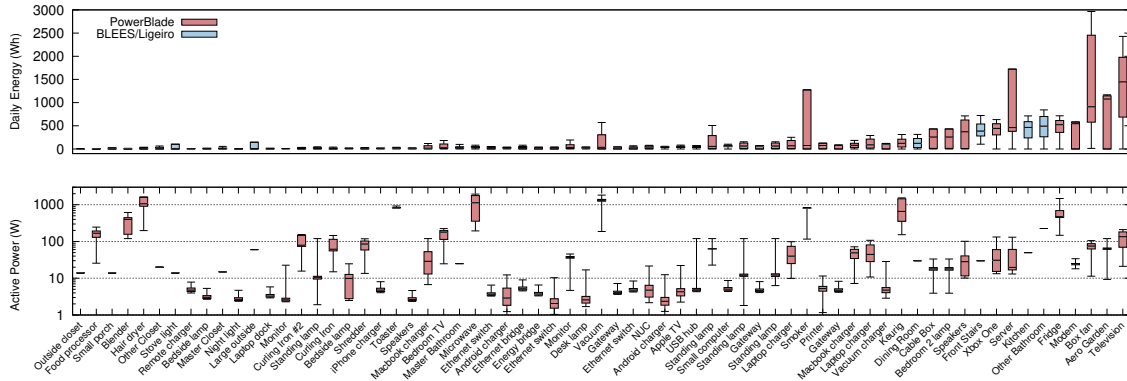
Figure 6.1: Device breakdown for one location. For each device we show the average daily energy and the average active power, as well as the minimum, first quartile, third quartile, and maximum. The highest-power devices are not among the largest total energy consumers, devices above 500 W account for only 2.96% of measured energy. Instead, the top three devices account for 35.8% despite an average power of only 91 W. Devices under 100 W account for 71.1% of measured energy and devices under 30 W account for 37.0%.

6.1 Energy Feedback and Reports

Organizing the devices by the total energy consumed allows the consumer or regulator to identify which devices have the largest contribution to the total energy budget, independent of their instantaneous power. A medium-power device that is left powered nearly constantly can draw significantly more energy than its low instantaneous power may indicate. A common example of this is a cable box, which typically draws about 30 W whether or not the device is 'on'. Additionally, a high powered device may draw significant power when in use, but it is used so rarely that it accounts for minimal energy. An example here is a microwave or vacuum. This system identifies energy spending that users, policymakers, or device manufactures do not understand, and bridge this understanding gap.

6.1.1 Cumulative Energy Breakdown

Figure 6.1 shows the energy and power breakdown for one sample location. The upper subplot shows the average daily energy consumption for each device, sorted by energy. The lower subplot shows the average active power. Notably, the highest power devices are not among the largest contributors to the total energy consumption. This breakdown can be used by the homeowner to understand the largest energy consumers. In this case, the six devices above 500 W combine for only 2.96% of the total measured energy. The largest contributors include some devices that are expected, like the television, refrigerator, and indoor garden, but others are surprising. The box fan draws a lot of energy due to a surprisingly high active
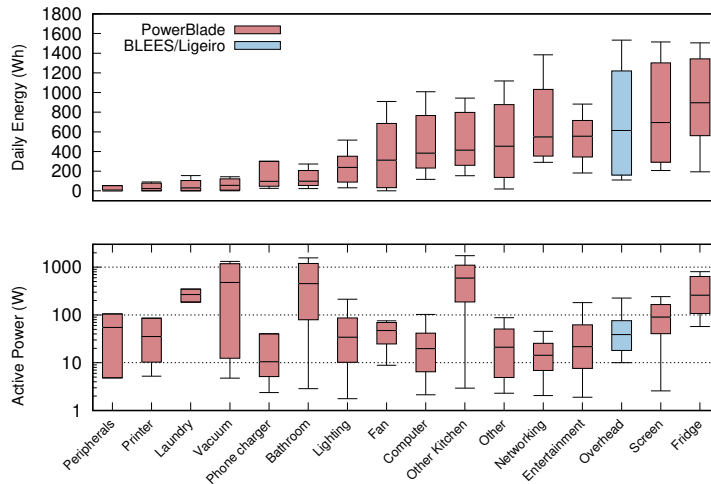
Figure 6.2: Category-based breakdown for all devices in the study. For each location and each category we take the sum of all devices, and average across all locations. Fridges are the single highest energy consumers, but the combination of overhead lighting and lamps are also significant. Screens, largely televisions, are the second largest consumer in our deployment and the single largest in the EIA results [34].

power. The modem draws more energy than expected, and there often exist opportunities to upgrade outdated equipment to higher-efficiency models. For each household in our study we provide a similar breakdown, and most result in similar findings to the sample.

At the deployment level, Figure 6.2 shows the category-based energy breakdown for all the devices across all eight residential locations. For each location we take the sum for each category, and average the locations together. This dataset can be used by regulators and policy makers, as well as device manufacturers, to direct optimization and efficiency efforts.

In the aggregate the single largest category in this dataset is fridges, each location has one and it is often a large consumer at that location. Televisions are the second highest, and the combination of "Lighting", which is PowerBlade-based lamps, and "Overhead", which is BLEES and Ligeiro lights, is also significant. Most locations in the deployment have not fully upgraded to high efficiency bulbs, and this result could be used to drive conservation efforts or to quantify the ROI of an upgrade. Although the fridge is considered a necessary expense, some fridges in the deployment cycle significantly more often than others. This can be used to target and motivate energy efficiency programs. Finally, "Entertainment" and "Networking" are driven by the cable boxes, routers, and modems that are common, and draw power even when not in use.

For some device types, this figure is a visualization of savings potential. Each category with a wide range represents possible savings - some households using significantly less

in that category than other households. Some locations spend over 1 kWh each day using Screen devices like televisions, while other locations use 0.2-0.3 kWh. Even the refrigerator category has an interquartile range of almost 0.8 kWh per day, likely due to a combination of old appliances and inefficient usage. This information could be used to target the largest consumers, and quantify the impact of a device upgrade or behavior change. In addition this helps policy makers determine which device types have the greatest possibility for improvement.

6.1.2 Cumulative Energy Distribution

In addition to the category-based breakdown, we can combine the measurement data with the ground truth aggregate energy use from the utility company. For this analysis we calculate an energy CDF, sorting every device from every location by active power, from low to high. For each unique power level we calculate the fraction of measured energy and fraction of total energy that is consumed by devices of that power level or lower. This aggregate data can be used by regulators and policy makers to target efficiency programs to the specific device types or classes that are consuming the most energy.

Figure 6.3 shows this representation of the deployment data. The intersection with the right hand axis shows that in this dataset, plug loads and lighting account for 50.8% of the total energy. Because two of the residential locations are not instrumented with full coverage (missing BLEES and Ligeiro), this is a conservative estimate of the actual figure. Further, 91.5% of this measured energy is devices under 250 W. This agrees with the sample apartment, that the highest power devices are not significant contributors to the total energy. Devices under 100 W account for 53.6% of the measured energy, 27.2% of the total. Devices under 30 W, which are typically the modems and cable boxes that are always on, account for 30.4% of the measured energy and 15.6% of the total.

This figure represents the three-way product of device quantity, power, and usage time. The significant weighting of certain power levels indicate that those devices are high in quantity, are used often, or draw high power. Some devices like televisions fit multiple of these categories. Understanding this weighting allows policymakers and device manufacturers to target efficiency programs where there is the most need. In this case, since 20-250 W devices are the biggest contributors, these devices should be targeted first. Microwaves, often the target of current efficiency and replacement programs, should almost be last.
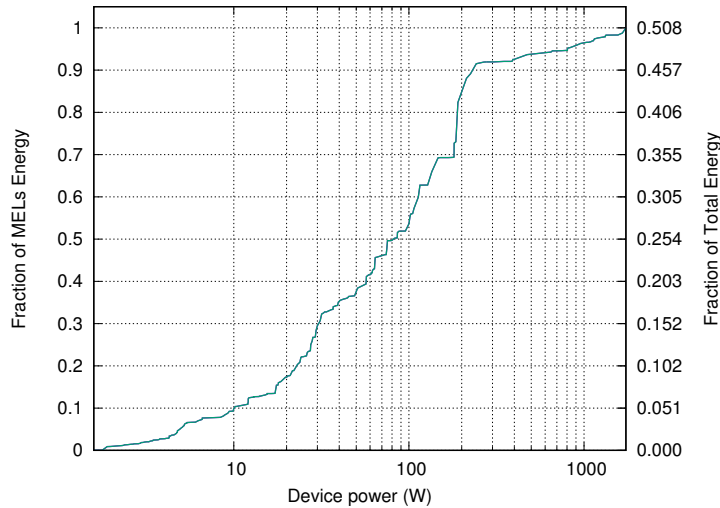
Figure 6.3: CDF of average daily energy and active power. All devices in the study are sorted by power, and energy for each level is summed with all lower power levels. This shows that MELs and lighting account for 50.8% of total energy, and devices under 250 W account for 91.5% of measured energy. This figure represents the three-way product of device quantity, power, and usage time.

6.1.3 Comparison with Prior Assumptions

The United States Department of Energy (DOE) has multiple sub-agencies that analyze energy across the country. In particular the Energy Information Administration (EIA) provides statistics and other data on energy in general, and the Building Technology Office (BTO) works with research technologies specific to buildings. The consensus among the EIA, BTO, and other DOE sub-agencies is that MELs account for 30%–35% of residential and commercial energy [29, 31, 35]. The EIA specifically lists residential/commercial at 29.67%/35.82%.

It is important to note that the user sample in the PowerBlade evaluation is not representative of the nation as a whole, but this provides a benchmark to compare our results. In our study the total contribution for MEL devices is 50.7%, somewhat higher than the national finding. There are several factors that likely contribute to the difference, most notably demographics. Smaller households or those of a lower socioeconomic status may have fewer non-MEL loads like clothes washers and dishwashers, and their contribution percentage may be higher. Larger households may have more non-MEL devices, but may also have more devices overall. These households may also have upgraded to ENERGY STAR® appliances, which increases the percentage of MELs.

Considering demographic differences, the PowerBlade system likely results in similar findings to those of the national agencies. This alone is a valuable comparison, as the
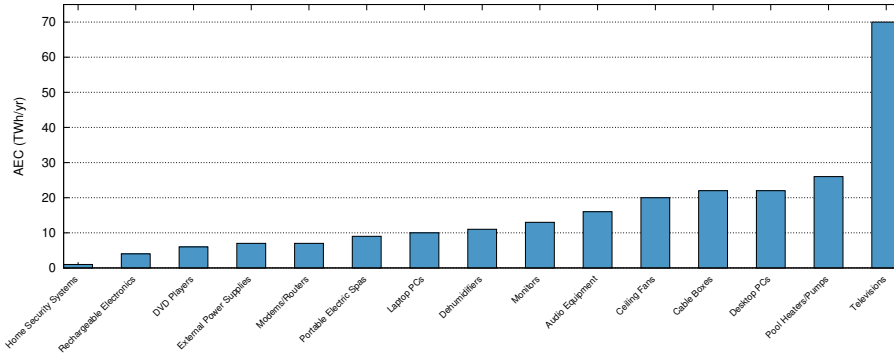
Figure 6.4: Device energy breakdown as reported by the U.S. Energy Information Administration (EIA) [34]. Pool heaters and ceiling fans are not included in our study, but otherwise the highest-consuming devices agree with our findings. We find that laptops draw more than desktops which is reverse in the EIA result, which indicates a more tech-savvy demographic in our sample. In addition, the EIA finds that routers and modems are not significant contributors to energy, which disagrees with our findings.

equipment and methodologies of the national research involve more expensive and more intrusive energy metering sensors. However, much of the national research is based on direct observation partially augmented by survey-based behavioral questionnaires to enhance the longitudinal relevance. For some devices, the possibility exists that this fails to capture some of the device behavior over time. This demonstrates the need for more studies that use a PowerBlade-like sensor, where direct activity can be studied for long periods of time.

At the level of the device breakdown there are also some disagreements between our results and the prior assumptions. Figure 6.4 shows the device breakdown as reported in the EIA *Analysis and Representation of Miscellaneous Electric Loads in NEMS* [34]. Similar results have been reported in private research [96, 109].

The largest energy consumer is televisions, which generally agrees with our findings. The fact that this category is higher than in our dataset is likely due to more televisions per household in the national average than in our deployment. In Section 2.1 the installed base is over two per household, and this is higher than in our experience. Fridges are not included in this EIA dataset, but are also significant household consumers [35]. Some of the other highest devices like pool heaters and ceiling fans are not included in our study, but others like cable boxes, computers, and audio equipment are consistent with our results as well.

