Journal of the Indian Institute of Science
A Multidisciplinary Reviews Journal
ISSN: 0970-4140 Coden-JIISAD

REVIEWS

Cyber Physical Systems for Smarter Energy Grids: Experiences at IBM Research—India

Deva P. Seetharam, Vijay Arya, Dipanjan Chakraborty, Zainul Charbiwala, Tanuja Ganu, Sunil Ghai, Jagabondhu Hazra, Palani Kodeswaran, Rajendu Mitra, Balakrishnan Narayanaswamy, Neha Sengupta and Shivkumar Kalyanaraman

Abstract | Power grid operators around the world are facing a number of critical challenges such as energy and peak power shortages, outages and the uncertainty introduced by intermittent renewable energy sources. To address these challenges, the research community has identified a few high-level objectives: alleviation of peak loads, minimization of grid losses, improving the energy efficiency of buildings and loads, and reducing the uncertainty about energy produced by renewable sources. Implementing these strategies would require a "smarter energy" system that is instrumented, with sensors and controls embedded into the fabric of its operations; it is interconnected, enabling the two-way flow of information—including pricing—and energy across the network; and it is intelligent, using analytics and automation to achieve the aforementioned objectives.

In this paper, we present a sampling of cyber physical systems we have designed to enable such a smarter energy supply chain. These systems depend on varying levels of instrumentation (sensing/actuation) and network connectivity. There remains a large opportunity to deepen these contributions and taking innovations to full market impact, which require overcoming commercial and regulatory challenges as well. A feature of our work was to consider and be informed of real world and client constraints in our work, and we have built prototypes and experiments for similar circumstances. We hope that these experiences will spark more experimental innovation activity that is critically important for being well grounded in research and indeed, for the success of smart grids worldwide.

1 Introduction

Power grid operators around the world are facing a number of critical challenges:

Energy and power shortages: While the developed countries usually suffer from power shortages during the peak load periods, many developing countries suffer from both power and energy deficits. For instance, the Indian electricity sector, despite having the world's fifth largest installed capacity, suffers from severe energy and peak power shortages. In February 2013, these shortages were 8.4% (7.5 GWh) and 7.9% (12.3 GW) respectively.[1] In fact, India has long struggled with electricity deficit and as a consequence, millions of consumers have been suffering from inadequate power supply.[2,3]

Power outages: Since the electricity infrastructure underpins the economic and social activities, it is important to ensure that there are no outages in the grid. For example, the electrical system in the United States is more than 99 percent reliable, but it still experiences power interruptions

IBM Research, India.
dseetharam@in.ibm.com
vijay.arya@in.ibm.com
cdipanjan@in.ibm.com
zainulcharbiwala@in.ibm.com
tanuja.ganu@in.ibm.com
sunilkrghai@in.ibm.com
jahazra1@in.ibm.com
palani.kodeswaran@in.ibm.com
rajenmit@in.ibm.com
Murali.Balakrishnan@in.ibm.com
neha.sengupta@in.ibm.com
shivkumar-k@in.ibm.com

that cost homes and businesses $150 billion a year [GAV10]. These power interruptions could get aggravated as the nation's entire energy infrastructure is vulnerable to severe and costly weather events driven by climate change.[4]

Intermittent Energy Sources: In many parts of the world, electricity generation is a significant source of green house gas emissions as fossil fuels are burnt to generate electricity. For instance, in the U.S., electricity generation is the largest single source of CO_2 emissions accounting for about 38% of total CO_2 emissions and 32% of total greenhouse gas emissions in 2011.[5] In India, generation plants, on an average emit 0.89 tons of CO_2 per MWh.[6] As a result, many governments are mandating that the proportion of renewable energy sources in the electricity generation portfolio must increase. For instance, as per the California Renewable Portfolio standard, all the electricity retailers must adopt the goals of 20 percent of retail sales from renewables by the end of 2013, 25 percent by the end of 2016, and the 33 percent requirement being met by the end of 2020.[7] Since some of those common renewable sources such as solar and wind are intermittent, novel techniques are required for dispatching them.

To address these challenges the research community has identified a few high-level objectives: alleviation of peak loads, minimization of grid losses, improvement of energy efficiency of buildings and loads, and reduction of uncertainty about energy produced by renewable sources. Implementing these strategies would require a "smarter energy" system that is instrumented, with sensors and controls embedded into the fabric of its operations; it is interconnected, enabling the two-way flow of information—including pricing—and energy across the network; and it is intelligent, using analytics and automation to achieve the aforementioned objectives.[8]

In this paper, we present a sampling of cyber physical systems we have designed to enable such a smarter energy system. As shown in Table 1, these systems depend on varying levels of instrumentation (sensing/actuation) and network connectivity. On the one end of instrumentation, the SoftGreen system uses data only from opportunistic sensors to enable occupancy based energy management in commercial buildings, and at the other end, YouGrid employs multiple types of sensors/actuators (motion sensor, diesel fuel monitor, battery sensors, etc) to improve the fuel efficiency of microgrids; on the one end of connectivity, nPlug is an autonomous peak load alleviation system, and at the other end, CPS-Net requires high speed network connectivity to communicate with Phasor Measurement Units (PMUs) that are geographically distributed across the grid. The following sections provide a brief overview of each of these systems.

2 nPlug: An Autonomous Peak Load Controller

Power utilities worldwide face a major challenge of *peak demand*. During the peak demand hours, the demand for electricity is usually more than the base and intermediate supply capacities. The extra demand can be met by starting "peaker" power plants or by buying energy from bulk power markets. But, both of these options are expensive for

Table 1: Illustrative cyber physical systems for enabling smarter energy solutions.

