

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

Nipun Batra, Manoj Gulati, Amarjeet Singh
Indraprastha Institute of Information Technology, Delhi
{nipunb,manojg,amarjeet}@iiitd.ac.in

Abstract

Residential buildings contribute significantly to the overall energy usage across the world. Residential deployments provide insights into home energy consumption and occupant behavior, thus paving the way for building systems that can reduce energy consumption. Much of the prior work in residential deployments is from the developed countries. However, 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 one home in Delhi, measuring electrical, water and ambient parameters.

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 unreliable nature of electrical grid and internet in such settings, we present *Sense-Local Store-Upload* architecture. While providing several unique aspects, our deployment further validates the common considerations from similar residential deployments done in the Western context. We believe that the discussion in this paper will help in the development of building monitoring and control systems, that can scale across diverse deployment scenarios, prevailing in both developed and the developing countries.

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

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Buildsys'13, November 11–15, 2013, Rome, Italy.
Copyright © 2013 ACM 978-1-4503-1169-4 ...\$10.00

1 Introduction

Buildings (both residential and commercial) account for significant proportion (more than 30%) of overall energy consumption globally [9]. Of this energy consumption, a large proportion is contributed by residential buildings. 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, upgrading the physical plant of the building with more energy efficient solutions, 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 (gas, water and electricity). These deployments also provide data sets which can be leveraged for developing and testing suitable control strategies, that are otherwise complex to undertake in a real occupied building. In the recent past, several datasets, such as REDD [13], BLUED [3], Smart* [5], monitoring household electricity and ambient parameters, have been publicly released. Several building monitoring and control research work has since used these datasets to prove the validity of the work for real life settings [16, 6].

However, most 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. 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 one of the homes in Delhi, India. This home is instrumented with - smart electricity and water meters, plug level load monitoring for major appliances in the home and ambient parameters across every room, among others.

To the best of our knowledge, this is the first such exten-

sive deployment outside any developed country. We discuss in detail the unique aspects of our deployment, that are characteristic of systems and infrastructure pertaining to building energy in developing countries. We further discuss common aspects in residential deployments, highlighted in previous work, and how we addressed them. 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, for wider applicability across different contexts.

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 previous work [1, 7] have shown the scope for significantly reducing HVAC and plug load energy consumption in respective settings. Residential deployments have been studied in the past highlighting the unique challenges involved in residential sensing.

Several residential deployments have specifically targeted residential sensing for modeling and inference, specifically pertaining to Non Intrusive Load Monitoring (NILM). Kolter et al. [13] 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 55 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 61 days at the time of writing) and uses a mix of COTS and customized hardware (due to non-availability of COTS, for everything we wanted to monitor, in the Indian context).

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 [15]. 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 and water consumption at different granularity.

Hnat et al. [12] 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 activities of daily life (ADLs). While providing unique aspects of our deployment, we further es-

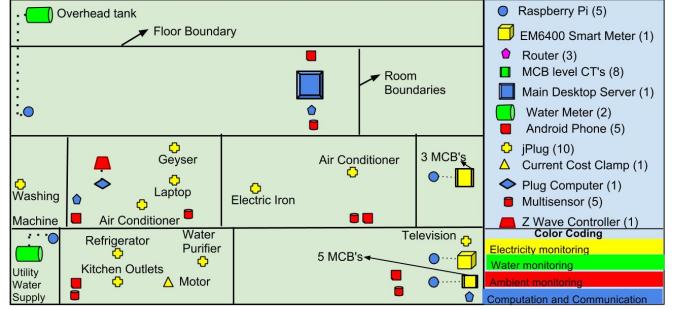


Figure 1: Schematic showing overall home deployment

tablish the commonalities in our deployment with the learning discussed in their work. All of the deployments discussed above are in the context of developed countries. While residential deployments anywhere across the world are challenging, our deployments highlight some unique challenges specific to 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 granularities, in a home in Delhi, India during May-July 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. The ecosystem of available sensing options that restrict the possible deployments; 2. Energy and water consumption patterns; and 3. Grid and network reliability. Figure 1 shows the deployment of these sensors in a 3 storey home, together with the required computing and communication infrastructure.