Two things stand out. First, "Laptop PCs" is a low-energy category while "Desktop PCs" is a high-energy category, which is generally reversed from our findings. This indicates a slightly younger or more tech-savvy demographic in our sample than in the national average. Very few locations in our study even have a desktop PC in the traditional sense, although many have embedded smarthome-style devices. Additionally, "Modems/Routers" are a very

Figure 6.5: Blink sensor used for occupancy sensing. This passive infrared (PIR) sensor can be added to the existing energy metering system to correlate device activity with occupancy. The sensor gives a binary occupancy indication, and operates on a similar wireless protocol to PowerBlade and BLEES. Periodically Blink sends new occupancy data in a BLE advertisement packet.

low category, about 3% of the total. This disagrees somewhat with our finding, which is that these devices are always drawing power and are moderate contributors overall. This possible disagreement indicates the need for further evaluation.

6.1.4 Energy Reporting Summary

In this section we demonstrate the potential of the PowerBlade sensor system, combined with the calculation methodology described in Section 5.2.6.2, to generate basic energy reports. These reports distill significant volume and velocity of energy measurements to generate concise output that can be used by homeowners or policymakers to direct energy savings efforts. These reports are similar to, and in many cases modeled on, reports and studies conducted by the EIA and others that require more time and resources than this system. The PowerBlade-based system is simply set and forget. In the following sections we explore other applications that can be built on this energy data. Although not an exhaustive list, this demonstrates the potential of this high-resolution, high-coverage dataset.

6.2 Correlation of Device Activity with Occupancy

Although the energy and power data describe significant detail about each location, what is missing is information about the human interaction with AC devices. In the previous section we demonstrate the ability to identify energy wastage based on device activity that is outside

of the normal bounds for that category. Adding an understanding of when occupancy is and is not associated with device activity can be used to more clearly identify energy wastage at the level of each location. In addition, using occupancy data we can determine the direct potential of occupancy-based demand response applications.

We augment our deployment system to include a sensor for measuring room-level occupancy. In each room at five of the eight residential locations in the deployment, we add a passive infrared (PIR) sensor called Blink that emits a binary signal if the room is occupied (shown in Figure 6.5). It operates by the same principle as PowerBlade and BLEES, where it transmits a periodic packet via BLE. This connects directly with the existing data sources of the deployment, and can be analyzed alongside.

When working with Blink data we convert the binary occupancy information to continuous values by integrating over hour-long periods. We then normalize this signal to the maximum value, so the occupancy signal contains hour-resolution data in the range [0,1]. In addition, when correlating with occupancy we integrate energy for each device for each hour and normalize to the maximum. This results in occupancy and power signals with the same resolutions and dynamic ranges to allow for direct comparisons.

6.2.1   Cross-Correlation and Conditional Probability

We consider two metrics when analyzing the relationship between occupancy and device activity. The first is the simple cross-correlation of the normalized occupancy signal with each device's normalized power signal. A high cross-correlation between occupancy and power means that whenever the room is occupied the device is active, and the device is never active when the room is not occupied. However, a low cross-correlation can fail to capture operational dependence on occupancy for certain devices. For this reason we calculate a second metric, the conditional probability of occupancy based on device activity $P(o|a)$. Unlike the cross-correlation, this directly captures the likelihood that the room is occupied given that the device is active.

To illustrate the difference, consider Figure 6.6 showing power and occupancy for several devices. For the television and the speakers, the cross-correlation is very high (0.82 and 0.88, respectively). These devices are always turned on when the room is occupied and turned off when it is vacated. However, the cross-correlations for the game system and for the laptop charger are much lower (0.38 and 0.34, respectively). Although these devices are only used during occupancy, they are not guaranteed to be in use when the room is occupied. To reflect the fact that all four are used only during occupancy, we calculate $P(o|a)$, the probability that the room is occupied when the device is in use.

100
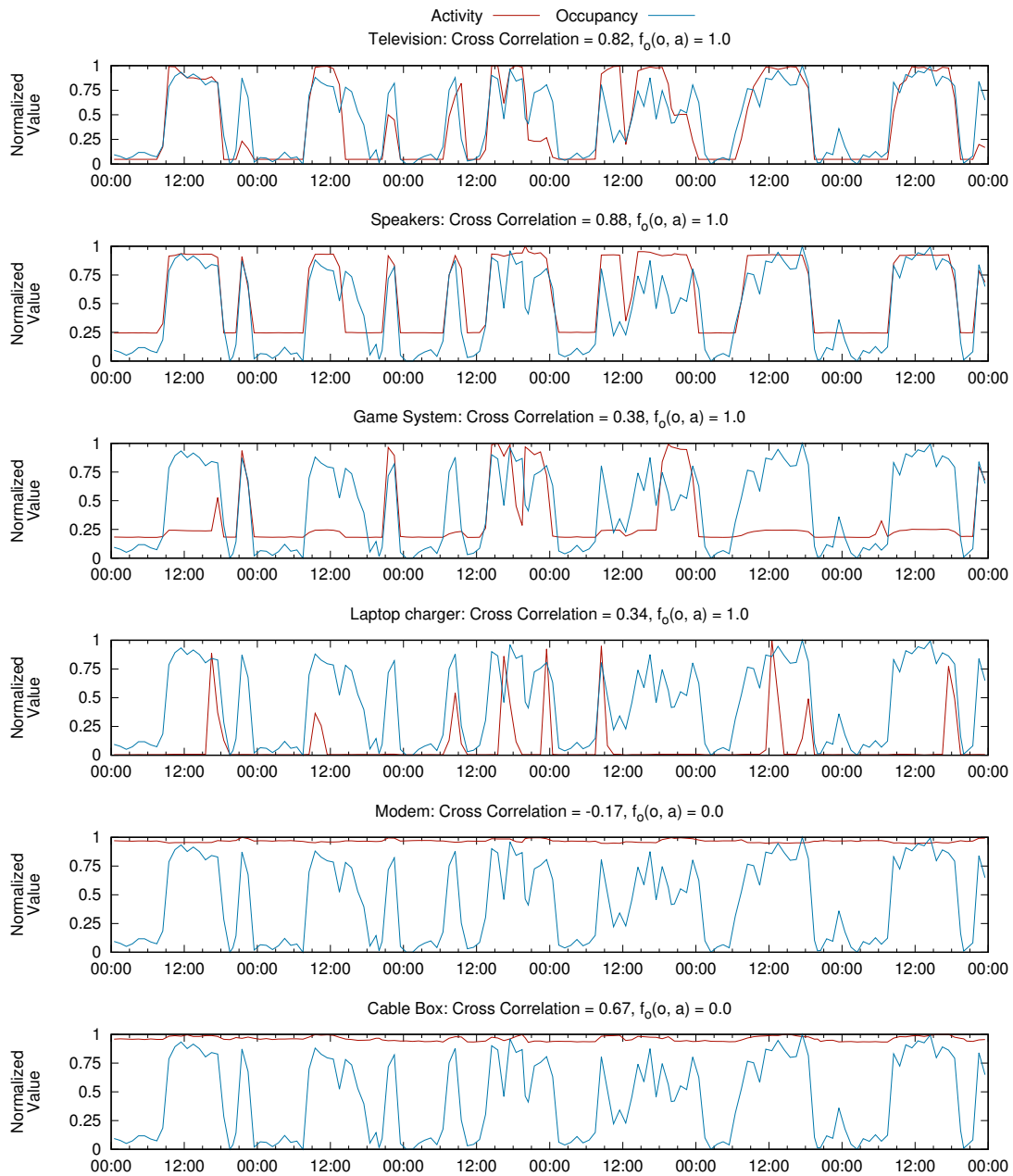
Figure 6.6: Demonstration of the difference between cross-correlation and $f_o(o,a)$, which is based on the conditional probability, $P(o|a)$, for several devices. The television and speakers have high cross-correlation, these devices are always turned on when the room is occupied and off when it is vacated. However, the game system and laptop charger are only used during occupancy, but have low cross-correlation. $f_o(o,a)$ captures the direct likelihood that the room is occupied during activity. The cable box and modem always draw power, so there is no indication positive or negative to occupancy.
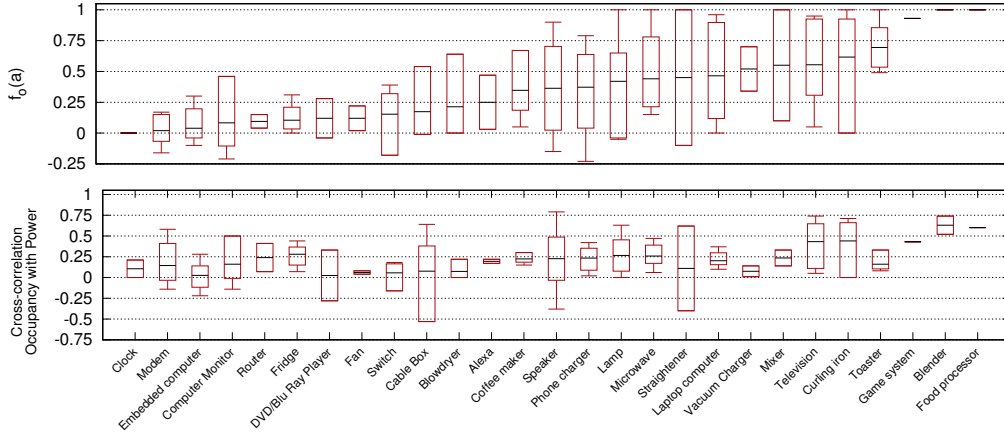
Figure 6.7: Category-based breakdown of $f_o(o,a)$ and cross-correlation for all the devices for which we have occupancy data (commercial included). The "Blender" and "Food Processor" categories have the highest conditional probability (one would not turn on a blender and leave the room). Some of the devices with a wide range represent savings potential. The "Lamp" and "Television" categories, for example, show some users turning off the device when the room is vacated and others not. This dataset could be used to quantify the savings potential for those that do not to change their behavior.

However, the probability of occupancy conditioned on activity, $P(o|a)$, is based in part on the probability of occupancy itself, $P(o)$. Further, if $P(a) = 1$ (as is the case for an always-on modem) the conditional probability $P(o|a)$ reduces to $P(o)$ alone. When comparing two rooms or two locations, frequently-occupied rooms will have artificially higher conditional probabilities which will make comparison difficult or impossible. To account for these differences, we define a new piecewise function, $f_o(P(o), P(o|a))$ which we refer to here as $f_o(o,a)$.

$$
f_o(o,a) = \begin{cases} \frac{P(o|a)}{P(o)} - 1 & \text{for} \quad 0 \leq P(o|a) \leq P(o) \\ \frac{P(o|a)}{1-P(o)} - \frac{P(o)}{1-P(o)} & \text{for} \ P(o) < P(o|a) \leq 1 \end{cases}
$$

$f_o(o,a)$ is a linear mapping of of $P(o|a)$ relative to $P(o)$, with the range of $[0, P(o)]$ mapping to [-1, 0] and the range of $[P(o), 1]$ to [0, 1]. If $P(o|a)$ is found to be 0 then the room is never occupied when the device is active and $f_o(o,a)$ gets the value -1. If $P(o|a)$ is found to be equal to $P(o)$ then device activity gives no indication positive or negative to occupancy, and $f_o(o,a)$ gets the value 0. Finally if $P(o|a)$ is found to be 1 then the room is always occupied when the device is active and $f_o(o,a)$ gets the value 1. Values between these endpoints are mapped to the range [-1,1]. Through this technique the conditional probability of each device can be compared without considering $P(o)$ for each room.

Figure 6.7 shows the category-based breakdowns of both conditional probability and cross-correlation, sorted by conditional probability. The devices with the highest conditional probability are food processors and blenders, which make sense. One would never leave a blender running and leave the room. Game systems and toasters are also high, as these types of devices are closely associated with use. Microwaves are lower than might be expected, and blow dryers are significantly lower. This might actually indicate a high idle power draw – i.e. the device is counted as active when it it only consuming background energy.

Several devices in Figure 6.7 have wide ranges of conditional probability. Lamps range from 0 to 1, which means that some users turn off the lights when they leave the room and some do not. Televisions are another example of some users turning off the device when they leave the room and others not. Other devices with wide ranges, in particular the straightener and curling iron, might have a high idle power even when the device is 'off' or may be left to heat up before use. Each of these wide ranges represents savings potential. If some users can turn off the device when not in the room, it is likely that others could as well, given motivation. Particularly if the system can directly quantify savings potential, and provide feedback over time to reinforce behavior, there is opportunity for savings.

