It's Different: Insights into home energy consumption in India

Nipun Batra[†], Manoj Gulati[†], Amarjeet Singh[†], Mani B. Srivastava[§]

[†]Indraprastha Institute of Information Technology
Delhi, India
{nipunb,manojg,amarjeet}@iiitd.ac.in

[§]University of California
Los Angeles, United States
mbs@ucla.edu

Abstract

Residential buildings contribute significantly to the overall energy usage across the world. Real deployments, and collected data thereof, play a critical role in providing insights into home energy consumption and occupant behavior. Existing datasets from real residential deployments are all from the developed countries. Developing countries, such as India, present unique opportunities to evaluate the scalability of existing research in diverse settings. Building upon more than a year of experience in sensor network deployments, we undertake an extensive deployment in a three storey home in Delhi, spanning 73 days from May-August 2013, measuring electrical, water and ambient parameters. We used 33 sensors across the home, measuring these parameters, collecting a total of approx. 400 MB of data daily. We discuss the architectural implications on the deployment systems that can be used for monitoring and control in the context of developing countries. Addressing the unreliability of electrical grid and internet in such settings, we present *Sense Local-store Upload* architecture for robust data collection. While providing several unique aspects, our deployment further validates the common considerations from similar residential deployments, discussed previously in the literature. We also release our collected data- Indian data for Ambient Water and Electricity Sensing (iAWE), for public use.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Design, Experimentation

Keywords

Deployment, Buildings, Smart Homes, Sensor Networks

Buildsys'13, November 11–15, 2013, Rome, Italy.

1 Introduction

Buildings account for more than 30% of overall energy consumption globally. Of this energy consumption, a large proportion (e.g. 93% in India) is contributed by residential buildings [10]. Information Technology (IT) such as cyber-physical systems, wireless sensor networks, embedded control, computational modeling, machine learning, and simulation tools can play a key role in reducing the net energy resources consumed in a building while maintaining the productivity of its occupants. These IT methods include guiding the occupants towards resource conserving behaviors, alerting for timely repair of energy-wasting degradations in building facilities, intelligent control of the building systems, and opportunistically harvesting energy from the environment.

Of particular importance are the systemic building deployments that can provide detailed insights about occupant behavior (specifically, Activities of Daily Living (ADLs)) and energy consumption. These deployments also provide data sets that can be leveraged for developing and testing suitable control strategies. These control strategies are otherwise complex to undertake in a real occupied building. In the recent past, several datasets, such as REDD [14], BLUED [3], AMPds [15], Smart* [5], monitoring household electricity and ambient parameters, have been released publicly. Several building monitoring and control research has since used these datasets to prove the validity of their work for real life settings [18, 6].

However, all of the previous deployments have been done in the context of developed countries. Developing countries, such as India, have higher electricity deficit, are adding new building space at a higher rate and constitute different infrastructure and energy consumption patterns. A deeper understanding of these different settings in developing countries can help in the development of systems that can scale across diverse settings in a robust manner. We have been involved in sensor network deployments in the Indian context for more than a year [7], whereby, we have instrumented 25 homes with smart meters, a smart campus with sensors for ambient monitoring in a research wing and 52 smart meters in the institute dorms. In this paper, we discuss an extensive ongoing deployment in a three storey home in Delhi, India, spanning 73 days from May-August 2013. Monitored parameters include electricity and water consumption at the meter level, plug level load monitoring for major appliances, and ambient parameters across every room. We used 33 sensors to measure these parameters, collecting approx. 400 MB data everyday.

To the best of our knowledge, this is the first such extensive deployment outside any developed country. We discuss, in detail, the unique aspects of our deployment that are also characteristic of buildings in the developing countries. Correspondingly, we provide insights into these aspects, of building systems, critical for robust data collection and control. We further discuss aspects of our deployment that were similar to those highlighted in the previous work on residential deployments. Our deployment was maintained as an open source project, clearly illustrating the issues faced and how these were addressed. Unlike many of the past deployments, detailed metadata logs, such as appliance make and mode of operation, are also provided. We believe that the unique aspects of the building energy infrastructure, as discussed in this work, will enrich the existing research in building energy domain, which has only leveraged deployments and data collection in the context of developed countries until now. Previous work [14] has established that dataset availability has spurred application areas in machine learning. With an aim of advancing research in building energy domain, we release our dataset iAWE.

2 Related Work

Building deployments have been studied in the past with the goal of improving the building energy efficiency. Office and campus deployments presented in the previous work [1, 7] have shown the scope for significantly reducing HVAC and plug load energy consumption in respective settings. Several other deployments target residential sensing for modeling and inference, specifically pertaining to Non Intrusive Load Monitoring (NILM). Kolter et al. [14] performed deployments across 6 homes in Boston (US) in 2011, with collected data spanning up to 19 days in some homes. They monitored household electricity at the meter, circuit and appliance level using Commercial Off-The-Shelf (COTS) devices. Their dataset (REDD) has been frequently used to validate NILM research and to date has been cited 57 times, clearly suggesting the impact of such deployments. Anderson et al. [3] performed a week long residential deployment, specifically focusing on collecting fully labeled high frequency electrical data and released BLUED dataset. Our data, in comparison, is for a much longer duration (spanning 73 days at the time of writing), uses a mix of COTS and customized hardware (due to non-availability of COTS, for everything we wanted to monitor, in the Indian context), and provides a unique combination of electricity, water consumption and ambient parameters.