3.1 Sensing Infrastructure

For sensing, we took a “leave no stone unturned” approach, where we chose to monitor as many physical parameters (ambient conditions, electricity usage and water usage) and non-physical parameters (such as network strength and network connectivity) 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 granularities- electricity meter 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 appliance (see Figure 3a for illustration).

1. **Meter level:** Schneider Electric EM6400¹ smart meter is used to instrument the main power supply (see Figure 2a). Necessary parameters including voltage, current, frequency and power (apparent, real, reactive) were read at 1 Hz using Modbus over RS485-serial link

¹www.goo.gl/01edPS

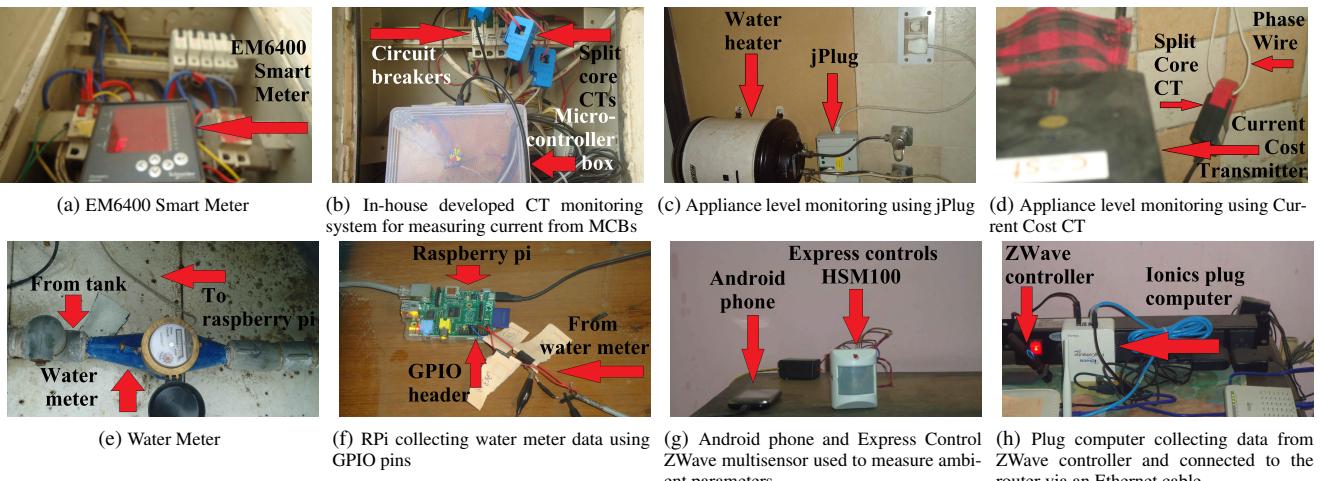


Figure 2: Sensing, computation and communication equipment used for deployment

provided by EM6400. A robust deployment was easily undertaken by a paid electrician thus demonstrating that the support required for large scale smart metering deployments exists in the Indian context.

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 low cost microcontroller platform to measure current output from CTs at 125 KHz, calculate the RMS value and communicate it over serial UART interface (Figure 2b illustrates CTs monitoring 3 MCBs on the first floor MCB box) to a single board computer. A total of 8 CTs were used to monitor different MCB circuits in the home.
3. **Appliance level:** Similar to circuit level monitors, there are no good commercial options for plug level monitors in Indian context. Correspondingly, we worked with our collaborators who in-house developed jPlug² to measure individual appliance power consumption. Ten jPlugs were used to monitor different plug-load based appliances across the home. jPlug sits in between an appliance and the socket and provides multiple parameters (including voltage, current, phase and frequency) at 1 Hz resolution. We also used Current Cost based CT 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). jPlug and Current Cost CT deployment are shown in Figure 2c and Figure 2d respectively. jPlug uploads measurement data using HTTP POST. Current Cost exposes apparent power data over the USB port.