6.2.2   Potential Improvements using an Automatic System

In addition to using occupancy to identify wastage, we can quantify the savings potential of direct load control applications. Examples of this type of program exist: lights turn off automatically when a room is vacated [97]. This is an important part of a total system, but the depth and resolution of PowerBlade data enable richer interaction with load control. In particular we examine two specific variations: targeting device *active* power during momentary absences and targeting device *idle* power during prolonged absences. Both involve using occupancy data alongside power data to change a device to a low or zero power state when not in use. Although this requires additional hardware and/or manufacturer buy-in to realize, we can quantify the savings potential using our dataset.

6.2.2.1   Active Power Savings Potential

Several categories in Figure 6.7 have high $f_o(o, a)$ across all devices in the category. Activity in a microwave, television, stove, toaster, game system, blender, or food processor all give high indications of occupancy (which also means that $P(\neg o|a)$ is low). These categories also pass a common sense check of devices that are typically only used when the user is in the same room. An automated system can detect the times when these devices are active and the room is *not* occupied, and change the device state.

103

As an example, the brightness on a television can be decreased, or the screen turned off, if the user temporarily leaves the room. One television in the study draws 40% less power simply by turning down the contrast, and the state can be instantly restored when the user leaves the room. This is the largest single category in the EIA dataset, and this technique alone could result in large savings nationwide. As another example, a stove could require an override to remain on if a user leaves the room, a safety feature similar to the tether on a treadmill. This system can be automated – with enough data certain high-probability categories like televisions can be automatically identified.

6.2.2.2 Idle Power Savings Potential

In Figure 6.6 we see that two devices, the game system and the speakers, have high conditional probability but do not return to zero power when occupancy is false. These devices are distinct from the cable box and modem, which also have a permanent/idle power draw, because the game system and speaker have strong dependence on occupancy. This dependence indicates that the device may provide less or no utility when the room is not occupied. This is an opportunity to power gate the device when not in use.

We calculate both the idle power (calculated as the average power when occupancy is false) and the percentage of time that the device is active, for all devices with a high conditional probability. For example the game system in Figure 6.6 has an idle power of 12.77 W and is actively in use 24% of the time. Power gating this device when occupancy is false has the potential to save 85 kWh per year.

Across all devices in the deployment where occupancy data is available, we calculate the idle power savings potential for devices with a high $f_o(o,a)$. We define two thresholds for $f_o(o,a)$, 0.5 and 0.75, for varying degrees of dependence on occupancy. For the $f_o(o,a) = 0.5$ threshold, the average savings is 164.55 kWh, approximately \$20, per location per year, and the maximum is 503.61 kWh, approximately \$60, per year. For the $f_o(o,a) = 0.75$ threshold, which captures slightly fewer devices to control, the average is 124.26 kWh, approximately \$15, and the maximum is 327.72 kWh, approximately \$40, per location per year. These savings are enabled by the addition of an simple occupancy sensor and load control, and can be combined with other applications and savings to further improve the ROI.

6.2.3 Repurposing the Conditional Probability

Finally, rather than enable the above savings applications, the high conditional probability of certain device categories can be used to replace the occupancy sensor itself. All device categories with a high $P(o|a)$ can be used as approximate occupancy sensors, and multiple

devices from the same room correlated to produce a more accurate result. This function is generally mutually exclusive with the average and idle power savings applications above, which require both an energy sensor and an occupancy sensor active at the same time. However, it is possible that all other devices in a room could be used together to produce an occupancy estimate, which could still be used to control the device in question. We do not evaluate this here, but it is an opportunity for future research.

6.2.4 Occupancy Summary

Augmenting the energy metering system with an occupancy sensor enables applications built on top of basic energy data. Analysis of the relationship between occupancy and device activity can be used to identify certain households that are using devices in a less efficient way than others, and provide usage-customized feedback of savings potential to those individuals. At the system level, we can also identify certain categories of device that are frequently tied to occupancy, and make automated decisions based on this information. Although this more advanced level of automation would require manufacturer buy-in, the PowerBlade metering system is able to demonstrate the viability, and quantify the potential, of certain improvements.

6.3 Automating Device Naming to Simplify Installation

Installing the sensor system in a residence typically takes about one hour, but for some locations this can be higher. The time is dominated by the manual process of recording the device type (fridge, television, etc) for each device and then uploading that information to a table in the cloud database. Not only is this a significant time overhead in the deployment process, it introduces the possibility for error. The human installer can read the MAC address incorrectly on the PowerBlade, or can mix up the tangled cords behind a television or entertainment center. A PowerBlade device can also be moved, either inadvertently or because the device is replaced. These issues are mitigated if the system recognizes each device that a PowerBlade is measuring, and eliminate, or at least reduce, manual labeling.

In automatic device identification the system itself analyzes the device and classifies it to a category or even to a particular make and model. Multiple systems exist for doing this, but these systems are geared toward NILM/disambiguation approaches. This typically means using high-precision sensors, either at the panel level and classifying disambiguated data or at the plug level to simulate disambiguated data. To determine the utility of the PowerBlade sensor for device identification, we compare our output to the results in *Non-Intrusive Load*

Identification for Smart Outlets by Barker et. al [45]. This approach uses 1 Hz outlet-level data similar to that of PowerBlade, and provides a straightforward platform for evaluating the PowerBlade system. Other systems exist that we do not evaluate against, the primary purpose of this evaluation is to determine the efficacy of our data compared to datasets collected using more expensive and obtrusive sensors.

Barker et. al. previously demonstrate the ability to classify devices based on the SMART dataset [44]. They utilize three steps in the classification process. In the preprocessing step they convert raw power data into a series of deltas between consecutive measurements. Consecutive deltas of the same direction are combined. Calculating deltas allows the classifier to operate on individual components of the energy signature (i.e. the 100 W fridge compressor and 30 W fridge light are treated separately rather than as a single 130 W measurement).

In the feature extraction step they calculate a data vector for each time period and for each device. Each vector contains statistical metrics like mean, variance, and device duty cycle, as well as 30 additional features describing the characteristics of the waveform. This step takes raw delta power data and converts to a single vector that can be used for classification.

Finally in the classification step they use three built-in classifiers as part of the Weka toolkit [79]: the Naive Bayes classifier, the J48 implementation of the C4.5 decision tree classifier, and the libSVM implementation of a support vector machine (SVM) classifier. They classify "seen" devices, splitting all data into a training set and a testing set, and "unseen" devices, which involves isolating entire households from the training set. Of these, the J48 and SVM implementations perform similarly for both seen and unseen devices. The Naive Bayes classifier performs worse for seen devices, but is able to generalize better and performs better on unseen devices.

We implement a similar system design to Barker et. al. to compare our results, but there are certain differences. First, the SMART dataset contains some devices like dishwashers and air conditioner units that PowerBlade does not have access to. Instead, PowerBlade measures the long tail of devices like phone chargers that are not typically considered in NILM. The lack of overlap demonstrates the need for concurrent systems - a PowerBlade system for plug-loads working alongside a NILM system for high-power loads - but it also means we evaluate on a different set of devices. In addition, the concept of seen devices is useful in disaggregation and NILM, wherein individual devices are continuously differentiated from the others, but less useful in PowerBlade where all devices are already disambiguated. Unseen devices, on the other hand, can be used to classify PowerBlade targets in new households, eliminating the need for manual labeling. For this reason we focus on unseen devices. Finally, we evaluate how adding occupancy data or combining multiple days improve the classification result.

6.3.1 A Machine Learning Prototype to Classify Deployment Data

We implement the same three steps as Barker et. al: preprocessing, feature extraction, and classification, but we perform a slightly different analysis for unseen devices. For unseen devices Barker et. al randomly select one entire household to call the test set. We find significant variation in locations both in which devices are present and in the characteristics of those devices, so the selection of a random location for testing has consequences on accuracy. Instead, we take each device in turn and generate a training model without that device, classify that device using the model, and sum all the classification results together.

We also evaluate two sizes of dataset. Barker et. al evaluate a dataset with 10 device categories, and to compare as closely as possible (although with different categories) we also evaluate a 10-label dataset. For this we select the 10 most common categories by device count: lamps, phone chargers, laptops, televisions, microwaves, refrigerators, routers or modems, toasters, cable boxes, and coffee makers. However this only covers 40% of the devices in our deployment. We also use a larger dataset that includes five additional categories: hair dryers, curling irons, blenders, fans, and exterior lighting. These are all the categories for which we have more than a few devices, and cover 55% of the devices.

Because the SVM performs similarly to J48, and due to the performance requirements of a new SVM model for each device, we evaluate only J48 and Naive Bayes. This allows us to test the PowerBlade dataset against the SMART dataset for device classification.
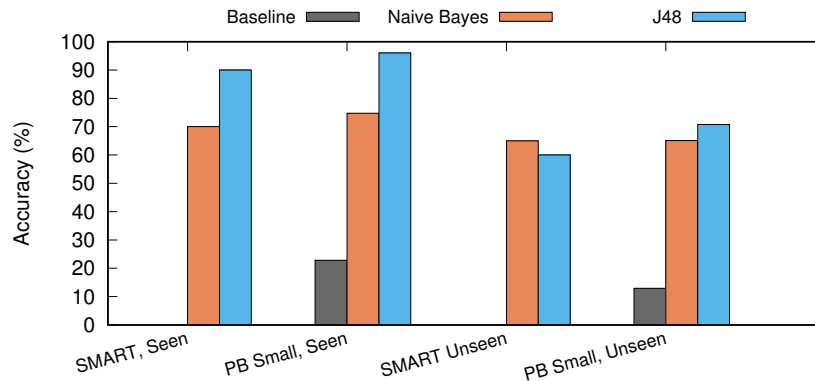
Figure 6.8: Basic comparison of results using the SMART dataset vs. the small PowerBlade dataset. Zero Rule baseline accuracy for PowerBlade is also shown. In the small PowerBlade dataset we use the same number of device categories (labels) as in the SMART dataset (10). Although the categories are different, this gives us a basic framework to evaluate PowerBlade data for device identification. The PowerBlade dataset is comparable to the SMART dataset for seen devices, and performs somewhat better for unseen devices. The difference is minimal, and may be attributable to the category selections. This shows that PowerBlade data are comparable to existing datasets for device identification.

107

Figure 6.8 shows the performance of the SMART dataset vs. the small PowerBlade dataset using Naive Bayes and J48 classifiers. Also shown are the baseline accuracies for the small PowerBlade dataset calculated using the Zero Rule (baseline accuracy for the SMART dataset is not published). The baseline accuracy is low due to the large number of classes and relatively even distribution between the classes. Of the two classifiers, performance is similar between SMART and PowerBlade for seen devices, and somewhat better on the PowerBlade dataset for unseen devices. The difference is likely attributable to category selection, but demonstrates that the PowerBlade output is comparable to existing datasets for device classification.

We also compare the performance of the small PowerBlade dataset vs. the large dataset. Figure 6.9 shows the change in performance that results from adding more device categories. Again, the baseline accuracies are low due to relatively even distribution among the classes. For the classifiers, on seen devices there is minimal change with the addition of more devices. On unseen devices there is a larger decrease, 17.8% for the Naive Bayes classifier and 10.2% for the J48 classifier. This result makes sense based on how much higher the seen accuracy is vs. unseen overall. Once a device is part of the training set the system can very accurately tell that device from others. Unseen devices require generalization which is more difficult, and the larger 15-label dataset requires even more extensive generalization than the smaller dataset. The accuracy of the J48 classifier on the large unseen dataset is 63.54%, which is a similar result to the SMART dataset results.
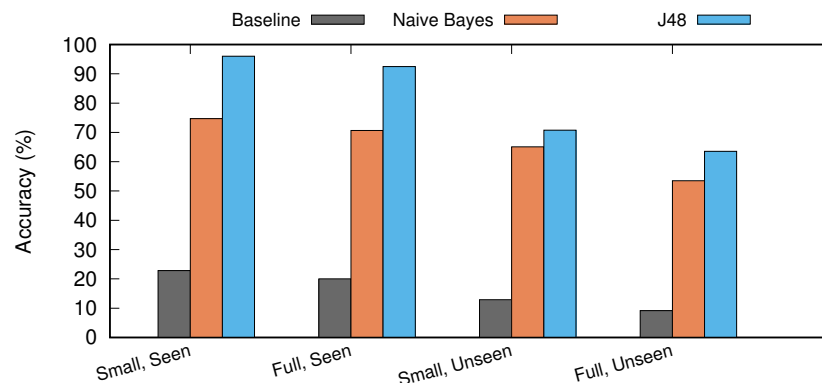
Figure 6.9: Comparison of classification results using a small 10-category dataset vs. a larger 15-category dataset. The larger dataset represents about 55% of the total devices in the deployment, and the small dataset covers about 40% of devices. For 'seen' devices there is only slight performance decrease on both classifiers from the small to the large dataset. For unseen devices the difference is slightly larger, 17.8% decrease for the Naive Bayes classifier and 10.2% for the J48 classifier. On the full 15-label dataset the J48 classification success is 63.54%. This technique is viable for a PowerBlade deployment system.