Barker et al. [5] performed deployments across three homes and have collected information across varying modalities including, but limited to, electricity consumption, occupancy, weather and renewable electricity generation. They illustrated the wide applicability of the data from their residential deployment, including peak demand flattening [6] and cost optimization using variable electricity pricing [17]. They further highlighted the value of additional information obtained by correlating across multiple sensing modalities. Motivated by this work, we decided to monitor the ambient parameters, in addition to monitoring electrical

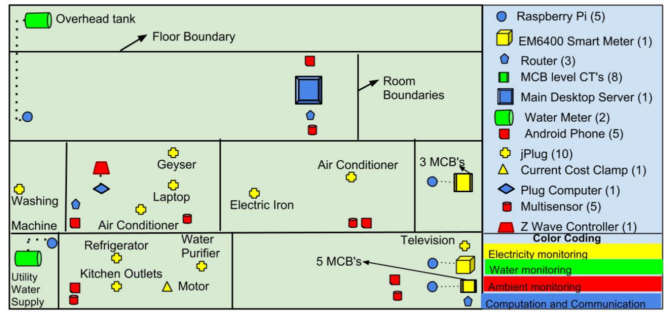


Figure 1: Schematic showing overall home deployment

and water consumption at different granularity.

Hnat et al. [13] provide a detailed guide into residential deployments and provided lessons learned from residential deployments across multiple homes for several years. They proposed various applications of such detailed sensing including identifying ADLs. While providing unique aspects of our deployment, we further establish the commonalities in our deployment with the learning discussed in their work. To the best of our knowledge, all the deployments discussed previously in the literature pertain to the developed countries. While residential deployments anywhere across the world are challenging, our deployments highlight some unique challenges specific to the developing countries. Some of these challenges in our setting include, but are not limited to, unreliable grid, unreliable internet and difficulty in procuring quality COTS.

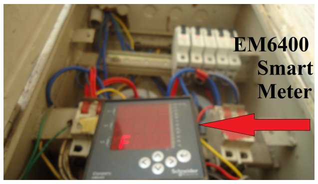
3 Deployment Overview

Our deployment constitutes 33 sensors measuring electricity, water and ambient parameters at different granularity, in a home in Delhi, India during May-August 2013. Primary objective for this deployment was to bring forth the differences in the Indian context, as compared to the context of developed countries along the dimensions of - 1. Grid and network reliability; 2. Energy and water consumption patterns; and 3. The ecosystem of available sensing options that restrict the possible deployments. Figure 1 shows the deployment of these sensors in a 3 storey home, together with the required computing and communication infrastructure. It must be noted that the intended purpose behind this deployment is not to develop a low cost, scalable Home Area Network (HAN), but, to do an extensive deployment for obtaining insights into energy consumption, which may be used for developing suitable HANs.

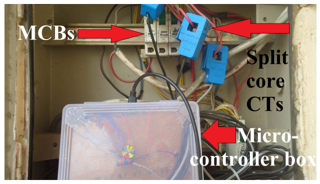
3.1 Sensing Infrastructure

For sensing, we took a “leave no stone unturned” approach, similar to SMART* [5], where we chose to monitor as many physical (ambient conditions, electricity usage and water usage) and non-physical (such as network strength and network connectivity) parameters as possible. We took care to deploy these sensors in a way that residents can continue their daily routines without added inconvenience. Constrained by the limited options available in the Indian context, our sensors constitute COTS (procured from both within and outside India) and custom built hardware.

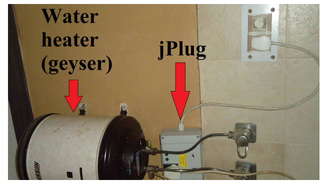
Electricity monitoring: Motivated by prior electricity consumption deployments, we also chose to monitor electricity consumption across different granularity - electricity me-



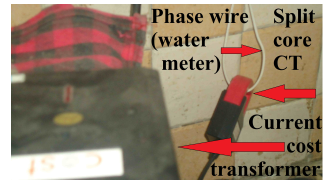
(a) EM6400 Smart Meter



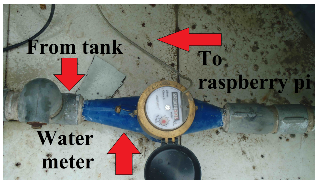
(b) CT based system for monitoring MCBs



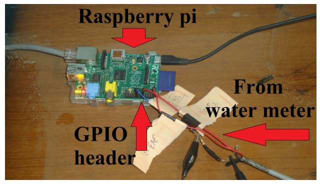
(c) Appliance level monitoring using jPlug



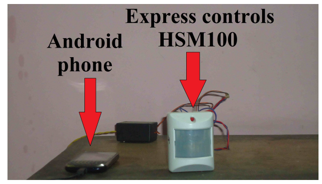
(d) Current Cost CT based monitoring



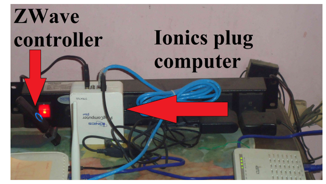
(e) Water Meter



(f) RPi collecting pulse outputs from water meter over GPIO



(g) Android phone and ZWave based multisensor (measuring ambient parameters)



(h) Plug computer collecting ZWave data and sending over network using Ethernet

Figure 2: Sensing, computation and communication equipment used in our home deployment

ter monitoring the consumption at the home aggregate level, current transformers (CTs) monitoring current for Miniature Circuit Breakers (MCBs) (each connected to a combination of appliances) and plug level monitors for monitoring plug load based appliances (see Figure 3a for illustration).