System	Point of operation	Objective	Instrumentation	Network connectivity
nPlug	Sockets	Peak load reduction	Local sensing	No network required
SoftGreen	Buildings	Building energy efficiency	No hardware instrumentation (only opportunistic sensing)	High-speed connectivity
Wattzup	Buildings	Load disaggregation	Smart meters	Low-speed connectivity
Connectivity models	Distribution networks	Network mapping	Smart meters and socket-level meters	Low-speed connectivity
CPS-Net	Transmission networks	Wide area situational awareness and control	Phasor Measurement Units and Phasor Data Concentrators	High-speed connectivity
YouGrid	Campus level microgrids	Fuel efficient microgrids	Motion sensor, Diesel fuel monitors, battery controllers, etc.	High-speed connectivity

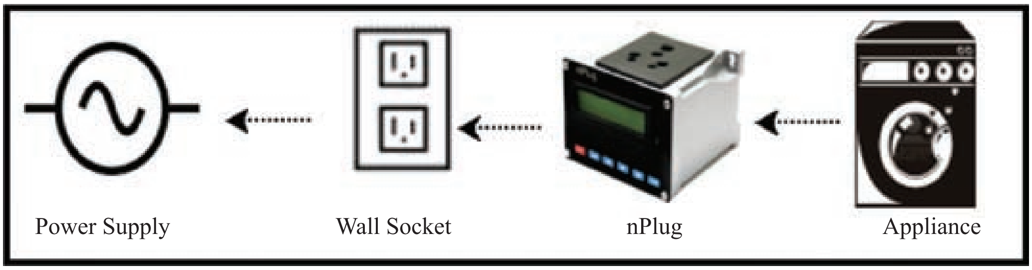


Figure 1: nPlug structure.

electricity suppliers who sell electricity to their consumers at flat rates.

The peak demand could be alleviated, if a large number of deferrable loads, particularly the high powered ones, could be moved from on-peak to off-peak times. However, conventional demand-side-management (DSM) strategies that depend on communication infrastructure may not be suitable for developing countries like India, as the local conditions usually favor only inexpensive solutions with minimal dependence on the pre-existing infrastructure.

To address this need, we have developed a completely decentralized demand side controller called the nPlug.[9] As shown in Figure 1, nPlug is positioned between wall sockets and the corresponding deferrable loads such as water heaters, electric vehicles, washing machines, etc. The nPlug can accept the following configuration inputs: 1. *Earliest start time*: the earliest time at which an appliance can be switched on; 2. *Latest end time*: the latest time at which the appliance must finish running; 3. *Duration*: the duration for which the appliance must be powered; and 4. *Hold time*: the minimum time an appliance must be run at a stretch when turned ON. For example, a residential consumer who leaves for work at about 8 AM may specify that her insulated water heater must be run for 30 minutes between 6 AM and 8 AM.

Using these inputs as guidelines, a nPlug schedules the attached appliance so that the peak load and supply-demand imbalance periods are avoided. nPlugs determine such periods by sensing the line voltage[a] and frequency[b] measured at the socket. The sensed time-series data is then pre-processed using the Piecewise Aggregate Approximation (PAA) and the peak load pattern is identified using the novel extension of Symbolic Aggregate Approximation (SAX) technique on the preprocessed voltage data. The situations of supply-demand imbalance can be identified using 2-SD statistical test on the line frequency data. The decentralized scheduling of nPlugs is performed by using the concept of Grid-Sense Multiple Access (GSMA), which is inspired by Carrier Sense Multiple Access (CSMA) approach of computer networks.

The embedded analytics approach of nPlug provides various benefits over the existing peak reduction methods—1. *Network free*—nPlug does not require any communication infrastructure for sensing and control, and hence can be completely autonomous. 2. *Brownfield innovation*—it does not require any changes to the grid or to the appliances and hence suitable for existing power grids and for the millions of appliances already in use. 3. *Incremental adoption*—since each nPlug has the potential to alleviate peak load, nPlugs can be introduced in small batches into the grid. This reduces the initial investment as well as the risks of introducing new technology into a pre-existing infrastructure. 4. *No policy changes required*—nPlugs don't depend on differential pricing schemes or smart meters and hence deploying them doesn't require any regulatory approvals. 5. *Inexpensive solution*—every hardware and software component in nPlugs are based on careful analysis of cost-performance trade-offs. The prototype we have built costs about USD 30 in small volumes (<100 units) and we estimate nPlugs in large volumes (>100,000 units) would cost about USD 15.

The experimental results indicate that nPlug could be an effective and inexpensive technology for addressing the peaking shortage. The embedded analytics approach of nPlug has the potential of further enhancement for the power management of large computing loads like electric vehicles, data centers and for the grid stabilization with distributed renewable energy sources.

[a] As per power systems theory and as reconfirmed by the data we have collected, grid voltage is inversely correlated with the load levels.

[b] As per power systems theory and as reconfirmed by the data we have collected, grid frequency is inversely correlated with the load levels imbalance.

3 SoftGreen: Building Energy Efficiency through Opportunistic Sensing

Reducing the energy consumption of commercial buildings has emerged as an important research focus for energy efficiency since those buildings account for 40%[10] of the global energy consumption. Primary energy consumers in a typical office building include HVAC-L (Heating, Ventilation, Air Conditioning and Lighting) and plug loads (such as laptops, desktops, water coolers etc.). However, a large fraction of electricity usage, nearly 70%,[11] can be attributed to HVAC-L infrastructure. Despite that, such loads are controlled using static policies (e.g. operate from 8 AM to 6 PM) and not through dynamic policies that are derived based on occupancy (people's presence). Although there have been research efforts in the past to develop occupancy-based load management, the focus has been on developing low cost occupancy sensing modules or changing occupants' behavior.[12]

On the other hand, we aim towards developing an occupancy detection system that can be built over existing building infrastructure and does not require any additional hardware instrumentation. Modern commercial office environments are equipped with various infrastructures such as communication networks, online collaboration tools and access card systems to make the work environment comfortable and safe. These infrastructures, in addition to performing their intended functions, can serve as opportunistic data sources or soft-sensors and provide context data about occupants' locations. As illustrated in Figure 2, data from common infrastructure elements such as, Wi-Fi access points, online calendars, instant messaging client, computing devices (e.g., laptops) etc. can be collected and fused to accurately detect occupancy in meeting rooms and cubicles of offices without installing any hardware sensors.