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	<- Classified as
88	0	11	9	0	1	0	0	0	0	0	0	0	0	0	a = Coffee maker - 0.81 correct
0	43	0	0	20	0	0	15	4	4	24	0	0	0	0	b = Television - 0.39 correct
7	0	36	14	0	1	0	0	0	0	0	0	0	0	0	c = Toaster - 0.62 correct
1	0	1	24	0	7	0	0	0	0	0	0	0	0	2	d = Blowdryer - 0.69 correct
0	0	0	0	119	7	0	4	0	0	0	0	0	0	0	e = Fridge - 0.92 correct
1	0	2	0	0	95	0	0	0	0	0	0	0	0	0	f = Microwave - 0.97 correct
0	0	0	0	0	0	90	0	17	0	0	0	0	0	0	g = Phone charger - 0.84 correct
0	21	0	0	0	0	2	31	4	2	0	0	1	0	0	h = Curling iron - 0.51 correct
0	0	0	0	0	0	20	4	77	12	9	1	8	7	0	i = Lamp - 0.56 correct
0	10	0	0	0	0	0	0	40	38	16	5	0	0	0	j = Fan - 0.35 correct
0	7	0	0	0	0	0	1	18	2	85	0	0	0	0	k = Laptop computer - 0.75 correct
0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	l = Exterior lighting - 1.0 correct
0	0	0	0	0	0	0	0	22	0	0	0	85	26	0	m = Router/Modem - 0.64 correct
0	0	0	0	0	0	0	0	0	0	0	0	84	0	0	n = Cable Box - 0.0 correct
3	0	1	1	1	9	0	1	0	0	0	0	0	0	3	o = Blender - 0.16 correct

Table 6.1: Confusion matrix of the J48 classifier on the large PowerBlade dataset. Cable boxes are often misclassified as routers or modems due to very similar activity characteristics. Also televisions are misclassified as fridges and laptop computers due to a wide range of television type and size in the deployment. Also these three types vary similarly over time, showing a spike at first use and a decrease over time. The total accuracy in this dataset on the J48 classifier is 63.54%.

Some devices are classified with higher accuracy than others. Table 6.1 shows the confusion matrix for the J48 classifier on the large 15-label dataset. The devices with the highest error are a cable box misclassified as a router/modem. Televisions are also often misclassified, in part due to a wide range of sizes and types in the deployment. Several devices are also misclassified as lamps, also because of wide variation in this category. Fridges, coffee makers, and microwaves all present very unique feature vectors, and are classified with very high success. The overall accuracy of about 65% represents this range of poorly classified device types to near-perfectly classified device types.

6.3.2 Improvements using Occupancy

In the basic identification system data are based only on the direct energy measurements, but we can also use data from the occupancy sensor. For this evaluation we calculate the classification vector as before, but we also calculate the occupancy cross-correlation and conditional probability on a day-by-day basis and add this to the classification data. This allows us to determine the utility of classifying using the relationship between devices and occupancy as well as just on energy data.

Figure 6.10 shows the result of this experiment for both datasets, with the baseline accuracies also shown. Despite the distinct correlation with some devices on occupancy, in both cases the addition of occupancy on seen and unseen devices adds minimal difference. Further, adding occupancy makes the system more likely to classify a "Router/Modem"
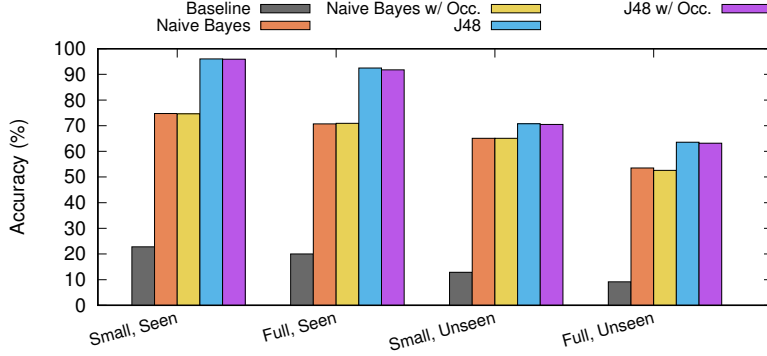
Figure 6.10: The addition of occupancy to the device classification system. Despite the distinct correlation with some devices on occupancy, on both datasets the addition of occupancy does not affect classifier performance. This is because the devices commonly misclassified, cable boxes with routers and televisions with laptops, present similar occupancy characteristics.

device or a "Cable Box" device to the "Lamp" category. This is because some routers and cable boxes in our study have correlation with occupancy but most do not, and the system fails on those outlier devices. Overall, occupancy data do not improve the system where it is really needed: televisions and laptops have similar occupancy characteristics as do routers and cable boxes. With more data collected as part of a larger deployment there may be more distinct relationships on occupancy.

6.3.3 Bayesian Inference Using Classification Results

In terms of probability, classification accuracy for each type t is the probability that the classifier will return the classification c_t when classifying the device D_t (i.e. $P(c_t|D_t)$). However, in a real-world system the actual label is unknown. Using Bayes Rule, we can calculate the posterior probability $P(D_t|c_t)$ from classification results.

For each t in the set of all device types $\mathbb{T}$, we define two probabilities. $P(D_t)$ is the *prior probability* for each D_t, the percentage of the test set of that type. $P(c_t)$ is the *marginal likelihood* for each c_t, the percentage of the result set (the types output by the classifier) of that type. Each cell in the confidence matrix gives us $P(c_{t_a}|D_{t_b})$ for all types t_a and t_b in $\mathbb{T}$. These values comprise Bayes Rule for inferring the posterior probability for each label.

$$\forall t \in \mathbb{T}, \quad P(D_t|c_t) = \frac{P(c_t|D_t) * P(D_t)}{P(c_t)} \tag{6.1}$$

Equation 6.1 shows Bayes Rule applied to device classification results. The posterior probability for any device type given any classification result is the weighting of that device,

Device Type t	Accuracy $P(c_t\|D_t)$	Confidence $P(D_t\|c_t)$	Device Type D_t	Accuracy $P(c_t\|D_t)$	Confidence $P(D_t\|c_t)$
Coffee maker	0.81	0.88	Television	0.39	0.53
Blowdryer	0.69	0.5	Toaster	0.62	0.71
Fridge	0.92	0.85	Microwave	0.97	0.79
Phone charger	0.84	0.80	Curling iron	0.51	0.55
Lamp	0.56	0.42	Fan	0.35	0.66
Laptop computer	0.75	0.63	Exterior lighting	1.0	0.87
Router/Modem	0.64	0.48	Cable Box	0.0	0.0
Blender	0.16	0.6			

Table 6.2: Accuracy and confidence for any single classification result c_t. The accuracy is simply the diagonal values from the confidence matrix, and gives an indication of the likelihood of this classification result when the system is observing this type. The confidence is the posterior probability for each classification, the likelihood that the device is actually that classification. Televisions and Blenders have higher confidence than accuracy because other devices are rarely misclassified to these labels. Similarly, an output of Lamp or Router/Modem has a higher likelihood to be an error because devices are often misclassified to these labels.

$P(D_t)$, multiplied by the ratio of the likelihood of that result given that device, $P(c_t\|D_t)$, divided by the likelihood of that result overall, $P(c_t)$. Using Equation 6.1 and the results in the confusion matrix shown in Table 6.1, we calculate the posterior probability, effectively the confidence of any one classification result, for each result value. For this and for the remainder of this section we use the larger PowerBlade dataset with 15 labels.

Table 6.2 shows the success rate in the confusion matrix for each type, $P(c_t\|D_t)$, as well as the posterior probability for that type, $P(D_t\|c_t)$. The television, blender, and fan all have much higher posterior probabilities than success rates. Classifications to each are rare, but misclassifications to these are *more* rare, so this is more credible than certain others. The lamp, router, and hair dryer all have lower posterior probabilities because there is a label commonly misclassified as each. The cable box is never output by the classifier (always misclassified as a router), so both success and posterior probability are zero. The system has high confidence in some classification outputs like microwaves, fridges, and coffee makers, but less than 50% confidence in others like lamps and routers. In most cases the system has reduced to two labels – only two labels have nonzero confidence. The user could pick between the two, but the system can be also improved by examining multiple days.

Up to this point, the device identification system operates on daily measurements independent of prior measurements on that device. We do this because PowerBlades can be moved, and the original label no longer applies. Without any system-level record of device movement, the classifier must re-classify every device on every day. However, the

classifier has high success on seen devices, meaning it can characterize a device once the device has been directly measured. This means the system is likely able to detect if the device is replaced with another device. Knowing that a device has not been moved allows us to combine multiple classifications on that device. C_t is now not a single classification but a distribution of individual classifications $C_t = c_1, ..., c_n$, but otherwise Equation 6.1 is applicable in the multi-day situation as well.

6.3.3.1 Detecting Device Movement

We determine the viability of this approach by attempting to use PowerBlade to detect device movement. For each device in our deployment we determine if the system could detect if that device was replaced with each other device. For N devices, this results in $N * (N - 1)$ classification problems. On our dataset the average classification success on this problem is 99.18% for a Naive Bayes classifier, and the 5th percentile is 94.73%, indicating high classification success across the full set of devices. For a J48 decision tree classifier the average success is 96.55% and the 5th percentile is 88.46%. Even if the switch is to a device of the same type, the 5th percentile successes for the Naive Bayes and J48 classifiers are 89.95% and 87.33%, respectively. The PowerBlade system can very accurately detect device movement.

6.3.3.2 Device-Level Automatic Energy Breakdown

Detecting device movement allows us to combine classifications from multiple days, and classify on a device level rather than a day-by-day level. Table 6.3 shows the device-level confusion matrix, the result of combining multiple days together. The most striking change is for the Cable Box. In this dataset the system never classifies a Cable Box correctly, but after 30 days every Cable Box has been identified correctly. Similarly, it performs poorly for Blenders on any single day, but identifies 66% of Blenders after 30 days. There are other improvements for Televisions, Curling Irons, and Fans as these devices are misclassified in expected ways.

In most cases the accuracy over time either improves moderately or stays consistent, and after 30 days 73.15% of devices are correctly classified (approximately a 15% improvement from the individual-day results). This value of 73.15% classification accuracy is of the same order of magnitude as the 75% benchmark for classification accuracy described in *Area of Interest 3.7* in the 2017 Department of Energy *Buildings Energy Efficiency Frontiers & Innovation Technologies (BENEFIT)* funding opportunity announcement (FOA) [36]. In addition, although a larger dataset may result in better performance, we can assess the impact

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	<- Classified as
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	a = Coffee maker - 0.75 correct
0	4	0	0	1	0	0	1	0	0	1	0	0	0	0	b = Television - 0.57 correct
1	0	4	0	0	1	0	0	0	0	0	0	0	0	0	c = Blowdryer - 0.67 correct
0	0	1	5	0	0	0	0	0	0	0	0	0	0	0	d = Toaster - 0.83 correct
0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	e = Fridge - 0.83 correct
0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	f = Microwave - 1.00 correct
0	0	0	0	0	0	8	0	2	0	0	0	0	0	0	g = Phone charger - 0.80 correct
0	0	0	0	0	0	0	4	1	0	0	0	0	0	0	h = Curling iron - 0.80 correct
0	0	0	0	0	0	3	1	13	4	1	0	2	0	0	i = Lamp - 0.54 correct
0	0	0	0	0	0	0	0	0	3	1	0	0	0	0	j = Fan - 0.75 correct
0	0	0	0	0	0	0	0	1	0	10	0	0	0	0	k = Laptop computer - 0.91 correct
0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	l = Exterior lighting - 1.00 correct
0	0	0	0	0	0	0	0	0	0	0	0	5	5	0	m = Router/Modem - 0.50 correct
0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	n = Cable Box - 1.00 correct
0	0	0	0	0	0	0	1	0	0	0	0	0	0	2	o = Blender - 0.67 correct

Table 6.3: Confidence matrix at the device level after 30 days. Because the system can accurately detect device movement, we use Bayes Rule to combine multiple days together to increase accuracy and confidence in the result. The cable box in particular sees an improvement, but there are also improvements in televisions and fans as these devices are typically misclassified in characteristic ways. The stark improvement for some devices is likely due in part to the relatively low number of devices, but even with a larger dataset this technique can be used to increase classification accuracy and confidence.

of this level of accuracy by re-computing the energy breakdown. Table 6.4 shows how the category-based energy breakdown shown in Figure 6.2 changes when using the classified labels rather than the actual labels. Because some of the 15 labels in the larger classification dataset combine into the same category, and also because we do not classify some categories like "Overhead" and "Other", classification changes the energy for 10 categories.