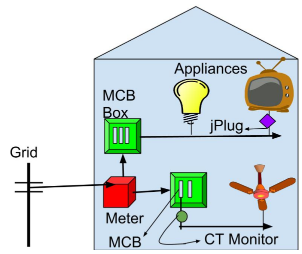
1. **Meter level:** Modbus-serial enabled Schneider Electric EM6400[1] meter was used to instrument the main power supply (see Figure 2a). We collected data including voltage, current, frequency, phase and power at 1 Hz.
2. **Circuit level:** Split-core CTs, clamped to individual MCBs, are used for monitoring circuit level current. Since no commercial solution was easily available in India for panel level monitoring, we used a custom built solution involving low cost microcontroller and Single Board Computer (SBC) platform. Figure 2b illustrates CTs monitoring 3 MCBs on the first floor MCB box in our home. A total of 8 CTs were used to monitor different MCB circuits in the home.
3. **Appliance level:** Since no good commercial options were available for plug level monitors, we worked with our collaborators and used their in-house developed jPlug[2] for monitoring individual appliance level power consumption. Ten jPlugs were used to monitor different plug-load based appliances across the home. jPlug measured multiple parameters including voltage, current, phase and frequency, that were uploaded to server using HTTP POST. Additionally, Current Cost (CC) based CT is used to measure the power consumption for electric motor (used to pump water), which is not a plug-load, but has a significant power consumption (approx. 700 Watts). CC exposes apparent power data over the USB port. jPlug and CC are shown in Figure 2c and Figure 2d respectively.

Water monitoring: To work around the short (only for a few hours a day) water supply in India, overhead water tanks (typically of 1000 liters capacity) are used to store water. Due to low water pressure, electric motors are used to pump the water for storage when the supply is available. Figure 3b illustrates the water flow distribution in the monitored home, together with the placement of water meters. One water meter is placed at the inlet (coming from the utility) and another one at the outlet from the water tank (flowing downwards).

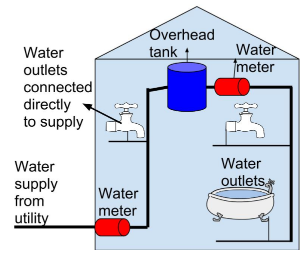
Due to prohibitive cost for digital water meters in India,

[1]www.goo.gl/01edPS

[2]A variant of nPlug [11]



(a) Different granularity of measuring electricity consumption in home: meter, circuit and appliance



(b) Different granularity of measuring water consumption in home: inlet supply from utility, outlet supply from tank

Figure 3: Electricity and water flow inside a home and different granularity at which these parameters can be monitored.

we chose to use Zenner Aquameter's multijet[3]. The multijet uses pulse output generated through a 4-20 mA current loop. Water meter connected to the utility, over a 0.5 inch diameter pipe, generates a pulse for every 1 liter of water consumption. Water meter connected to the outlet of storage tank, with 1.25 inch diameter, generates a pulse every 10 liters of water consumption. Figure 2e shows the water meter deployed inline at the overhead tank.

Ambient monitoring: ZWave based Express Controls HSM100[4] multisensors were used for monitoring motion, light and temperature across 5 rooms in the home. To the best of our knowledge, at the time of deployment, no commercial ZWave based sensor working on Indian frequency (865.2 MHz) was available. We correspondingly imported EU frequency (868.4 MHz) devices and used them for ambient monitoring. For these HSM100, motion is reported in an event-driven manner (i.e. whenever there is change in motion status, a reading is reported) and temperature and light are polled at 1 Hz. An Android phone, running FunF journal application[5], was placed at a fixed location in each room to log ambient parameters such as light and sound level every 30 seconds for 5 seconds.

Miscellaneous: Android phones, in addition to measuring ambient conditions, were also used to scan and log Bluetooth, WiFi and GSM networks. All the home occupants were requested to keep the Bluetooth, for their personal phone, on during the duration of the experiment. The network scanning was done every 1 minute and is stored locally

[3]www.aquametwatermeters.com/multijet.html

[4]http://goo.gl/Bszg0u

[5]http://www.funf.org/journal.html

Table 1: Details of sensing infrastructure used in our deployment

Sensor name	Procurement	Sampling frequency	Granularity	Quantity	Communication	Observed parameters
EM6400	COTS (India)	1 Hz	Home	1	RS 485 Serial	Voltage, Current, Frequency, Phase, Power (Active, Reactive and Apparent), Energy
Aquamet multijet	COTS (India)	5 Hz	Main supply and tank	2	4-20 mA output to GPIO	10 liter pulse for tank output and 1 liter pulse for main supply
Express Controls HSM100	COTS (Imported)	Light, temperature: 1 Hz; Motion: event based	Room	6	ZWave	Light, temperature and motion
Android phones	COTS (India)	Audio, light: 5 seconds every 30 seconds; Network scanning: once every 60 seconds	Room	5	Manual transfer	Audio features, light, nearby Bluetooth, cell-tower, WiFi
CT Monitor	Prototype	20 Hz	MCB	8	Serial	RMS Current
jPlug	Prototype	1 Hz	Appliance	10	WiFi	Voltage, Current, Frequency, Power (Active and Apparent), Energy, Phase
Current Cost	COTS (Imported)	Once every 6 seconds	Appliance	1	Serial	Apparent power

on the SD card. External weather conditions, such as temperature, humidity and wind speed, were also logged every 10 minutes using publicly available weather monitoring APIs[6].

