Home Energy Simulation for Non-Intrusive Load Monitoring Applications

Krishnan Srinivasarengan
Innovation Labs
Tata Consultancy Services
Bangalore, India
krishnan.srinivasarengan@tcs.com

Goutam Y G
Innovation Labs
Tata Consultancy Services
Bangalore, India
goutam.yg@tcs.com

M Girish Chandra
Innovation Labs
Tata Consultancy Services
Bangalore, India
m.gchandra@tcs.com

ABSTRACT

Home Energy Management (HEM) is a vital component of smart grid, which can be considered as a distributed cyber physical system. HEM involves appropriate management of home appliance usage through deliberate efforts from the end-user. This can enable a stable operation of the grid as well as reduce energy usage and bills for the end-user. The installation of smart meter has led to a number of analytics and applications developed on top of its data. However, the algorithms are evaluated over a very small subset of experimental or open dataset. To mitigate this problem, a bottom-up data generation approach is proposed in this paper. The appliances are considered as combination of fundamental electrical components. The appliance characteristics and operations are modeled through stochastic parameters, which are available as prior information or through learning from existing meter data. Preliminary results of generating data for the application of Non-Intrusive Load Monitoring is presented.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Home Energy Management, Data Generation, Non Intrusive Load Monitoring, Pattern recognition, Machine Learning

1. INTRODUCTION

Smart grid shall be considered as an instance of a distributed Cyber Physical System (CPS) with the various components like buildings, industries, generation and transmission infrastructures forming the cyber physical system nodes. One of the key emerging CPS element is the individual homes, primarily due to the installation of smart meters and availability of mobile computing resources. A number of application driven works consider a house as a CPS, Home Energy Management (HEM) being the chief among them. The aim here is multi-fold, from peak load management and better demand response from a utility perspective to energy and cost savings from a end-user perspective.

The key to application development is on the algorithms that perform analytics on the smart meter data.The testing of such algorithms, however, is not comprehensive. This is due to the cost involved in a setup that can provide smart meter and appliance level measurement and the limited open data set that are available. For instance, the public datasets REDD [7] provides both appliance-level and whole-house consumption data, and Tracebase [9] provides appliance level plug data. A number of other public datasets are also available (e.g. [1]). The key disadvantages of present scenario are:

- Lack of all-weather energy usage behaviour which requires at least a full year data capture
- Some appliance combination pattern will never be seen in a house because of household usage characteristics. The algorithm would not have been trained for this pattern and hence affecting its performance in test data.
- The demography specific appliances (e.g. water pumps in India) would not be captured if the open data set is restricted to a specific neighbourhood or a country.

This paper proposes a bottom-up simulation approach to generate near-realistic data for household consumption, considering the characteristics of the appliances' internal components. The key contributions are:

- A modular internal component based approach to generate appliance characteristics, making the approach sampling rate independent.
- Models that can learn appliance characteristics and contextual information from available data. These models will allow for generating data for a longer duration.

The approach is an outcome of the data generation effort to mitigate the lack of annotated data for a disaggregation application. The paper would discuss the overall architecture of the simulation approach as well as the application which led to the proposed generalized framework. To enable that, the paper is organized as follows: Sec 2 discusses some of the literature that have relevance and overlap to this work.

In Sec 3, the Simulation approach is detailed, and an example application for Non-Intrusive Load Monitoring (NILM) is discussed in Sec 4. The future efforts with respect to this simulation tool is described in Sec 5.

2. RELEVANT WORKS

The simulation of buildings from the energy system perspective can be broadly classified into two categories: (1) Based on a physical thermodynamic model (2) Based on the constituent generators and loads in the house. The former case (see for instance, [5]) focuses mainly on the comfort aspect of the occupant (e.g. a human in a house or servers in a data center). Hence, it is not much suitable for the home energy management applications that encompass many types of appliances. We would review some key simulation works which consider the latter approach.

In [11], DRSim, a cyber physical Simulator for Demand Response is proposed. The simulator uses an agent oriented modeling which can fuse data from different sources and infer a community's consumption behaviour. The Energy demand model component of DRSim projects the appliance, activities and their relationships into spatial dimension like room. While this approach of energy generation is useful to capture the average consumption of a household, applications requiring finer granular data may not benefit. In [8], the authors illustrate a bottom up approach of generating hourly consumption data from the aggregation of hourly consumption of individual appliance in a household, which is however too coarse for many applications. The HomeSim simulator proposed in [10], has a scheduler-node approach to model a household. The node can represent any of energy sources, loads, storage and hybrid element of a house. The scheduler determines which nodes are active at what time. The paper illustrates the use of this simulator to validate different technolgies and scenarios in a grid.