In Table 6.4, "Screen" and "Networking" are both shifted downward from the correct values, and other categories are shifted up in response. The device-level confusion matrix in Table 6.3 shows that almost half of the televisions ("Screen" category) are misclassified, as are many of the routers ("Networking"). This results in the downward shift of those categories, and in particular "Entertainment" is increased as routers are often misclassified as cable boxes. With the exception of "Screen" and "Networking" most other categories are in the correct order, only "Computer" and "Other Kitchen" are reversed. If users are able to label their own televisions and routers, typically 2-3 devices per household, the classified energy breakdown is an accurate approximation of the true values.

6.3.4 Dataset Limitations and Other Devices

The size of the dataset is a limiting factor for this application. The small size of the current dataset makes it brittle for device classification. With only a few of each device type, a certain household using the television in a unique way, or a having a router that is more

Category	Actual Daily Energy (Wh)	Energy Change (%)	Rank Change
Fridge	896	-10.4%	–
Entertainment	556	+35.6%	+1
Other Kitchen	413	+5.8%	+2
Computer	384	+9.9%	+2
Fan	312	+6.4%	+2
Screen	695	-59.2%	-4
Networking	547	-61.2%	-3
Lighting	238	-15.9%	–
Bathroom	98	-18.3%	–
Phone Charger	97	-53.6%	–

Table 6.4: Energy breakdown using classification results. Televisions (in the "Screen" category) and Routers ("Networking") are commonly misclassified, and have significant change in average daily energy. Other categories are shifted up in response, and "Entertainment" is increased from Routers being misclassified as Cable Boxes. If users are able to label their own television and router, typically 2-3 devices per household, classification results in a very accurate energy breakdown. Only the "Computer" and "Other Kitchen" categories are reversed in this case.

advanced than most, presents an outlier. It operates in that way for weeks, and generates a large quantity of outlier data that skews the results. The only solution here is more locations, not more time, as devices generate consistent feature vectors over time.

In addition, this classifier only looks at the 15 labels in the large dataset. Future versions of this system could examine the remainder of the devices not covered by the full dataset. There are too many individual and unique types to classify these devices, but adding the category of 'other' would give the system an opportunity to express doubt to the user even after multiple repeated and consistent classifications.

6.3.5 Classification Summary

In this section we evaluate PowerBlade against an existing dataset and methodology for device identification based on power data. This existing dataset was collected using more expensive and invasive sensors, and this allows us to evaluate the less invasive PowerBlade data for comparison. We utilize two different sizes of PowerBlade dataset: a small dataset covering 10 device types and 40% of the total devices measured, and a large dataset covering 15 device types and 55% of the total devices. A J48 classifier is able to classify each day in the large dataset with 63.54% accuracy, and after 30 days the device-level accuracy is 73.15%. This represents similar results to the existing systems using a less invasive sensor. Going forward the PowerBlade device could be used to collect significantly more data, and

the classification results improved by a broader test set, but this demonstrates a proof of concept device identification system using PowerBlade data.

6.4 Automatically Detecting Device Faults and Failures

Studies have shown that an effective tool for utilizing this type of energy data is fault detection, either from the perspective of Monjolo at the panel level or PowerBlade at the plug level. The concept is that if something goes wrong, whether a compressor failing or a piece of malware infecting an embedded system, the power signature of the device will change and this will be detectable by a power meter.

Related work in this field has demonstrated detecting malware on medical devices using the WattsUp energy meter [58]. The concept is that an embedded medical device performs actions that are repetitive and discrete like a pump or a sensor measurement. These repetitive actions will appear as a pattern in the electrical signature of the device. Deviations from that pattern, either over time or from among a group of similar devices, can indicate that the device is potentially infected with malware and is in need of servicing. This does not require knowledge of the device, and can be done by training on similar devices or over time.

6.4.1 Residential Fault Detection Conceptual Design

In a similar way, PowerBlade can detect household device faults. However, unlike in industrial systems where machinery is long-lasting and failures are very costly, residential devices have few real failures. Most devices, if broken, are simply replaced. There are, however, some cases in residential and commercial situations where detecting a failure is important. First, networking equipment, from servers to routers, can lock up and need to be reset. Although this is particularly acute in commercial, many users will be familiar with the home router needing to be power cycled as well. In addition, dust buildup can cause significant performance issues in devices from electronics to vacuum cleaners. A user might realize, for example, that the vacuum filter needs cleaning after extensive unsuccessful vacuuming. Detecting the gradual buildup, and altering the user before a performance issue occurs, would provide value in both residences and offices.

To determine if PowerBlade data is effective in detecting these types of faults we instrument certain machines for several weeks, and generate classification vectors similar to those used for device identification. However, rather than labeling the data with the type of device, the label is either *working* or *not working*. Since this would be run on the output from a single device, identifying certain devices relative to others is no longer required. Because

115
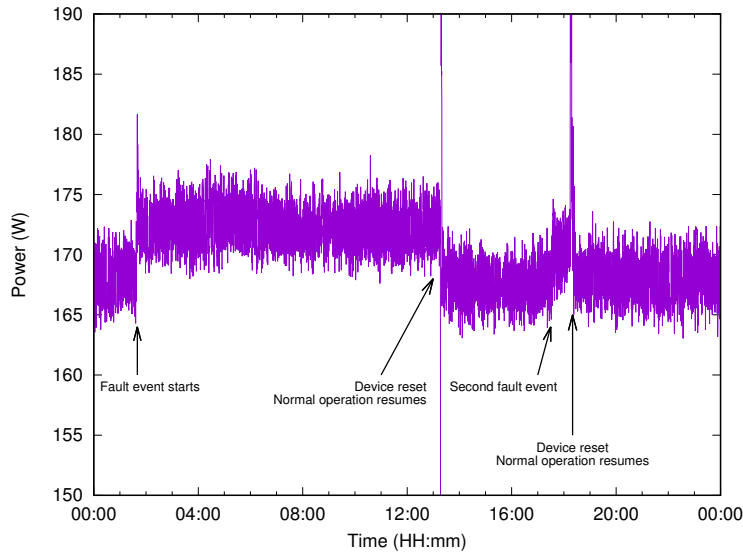
Figure 6.11: Two fault events in a working data server. Using a PowerBlade sensor, we measure a data server during use and record two times that it stops responding to network requests. This event is only recovered from when a user detects the problem and sends a reset command to a nearby IP-controlled AC switch. PowerBlade can be used to detect this event with 99.24% classification accuracy (Zero Rule baseline accuracy is 50.75%), and this can be used to automatically control the AC switch.

the mean time between failures in the general case is unknown, we subsample the working and not working conditions such that there is an equal number of samples for each label. This allows us to determine the viability of fault detection in the general case, but it does not necessarily reflect the underlying distribution of working and not working conditions in reality.

6.4.2 Evaluation on Real-World Devices

We implement a prototype of this system to determine the effectiveness of using power to detect faults in household devices. To detect faults in networking equipment we select a functioning data server in a university data center. Once or twice per week the device stops responding to network commands and must be physically reset. The source of the problem is unknown, and an employee periodically sends a *ping* request to the device to detect the error and, if necessary, resets the device either manually or using a nearby IP-controlled AC switch.

Our data capture on this machine records multiple failure events, which give ample opportunity to fully train a classifier, but the difference is apparent after the very first occurrence. Figure 6.11 shows the power in watts during two failure events on this machine
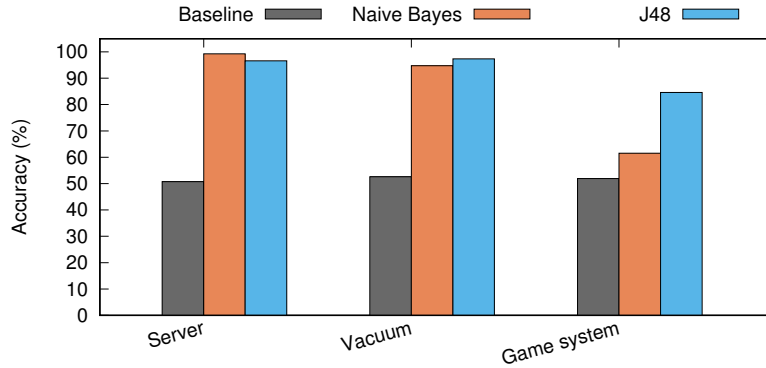
Figure 6.12: Classification success and baseline accuracy of working/not working for a data server and clean/dusty for a vacuum and a Xbox game system. The server has very low variability in both working and not working states, and a clear distinction in power between the two. The vacuum also has a low variability in operating power, but there is less distinction between clean (working) and dusty (broken). Finally the Xbox has significant variability in power states, so classification success is much lower when introducing the dirty/broken states.

with the key events labeled. Some events are provided by the computer's operators, and others can be reconstructed from logs on the machine itself. The Zero Rule baseline accuracy is 50.75%, and a Naive Bayes classifier is able to distinguish these events with 99.24% accuracy. A J48 decision tree results in 96.59% accuracy. This high level of accuracy shows that for this particular failure a PowerBlade-based detection method is very effective.

The low variability in power over time, both in the working and not working states, makes detecting the failure on this device very straightforward, but this result can extend to other networking devices and other household devices at large. In particular any device that performs a consistent function over time, which applies to servers and routers/modems but also vacuums and blenders, should be classifiable as either *normal* or *other*. As the variability of the device is increased, for example a television has a large range of valid power states depending on activity, this basic detection becomes more difficult.

To study the effect of variability, as well as to study dust in household device performance, we instrument a vacuum and an Xbox video game system from our study. These devices are selected in particular because they were cleaned of dust during the deployment period. This means we have data from before and from after cleaning each device. After cleaning, the vacuum draws slightly more power (about 30 W across a range of usage) but also performs better. The game system has multiple power states depending on activity, but the overall average power is about 10 W to 15 W lower after cleaning. The vacuum going *up* after cleaning and the game system overall going *down* are likely due to the different functions of the devices, but each may be detectable using a classifier.

Because the vacuum performs a more consistent single function, deviation from this state is apparent, and the performance of the classifier is significantly higher. Figure 6.12 shows the working/broken classification accuracy with baseline (Zero Rule), Naive Bayes, and J48 classifiers for the server as well as the clean/dusty classification accuracy for the vacuum and the game system. The vacuum baseline accuracy for detecting clean or dusty is 52.63%, and the Naive Bayes and J48 classifiers result in 94.73% and 97.36%, respectively. The Xbox baseline is 51.92% and the Naive Bayes and J48 classifiers are 61.53% and 84.61%, respectively. Although the Xbox accuracy is significantly lower, it is actually somewhat higher than might be expected given the wide range of things the game system does.