This approach was evaluated through a pilot deployment for five users over a month in a commercial office. The context data was collected from Wi-Fi access points, system activity, instant messaging, and online calendar. Volunteers marked their locations during the pilot period to establish ground truth. Decision tree based classification and regression approaches were evaluated to fuse the data for occupancy. As presented in Figure 3, the classification approach yields more accurate results than the regression based one. Another benefit of using classification is that it can process both numeric and nominal input variables, therefore, there is no need of mapping the context cues to numeric values. Further details are documented in.[13,14]

We are conducting a large-scale pilot deployment with more number of users and hardware-based ground truth collection. In addition to devising unsupervised approaches for data fusion, handling occupants' privacy concerns and finding other potential application areas are the focus of our future work.

4 Smart Meter Analytics

As smart meter deployments are on the rise,[15] the data collected from those meters enables a number

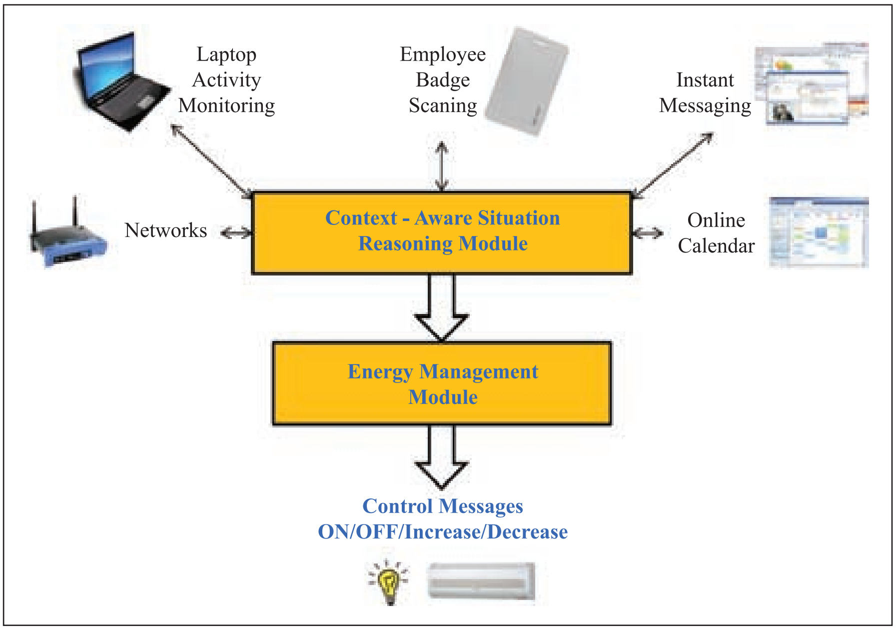


Figure 2: SoftGreen approach.

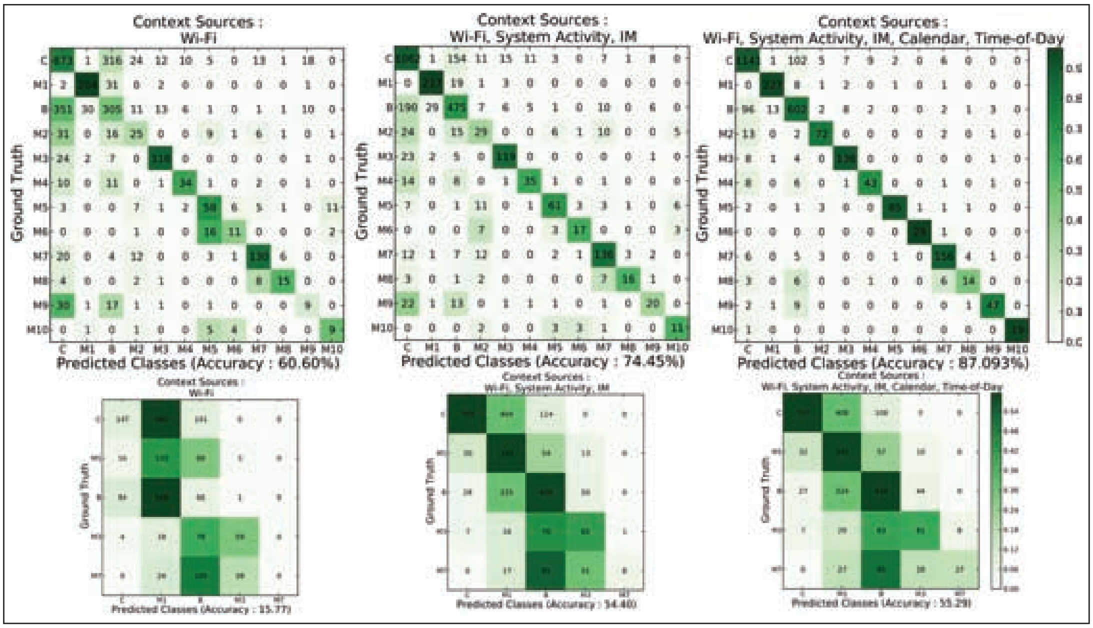


Figure 3: Confusion matrices representing accuracy of inferring one volunteer's situation using classification (top three) and regression (bottom three) approaches using data from subsets of context-sources. Higher numbers in the primary diagonal indicate higher accuracy of inferences. On the other hand, higher numbers in the non-primary diagonal indicate lower accuracy of inferences. It is important to note that both precision and recall for classification-based approach is much higher than that for regression-based one.

of analytics applications such as load forecasting and theft detection. We have studied two potential use cases of the smart meter data:

1. A discriminative method for identifying individual appliance consumption from the aggregate consumption data reported by the smart meters.
2. Inferring the interconnections between various customers and assets in the grid downstream of a substation using the consumption data reported by the smart meters at the customer premises and the load data reported by smart meters deployed on distribution grid elements such as transformers and feeders.