Water monitoring: There are a few differences in home water distribution in India when compared to developed countries. Due to non availability of 24x7 water supply, overhead water tanks (typically 1000 liters capacity) are used to store water, when available, and supply it for home activities throughout the day. Due to low water pressure, electric motors are used to pump the water for storage for the times when the supply is available. Figure 3b illustrates the water flow distribution, together with the placement of water meters. One water meter is placed at the inlet (coming from

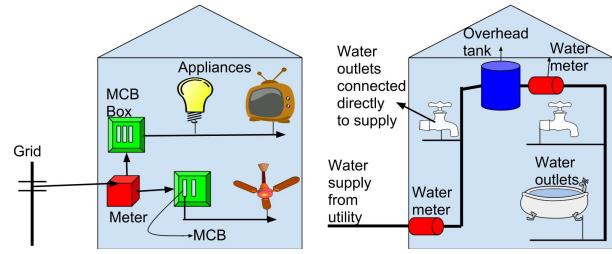


Figure 3: Electricity and water flow inside a home and different granularities at which these parameters can be monitored.

the utility) and another one at the outlet from the water tank (flowing downwards).

Due to prohibitive cost and procurement difficulty for digital water meters in India, we chose to use Zenner Aquameter's multijet³. The multijet uses pulse output generated through a 4-20 mA current loop. For the water meter connected to the utility, over a 0.5 inch diameter pipe, a pulse is generated 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: We used ZWave based Express Controls HSM100⁴ multisensors for monitoring motion, light and temperature across 5 rooms in the home. Till date, to the best of our knowledge, no commercial ZWave based sensor is available working on Indian frequency (865.2 MHz). We correspondingly imported EU frequency (868.4 MHz) devices and used it 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. We also placed 1 Android phone at a fixed location in each room and ran FunF journal application⁵ to log ambient parameters such as light and sound level every 30 seconds for 5 seconds.

Miscellaneous: Android phones, in addition to measuring

³www.aquametwatermeters.com/multijet.html

⁴<http://goo.gl/Bszg0u>

⁵<http://www.funf.org/journal.html>

²A variant of nPlug [10]

Table 1: Details of sensing infrastructure used in our deployment

Sensor name	Procurement	Sampling frequency	Resolution	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

ambient conditions, were also used to scan and log Bluetooth, WiFi and GSM networks. All 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 on the SD card. External weather conditions, such as temperature, humidity and wind speed, were also logged every 10 minutes using publicly available APIs from weather monitoring stations⁶. Additionally, network traffic was sensed to monitor connectivity and signal strength.

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

3.2 Communication and Computation

Our sensing infrastructure, described in Section 3.1, involves multitude of communication channels. Different computing platforms - microcontrollers, single board computers (SBCs) (e.g. Raspberry Pi⁷ (RPi) and Ionics Stratus⁸) and desktops are used for data collection. SBCs served as a good low cost alternative (25 \$) constituting 700 MHz ARM processors, support for Linux based distributions and allow for coding in higher level languages such as Python. We used 5 RPi and 1 Ionics Stratus plug computer as SBCs and a 2 GHz Desktop PC running Linux, as the main local server.

One RPi was connected to EM6400 using RS485-USB converter. We developed a custom Python program based on pyModbus⁹ to collect the desired parameters. All the collected parameters, together with the UTC timestamp is collected using *Sense-Local Store-Upload* (SeLStUp) model, described in detail in section Section 4. On similar lines, USB output from Current Cost (providing data in xml format) is received on another RPi and is communicated to the desktop server. Although Current Cost has its own cloud based APIs, we chose to collect data from it locally, in order to avoid data losses occurring due to network failures.