To examine how the J48 classifier is able to have any success classifying the Xbox as clean or dirty, we plot the decision tree generated by the J48 algorithm. In Figure 6.13 the first decision is whether the minimum power is greater than a threshold. If it is, then only the highest-power device activity (playing a game) is present in that feature vector. This reduces to a greater/less than question similar to the server or vacuum. If the minimum is less than the first decision's threshold, then the remainder of the activities (watching a
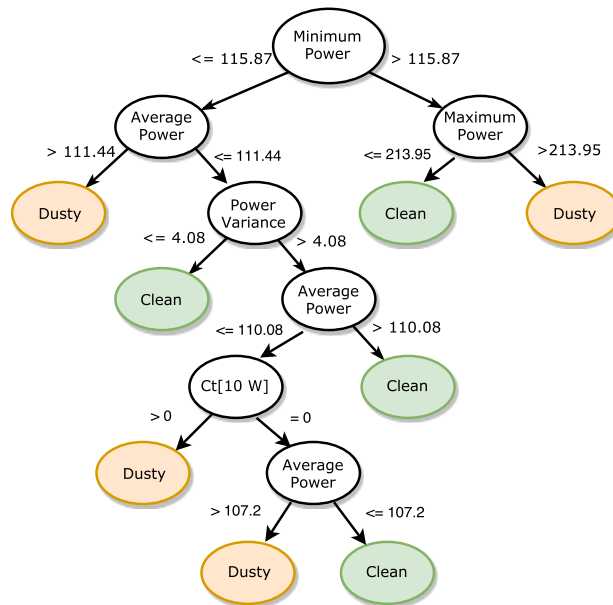
Figure 6.13: J48 decision tree to determine if the vents on an Xbox game system are clogged with dust. The device draws slightly more power when dusty, but there are numerous valid power states based on different activities performed by the system. If the minimum power is greater than 115.87 W then only the highest-power activity is present. This reduces the problem to a greater/less than comparison. If the minimum is less than 115.87 W then the system attempts to distinguish which activity/activities are present. In general dusty is the higher power or higher variance, and the system is able to slightly separate the different states.

movie, displaying the home screen, etc) are grouped together and the classification process becomes harder to visualize. However, on this side of the tree, the state with the higher power or higher variance is typically classified as dirty. Higher power has been previously observed, and the higher variance may be due to fans cycling more often or to a faster speed to prevent overheating when the system is dusty. 84% success is not ideal, but this system could provide a simple warning to the user that something may be wrong before the user becomes too invested in using the device.

6.4.3   Fault Detection Summary

In this section we repurpose the device classification system to detect faults in AC devices. For devices that perform consistent repeatable functions, even slight deviations from that function are observable in the energy signature. Although these systems already exist in some industrial environments for detecting machine failures, there are limited examples in residential due to a perceived lack of value.

We demonstrate multiple applications where this extra bit of information can be used to save users significant time or overhead. These applications include monitoring the state of networking equipment like servers and routers to detect drop-outs, and observing the performance decrease as household devices accumulate dust. Beyond what we illustrate here, other possibilities exist. Refrigeration systems could self-diagnose if a compressor fails and provide an early warning before food is spoiled, or hair dryers could detect blockages before potentially dangerous events. For safety reasons we do not evaluate these additional applications here, but with manufacturer participation both device types fit the criteria of having minimal power states to make detection possible.

In some cases the effect of this fault detection system is more difficult to quantify than in other applications, but the savings in user time is, to some, a more valuable metric. Empowering the PowerBlade system to send the reset command to the IP-controlled switch when the server is broken could free the employee to do other tasks. Alerting the user that the vacuum suction may be compromised could prevent vacuuming that will need to be re-done after the filter is cleaned. This makes fault detection another valuable tool in the PowerBlade toolkit.

6.5   Data Processing and Storage

These reports and applications collectively describe the manner in which the system's measurement data must be stored. Throughout this process we store raw sensor measurements
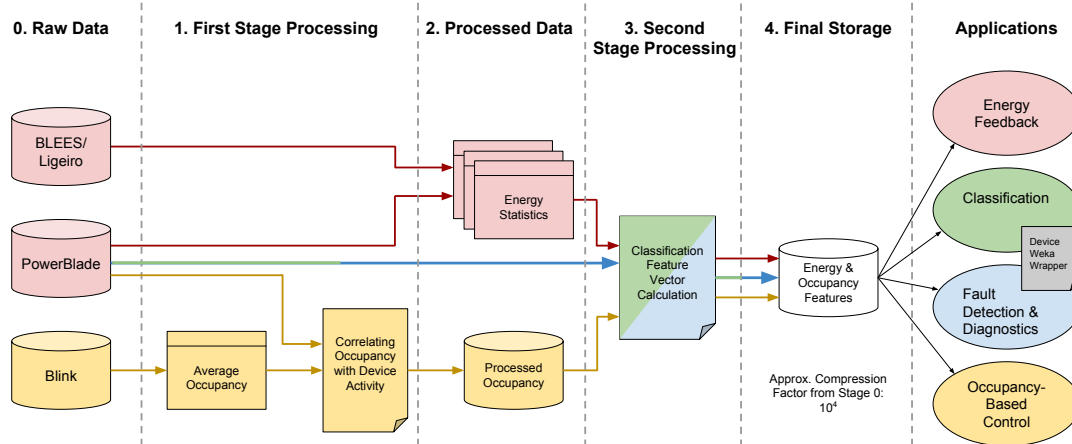
Figure 6.14: Data processing steps for energy applications. In red, PowerBlade, BLEES, and Ligeiro data are combined in SQL views of energy statistics. This step is shown in more detail in Figure 5.3. In yellow, A SQL view showing average occupancy is combined with device activity data to correlate activity with occupancy. These values are stored as intermediate processed data, which are then used along with additional waveform information to calculate device feature vectors (shown in green/blue). These data are then used for all of the applications (classification and fault detection also use the common Weka toolkit). The final storage in stage 4 enables each application, and is compressed from stage 0 by a factor of approximately 10^4.

for all PowerBlade, BLEES, Ligeiro, and Blink data. This allows for flexibility as new applications are developed that use these data in different ways, but it presents a storage problem. This relatively small deployment of modest scope and duration results in a dataset that occupies several hundred gigabytes of cloud database storage. In order for this system to scale to a size that increases the effectiveness and reduces the brittleness of each application listed above, data must be compressed.

Each application presented here runs on a processed version of the raw dataset(s), and in many cases these processed data stages can be combined and reused. Figure 6.14 show the data processing flow used in this system. In red are the energy calculations from Section 5.2.6.2, resulting in several SQL views describing energy and power statistics. In yellow, a SQL view for average occupancy can be combined with PowerBlade data to correlate device activity with occupancy. This processing is done locally, and re-uploaded to the cloud database for reuse. Finally in blue/green, another script calculates the device feature vectors used for classification and fault detection, and re-uploads this data for reuse.

The *Energy & Occupancy Features* table directly stores the feature vectors used for device classification and fault detection, but it also contains the information for the other applications. These data can be aggregated at the day or year level to enable energy feedback, and the occupancy features are directly used in the occupancy applications. This method of

storage is approximately four orders of magnitude more compressed than the raw data, and is sufficient for most applications.

In addition, this diagram shows multiple local steps interleaving with cloud storage, but this processing could all happen at the gateway level. Because the gateway collects data from all three types of sensor, it could perform up to stage 3 locally, and only upload the processed data for stage 4 storage. This would reduce the network overhead significantly, and may even make cellular upload a viable option which would greatly simplify the deployment.

6.6 Summary

In this chapter we explore whether the PowerBlade-based system with regard to producing human-understandable reports, yielding energy management actions, and enabling usable energy-based applications. To explore the scope of the reports and applications that the PowerBlade system enables, we select four in particular and demonstrate each with an example. We start with the energy breakdown at a device level, at a category level, and as an overall portion of the total energy consumed. This type of analysis is similar to, and based in part on, the energy analyses performed by the EIA, BTO, and others, but is produced by a much less invasive sensing system. Next we add a simple occupancy sensor to each room, and demonstrate how this small addition enables examining the power and energy data from a new perspective, and with new savings potential. We then address a problem with the system itself: that device naming during installation is a burden to the homeowner. In fact, for many it is their only interaction with the system, and if it could be reduced or eliminated so, too, would be one of the larger barriers to large-scale adoption. We demonstrate a classifier that performs comparably to existing device classification datasets. Finally we demonstrate that detecting faults in residential devices, a similar process to the classification algorithm, can save users significant time and headache when interacting with AC devices. These applications are likely only a portion of the total set that could be built on PowerBlade data.

In addition these reports and applications provide insight into which data must be stored, and how, and which can be discarded to make room for the growing dataset. Although for our limited evaluation we store all the data, this decision is clearly intractable even for a system deployed across a few dozen homes. However, most applications utilize the raw data in similar ways, and a few common usage types emerge. This indicates that these and even future applications could be enabled by a few common data building blocks. Computing the components of these building blocks in real time and discarding the raw data, even at the level of the gateway, reduces networking overhead, cloud storage expenses, and downstream processing complexity.

CHAPTER 7

Conclusion

High-coverage electrical submetering has the potential to change the way homeowners understand and consume energy. From simple reports and feedback that homeowners and other consumers use to save energy to high level applications that take action, and yield savings, on their own, the common requirement is data. Existing meters for collecting device-level energy data are too large and too expensive to deploy at scale, and other top-down approaches do not scale to the long tail of energy consumers. In this work we explore the designs and methodologies of an unobtrusive metering system that can be used to collect household data on an unprecedented scale, thereby facilitating deeper understanding of energy consumption and enabling energy savings.

This focus is in response to recognition in the Federal Government of the need for building data [34]. Buildings account for 40% of energy and 70% of electricity use in the United States. Although this is been relatively consistent for over a decade, the makeup of devices is shifting as high-efficiency devices and improvements to the building envelope reduce the energy use intensity (EUI) of HVAC, clothes and dish washing, and other high-powered loads. Alongside, Miscellaneous Electrical Loads (MELs) are growing in quantity and EUI, and may soon come to dominate energy consumption. This shift, and the new attention paid to the problem, are evident from solicitations like the Department of Energy Wireless Metering Challenge [33]. Building data, including data on the long tail of electrical loads, will be an important part of the building energy solution.

Making use of what building data is available has received attention from both research and industry, with compelling results. Energy consumers lack the tools and knowledge, but often not the motivation, to save energy, and studies have shown that simply providing direct energy feedback can enable to user to shed 10% of energy or more. These savings can be long-lasting if applied repeatedly and customized to the individual household. Other applications make use of this data to automatically identify devices, detect faults, and provide additional analytics by combining energy data with data from other sensors.

Collecting this energy data has typically been done using one of two approaches. The top-down approach is to measure household energy consumption at a single location using a high-precision meter, and disambiguate the individual energy components of each device. This scales inexpensively to new households, as the data is often already available through the utility-installed smart meter, but typically does not accurately disambiguate beyond the 10-20 highest-power loads in the household. The alternative bottom-up approach is to directly meter each AC device. This is more accurate particularly on lower-power devices, but presents a scaling issue. Existing meters are too large, expensive, and high-power, to scale up to the needs of full-coverage metering.

To respond to the gaps in metering coverage of devices in buildings, we leverage novel metering techniques and technologies to re-think the power and size point for individual plug load metering. This results in two metering systems: an energy harvesting energy meter called Monjolo that can be easily and safely installed in the electrical panel to measure high-power loads, and an unobtrusive plug-load sensor called PowerBlade that can be scaled to the size, cost, and power point required by the Wireless Metering Challenge. In particular the PowerBlade sensor enables residential and commercial metering with density and coverage previously impossible with existing systems.

However, the energy harvesting nature of Monjolo and the small form factor of PowerBlade introduce new challenges regarding the power available on the device. Fewer options are available for radio and wireless topology, and communicating data to the user becomes the limiting factor. Both devices communicate directly with a central gateway in an uncoordinated star topology. This limits both the size and the density of the networks that these sensors can support. In order to evaluate PowerBlade and Monjolo in residential situations we combine these sensors with others measuring occupancy and light, and deploy them alongside gateways that relay measurements to a cloud database.

To test the performance of the unobtrusive energy metering system, as well as to generate a high-coverage electricity dataset, we deploy this system in eight residential locations and two other locations. This allows us to test the accuracy of the metering system, percent of plug-loads and built-in lighting covered, and to evaluate the wireless topology in a real-world situation. We record data for an average of 62 days at each location, and deploy on average 34 PowerBlade sensors in each, 338 total, as well as a total of 43 light sensors, 27 PIR occupancy sensors, and 16 gateway devices. In controlled experiments, this system achieves full-coverage energy metering with 5%-10% error. Approximately 80% of PowerBlade measurements are received with the nominal 1 s latency, and only a few devices have an average latency of greater than 10 s. These results indicate that this system is viable for residential and commercial energy metering.

In addition to evaluating the metering system, our deployment generates a dataset that can be used to explore applications and analyses of building energy data. To provide energy feedback to users that enables significant, lasting savings, we produce an energy breakdown for each location and provide this report to the homeowner. We also aggregate data across locations to produce policy-level analyses and understanding. Next, we combine our energy data with data from an occupancy sensor to enable additional savings. Then we turn to a problem with the system itself, and demonstrate a process of automatic device naming that can be used to simplify deployment installation. Finally we explore fault detection in residential devices, and the time and user effort that can be saved with this additional information. These applications demonstrate the potential of the PowerBlade system, and other applications are possible.