4.1 *Wattzup—A discriminative load disaggregation method*

A deep and accurate understanding of user's energy consumption patterns in correlation to their normal activities is paramount for the next generation of Auto Demand-Response Systems as well as for undertaking studies that focus on inducing long-term sustainable behaviors at a societal level.[16] While obtaining this information through plug level monitoring is sometimes feasible, particularly in commercial settings, such intrusive methods are not scalable to large deployments due to both hardware costs—which can be more than USD 1800 per home—and the inconveniences caused to home owners.

These factors, the value of detailed knowledge of consumption and the limitation of available metering, have resulted in a need for systems for Non-Intrusive Appliance Load Monitoring (NIALM). As defined in the seminal work of[17] "A nonintrusive appliance load monitor determines the energy consumption of individual appliances turning on and off an electric load, based on detailed analysis of the current and voltage of the total load, as measured at the interface to the power source. " The design of these load disaggregation systems has recently become the focus of much research.

Load disaggregation problem is defined as given $y_t^{(i)} = 0, y_t^{(i)} \in y^{(i)}$, a discrete sequence of observed aggregate power readings of n number of individual appliances $\bar{y} = \{\bar{y}_1, \bar{y}_2, \bar{y}_3, \ldots, \bar{y}_T\}$, $\bar{y}_t = \sum_{i=1}^{n} \bar{y}_t^{(i)}$, $i = 1, \ldots, n$, $t = 1, \ldots, T$, determine the operational state/load of each appliance in a specific time $\hat{y}_t^{(i)}$. In real condition, the total household consumption data $\bar{y}$ may include power consumption from unmonitored household appliances since appliance level sensors deployment does not cover all appliances inside the premises due to cost and installation issues. This constraint has significantly increased the challenge of providing high accuracy load disaggregation models when compared with other research in the area that assumes that aggregate data is the sum of appliance level data.

We have developed a discriminative load disaggregation approach that predicts the most likely or possible appliance state configuration from total power consumption in a specific time period using simple non-parametric classification algorithms.[18] In our model, we first automatically discretize each appliance's load $y^{(i)}$ into z states where $z \geq 2$, since each appliance has at least 2 states, on and off. When real power consumption of an appliance at specific time is 0, we can consider this appliance is not in use or in off state. So, for any number of z, we assign all $y_t^{(i)} = 0$, $y_t^{(i)} \in y^{(i)}$ as off. We then use Expectation-Maximization (EM) algorithm[19] to cluster the non-zero appliance load into $z-1$ clusters and label those clusters sequentially based on the cluster mean.

Having discretized our training data set, we then run kNN algorithm to build the predictive model. For each aggregated load $\bar{y}_t$ in the test data we find k nearest neighbors from the training data based on their Euclidean distance of aggregate load, day and time of the day attributes. Each neighbor has a combination of appliances' state as a result of the previous discretization process. We decide the most probable state combination from those neighbors based on the majority/voting approach.

The primary characteristics of our method that are important to emphasize are: (i) The discriminative machine learning approach we use is based directly on historical consumption data which consists of each monitored appliance's consumption data and total household consumption that includes both monitored and unmonitored appliances' energy consumption; (ii) The method does not directly predict consumption levels, but first predicts activity or state, and then uses historical data to predict usage level. This seems to improve accuracy substantially compared to previous approaches that directly estimate consumption levels; (iii) The method is simple and computationally efficient while offering high enough accuracy for the next steps of a complete DR system.

Our approach has been evaluated on the publicly available REDD data set[20] and it achieves 88% accuracy[c] in state prediction and 82% accuracy in total energy assignment [RAH12].

Our immediate focus is on analyzing the level of accuracy required to have reliable DR system, developing the unsupervised load disaggregation method and discovering effective techniques for leveraging situation awareness through external/rich context information.

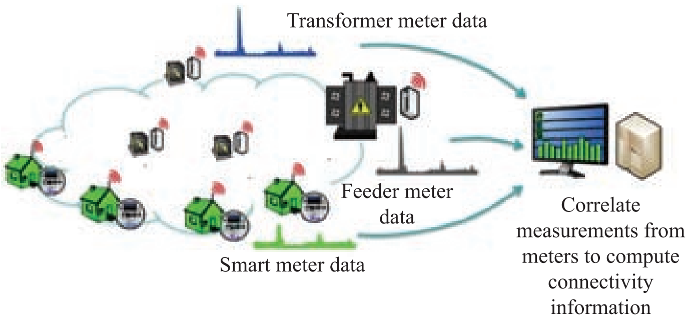
4.2 *A data driven inference of connectivity models*

The connectivity model (CM) of a distribution grid gives the underlying interconnections between various customers and assets in the grid downstream of a substation. The connectivity includes which customer is powered by which distribution transformer, which distribution transformer is powered by which phase of the feeder, and so on. A common problem faced by energy utilities is that they have an out-of-date or inaccurate CM of their distribution network i.e. the CM is inconsistent with the actual connectivity relationships that exist in the field. The accuracy of the CM generally deteriorates over time due to maintenance, repairs, and restoration activities following faults and outages.