Two RPi were used to connect to the microcontroller based circuit monitor to read the data over the serial port and process it further using SeLStUp. GPIO headers on two more RPis were used to interface with each water meter respectively for collecting pulse output over the 4-20 mA current loop interface. We initially wrote an interrupt driven program in Python to detect 10 liter and 1 liter events for the tank and supply water meters respectively. 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 a frequency of 5 Hz to obtain GPIO status. The RPi collecting Current Cost data was also used to collect data from the water meter attached to the supply.

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 OpenZWave¹⁰ to collect temperature and light information on per second basis, and motion information based on events. 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 locally collected data on each Android phone was performed every 15 days. Electricity failure and unreliable internet are well known problems in India and are highlighted in Section 5. Owing to these problems, we chose to collect weather data from 3 different weather stations, in one of the servers hosted by our collaborators in the USA.

A lot of issues, pertaining to SBCs, were observed in our deployment. As an example, the OpenZWave program that we used created log files for its own diagnostics, that eventually consumed the 512 MB flash drive space on the plug computer. This was fixed eventually by deleting older logs. Such problems encouraged us to develop soft-sensor [17] streams, whereby we collected hard disk space, ping success, CPU utilization, remaining RAM space and temperature of the processor, for all the computing devices including the server at regular intervals. These soft-sensor streams can be used for offline analysis as well as for real time alerting and fault diagnosis. Similar soft-sensor streams have also been used in the previous work [12] for raising alerts.

Similar to prior research, reporting WiFi discontinuity in homes in the USA [12], we also observed that one WiFi router did not provide complete coverage for our 3 storey home. We thus used 3 Netgear JNR1010¹¹ 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.