The approach proposed in this paper models appliances so as to capture both their deterministic power characteristics as well as contextual usage characteristics. The usage of internal components as the fundamental element will allow for flexible sampling rate for data generation application which is not straight forward in the existing literature. The key application under focus is the Non-Intrusive Load Monitoring, whose fundamental limitations are analyzed to some extent in [3], but in practice a simulation tool would be for a thorough analysis. The recent work [2] has a similar approach, but focuses more on obtaining data from extensive appliance level sensing to extract models and then generating simulated data.

3. SIMULATION APPROACH

Appliances in a household are made up of fundamental electrical components. The load as seen by the electric grid and measured by the smart meter will be seen as a combination of these fundamental electrical components. A non-exhaustive list of this could be: (1) Heater (Resistive, Inductive) & Lighting (2) Motor (DC, Induction) (3) Electro mechanical (Solenoids) (4) Switch Mode Power Supply (SMPS) for electronic devices. Each of these fundamental components have a different response as loads to the AC power input to them and are also used in different ways across different appliances (e.g., the induction motor used in a Washer is different from that of the compressors used in a refrigerator). A pictorial representation of this approach of modeling a household is shown in the Fig 1. In [6], the authors propose that internal component analysis of appliances could be helpful in analysing complex waveforms of multi-state appliances. This work also attempts to use a similar approach, but for data generation purposes.

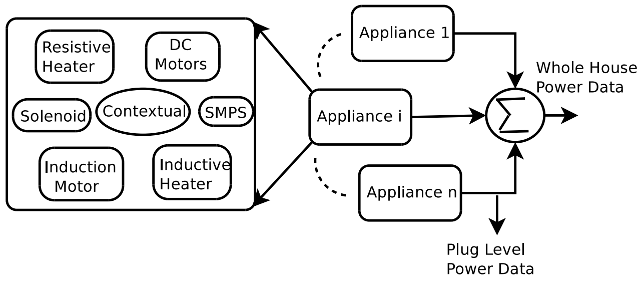

Figure 1: House Model as Appliance Combination

3.1 Appliance Model Characteristics

The appliance representation is driven by the response of the internal components to the ON/OFF and mode select inputs from the user when powered by the household electricity supply. These could be broadly classified into two:

- Transient Characteristics: Capture the dynamic response of the internal components to user inputs and are visible only in high frequency data (at least 120Hz).
- Steady State Characteristics: Those characteristics visible at low sampling rates. (e.g. data from smart meters being installed across the world).

There are multiple electrical parameters that could be extracted that can give data revealing underlying phenomenon. Active & Reactive Power, Current Waveforms (Startup & Harmonics) and Transient voltage noise are the chief among them [4], apart from the average consumption.

All the electrical components can be summarized as an RLC (Resistor-Inductor-Current) network. However, the models of each appliance can be summarized based on the response of the fundamental components to incident power. The usage characteristics of the appliance, attempts to model the influence of external factors. This can be summarized as:

$$A = (\Theta, \Gamma_1, ...\Gamma_n, \tau, \phi, \eta) \quad (1)$$

where Θ refers to the set of internal components that are relevant to a particular appliance A and its type of usage (number of times a day an appliance is used, always ON or intermittent etc.). $\Gamma_1 ... \Gamma_n$ are vectors of random variables, which represent the electrical parameters corresponding to each of the n internal components. A Gaussian distribution accommodates the variation in the electrical characteristics due to various unmodeled nonlinear factors. τ captures the temporal correlation of the appliance usage (time of day, day of week, month of year etc.), whereas ϕ captures the relationship with the external factors (e.g. temperature). η represents the correlation between appliances based on their usage (e.g. washer usage precedes dryer).

The model parameters can be defined in a number of ways. For a simulation-only approach, this shall be based on the

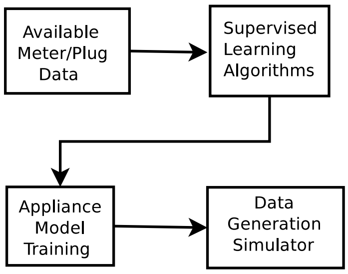

Figure 2: An Analysis-Synthesis approach for Parameter Identification

generic information available in literature and data bases [1]. The databases can be automatically accessed and missing information could be manually entered. The usage characteristics of τ and ϕ are represented through a time series with the probability of usage as the ordinate. η is a matrix of 2-tuple with the entry η_{ij} representing the correlation coefficient between the appliances i and j as the first entry and the time delay as the second entry. While this form of model parameters are efficient to represent the characteristics of interest, this can be developed further to capture more complex characteristics.