Despite these results, the scope of this deployment is limited compared to nationwide surveys conducted by the EIA, BTO, and others, and is not statistically representative of the nation as a whole. Expanding this system to a representative sample of households, and in particular more geographic regions, more household sizes and socioeconomic levels, and more seasons of the year, would enable more meaningful energy conclusions. The unobtrusive nature of the system, and in particular the PowerBlade sensor, enable this expansion in a way not previously possible. A larger dataset would also improve the quality of energy applications, and in particular the device classification and fault detection.

To facilitate this expansion we can also improve the system at a technology level. The thickness of PowerBlade can be reduced by switching back to a flexible system but without the thick rigid layer of FR4. This requires system structure to come from plastic overmolding, which also improves the safety of the system and simplifies incorporation of a MuMetal shield. This also opens the possibility that the system can be UL certified, expanding opportunities for distribution. Installation of the the system can also be simplified by directly integrating the automatic device identification procedure. Finally the addition of voltage measurement on Monjolo enables true power metering in the electrical panel (as opposed to the current system's apparent power measurement), which can directly integrate data from the HVAC, oven, and others into the deployment data. These improvements target the most noticeable weaknesses of the system, and result in a system that can scale to residential metering.

Finally, we identify two areas for further exploration at a research level: reevaluating the networking technique and adding time to the system. The broadcast-without-acknowledgment BLE methodology enables low-power operation and a very simple network setup, but it leads to unavoidable data loss unless some kind of forward error correction scheme is employed [103]. There is also no opportunity to respond to the PowerBlade,

whether for device configuration or for data acknowledgment. We have already evaluated creating a BLE connection between the gateway and PowerBlade, this technique is currently used for PowerBlade calibration, and a BLE service could also be used for device re-programming or configuration. The BLE scan response can also be utilized for data acknowledgment of the advertisement packet. Both techniques increase overhead, and therefore decrease the maximum deployment density, but address configuration and data quality issues in the current system. Finally, we can increase interaction with smartphones, even potentially leveraging the smartphone in place of the gateway. This intriguing option is currently limited by the handset OS, which does not permit background data collection, but this issue could potentially be overcome in the future.

Time in the deployment system is captured at the gateway, where measurements are timestamped upon receipt. If a PowerBlade only achieves a successful packet transmission occasionally, there is no loss of overall accuracy but there is loss of resolution. Interval metering requires accurate timekeeping, typically down to the 15-minute level [30]. To enable interval metering the PowerBlade device can keep time, and two aspects of the network could make this possible without a real time clock (RTC) that requires long-term energy storage to maintain the time record. First, each PowerBlade device has access to the same 60 Hz zero-crossing signal, and this can be used for network synchronization and accurate time keeping. Second, although some devices are added and removed, there is typically at least one device in each room that is plugged in long-term. This device could be the timekeeper, and synchronize the other devices when they come online. Through this technique PowerBlade could be used as an interval meter.

Solving the building energy problem will require a multi-faceted approach of sensor hardware, processes and software, and buy-in from the people who are the eventual energy consumers. In this dissertation, we design and evaluate novel sensors and processes that offer high-coverage insight into energy consumption and provide a blueprint for larger scale deployments of this technology.

BIBLIOGRAPHY

[1] ADE7753 energy metering IC.
http://www.analog.com/en/products/analog-to-digital-converters/
integrated-special-purpose-converters/energy-metering-ics/ade7753.html.

[2] APS 3B series datasheet. http://www.adaptivepower.com/Resources/Documents/
APS_3B%20Series_Datasheet-1014.pdf.

[3] Beaglebone black. https://beagleboard.org/black.

[4] Belkin conserve insight energy use monitor.
https://www.belkin.com/conserve/insight/.

[5] Electronic product design Inc. PLM-1.
http://www.epd.com/power_meters.html#PLM-1.

[6] Energy star, the simple choice for energy efficiency. https://www.energystar.gov.

[7] Flipit! plug-less usb charger. https://www.amazon.com/
Ideative-PP1009W-Flipit-Plug-Less-Charger/dp/B004LKRY4K/ref=lh_ni_t.

[8] Generic access point. https://github.com/lab11/gap.

[9] Insignia wi-fi smart plug.
https://www.insigniaproducts.com/pdp/NS-SP1X7/5529012.

[10] International electrotechnical commission world plugs.
http://www.iec.ch/worldplugs/.

[11] Kill-a-watt electricity use monitor.
http://www.p3international.com/products/p4400.html.

[12] Leadership in energy and environmental design (LEED). http://www.usgbc.org/leed.

[13] LEM TT 100-SD current transformer.
http://www.lem.com/docs/products/tt100sd_e.pdf.

[14] Linear Technology LTC5800 wireless mote-on-chip.
http://www.linear.com/product/LTC5800-WHM.

[15] LTC3588 piezoelectric energy harvesting power supply.
http://cds.linear.com/docs/en/datasheet/35881fa.pdf.

[16] Mill-Max discrete spring loaded contacts.
http://www.mill-max.com/assets/pdfs/024.pdf.

[17] MQTT machine-to-machine (M2M)/"Internet of Things" connectivity protocol.
http://mqtt.org/.

[18] Nema standards publication ansi/nema wd 6-2002 (r2008).
https://www.nema.org/Standards/ComplimentaryDocuments/NEMA%20WD%
206%20-%20Dimensions%20for%20Wiring%20Devices%20-%20Excerpt.pdf.

[19] Netzero energy commercial building initiative. Technical report, Building
Technologies Program, U.S. Department of Energy.

[20] Nordic Semiconductor Bluetooth Smart SoC. https://www.nordicsemi.com/eng/
Products/Bluetooth-Smart-Bluetooth-low-energy/nRF51822.

[21] Ramtron FM25L04B. http://www.cypress.com/file/125066/download.

[22] Supertex SR10.
http://ww1.microchip.com/downloads/en/DeviceDoc/SR10%20B080613.pdf.

[23] Texas instruments CC2420. http://www.ti.com/lit/ds/symlink/cc2420.pdf.

[24] Texas Instruments MSP430F1611.
http://www.ti.com/lit/ds/symlink/msp430f1611.pdf.

[25] Texas Instruments MSP430FR5738 mixed signal FRAM microcontroller.
http://www.ti.com/product/msp430fr5738.

[26] Texas Instruments TPS62122 step-down converter.
http://www.ti.com/product/tps62122.

[27] Watts Up? Pro plug load meter.
https://www.wattsupmeters.com/secure/products.php?pn=0.

[28] Wemo home automation.
http://www.belkin.com/us/Products/home-automation/c/wemo-home-automation/.

[29] Annual energy outlook 2015. Technical report, U.S. Energy Information
Administration, 2015.

[30] An assessment of interval data and their potential application to residential electricity
end-use modeling. Technical report, U.S. Energy Information Administration, 2015.

[31] Department of energy quadrennial technology review (2015). Technical report, U.S.
Department of Energy, 2015.

[32] Electric power monthly with data for June 2015. Technical report, U.S. Energy Information Administration, aug 2015.

[33] Low cost wireless electric energy meter specification. Technical report, U.S. Department of Energy, 2016.

[34] Analysis and representation of miscellaneous electric loads in nems. Technical report, U.S. Energy Information Administration, 2017.

[35] Annual energy outlook 2017. Technical report, U.S. Energy Information Administration, 2017.

[36] DE-FOA-0001632 - buildings energy efficiency frontiers & innovation technologies (benefit). Technical report, U.S. Department of Energy, 2017.

[37] W. Abrahamse, L. Steg, C. Vlek, and T. Rothengatter. A review of intervention studies aimed at household energy conservation. *Journal of Environmental Psychology*, 25(3):273–291, 2005.

[38] N. Abramson. The Aloha system: another alternative for computer communications. In *Proceedings of the November 17-19, 1970, fall joint computer conference*, pages 281–285. ACM, 1970.

[39] C. Aguayo Gonzalez and J. Reed. Power fingerprinting in SDR & CR integrity assessment. In *Proceedings of the IEEE Military Communications Conference MILCOM*, 2009.

[40] J. Alcal, O. Parson, and A. Rogers. Detecting anomalies in activities of daily living of elderly residents via energy disaggregation and cox processes. In *BuildSys 2015 - Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, pages 225–234, 2015.

[41] I. Ang, F. Dilys Salim, and M. Hamilton. Human occupancy recognition with multivariate ambient sensors. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops 2016*, 2016.

[42] P. Armstrong, C. Laughman, S. Leeb, and L. Norford. Detection of rooftop cooling unit faults based on electrical measurements. *HVAC and R Research*, 12(1):151–175, 2006.

[43] A. Bandura. Self-efficacy mechanism in human agency. *American Psychologist*, 37(2):122–147, 1982.

[44] S. Barker, A. Mishra, D. Irwin, E. Cecchet, P. Shenoy, and J. Albrecht. Smart: An open data set and tools for enabling research in sustainable homes. *SustKDD*, 2012.

[45] S. Barker, M. Musthag, D. Irwin, and P. Shenoy. Non-intrusive load identification for smart outlets. In *2014 IEEE International Conference on Smart Grid Communications, SmartGridComm 2014*, pages 548–553, 2015.

[46] K. S. Barsim, L. Mauch, and B. Yang. Neural network ensembles to real-time identification of plug-level appliance measurements. In *3rd International Workshop on Non-Intrusive Load Monitoring (NILM 2016)*, 2016.

[47] N. Batra, A. Singh, and K. Whitehouse. If you measure it, can you improve it? exploring the value of energy disaggregation. In *BuildSys 2015 - Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built*, pages 191–200, 2015.

[48] N. Batra, A. Singh, and K. Whitehouse. Gemello: Creating a detailed energy breakdown from just the monthly electricity bill. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, volume 13-17-August-2016, pages 431–440, 2016.

[49] L. Becker. Joint effect of feedback and goal setting on performance: A field study of residential energy conservation. *Journal of Applied Psychology*, 63(4):428–433, 1978.

[50] B. Campbell and P. Dutta. Gemini: A non-invasive, energy-harvesting true power meter. In *Real-Time Systems Symposium (RTSS), 2014 IEEE*, pages 324–333. IEEE, 2014.

[51] K. Carrie Armel, A. Gupta, G. Shrimali, and A. Albert. Is disaggregation the holy grail of energy efficiency? the case of electricity. *Energy Policy*, 52(C):213–234, 2013.

[52] H.-H. Chang, C.-L. Lin, and J.-K. Lee. Load identification in nonintrusive load monitoring using steady-state and turn-on transient energy algorithms. pages 27–32, 2010.

[53] D. Chen, S. Barker, A. Subbaswamy, D. Irwin, and P. Shenoy. Non-intrusive occupancy monitoring using smart meters. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, BuildSys'13, pages 9:1–9:8, New York, NY, USA, 2013. ACM.

[54] V. Chen, M. Delmas, and W. Kaiser. Real-time, appliance-level electricity use feedback system: How to engage users? *Energy and Buildings*, 70:455–462, 2014.

[55] T. Chiang, S. Natarajan, and I. Walker. A laboratory test of the efficacy of energy display interface design. *Energy and Buildings*, 55:471–480, 2012.

[56] J.-S. Chou and A. Telaga. Real-time detection of anomalous power consumption. *Renewable and Sustainable Energy Reviews*, 33:400–411, 2014.

[57] D. Christensen, L. Earle, and B. Spam. NILM applications for the energy-efficient home. Technical report, United States National Renewable Energy Laboratory (NREL), 2012.

[58] S. S. Clark, B. Ransford, A. Rahmati, S. Guineau, J. Sorber, W. Xu, and K. Fu. Wattsupdoc: Power side channels to nonintrusively discover untargeted malware on embedded medical devices. In *Presented as part of the 2013 USENIX Workshop on Health Information Technologies*, Washington, D.C., 2013. USENIX.

[59] R. Condit. Transformerless power supplies: Resistive and capacitive. http://ww1.microchip.com/downloads/en/AppNotes/00954A.pdf, 2004.

[60] M. Cottrell, P. Gaubert, C. Eloy, D. Franois, G. Hallaux, J. Lacaille, and M. Verleysen. Fault prediction in aircraft engines using self-organizing maps. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5629 LNCS:37–44, 2009.

[61] G. Crabtree. Energy future: think efficiency. Technical report, American Physical Society, 2008.