⁶Forecast, World Weather, Open Weather Map

⁷www.raspberrypi.org

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

⁹www.github.com/bashwork/pymodbus

¹⁰www.code.google.com/p/open-zwave

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

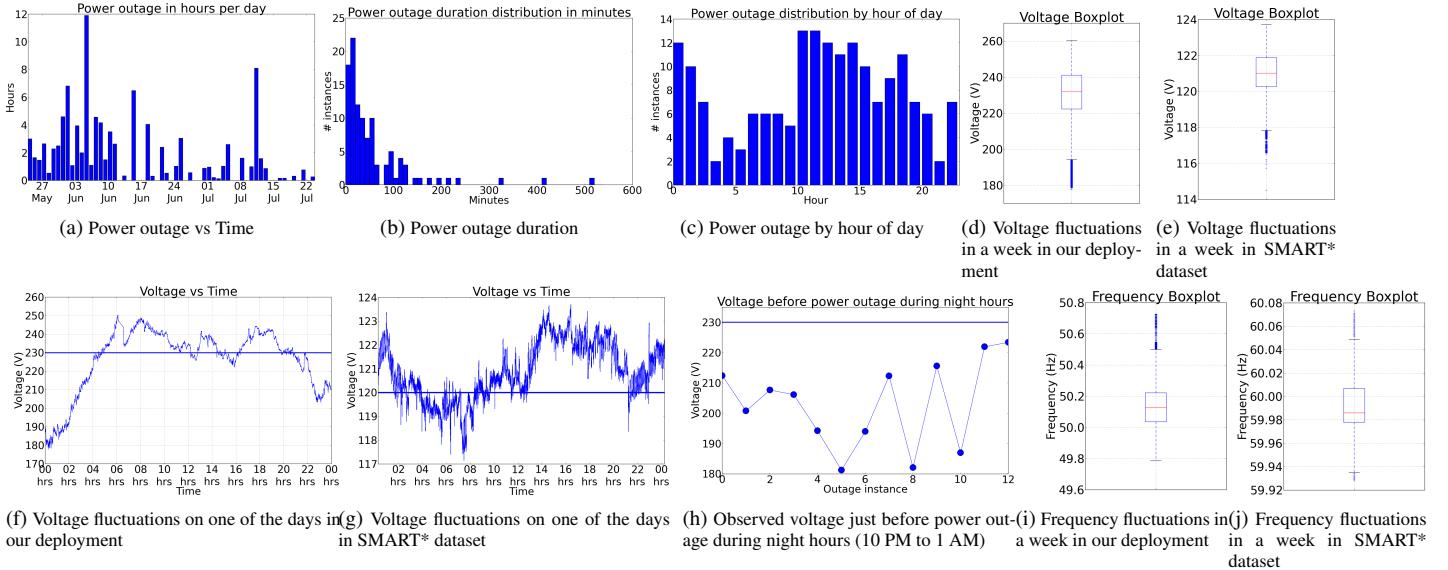


Figure 4: Unreliable grid. Rated voltage in India and US are 230V and 120V respectively. Rated frequency in India and US are 50Hz and 60Hz respectively

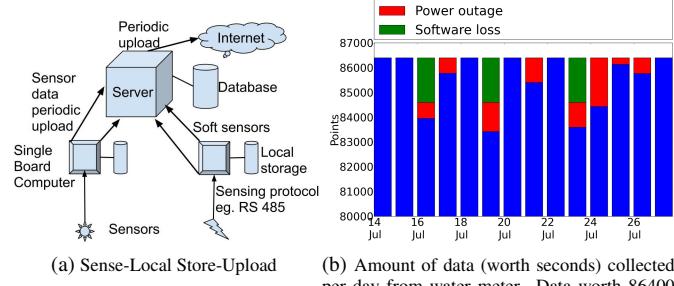


Figure 5: Sense-Local Store-Upload architecture and its impact on ensuring no data loss

4 System Architecture

Middleware systems such as sMAP [8], BuildingDepot [2] and SensorAct [4] have been proposed in the past for sensor data collection. However, we found that they do not sufficiently address the requirements of our deployment context e.g. faulty internet and repeated power failures. Motivated by our experience as well as previous work from other researchers [12], where importance of simplifying the architecture are proposed, we propose Sense-Local Store-Upload (SeLStUp) model. SeLStUp involves two main ideas- local storage (on SBCs) and periodic data upload (from SBC to server, and from server to cloud). 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 case the 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.

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 5a illustrates the SeLStUp archi-

ecture. The salient features of our architecture are:

Decoupled sensing and data uploading: ensuring that an error in data uploading does not impact the sensing and vice versa, avoiding data loss due to network (even local in-home WiFi) failure. The design choice of decoupling comes from the well established principles of software engineering.

Minimal internet requirement: Internet is required **only** when outside researchers wish to view and analyze collected data in near-realtime. Internet failure does not have any impact on robust sensor data collection. The periodic nature of our uploads ensured that data would be uploaded 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. While it comes at the cost of increased storage requirements on SBCs, our analysis shows that 1 GB of local storage will suffice for most of the practical scenarios.

In our experience, we found that the novel features of Sense-Local Store-Upload prevented data loss due to internet, server downtime, etc. We had accidentally killed the server script for taking CSV data from RPis collecting water data. However, when we rectified that, within 1 hour all the data for previous 1 week, which had been locally stored on the RPi was eventually collected on the server. Figure 5b shows the amount of data collected from the water meter during that week.