The CM is important, as it is needed in the operations and maintenance of distribution networks. Most solutions that automate the management of distribution networks require the CM as input. For instance, it is required by the outage management system (OMS) to accurately record and respond to outages. When a fault occurs or an asset (e.g. a transformer) fails in the distribution network, the CM is required to assess the magnitude of the fault and the customers that may have been impacted by it. Additionally, energy distributors may have obligations to accurately report customer outages to regulatory bodies, which is difficult without an accurate connectivity model. The distribution management system (DMS) needs the connectivity model for fault detection, isolation, and service restoration (FDIR) and also to conduct accurate power-flow calculations. An inaccurate CM may lead to incorrect diagnosis of faults, suboptimal response, and faulty voltage profiles of the distribution grid, which may affect the reliability of energy delivery to the customer.

A significant fraction of all losses in a power system occur in the distribution network and the growing imbalance between supply and demand is driving the deployment of solutions that can improve the overall efficiency of energy delivery. An accurate connectivity model can enable many of these solutions. For instance, solutions for energy auditing and loss localization use the connectivity model to localize energy losses from theft and inefficiencies in the distribution network. The phase balancing solution requires the connectivity model in order to balance the load on the three phases of a feeder so that losses incurred

[c] We use the performance evaluation formula provided by Kolter and Johnson [KOL11].



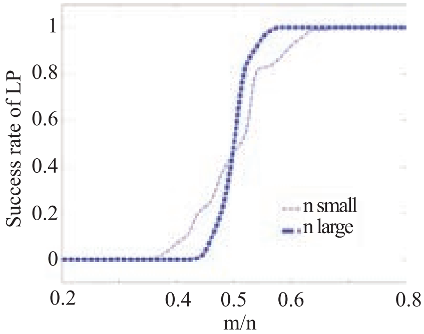


Figure 4: (a) Smarter Meter Measurement Time-Series used to estimate a tree connectivity model between the metered assets in the distribution grid. (b) Benchmark performance of linear programming relaxation techniques.

while delivering energy to the customer are minimized. Additionally when customers have behind the meter resources such as distributed generation and storage, the connectivity model is required to ensure a balanced and reliable infusion of energy back into the grid via the distribution network.

Manual identification of the CM is expensive and unsustainable as the field configuration changes over time. Existing solutions based on power line communication are generally capital intensive and may not be suitable for regions where the signal needs to propagate over long distances.

In this work,[21,22] we present a novel analytics approach (Figure 4(a)) to infer the connectivity model of a radial distribution network. Our techniques are novel as they are purely based upon a time series of synchronized power measurements collected by various meters in the distribution grid. The time series measurements from all meters are used to set up a system of linear equations based upon the principle of conservation of energy i.e. during any time interval, the load (watt-hours) measured by a feeder meter must be equal to the sum of loads measured by all customer meters connected to that feeder plus any errors. The errors arise due to imperfect synchronization of measurements at different meters, different sampling rates, unmetered loads such as streetlights, and unknown and time-varying transmission line losses. The equations are analyzed to regress a tree network between the meters, which is consistent with the observed time series measurements.

Our work can be regarded as an early example of the tomography technique applied in the smart grid context, where solution to a linear inverse problem of the form $Ax + e = b$, is used to recover the underlying topology of a distribution network. We propose a number of different optimization formulations (mathematical programs and their relaxations) for noiseless and noisy variants of the tree inference problem and study the conditions for uniqueness of solution as a function of the number of meter measurements. Figure 4(b) shows the benchmark performance of the proposed linear programming relaxation techniques, which retrieve the true solution as the number of meter measurements (m) exceeds half the number of meter nodes (n) in the system.

5 CPS-Net: In-Network Aggregation for Synchrophasors

A phasor measurement unit (PMU) or synchrophasor is a device that measures the electrical waves on an electricity grid, using a common time source for synchronization. Time synchronization allows synchronized real-time measurements of multiple remote measurement points on the grid. These devices can sample the electricity grid at the rate of up to 120 Hz and publish these measurements as streams that need to be delivered reliably and in real-time to a number of synchrophasor applications.

A wide variety of synchrophasor applications have been proposed and the Quality of Service (QoS) requirements of these applications have been classified by the North American Synchrophasor Initiative (NASPI) into a set of classes.[23] While building a network that satisfies the basic QoS requirements is a well studied problem, the dimensions that differentiate PMU-specific networks are:

- Application requirements need to be mapped onto a real-time wide-area publish-subscribe architecture requiring QoS support beyond simple point-to-point QoS.
- During overloads or critical events when sampling rates increase or more PMUs are active, it is important to gracefully degrade

performance and data stream delivery in an application-specific manner.

Graceful degradation of performance and QoS for many-to-many real-time, wide-area streams is hence an important requirement for enabling PMU applications. To address this need, we proposed Cyber-Physical Network (CPS-Net), a flexible 3-layered architecture that leverages the benefits of layering and point-to-point QoS, while allowing application-specified in-network aggregation of data streams during overload.[24] The bottom layer provides basic path-specific QoS. The middle layer provides real-time wide-area publish-subscribe capabilities, integrated with traffic engineering of data streams across multiple lower level paths and trees. The top layer provides a distributed stream-processing infrastructure for application-specified aggregation that helps in graceful degradation during network overload.