[62] S. Darby. The effectiveness of feedback on energy consumption. Technical report, Environmental Change Institute, University of Oxford, Oxford, England, 2006.

[63] S. Dawson-Haggerty, S. Lanzisera, J. Taneja, R. Brown, and D. Culler. @scale: Insights from a large, long-lived appliance energy wsn. In *Proceedings of the 11th international conference on Information Processing in Sensor Networks*, pages 37–48. ACM, 2012.

[64] S. DeBruin, B. Campbell, and P. Dutta. Monjolo: An energy-harvesting energy meter architecture. In *SenSys 2013 - Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, 2013.

[65] S. DeBruin, B. Ghena, Y.-S. Kuo, and P. Dutta. Powerblade: A low-profile, true-power, plug-through energy meter. In *SenSys 2015 - Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pages 17–29, 2015.

[66] S. DeBruin, J. Grunnagle, and P. Dutta. Demo abstract: Scaling the wireless ac power meter. In *11th International Conference on Information Processing in Sensor Networks (IPSN)*, 2012.

[67] K. Ehrhardt-Martinez, K. A. Donnelly, and J. A. Laitner. Advanced metering initiatives and residential feedback programs: A meta-review for household electricity-saving opportunities. Technical Report E105, American Council for an Energy-Efficient Economy, Washington, D.C., jun 2010.

[68] F. Englert, T. Schmitt, S. Kler, A. Reinhardt, and R. Steinmetz. How to auto-configure your smart home? high-resolution power measurements to the rescue. In *e-Energy 2013 - Proceedings of the 4th ACM International Conference on Future Energy Systems*, pages 215–224, 2013.

[69] V. Erickson, S. Achleitner, and A. Cerpa. Poem: Power-efficient occupancy-based energy management system. In *IPSN 2013 - Proceedings of the 12th International*

Conference on Information Processing in Sensor Networks, Part of CPSWeek 2013, pages 203–216, 2013.

[70] A. Faruqui, S. Sergici, and A. Sharif. The impact of informational feedback on energy consumption - a survey of the experimental evidence. *Energy*, 35(4):1598–1608, 2010.

[71] A. M. Fawaz, M. A. Noureddine, and W. H. Sanders. Poweralert: An integrity checker using power measurement. *CoRR*, abs/1702.02907, 2017.

[72] C. Fischer. Feedback on household electricity consumption: A tool for saving energy? *Energy Efficiency*, 1:79–104, 2008.

[73] J. Froehlich, E. Larson, S. Gupta, G. Cohn, M. Reynolds, and S. Patel. Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, 10(1):28–39, 2011.

[74] T. Fujimi, Y. Kajitani, and S. Chang. Effective and persistent changes in household energy-saving behaviors: Evidence from post-tsunami japan. *Applied Energy*, 167:93–106, 2016.

[75] J. Gao, E. Kara, S. Giri, and M. Berges. A feasibility study of automated plug-load identification from high-frequency measurements. In *2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015*, pages 220–224, 2016.

[76] J. Gillis, S. Alshareef, and W. Morsi. Nonintrusive load monitoring using wavelet design and machine learning. *IEEE Transactions on Smart Grid*, 7(1):320–328, 2016.

[77] G. Guerassimoff and J. Thomas. Enhancing energy efficiency and technical and marketing tools to change people's habits in the long-term. *Energy and Buildings*, 104:14–24, 2015.

[78] S. Gupta, M. Reynolds, and S. Patel. Electrisense: Single-point sensing using emi for electrical event detection and classification in the home. In *UbiComp'10 - Proceedings of the 2010 ACM Conference on Ubiquitous Computing*, pages 139–148, 2010.

[79] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. Witten. The weka data mining software: An update. *SIGKDD Explorations*, 11(1):10–18, 2009.

[80] G. W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.

[81] C. Horne and E. Kennedy. The power of social norms for reducing and shifting electricity use. *Energy Policy*, 107:43–52, 2017.

[82] A. Iwayemi and C. Zhou. SARAA: Semi-supervised learning for automated residential appliance annotation. *IEEE Transactions on Smart Grid*, PP(99), 2015.

[83] R. Jain, J. Taylor, and P. Culligan. Investigating the impact eco-feedback information representation has on building occupant energy consumption behavior and savings. *Energy and Buildings*, 64:408–414, 2013.

[84] X. Jiang, S. Dawson-Haggerty, P. Dutta, and D. Culler. Design and implementation of a high-fidelity AC metering network. In *Information Processing in Sensor Networks, 2009. IPSN 2009. International Conference on*, pages 253–264. IEEE, 2009.

[85] X. Jiang, M. Van Ly, J. Taneja, P. Dutta, and D. Culler. Experiences with a high-fidelity wireless building energy auditing network. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, SenSys '09, pages 113–126, New York, NY, USA, 2009. ACM.

[86] J. H. Jiménez, Q. Chen, J. Nichols, C. Calhoun, and S. Sykes. Towards a cyber defense framework for SCADA systems based on power consumption monitoring. In *HICSS*, 2017.

[87] D. Jung, H. Nguyen, and D. Yau. Tracking appliance usage information using harmonic signature sensing. In *2015 IEEE International Conference on Smart Grid Communications, SmartGridComm 2015*, pages 459–465, 2016.

[88] A. Kamilaris, B. Kalluri, S. Kondepudi, and T. Kwok Wai. A literature survey on measuring energy usage for miscellaneous electric loads in offices and commercial buildings. *Renewable and Sustainable Energy Reviews*, 34:536–550, 2014.

[89] S. Karjalainen. Consumer preferences for feedback on household electricity consumption. *Energy and Buildings*, 43(2-3):458–467, 2011.

[90] A. Katunin, M. Amarowicz, and P. Chrzanowski. Faults diagnosis using self-organizing maps: A case study on the damadics benchmark problem. In *Proceedings of the 2015 Federated Conference on Computer Science and Information Systems, FedCSIS 2015*, pages 1673–1681, 2015.

[91] J. Kelly and W. Knottenbelt. Neural nilm: Deep neural networks applied to energy disaggregation. In *BuildSys 2015 - Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built*, pages 55–64, 2015.

[92] Keystone Electronics. *Positive Battery Contact 120*, August 2015.

[93] A. Khosrowpour, Y. Xie, J. Taylor, and Y. Hong. One size does not fit all: Establishing the need for targeted eco-feedback. *Applied Energy*, 184:523–530, 2016.

[94] W. Kleiminger, C. Beckel, and S. Santini. Household occupancy monitoring using electricity meters. In *UbiComp 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 975–986, 2015.

[95] R. Kravets, A. F. Harris III, and R. Want. Beacon trains: blazing a trail through dense ble environments. In *Proceedings of the Eleventh ACM Workshop on Challenged Networks*, pages 69–74. ACM, 2016.

[96] S. Kwatra, J. Amann, and H. Sachs. Miscellaneous energy loads in buildings. Technical report, American Council for an Energy-Efficient Economy, 2013.

[97] T. Labeodan, C. De Bakker, A. Rosemann, and W. Zeiler. On the application of wireless sensors and actuators network in existing buildings for occupancy detection and occupancy-driven lighting control. *Energy and Buildings*, 127:75–83, 2016.

[98] J. Laustsen. Energy performance certification of buildings. Technical report, International Energy Agency, 2010.

[99] I. Leontiadis, C. Efstratiou, C. Mascolo, and J. Crowcroft. *SenShare: Transforming Sensor Networks into Multi-application Sensing Infrastructures*, pages 65–81. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[100] J. Lifton, M. Feldmeier, Y. Ono, C. Lewis, and J. A. Paradiso. A platform for ubiquitous sensor deployment in occupational and domestic environments. In *IPSN '07: Proceedings of the 6th international conference on Information processing in sensor networks*, Cambridge, Massachusetts, Apr. 2007.

[101] M. C. Lorek, F. Chraim, K. S. J. Pister, and S. Lanzisera. Cots-based stick-on electricity meters for building submetering. In *SENSORS, 2013 IEEE*, pages 1–4, Nov 2013.

[102] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse. The smart thermostat: Using occupancy sensors to save energy in homes. In *SenSys 2010 - Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pages 211–224, 2010.

[103] D. MacKay. Fountain codes. In *IEE Proceedings - Communications*, 2005.

[104] A. Molina-Markham, P. Shenoy, K. Fu, E. Cecchet, and D. Irwin. Private memoirs of a smart meter. In *BuildSys'10 - Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pages 61–66, 2010.

[105] National Science and Technology Council–Committee on Technology. Federal research and development agenda for net-zero energy, high-performance green buildings. Technical report, oct 2008.

[106] National Science and Technology Council–Committee on Technology. Submetering of building energy and water usage: Analysis and recommendations of the subcommitte on buildings technology research and development. Technical report, oct 2011.

[107] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd. At the flick of a switch: Detecting and classifying unique electrical events on the residential power line. In *Proceedings of the 9th International Conference on Ubiquitous Computing*, UbiComp '07, pages 271–288, Berlin, Heidelberg, 2007. Springer-Verlag.

[108] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz. On the accuracy of appliance identification based on distributed load metering data. In *2012 Sustainable Internet and ICT for Sustainability, SustainIT 2012*, 2012.

[109] K. Roth, K. Mckenney, C. Paetsch, and R. Ponoum. U.S. residential miscellaneous electric loads electricity consumption. In *2008 ACEEE Summer Study on Energy Efficiency in Buildings*, 2008.

[110] A. Rowe, M. Berges, and R. Rajkumar. Contactless sensing of appliance state transitions through variations in electromagnetic fields. pages 19–24, 2010.

[111] J. Schleich, C. Faure, and M. Klobasa. Persistence of the effects of providing feedback alongside smart metering devices on household electricity demand. *Energy Policy*, 107:225–233, 2017.

[112] J. C. Seabra, M. A. Costa, and M. M. Lucena. Iot based intelligent system for fault detection and diagnosis in domestic appliances. In *2016 IEEE 6th International Conference on Consumer Electronics - Berlin (ICCE-Berlin)*, pages 205–208, Sept 2016.

[113] J. Stankovic. Research directions for the internet of things. *IEEE Internet of Things Journal*, 1(1):3–9, 2014.

[114] M. Svensson, S. Byttner, and T. Rgnvaldsson. Self-organizing maps for automatic fault detection in a vehicle cooling system. In *2008 4th International IEEE Conference Intelligent Systems, IS 2008*, volume 2, pages 248–2412, 2008.

[115] A. S. Tanenbaum. Computer networks, 4-th edition. *ed: Prentice Hall*, 2003.

[116] S. Tumanski. Induction coil sensors a review. *Measurement Science and Technology*, 18(3):R31, 2007.

[117] U.S. Department of Energy–Building Technologies Office. Multi-Year Program Plan 2016-2020, Feb. 2016.

[118] L. Yang, X. Chen, J. Zhang, and H. Poor. Cost-effective and privacy-preserving energy management for smart meters. *IEEE Transactions on Smart Grid*, 6(1):486–495, 2015.

[119] L. Yerva, B. Campbell, A. Bansal, T. Schmid, and P. Dutta. Grafting energy-harvesting leaves onto the sensornet tree. In *Proceedings of the 11th International Conference on Information Processing in Sensor Networks*, IPSN '12, pages 197–208. ACM, 2012.

[120] T. Zachariah, N. Klugman, B. Campbell, J. Adkins, N. Jackson, and P. Dutta. The internet of things has a gateway problem. In *HotMobile 2015 - 16th International Workshop on Mobile Computing Systems and Applications*, pages 27–32, 2015.

[121] B. Zhao, L. Stankovic, and V. Stankovic. Blind non-intrusive appliance load monitoring using graph-based signal processing. In *2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015*, pages 68–72, 2016.

[122] K. Zhou and S. Yang. Understanding household energy consumption behavior: The contribution of energy big data analytics. *Renewable and Sustainable Energy Reviews*, 56:810–819, 2016.

APPENDIX A

Design Resources

Design files and other resources related to the projects in this dissertation are available on the Lab11 website at http://lab11.eecs.umich.edu. Specific code and PCB design files for the Monjolo and PowerBlade projects can be found at the locations below.

Monjolo Energy Harvesting Energy Meter
> http://github.com/lab11/monjolo

PowerBlade Plug-Through Plug-Load Meter
> http://github.com/lab11/powerblade

Information on the database used to collect, store, and process deployment data can be found at http://github.com/lab11/powerblade/tree/master/sql. This directory contains a Readme file that describes the individual SQL scripts used for database management. This includes table creation and other setup, deployment status, energy and power calculations, and applications like fault detection and device identification.