5 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 black-out is 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 several sources, such as, Unix *last* command (which provides a history of boot times) on the desktop server, common missing data duration from multiple sensors, to find power outages reliably. Figure 4a shows power outages in number of hours per day during June-July 2013. One of the days experienced power outage for up to 12 hours (the situation is worse outside of the major cities in India, with electricity coming only for a few hours every day). Figure 4b shows the distribution of power outages duration. A total of 107 power outages were reported in the 61 day period. While the average power outage lasted about 1 hour, there were outages lasting upto 9 hours. Figure 4c shows the power outage distribution by hour of the day, showing maximum outages around 10 AM in the morning and around midnight. These times correspond to office time and night time when air conditioners in offices or homes are turned on leading to excessive demand on the grid. Figure 4i and Figure 4d show frequency and voltage fluctuations for a week in June from our deployment. We compare them with the frequency and voltage fluctuations for a week from SMART* dataset shown in Figure 4j and Figure 4e respectively. We can observe that there is a lot more fluctuation in both these parameters in the Indian setting. Figure 4f and Figure 4g show voltage fluctuations on one of the days from our deployment and the SMART* dataset respectively.

Significant amount of NILM literature uses current data, as measured at different granularities, for disaggregation. This has an inherent assumption of almost fixed voltage from the grid. *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 [11]) is important.* Figure 4h shows that voltage just before the power outage (for a selected sample of outages occurring in the from 10 PM to 1 AM). We observe that the voltage is well below the rated voltage of 230 V. This is in coherence with previous work [10], which hypothesized that frequency and voltage measured at the home level are potential indicators of the load on the grid.

Due to unreliable nature of the grid, we wanted to ensure that all our systems were capable of automatically restarting after a power outage. Data collection and upload during startup. This feature further provided us with another advantage - when the system was observed to be down, we just asked the home occupant to cycle power to the system. This ensured that there was minimal data loss till the time researchers could visit the site and diagnose the fault. Appropriate logs in the scripts ensured that the fault could be diagnosed in an offline manner. With several devices, each with its diverse sensing, computation and communication requirements, ensuring this reliability at the system level was observed to be non-trivial. Our experiences demonstrate how *a robust building monitoring and control system can ensure system recovery after power failure.*

Unreliable network connectivity: While India has one of the fastest growing internet user base, only 11% of the total

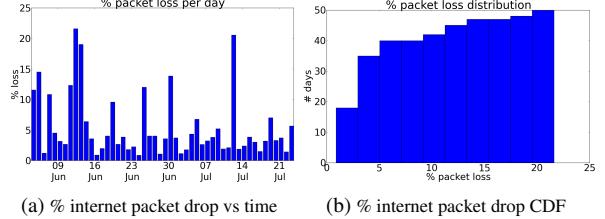


Figure 6: Unreliable internet

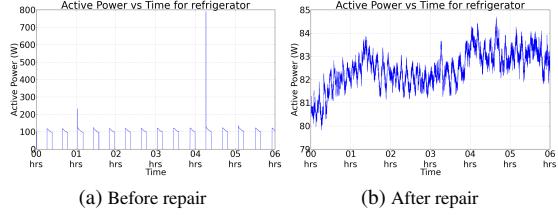


Figure 7: Refrigerator power consumption

population is connected to internet (the corresponding figure in USA is 78%) [14]. We observed internet was either unavailable or had a slow intermittent connectivity periodically throughout the day. This was the prime motivation behind the proposed SeLStUp model, discussed in Section 4. Based on our experience with the proposed model, we believe that for a building monitoring and control system to scale up for the context in developing countries, *the SeLStUp model should be adopted rather than completely relying on good internet connectivity.*

To measure the network connectivity, we collected network statistics by performing 15 internet ping requests every 15 seconds and collected the statistics on 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.

Importance of meta data collection: We collected metadata associated with electrical appliances throughout our deployment. We collected parameters such as appliance name, age, mode of usage (eg. air conditioner set temperature). 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 2nd 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 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, after which, its power consumption returned to previous levels. This metadata can be useful to NILM approaches which model duration of appliance usage. **To make the case even stronger, i am thinking of putting the need for measurement calibration for different h/w**

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

mated loads (e.g. dishwasher that work without any human intervention) is fewer and temperature control is often decentralized (i.e. a separate air conditioner at the room level and separate water heating, using a device commonly called as geyser, at the bathroom level). From our deployments, we observed that 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 very simple to disaggregate due to their high power consumption and repeated patterns (shown by the compressor in the air conditioner). Correspondingly, even a simple NILM approach can potentially provide useful insights towards energy reduction in the Indian context.