During underload, the top layer is quiescent, and all the PMU data from publishers is sent to subscribers. But during network overload, there may not be sufficient capacity to deliver all the PMU data. One response to overload would be to randomly discard data: however this could degrade performance in unpredictable ways. From video streaming literature, we know that if information can be dropped in an application-sensitive manner, then the quality of experience for multimedia applications can be gracefully degraded as a function of the level of overload. Analogously, we aim to provide application-sensitive in-network aggregation functions that could be used during overload periods to achieve graceful degradation of synchrophasor applications. Specifically, the lower layer of our three-layer architecture during overload triggers the co-optimization of higher layers, and application-specific filtering and/or aggregation of data is performed.

The subscribers, while subscribing to specific content, can specify the data aggregation and filtering mechanisms that they are willing to accept, types of data to which such mechanisms can be applied, and the timeframes during which those mechanisms are acceptable. The application writer would be best situated to express such aggregation and filtering functions, but would need a convenient API to express them. We propose to provide the application writer with a simple declarative API, based upon a stream computing programming model such as streamIt[25] or Spade and Infosphere Streams.[26] The declarative view is expressed in the stream programming language while the details of the network, placement and composition of operators are abstracted away, as part of the distributed stream computing system.

Initial simulation results show that the CPS-Net architecture can gracefully degrade data streams for real-time synchrophasor applications during network overload. As the examples given in our paper[24] indicate, there are a variety of application-specific aggregation methods as well as policies for specifying aggregation. Such policies could be a function of transient network overload, or based upon other factors such as price of power, or spatiotemporal and administrative considerations. CPS-Net needs to be extended to robustly incorporate some of these considerations in specific application domains. For instance, one open problem to consider is if different applications (consumers of sensing data) specify aggregation of data in different ways, how would a network-level operator combine these requirements, and handle conflicts, while meeting network-capacity constraints?

6 YouGrid: An Operating System for CPS-enabled Microgrids

Microgrids are considered the modern, small-scale versions of the centralized electricity system.[27] Microgrids have the advantage of reduced transmission and distribution losses and can be islanded from the main grid. They are ideal for remote locations and for rural electrification but can be used more generally for increased reliability through independence from macrogrid disturbances, fossil fuel, cost and carbon footprint reduction and energy diversification. Microgrids ease local integration of renewable sources and could encourage conservative practices through community participation. Microgrids are operationally challenging, however, due to the heterogeneous characteristics of sources and loads. IBM, in collaboration with the Universiti Brunei Darussalam (UBD), has taken a holistic approach to energy efficiency and conservation measures for microgrid operation. At the heart is YouGrid, a microgrid operating system that employs a cyber physical systems approach to interface with disparate energy sources, loads and actuators. YouGrid uses data-driven predictive analytics, advanced optimization and persuasive consumer interfaces to improve the availability of power, increase the efficiency of microgrid operation and enhance system reliability.

To evaluate this operating system, we have established a first of its kind microgrid at Kuala Belalong Field Studies Centre (KBFSC). KBFSC is situated in the Temburong district of Brunei in South East Asia. KBFSC is located deep within a

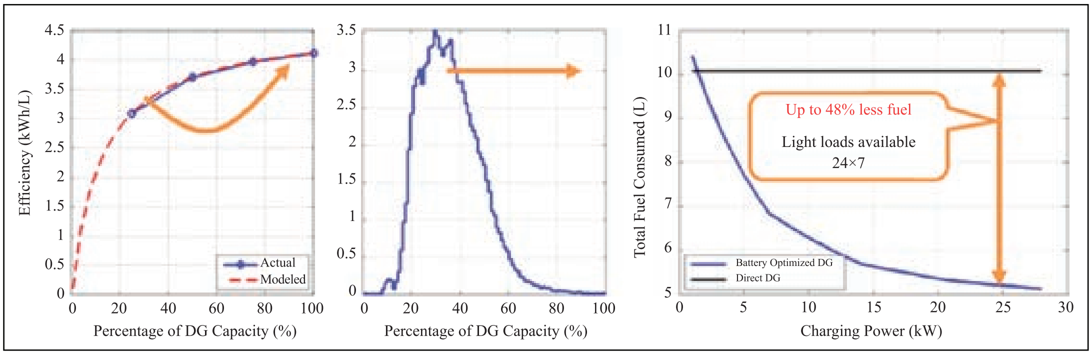


Figure 5: (a) DG Efficiency (b) DG Load Histogram (c) Projected savings by load-shifting.

primary rain forest as a research facility to study ecological diversity in the region. The remoteness of the centre makes it impractical to extend the national grid to the centre and since KBFSC is nestled in a valley, solar and wind harvesting for renewable power generation are extremely challenging. KBFSC, therefore, relies on a diesel generator (DG) for its power needs. Currently, the centre is powered for about 10 hrs each day, which corresponds to using about 20,000 litres of diesel every year. Transporting this diesel, about 1.4 metric tons a month, is done by canoe through narrow, serpentine waterways. Furthermore, due to the hilly terrain, diesel has to be transported in cans that have to be hand carried over relatively large distances. Needless to say, this is not a job anyone envies.

Our CPS-enabled microgrid uses flow sensing at the fuel intake of the DGs, energy metering at their outputs and energy metering at the building and socket level as well. The richness of this dataset enables data-driven modeling of the fuel conversion efficiency of the DG, as shown in Figure 5(a). The fuel conversion efficiency is the amount of energy the DG produces for every litre of diesel. The curve is non-linear and is heavily load dependent. Analyzing the data also reveals each appliance's contribution to the energy consumption and corresponding fuel consumption. Some appliances (e.g. lights) need relatively little power but require the DG to remain turned on at low fuel conversion efficiency. The histogram of energy produced by the generator over a 3 month period is illustrated in Figure 5(b) and shows how lightly the generator is loaded most of the time.