Complete sensing infrastructure, used in our deployment, is summarized in Table 1.

3.2 Communication and Computation

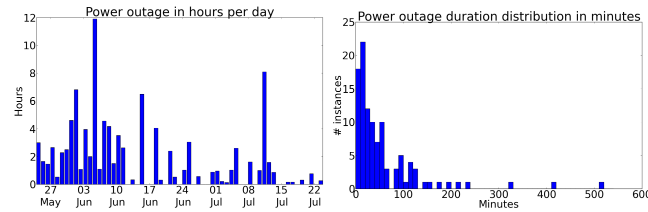
Different computing platforms - microcontrollers, SBCs and desktops are used for data collection. We used 5 RPis[7] and 1 Ionics Stratus plug[8] computer as SBCs and a 2 GHz Desktop PC running Linux, as the main local server.

One RPi, connected to EM6400 using RS485-USB converter, collected meter data using a custom program based on pyModbus[9] and communicated it to the desktop server. USB output (XML formatted) from CC is collected on another RPi and is communicated to the desktop server.

Separate RPis were used for prototype circuit level monitoring and for collecting data from water meter. We initially wrote an interrupt driven program to detect GPIO events corresponding to pulse output from water meters. We observed that noise introduced in the circuit due to long cable lengths led to a lot of false events. Correspondingly, we modified our program and polled at 5 Hz to obtain GPIO status.

A web daemon, running on the server, listened to the HTTP post request from jPlugs and dumped the data in MySQL. Ionics Plug Computer was used to collect data from all the ZWave based sensors. We wrote custom wrappers around OpenZWave[10] to collect temperature, light and motion data. While the plug computer had an internal ZWave (the reason for which it was selected), its range was limited and did not cover all the ZWave sensors. Correspondingly, a ZWave controller was connected over USB with Ionics, that provided reachability to all the ZWave devices. Figure 2h shows the plug computer collecting ambient sensor data from ZWave controller. A manual dump of collected data on each Android phone was performed every 15 days.

In the course of our deployment we observed several issues pertaining to SBCs. As an example, the OpenZWave based program, used to collect data, created log files for its own diagnostics. These log files eventually consumed the 512 MB flash drive space on the plug computer. This was fixed by deleting the older logs. Such problems encouraged us to develop soft-sensor [19] streams, whereby we periodically collected hard disk space, ping success, CPU utilization and available RAM, for all the computing devices. These soft-sensor streams can be further used for offline analysis as well as for real time alerting and fault diagnosis.



(a) Power outage vs Time (b) Power outage duration

Figure 4: Power outages observed during deployment

Similar to prior literature, reporting WiFi discontinuity in the homes in the USA [13], we also observed that one WiFi router did not provide complete coverage for our deployment. We thus used 3 Netgear JNR1010[11] routers, where the router on the first floor acted as the host and the routers on the ground and the second floor were bridged to it.

It must be noted that theoretically our deployment could have been done with fewer RPis. However, in order to ensure that sensing systems are independent of each other and for home aesthetics, we chose to use additional RPis.

4 How is this deployment different?

We now discuss some of the key unique aspects brought forward from our deployment.

Unreliable electrical grid: Load shedding or rolling blackout is a commonplace in the developing countries. Specifically in India, power outages are common in summers when the load is high due to excessive usage of air conditioners. Excessive load and poor infrastructure also leads to significant fluctuations in the supply voltage. Various statistics, collected from our deployment, further establish these aspects. We used multiple sources, e.g. Unix *last* command (providing a history of boot times) on the desktop server and common missing data duration from multiple sensors, to find power outages reliably.

Figure 4a shows power outages in aggregated number of hours per day during May-July 2013. One of the days experienced power outage for approx. 12 hours. Figure 4b shows the distribution for duration of all power outages. A total of 107 power outages were reported in the 61 day period reported here, with average power outage of approx. 1 hour.

Figure 5a and 5e show voltage and frequency fluctua-

[6] Forecast, World Weather, Open Weather Map

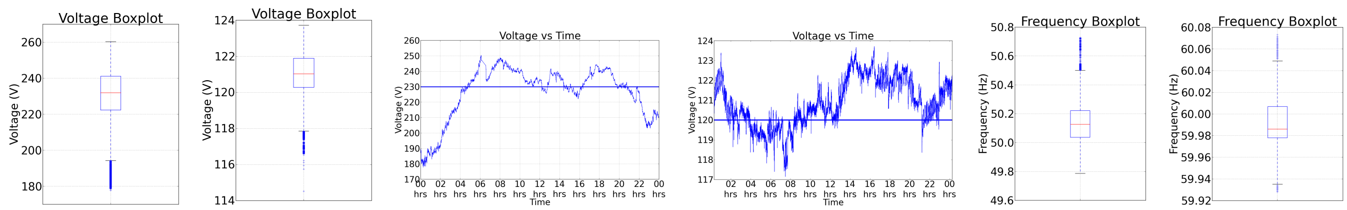
[7] www.raspberrypi.org

[8] www.ionics-ems.com/plugtop/stratus.html

[9] www.github.com/bashwork/pymodbus

[10] www.code.google.com/p/open-zwave

[11] www.support.netgear.com/product/JNR1010



(a) Voltage fluctuations in a week (ours) (b) Voltage fluctuations in a week (Smart*) (c) Voltage fluctuations on one of the days (ours) (d) Voltage fluctuations on one of the days (Smart*) (e) Frequency fluctuations in a week (ours) (f) Frequency fluctuations in a week (Smart*)