Additional energy 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 (during 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 the consumed water is the ground water and not the utility supplied water), Reverse Osmosis(RO) based water filters are 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. Thus, the water filter would consume 20 Wh energy for filtering 30 liters of water (typical drinking water usage).

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 takes roughly a minute to establish WiFi connectivity and then starts data collection, and before it could do so, the appliance was turned off. Correspondingly, *building monitoring and control systems for the Indian context should account for the short appliance usage and power off from the switch to ensure robust and reliable data collection.*

Lack of COTS devices made for Indian settings: As reported earlier in Section 3.1, due to non availability of ZWave sensors for Indian frequency, we imported European frequency based multisensors and plug load monitors. Import duty and shipment delays make them an expensive choice. Further, to avoid the peak switching current, the default state of the plug monitors was off (when powered manually from the switch). This implied that 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 plug sockets, such plug sockets did not

result in seamless usage. Correspondingly, we did not use ZWave based plug load monitors in our deployments as they had to be manually turned on after a power outage, which is common. Correspondingly, we had to resort to prototype solutions for appliance and circuit level monitoring.

6 Hitchhiker’s guide revisited

While there are several unique observations from our deployment, there exists multiple similarities with prior deployment experiences, most specifically- “The Hitchhiker’s Guide to Successful Residential Sensing Deployments” [12]. Some prominent similarities are:

Homes are hazardous environments: We observed that one of our multisensors would fail after power outage. We, eventually, figured that this behavior was due to the fact that this multisensor was put on the inverter point (battery backup, as is common in many homes to guard against intermittent power supply) and would not fail during a power outage. Thus, when the main power resumed, ZWave controller was not able to add this multisensor to its network. We resolved this by putting the multisensor at a non-inverter point. We also observed that one of the multisensor would always indicate motion. Initially we solved this problem by replacing its power supply. However, within a week, this problem resurfaced. We believe it is due to the faulty power supply in that room. The home occupants did not allow a thorough investigation of that electrical socket and we had to do away with that data. Although we used zip-ties extensively throughout the deployment to prevent hanging wires, we observed data loss in one of our multisensor and Android phone, which went out of power due to wire snag (shown in Figure 9b). Even after one month rigorous testing in the lab, before we started the deployment, we raised 60 new service complaints, when we moved the deployment to this home. This is a testimony that homes present unique challenges which can not be exactly emulated in laboratory settings.

Aesthetics matter: As stated in the previous work, sensor LEDs can be bothersome to home occupants, particularly in the night. Our 33 sensor deployment had introduced 63 LEDs in the home. Figure 9a shows our sensor LEDs blinking in the night. Choosing appropriate sensor location sufficed for our present deployment. However, for future such deployments, we intend to build better sensor casing to ensure that home occupants are not disturbed. Also, the home occupants complained of buzz like sound coming from the desktop that we had installed. This noise was due to the dust clogging in the desktop. Dust is a uniquely common aspect in the Indian setting. Correspondingly, 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 that the data collected from our ground floor MCBs (5 in number) was noisy. This was due to the fact that these MCBs are close (shown in Figure 9c) 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. For the 3 MCBs on the first floor, there was adequate gap amongst them and the data was of

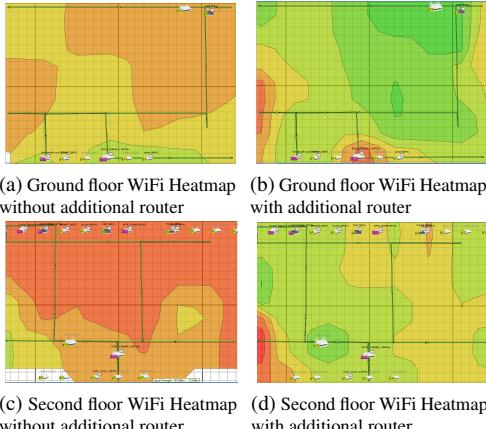


Figure 8: WiFi signal without placing additional routers on different floors is poor. The greener the region, the better the signal strength for the heatmap.