Our analysis of the DG efficiency and appliance profiles suggest peak load *aggregation*. That is, aggregate power consumption of appliances over shorter durations (say 4–6 hrs, instead of 10 hrs) at a higher power level to improve overall fuel conversion efficiency. The reduced DG timings would be offset using an energy buffer (battery bank) that would provide power to appliances identified as being low energy consumers but high fuel consumers. The battery bank also serves to increase the load on the DG when it is being recharged. By using a high efficiency, high capacity charger, we estimate diesel fuel savings of 20% to 50% for the same amount of energy produced.

Apart from modeling the DG and loads, our microgrid energy management system also builds models of the batteries that are then fed to an optimization system that minimizes fuel cost. The optimization determines the optimal schedule to run the DG based on the current and predicted load profiles and YouGrid closes the CPS actuation loop by automatically charging batteries and providing power according to this schedule.

7 Discussions

This paper has presented a sampling of our experiences and technical progress in applying cyber physical systems to the realm of energy grids, in different parts of the value chain. There remains a large opportunity to deepen these contributions and taking innovations to full market impact, which require overcoming commercial and regulatory challenges as well. A feature of our work was to consider and be informed of real world and client constraints in our work, and we have built prototypes and experiments for similar circumstances. We hope that these experiences will spark more experimental innovation activity that is critically important for being well grounded in research and indeed, for the success of smart grids worldwide.

Received 14 August 2013.

References

1. Central Electricity Authority, Ministry of Power, Government of India. Monthly Power Supply Position, http://www.cea.nic.in/monthly_power_sup.html, February 2013.

2. V. Bajaj, "India Struggles to Deliver Enough Power", New York Times, April 20th, 2012. http://www.nytimes.com/2012/04/20/business/global/india-struggles-to-deliver-enough-electricity-for-growth.html?ref = asia
3. S. Biswas, "Ten Interesting Things about India Power", BBC, July, 2012. http://www.bbc.co.uk/news/world-asia-india-19063241
4. U.S. Department of Energy, "U.S. Energy Sector Vulnerabilities to Climate Change and Extreme Weather", July 2013. http://energy.gov/sites/prod/files/2013/07/f2/20130716-Energy%20Sector%20Vulnerabilities%20Report.pdf
5. U.S. Environmental Protection Agency, "Overview of Greenhouse Gases," http://www.epa.gov/climatechange/ghgemissions/gases/co2.html, 2013
6. Central Electricity Authority, Ministry of Power, Government of India, "CO2 Baseline Database for the Indian Power Sector," http://www.cea.nic.in/reports/planning/cdm_co2/user_guide_ver6.pdf, 2013
7. California Energy Commission, "California Renewable Energy Overview and Programs". http://www.energy.ca.gov/renewables/
8. IBM Corporation. Transforming the energy value chain. http://public.dhe.ibm.com/common/ssi/ecm/en/euw03060usen/EUW03060USEN.PDF
9. T. Ganu, D.P. Seetharam, V. Arya, R. Kunnath, J. Hazra, S.A. Husain, L.C.D. Silva, and S. Kalyanaraman, "nplug: A smart plug for alleviating peak loads," in Third International Conference on Future Energy Systems, e-Energy, May 2012.
10. P. Lombard, Luis, J. Ortiz, and C. Pout. "A review on buildings energy consumption information." Energy and buildings, 40(3):394–398, 2008.
11. Y. Agarwal, T. Weng, and R.K. Gupta. "The energy dashboard: improving the visibility of energy consumption at a campus-wide scale." Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings. ACM, 2009.
12. V.L. Erickson and A.E. Cerpa, "Thermovote: participatory sensing for efficient building HVAC conditioning," Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings, pp. 9–16, 2012.
13. S.K. Ghai, L.V. Thanayankizil, D.P. Seetharam, and D.P. Chakraborty, "Occupancy detection in commercial buildings using opportunistic context sources," IEEE PERCOM Workshops, pp. 463–466, 2012.
14. L.V. Thanayankizil, S.K. Ghai, D.P. Seetharam, and D.P. Chakraborty, "Softgreen: Towards energy management of green office buildings with soft sensors," Proceedings of IEEE COMSNETS, pp. 1–6, 2012.
15. U.S. Energy Information Administration. Smart Meter deployments continue to raise. http://www.eia.gov/todayinenergy/detail.cfm?id = 8590.
16. S. Darby, "The effectiveness of feedback on energy consumption. A Review for DEFRA of the Literature on Metering, Billing and Direct Displays," 2006.
17. G.W. Hart. Nonintrusive appliance load monitoring. Proceedings of the IEEE, 80(12):1870–1891, 1992.
18. D. Rahayu, B. Narayanaswamy, S. Krishnaswamy, C. Labb' e, and D.P. Seetharam. Learning to be energy-wise: discriminative methods for load disaggregation. In Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy), 2012 Third International Conference on, pp. 1–4. IEEE, 2012.
19. T. Moon. "The expectation-maximization algorithm" in IEEE Signal Processing Magazine, 13(6):47–60, 1996.
20. J. Kolter and M. Johnson. REDD: A public data set for energy disaggregation research. In Workshop on Data Mining Applications in Sustainability (SIGKDD), San Diego, CA, 2011.
21. V. Arya, T.S. Jayaram, S. Pal, and S. Kalyanaraman, "Inferring connectivity model from meter measurements in distribution networks," ACMe-Energy 2013, pp. 173–182, 2013.
22. V. Arya, D.P. Seetharam, S. Kalyanaraman, K. Dontas, C. Pavlovski, S. Hoy, and J. Kalagnanam, "Phase Identification in Smart Grids," In Proceedings of IEEE SmartGridComm, 2011.
23. D.E. Bakken, A. Bose, C.H. Hauser, E.O. Schweitzer, D.E. Whitehead, and G.C. Zweigle, "Smart generation and transmission with coherent, real-time data," Washington State University, Tech. Rep., August 2010.
24. V. Arya, J. Hazra, P. Kodeswaran, D.P. Seetharam, N. Banerjee, and S. Kalyanaraman, "CPS-Net: In-Network Aggregation for Synchrophasor Applications", in COMSNETS 2011.
25. W. Thies, M. Karczmarek, and S.P. Amarasinghe, "Streamit: A language for streaming applications," in Compiler Construction, 2002, pp. 179–196.
26. B. Gedik, H. Andrade, K.-L. Wu, P.S. Yu, and M. Doo, "SPADE: the system S declarative stream processing engine," in SIGMOD '08: Proceedings of the 2008 ACM SIGMOD international conference on Management of data, 2008, pp. 1123–1134.
27. R. Galvin, K. Yeager, J. Stuller. Perfect power. McGraw Hill, 2009.