An elegant approach to identify model parameters is by learning from the available data, which is briefed in the next section.

3.2 Learning From Available Data

As was discussed in Sec 1, there is a small number of available open data sets. Instrumenting a few houses with multiple sensors for a short period of time is also feasible in pilot deployment. Such data availability shall be utilized to extract the model characteristics of the appliances to train the model. This learned model can then be used to generate the data. This analysis-synthesis approach is represented in Fig 2. The extraction of characteristics can be broadly into three types:

- Learning from plug data: In this case, the model parameters are learned directly from the data.
- Learning from annotated meter data: In this case, the annotation of the appliance usage will help extract relevant data and train the model parameters.
- Learning from partially annotated meter data: In this case, the annotation is inaccurate and appropriate measures should be taken up (e.g. locating appropriate high power transitions to identify an appliance around the annotation) and may need some manual intervention as well.

3.3 Data Generation

The models can only represent the individual appliance operation. However, the household data generation would need a background controller which will use the characteristics (especially τ, ϕ and η) of the appliance to invoke their operation. In a simplistic sense, the controller represents the human intervention which will utilize the contextual information captured by these model parameters to switch the appliance ON/OFF. Some sample tasks the controller needs to perform shall be summarized as:

- To choose the number of times an appliance should be invoked for that day (from η)
- To decide on the precise time of the day when the appliance should be used (from τ)
- To utilize the external temperature of the day to decide on the precise usage characteristics (from ϕ)

Apart from these, the controller also needs to tune the data generation process based on whether it is for a single day or for a few days or for a prolonged period of time. In case of the timeline being more than one day, the controller needs to track the past generated data to ensure the necessary long term behaviour (e.g. washer usage of once in 3 days) are met. In this case, the controller needs to have additional capabilities which are beyond the scope for this paper.

4. APPLICATION: SIMULATION FOR NILM

The smart meters that are being installed across the world have two types of configurations:

- Home Area Network (HAN): Active Power and Cumulative Average consumption sampled at once in 10 seconds (typically 1-10 seconds).
- Advanced Metering Interface (AMI): Average power consumption sampled at once in 15 mins.

An example of the generated HAN data and the actual data for a single European household is shown in Fig 3. It is to be noted that the simulated waveform does not aim to reproduce the original data. However, it captures all the necessary appliance characteristics based on the simulation models whose parameters were trained from a larger set of real data. The house under focus has partial annotation for washer, dryer, and refrigerator. The broad characteristics extracted from the meter data were as follows:

- Washer: Heater power and ON time, Motor power and ON time, Time of the day usage, Usage relationship with dryer.
- Dryer: Heater power, Heater ON/OFF cycle time, Nonheater power, Time of day usage, Usage relationship with washer.
- Refrigerator: Compressor power, duty cycle.

A semi-supervised learning algorithm was employed to learn these characteristics, where, the algorithm will search for known prior pattern for these appliances and train the model parameters. For instance, the washer pattern involves two distinct components and hence their characteristics were extracted through appropriate pattern search. Similarly, refrigerator is operated throughout the day with a fairly constant cycle pattern which is captured by the learning algorithm. Apart from the known annotated appliances, the data generation algorithm also created other components: the phantom load which represents the non-changing loads (manifest as offset in the aggregate data); the evening consumption due to the use of appliances like TV, lights etc.; and unknown appliances which were represented by amplitude characteristics with high variance to accommodate for the uncertainty. This data was used to test the NILM algorithms developed to disaggregate smart meter data into