Figure 5: Comparison of our data with Smart* deployment done in the USA

tions for a week in June from our deployment. Comparing these observations with the voltage and frequency fluctuations for a week from Smart* dataset, shown in Figure 5b and 5f respectively, we observe that our deployment shows a lot more variations in both of these parameters. Figure 5c and 5d show voltage fluctuations on one of the days from our deployment and the Smart* dataset respectively. In our deployment, we observed that the voltage was usually well below the rated voltage around 10 AM in the morning and around midnight. Significant amount of NILM literature uses current data for disaggregation, inherently assuming almost fixed voltage from the grid.

Learning: *Observed voltage fluctuations motivate two important aspects - 1. Load measurement devices should measure both current and voltage and not only current as is done in many of the CT based devices; and 2. When performing disaggregation, normalization to account for voltage fluctuations (as was proposed in the original NILM work [12]) is important.*

Due to unreliable nature of the grid, we wanted to ensure that all our systems were capable of automatically restarting after a power outage and the complete system achieves the same state as it was in before the outage. Correspondingly, data collection and upload scripts were executed as part of system startup process. This feature further provided us with another advantage - when the system was observed to be down, we just asked the home occupant to power cycle the system. This ensured that there was minimal data loss till the time researchers could visit the site and diagnose the fault. With several devices, each with its diverse sensing, computation and communication requirements, ensuring that the system recovers to the same state, as before the outage, was observed to be non-trivial.

Learning: *A robust building monitoring and control system should be tested for appropriate system recovery after power failure.*

Unreliable network connectivity: While India has one of the fastest growing internet user base, only 11% of the total population is connected to internet (the corresponding figure in the USA is 78%) [16]. We observed internet to be either unavailable or having slow intermittent connectivity throughout our deployment. We collected network statistics by performing 15 internet ping requests every 15 seconds and computed the corresponding packet drop. Figure 6a shows that packet drop of up to 22% was observed on certain days. The average packet drop per day was approx. 6%. Figure 6b shows a CDF plot of % packet drop. It can be seen that approx. one-fifths of total days reported greater than 10% packet loss.

Learning: *For a building monitoring and control system to scale up for the context of developing countries, with unreliable internet connectivity, an architecture that does not completely rely on good internet connectivity is important.*

We correspondingly propose Sense Local-store Upload architecture, as discussed in Section 5, to address for unreliable internet connectivity.

Importance of meta data collection: We collected metadata associated with electrical appliances, such as appliance name, age, mode of usage (eg. air conditioner set temperature), throughout our deployment. We believe this detailed metadata can enhance NILM and can provide useful insights for conserving electricity. An anecdotal evidence illustrates the utility of meta data collection. The home refrigerator was repaired on 2^{nd} July. Figure 7a and Figure 7b show the active power consumption before and after the repair. We observed that after repair, the refrigerator was set to the lowest temperature setting by the service professional, while before repair it was set to the highest temperature setting. After the repair, the refrigerator was found to be consuming 1KWh more per day (which is 140% above the normal). The residents configured their refrigerator again to the lowest temperature setting after we informed them about the increased energy usage, resulting in normal power consumption.

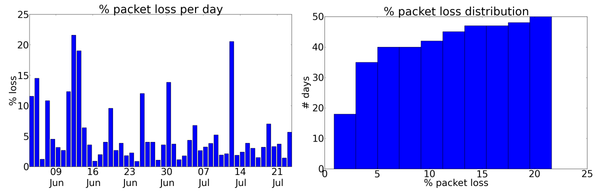
Load specifics: Appliance usage varies significantly in India compared to the USA and the Europe.

Decentralized control: Temperature control is often decentralized in the Indian settings i.e. a separate air conditioner is used for every room and a separate geyser (a water heating device) is used for each bathroom. From our deployments, we observed that these air conditioners and geysers account for up to 70% and 50% of the overall home electricity in summers and winters respectively. Thus, small improvements in efficiency of these two appliance can significantly lower the home electricity consumption. From NILM perspective, these loads are simpler to disaggregate due to their high power consumption and repeated patterns (shown by the compressor in the air conditioner).

Learning: *Even a simple NILM approach can potentially provide useful insights towards energy reduction in the Indian context.*

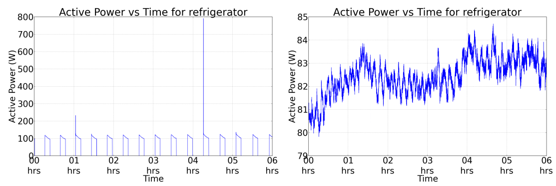
In our recent NILM work- INDiC [8], we illustrate how simple approaches such as Combinatorial optimization can give good NILM results when electricity consumption at different granularities is measured.