Deva P. Seetharam leads IBM India's Smarter Energy research initiatives. Previously, he led France Telecom's electronic paper display research group in Boston. He has also worked on various embedded wireless networking technologies including active RFID and wireless sensor networks. He holds a MS degree from the MIT Media Laboratory and a BE degree from the Bharathiar University

Vijay Arya is a Research Staff Member at IBM Research—India. Prior to joining IBM, he worked at National ICT Australia (NICTA) and completed his doctoral studies in Computer Science from INRIA Sophia Antipolis, France. His research interests include measurements and modeling, computer networks, and smart grids.

Dipanjan Chakraborty is a Senior Researcher at IBM Research India. He has 12+ years of experience in Research and Innovation. His primary research interests are in Mobile and Context-aware Computing and Services, Social Networks; Sensor Networks, and more recently on cloud-coordinated mobile sensing, semantic knowledge inferencing and behavior models of community sensing. Till date he has co-authored around 70 peer-reviewed publications. His work has received around 2500 citations. He has filed 21 patents and has 8 issued patents. At IBM, he has received multiple research division accomplishment awards for his work in the area of Context-aware services. He has been involved in organizing international conferences and workshops such as Mobile Data Management, Percom, Comsnets, International Conference on Pervasive Services.

Zainul Charbiwala is a researcher in the Smarter Energy Systems group at IBM's India Research Lab. He is currently interested in building a system level framework for smart microgrids. Zainul has a Ph.D. (2012) from the EE Department at UCLA working on energy efficient sampling and inferencing in wireless health and smart buildings and an MTech (2002) from IIT Bombay.

Tanuja Ganu is a Research Software Engineer at IBM Research—India. Prior to joining IBM, she completed her Masters in Computer Science at Indian Institute of Science (IISc), Bangalore. Her research interests are in machine learning, optimization and smart grids.

Sunil Ghai works as a research software engineer in the Smarter Energy group at IBM's India Research Lab. His background is in building energy management, context aware applications, usable security etc. He received a best paper award at the E6 workshop, IEEE COMSNETS 2012 and a best paper nomination at the ACM eEnergy 2013. In his past, he interned with Google and the Technical University of Munich. He holds a bachelors degree from the Delhi College of Engineering, Delhi University.

Jagabondhu Hazra is a Research Staff Member at IBM Research—India. Prior to joining IBM, he received his Ph.D. from IIT Kharagpur, India and worked as a post doctoral researcher at SUPELEC, France. His research interest includes smart grid, network congestion management, blackout prediction, HVDC, HVAC, renewable energy, etc.

Palanivel Kodeswaran is a researcher at the Telecom Research and Innovation Center at IBM Research, India. His research interests are in the broad areas of systems and networking, privacy, context aware computing and emerging systems such as Machine-to-machine (M2M) communication. His current research focusses on systems security and privacy challenges that arise in smart phones and M2M systems. He obtained his Ph.D. in Computer Science from the University of Maryland, Baltimore County, and bachelors from the College of Engineering Guindy.

Rajendu Mitra works as Software Engineer with Smarter Energy Group at IBM Research-India. He completed his Masters in Computer Science from Indian Institute of Science (IISc), Bangalore.

Balakrishnan Narayanaswamy received his Ph.D. from the department of Electrical and Computer Engineering (ECE) at Carnegie Mellon University in 2011, after which he joined the IBM Research Lab in Bangalore, India. His research interests lie at the intersection of AI, optimization, learning and inference particularly using them to understand, model and combat noise and uncertainty in real world applications.

Neha Sengupta works with the Smarter Energy team at IRL. Currently, she is working on Microgrid Management Systems, In particular, systems that employ sensors and use the sensor information in order to maximize energy efficiency, utilization and minimize wasteful consumption. She completed her Masters degree in Computer Applications from University of Delhi in 2011.

Shivkumar Kalyanaraman (S'93–M'97–SM'07–F'10) received the B.Tech. degree in computer science from the Indian Institute of Technology, Madras, India, in 1993, the M.S. and Ph.D. degrees in computer and information sciences from the Ohio State University, Columbus, in 1994 and 1997, respectively. He also received the Executive M.B.A. degree from Rensselaer Polytechnic Institute, Troy, NY, in 2005. He is the Chief Scientist at IBM Research—Australia, and Co-Director of the UBD | IBM Centre, Brunei. Previously he was Senior Manager, Smarter Planet Solutions with IBM Research—India, Bangalore, India, and was a Professor with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute. His research in IBM is at the intersection of emerging wireless technologies, smarter energy systems, and IBM middleware and systems technologies. Dr. Kalyanaraman is an ACM Distinguished Scientist. He is a visiting Professor at Universiti Brunei Darussalam.