[1] e.g. Open ICEcat

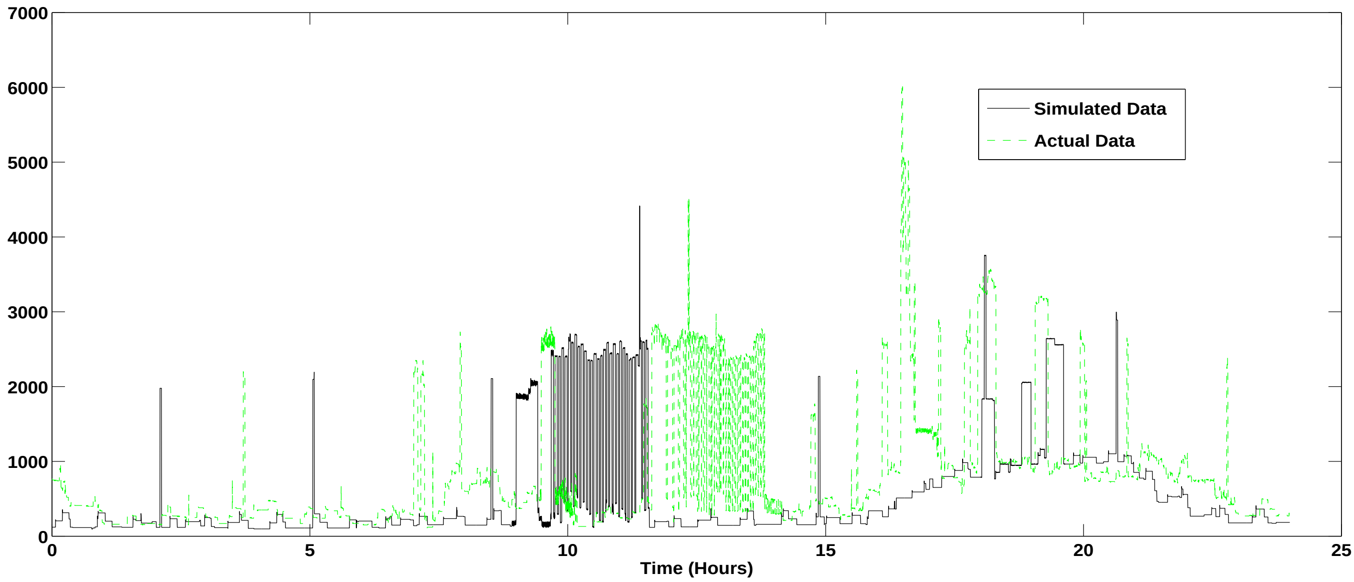

Figure 3: An Illustration of Simulated Data in comparison with the actual Meter Data

its appliance level consumption. The ready availability of ground-truth enabled understanding the accuracy and identify areas of improvement for the algorithm.

For the case of AMI, the annotated HAN data is averaged over a 15-min window with appropriate scaling so that the generated data represents the smart meter output. This data enabled testing some NILM algorithms which were intended to disaggregate the data into appliance categories (e.g. laundry, cold storage).

5. FUTURE WORKS

The paper has proposed a simulation to generate the appliance-level and house-level electricity data. This approach's usefulness was illustrated for NILM applications where it could be used to compare candidate algorithms. Moving forward, the approach could be extended to simulate the other household electric component categories like storage (batteries) and energy sources (e.g. solar panels). This would enable to aid in developing a more complete energy management simulation platform. A network of such bottom-up generated houses could also help understand user behaviours and articulate better Demand Response programs from a utility point of view.

6. REFERENCES

[1] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges. 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 (SustKDD)*, Beijing, China, Aug. 2012.

[2] S. Barker, S. Kalra, D. Irwin, and P. Shenoy. Empirical characterization and modeling of electrical loads in smart homes. In *Green Computing Conference (IGCC), 2013 International*, pages 1–10. IEEE, 2013.

[3] R. Dong, L. Ratliff, H. Ohlsson, and S. S. Sastry. Fundamental limits of nonintrusive load monitoring. *arXiv preprint arXiv:1310.7850*, 2013.

[4] S. Gupta, M. S. Reynolds, and S. N. Patel. Electrisense: single-point sensing using emi for electrical event detection and classification in the home. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, pages 139–148. ACM, 2010.

[5] J. Gustafsson, J. Delsing, and J. Van Deventer. Thermodynamic simulation of a detached house with district heating subcentral. In *Systems Conference, 2008 2nd Annual IEEE*, pages 1–8. IEEE, 2008.

[6] J. Kelly and W. Knottenbelt. Disaggregating multi-state appliances from smart meter data. In *SIGMETRICS'12*. ACM, 2012.

[7] 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.

[8] J. Paatero and P. Lund. A model for generating household electricity load profiles. *International journal of energy research*, 30(5):273–290, 2006.

[9] A. Reinhardt, P. Baumann, D. Burgstahler, M. Hollick, H. Chonov, M. Werner, and R. Steinmetz. On the Accuracy of Appliance Identification Based on Distributed Load Metering Data. In *Proceedings of the 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability (SustainIT)*, pages 1–9, 2012.

[10] J. Venkatesh, B. Aksanli, J.-C. Junqua, P. Morin, and T. Rosing. Homesim: Comprehensive, smart, residential electrical energy simulation and scheduling. In *Green Computing Conference (IGCC), 2013 International*, pages 1–8, 2013.

[11] T. K. Wijaya, D. Banerjee, T. Ganu, D. Chakraborty, S. Battacharya, T. Papaioannou, D. P. Seetharam, and K. Aberer. Drsim: A cyber physical simulator for demand response systems. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2013.