Energy embedded water: Additional energy, in the Indian context, is embedded into the water at the home level due to its low pressure and poor quality. Water pumping and filtering are the two activities whose scope spans across both water and electricity dimensions. Due to limited supply and line pressure, a water motor is used to pump the water up to the water tank on the roof. We observed that to fill 1 liter of water into the tank, it took 8 seconds without the motor (dur-



(a) % internet packet drop vs time (b) % internet packet drop CDF

Figure 6: Unreliable internet



(a) Before repair (b) After repair

Figure 7: Refrigerator power consumption

ing the times of maximum pressure) and 4 seconds when the motor was used. With power consumption of 700 W for the electric motor, every one hour usage will result in additional energy being embedded into the water due to its intermittent supply. Due to poor quality of supplied water (and often usage of ground water for drinking purposes), Reverse Osmosis based water filters are a commonplace in big cities in India. We observed that water filter takes approx. 1 minute to filter 1 liter of water and consumes 40 W in the process.

Learning: *Observing water consumption, together with the electricity consumption, can provide additional useful insights in usage and consumption patterns.*

Appliance switching from mains: Another interesting distinction in the Indian context is that each plug point has an associated switch and people are often conscious about turning the appliance off from the switch rather than keeping them in the standby (as is the usual practice in the USA). We observed that the jPlugs attached to the kitchen appliances such as microwave, when used for less than 1 minute, did not report data. This was due to the fact that jPlug setup takes roughly a minute to establish WiFi connectivity before starting the data collection. For small usage, before jPlug could start data collection, the appliance was turned off.

We also imported ZWave based plug monitors and controllers (with EU frequency) for plug level monitoring. After their initial deployment, we realized that the default state of the plug monitors was chosen as off (when powered manually from the switch), possibly to avoid the peak switching current. This implied that even after switching them on from the mains, unless they are switched on from the software (or with a separate ZWave based switch), they will not turn on the appliance. Since many of the loads in the Indian context are not always on and are controlled via mains, such plug sockets did not result in seamless usage.

Learning: *Plug level monitoring should account for the short appliance usage and power off from the main switch to ensure robust and reliable data collection, together with seamless usage.*

5 Sense Local-store Upload Architecture

Middleware systems such as sMAP [9], BuildingDepot [2] and SensorAct [4] have been proposed in the past for sensor data collection from deployments pertaining to buildings. However, we found that they do not sufficiently address the requirements of our deployment context e.g. intermitted network connectivity and repeated power failures. Motivated by our experience as well as previous work from other researchers [13], where importance of simplifying the architecture are proposed, we propose Sense Local-store Upload (SLsU) model. SLsU involves two main ideas - association of local storage (using SBCs) distributed across each sensing point and periodic data upload (from SBC to server, and from server to cloud). As discussed in Section 3.2, we used 6 SBCs (and local storage on the Android phones) to connect to multiple sensors spread through our deployment. Data collected from the sensors was **locally stored** in the form of comma separated value files (CSV), in SBCs and **periodically uploaded** to the main desktop server. In the case when upload failed, it was retried after a fixed time duration. Each SBC was provisioned with sufficient flash based local storage to accommodate sensor data for a few days, to account for persistent upload failure.

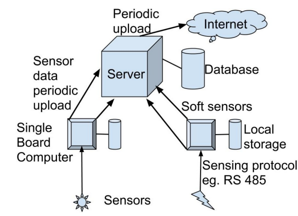


Figure 8: Sense Local-store Upload architecture

Web applications running on the server allowed residents to locally visualize their data from multiple sensing streams. Data from the server was periodically replicated to the cloud, allowing researchers to remotely visualize the data and maintain the deployment. Figure 8 illustrates the SLsU architecture. The salient features of SLsU architecture are:

Decoupled sensing and data upload: ensuring that an error in data upload does not impact the sensing and vice versa, thus avoiding data loss due to network (even the local inhome WiFi) failure.

Reduced dependence on always-on connectivity: Internet is required **only** when outside researchers wish to view data in near-realtime. Internet failure does not have any impact on the deployment data collection. The periodic nature of our uploads ensured that data would be uploaded when internet connectivity is re-established. Local storage, on SBC, further ensures reliable data collection, even in the cases of server failure.

Reduced load on server: Periodic upload of data (in larger volumes) results in reduced computation and bandwidth requirements for the SBCs and the server.

We provide anecdotal evidence to illustrate utility of SLsU in preventing data loss. One of the researchers involved, accidentally killed the server script responsible for collecting water consumption data. However, when the problem was rectified a week later, all the data for the previous week, which had been locally stored on the RPi, was collected within an hour on the server.

6 Hitchhiker's guide revisited

We now present some of the prominent similarities, albeit with some additional unique perspectives, with prior deployment experiences, most specifically - "The Hitchhiker's

Guide to Successful Residential Sensing Deployments" [13].

Homes are hazardous environments: We observed that one of our multisensors repeatedly failed after every power outage. We, eventually, figured that this behavior was due to the fact that this multisensor was put on the battery backup plug (commonly available in many homes to guard against intermittent power supply). During a power outage, this multisensor went to *sleep* state, in absence of communication with the ZWave controller (which went down during outage). Correspondingly, when power resumed, the ZWave controller assumed the node to be *dead*. We resolved this by putting the multisensor on the main plug as well. Although we used zip-ties extensively throughout the deployment to prevent hanging wires, we observed data loss in one of the ZWave multisensor and an Android phone, which went out of power due to wire snag (shown in Figure 10b). Even after a month of rigorous testing in the lab before we started the deployment, we raised 60 new service complaints, when we moved the deployment to the home.