Figure 9: Common problems in residential deployments good quality (verified using manual experiments involving switching appliances on those circuits).

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

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 Mapper¹² to map WiFi signal strength. Figure 8a and Figure 8c show the WiFi heatmap produced with the home router placed on the first floor. We observed that a lot of regions inside the home show poor signal strength. We bridged an additional router on the ground and the second floor with the existing first floor router. Figure 8b and Figure 8d show the corresponding WiFi heatmaps produced after the introduction of bridged routers. A significant improvement in the signal strength with the introduction of additional routers was observed.

7 Dataset and code release

Inspired by previous work [3], we release fully labeled data for 1 day. This dataset is of the first of its kind, across the world, where events corresponding to electricity, water and ambient sensors are timestamp recorded. We manually annotate the power consumed for each of the x appliances in their different states in the home. We similarly measure the amount of water consumed in 1 minute by each of the y water fixtures. We further provide a detailed metadata log of all electrical appliances, including, but not limited to, estimate date of purchase, mapping to MCB, star-rating, rated

Table 2: First floor water fixtures labeled metadata

Sno.	Type	Consumption	Purpose	Location
1	Flush	10 liters per flush; 210 seconds for flush tank to refill	Toilet flushing	Bathroom cum toilet
2	Tap	17.5 l/minute	Washing clothes	Washing area
3	Tap	5 l/minute	Washing hands	Washbasin
4	Tap	7 l/minute	Bathing	Bathroom
5	Tap	1 l/minute	Bathroom cleaning	Bathroom
6	Tap	13.2 l/minute	Gardening	Veranda

power. We intend to release more such fully labeled data in the future, but owing to privacy concerns, we limit the current fully labeled dataset to 1 day from this home. While the accompanying switch had been to our disadvantage in measuring short duration appliance usage, it provides us an advantage over the data collected in developed countries setting. By virtue of switch control, at least the ON-OFF events can be clearly demarcated. Thus, in some sense, our entire electrical dataset is event labeled.

We also publicly release our codebase which includes an implementation of Sense-Local Store-Upload, scripts for collecting data from various sensors, database schemas, soft-sensors, startup scripts and the fixes we developed for common networking problems on RPi.

8 Conclusions and Future Work

In this paper, we present our experiences with an extensive residential deployment monitoring electrical, water 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, differences in electrical loads, water-embedded energy within a home and limited availability of robust COTS infrastructure. 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: SeLStUp, which accounts for these pitfalls by introducing local storage and periodic upload. Such an architecture can be of particular importance in the context of building monitoring and control systems suitable for the developing countries.

We are in the process of installing our sensors across multiple other homes in Delhi. Data from the deployment will be released for public use after necessary post processing and filtering, to ensure privacy and appropriate usage. We further plan to carry out longer duration deployments capturing seasonal variations in energy usage patterns.

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, ac-

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

- cess 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] 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.
- [9] M. Evans, B. Shui, and S. Somasundaram. *Country Report on Building Energy Codes in India*. Pacific Northwest National Laboratory, 2009.
- [10] T. Ganu, D. P. Seetharam, V. Arya, R. Kannath, J. Hazra, S. A. Hussain, 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.
- [11] G. W. Hart. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, 80(12):1870–1891, 1992.
- [12] 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.
- [13] 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.
- [14] M. Meeker. Internet trends at stanford bases. KPCB, 2012.
- [15] 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.
- [16] 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.
- [17] 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.