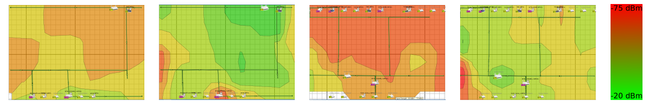
Aesthetics matter: As stated in the previous work, sensor LEDs can be bothersome to the occupants, particularly in the night. Our deployment introduced 63 LEDs in the home. Figure 10a shows our sensor LEDs blinking in the night. Choosing appropriate sensor location sufficed for the current deployment. However, for the future, we intend to case the sensors appropriately to ensure that home occupants are not disturbed. The residents also complained of buzz like sound coming from our desktop server. This noise was due to the dust clogging in the desktop. Dust is a uniquely common aspect in the Indian setting.

Learning: *Monitoring and control systems, aiming for long life deployments should include routine maintenance, to guard against dust and other environmental problems.*

Homes are not designed for sensing: We observed much more noise in the data collected from our ground floor MCBs than from the MCBs on the first floor. This was attributed to the fact that the MCBs on the ground floor were close (as shown in Figure 10c) to each other causing interference in our CT monitoring circuit. A workaround could have been to get additional cabling done, but the residents were not inclined for such changes.

Redundancy-Accounting for sensor failure: During our deployment 3 jPlugs and 1 multisensor stopped functioning. We had accounted for such failure keeping reserve sensors.

Homes have poor connectivity: During the preliminary phase of our deployment, we first tried to connect our sensors to the existing networking infrastructure in the home. Already existing WiFi router was on the first floor and we observed poor signal strength on the ground and the second floor. We used Ekahau Heat Mapper[12] to map WiFi signal strength. Figure 9a and 9c show the WiFi heatmap produced with the home router placed on the first floor. We observed that large regions inside the home show poor signal strength. We bridged additional routers on the ground and the second floor with the existing first floor router. Figure 9b and 9d show the corresponding WiFi heatmaps produced after the introduction of bridged routers. Additional routers signifi-



(a) Ground floor (without additional router) (b) Ground floor (with additional router) (c) Second floor (without additional router) (d) Second floor (with additional router) (e) Scale

Figure 9: WiFi Heatmap, with and without the additional routers, for the ground and the second floor (Best viewed in color)

(a) Glowing LEDs in night (b) Wire snag leading to data loss (c) Closely placed MCBs causing interference

Figure 10: Illustration of common problems

cantly improved WiFi coverage across the home, shown by increased green regions (signifying better signal strength as per the scale shown in Figure 9e).

7 iAWE dataset and code release

We are releasing our dataset **iAWE** for open use. iAWE consists of sensor data (ambient, water and electricity) worth 73 days. It also contains fully labeled data for 1 day for 63 electrical appliances, 18 water fixtures and ambient conditions across 6 rooms. We further provide a detailed metadata log for all the electrical appliances, including, approx. date of purchase, mapping to MCB, star-rating and rated power. All the appliance ON-OFF events can be easily captured using the plug level data collected from jPlug and Current Cost CT. To facilitate engaging research, we provide a web portal[13] for researchers to visualize different data streams.

Figure 11 illustrates the advantage of detailed labeling together with use of multi-modal sensing. One of the occupants returned back home around 7:15 PM, an event captured synchronously by ground floor motion sensor, electricity meter and ground floor light sensor. The occupant (who was alone in the home) then heated up some food in the oven. While the oven jPlug failed to record it (due to the small usage time), this event is captured in the electricity meter data (spikes reaching up to 1800 watts). Thereafter, the occupant went to the first floor, turned on lights and switched on the laptop (illustrated through measurements from motion, light sensor on the first floor and laptop jPlug. After approx. an hour, the occupant switched on the air conditioner which is captured by the jPlug connected to the air conditioner as well as a spike in the meter measurements. Correspondingly, the temperature sensor in the room started observing reduced temperature values.

We also publicly release our codebase which includes SLsU implementation, scripts for collecting data from different sensors, database schemas, soft-sensors, startup scripts and the fixes we developed for common problems on RPi. Our codebase and dataset is available on Github[14].

8 Conclusions and Future Work

In this paper, we present our experiences with an extensive residential deployment monitoring electrical, water

[12] www.ekahau.com/products/heatmapper/overview.html

[13] http://www.energy.iiitd.edu.in/iawe

[14] http://github.com/nipunreddevil/Home_Deployment

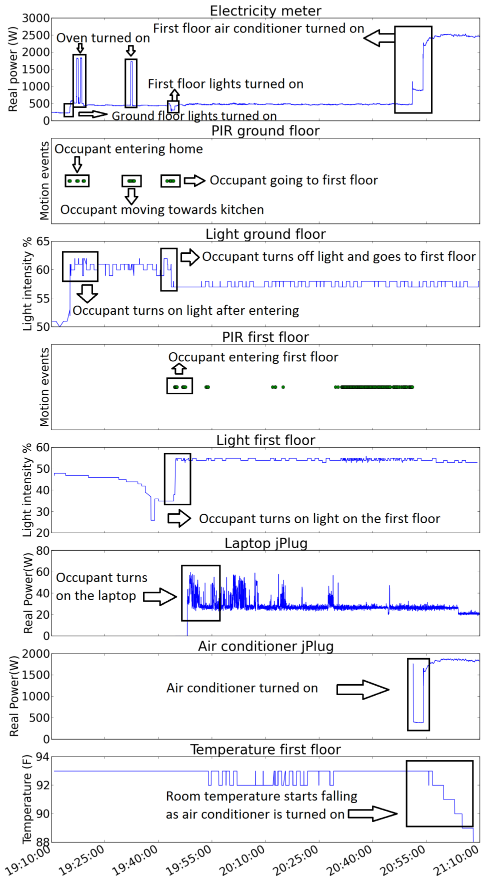


Figure 11: iAWE: Labeled multi-modal dataset

and ambient parameters in Delhi, India. To the best of our knowledge, this is the first extensive residential deployment in a developing country. We present key aspects of our deployment and discuss the corresponding impact on the design of building monitoring and control systems that aim to scale across diverse contexts offered in the developing and the developed countries. Some of the unique aspects, impacting the systems development in building energy domain, include - unreliable electrical grid, unreliable network connectivity, decentralized electrical loads and energy-water nexus within a home. We further discussed the similarities in our learning with prior work (done in the USA), demystifying the home environment for energy and water related deployments in the Indian context.

Frequent power outages and unreliable internet motivated us to develop the proposed sensing architecture: SLsU, which accounts for these pitfalls by introducing local storage and periodic upload. Such an architecture can be of particular importance for scaling the building monitoring and control systems for applicability across diverse contexts. We are in the process of installing our sensors across multiple other homes in Delhi. Detailed, annotated dataset from the deployment (iAWE) is released for public use.

Acknowledgment

The authors would like to thank TCS Research and Development for supporting the first author through PhD fellowship. We would also like to thank Department of Electronic and Information Technology (DEITy), Government of India for funding the project (Grant Number DeitY/R&D/ITEA/4(2)/2012).

9 References

[1] Y. Agarwal, B. Balaji, S. Dutta, R. K. Gupta, and T. Weng. Duty-cycling buildings aggressively: The next frontier in hvac control. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, pages 246–257. IEEE, 2011.

[2] Y. Agarwal, R. Gupta, D. Komaki, and T. Weng. Buildingdepot: an extensible and distributed architecture for building data storage, access and sharing. In *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pages 64–71. ACM, 2012.

[3] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Bergés. Blued: a fully labeled public dataset for event-based non-intrusive load monitoring research. In *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability, Beijing, China*, pages 12–16, 2012.

[4] P. Arjunan, N. Batra, H. Choi, A. Singh, P. Singh, and M. B. Srivastava. SensorAct: A Privacy and Security Aware Federated Middleware for Building Management. In *Fourth ACM Workshop On Embedded Sensing Systems For Energy-Efficiency In Buildings*, BuildSys, 2012.

[5] 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. In *The 1st KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, 2011.

[6] S. Barker, A. Mishra, D. Irwin, P. Shenoy, and J. Albrecht. Smartcap: Flattening peak electricity demand in smart homes. In *Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*, pages 67–75. IEEE, 2012.

[7] N. Batra, P. Arjunan, A. Singh, and P. Singh. Experiences with occupancy based building management systems. In *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*, pages 153–158, 2013.

[8] N. Batra, H. Dutta, and A. Singh. Indic: Improved non-intrusive load monitoring using load division and calibration. In *Machine Learning and Applications (ICMLA), 2013 Twelfth International Conference on*. IEEE, 2013.

[9] S. Dawson-Haggerty, X. Jiang, G. Tolle, J. Ortiz, and D. Culler. smap: a simple measurement and actuation profile for physical information. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems*, pages 197–210. ACM, 2010.

[10] M. Evans, B. Shui, and S. Somasundaram. *Country Report on Building Energy Codes in India*. Pacific Northwest National Laboratory.

[11] T. Ganu, D. P. Seetharam, V. Arya, R. Kunnath, J. Hazra, S. A. Husain, L. C. De Silva, and S. Kalyanaraman. nplug: a smart plug for alleviating peak loads. In *Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet*, page 30. ACM, 2012.

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

[13] T. W. Hnat, V. Srinivasan, J. Lu, T. I. Sookoor, R. Dawson, J. Stankovic, and K. Whitehouse. The hitchhiker's guide to successful residential sensing deployments. In *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*, pages 232–245. ACM, 2011.

[14] J. Z. Kolter and M. J. Johnson. Redd: A public data set for energy disaggregation research. In *proceedings of the SustKDD workshop on Data Mining Applications in Sustainability*, pages 1–6, 2011.

[15] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic. AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research. In *Electrical Power and Energy Conference (EPEC), 2013 IEEE*, pages 1–6, 2013.

[16] M. Meeker. Internet trends at stanford bases. KPCB, 2012.

[17] A. Mishra, D. Irwin, P. Shenoy, J. Kurose, and T. Zhu. Smartcharge: cutting the electricity bill in smart homes with energy storage. In *Proceedings of the 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet*, page 29. ACM, 2012.

[18] O. Parson, S. Ghosh, M. Weal, and A. Rogers. Non-intrusive load monitoring using prior models of general appliance types. In *26th AAAI Conference on Artificial Intelligence*, 2012.

[19] L. V. Thanayankizil, S. K. Ghai, D. Chakraborty, and D. P. Seetharam. Softgreen: Towards energy management of green office buildings with soft sensors. In *Communication Systems and Networks (COMSNETS), 2012 Fourth International Conference on*, pages 1–6. IEEE, 2012.