Available online at www.sciencedirect.com

SciVerse ScienceDirect

Energy Procedia 00 (2015) 000–000

Energy Procedia

www.elsevier.com/locate/procedia

The 7th International Conference on Applied Energy – ICAE2015

CBSF: A Framework for Accurate Simulation of Appliance Data for Future Smart Grid Applications

Maria Zaffar[a,b], Fahad Javed[b], Naveed Arshad[c]

[a]Kinnaird College, Lahore , Pakistan
[b]GIFT University, Gujranwala, Pakistan
[c]LUMS University, Lahore, Pakistan

Abstract

Efficient energy management of computing infrastructure has been shown to reduce the operational expenses of enterprises as well as contribute in reducing the carbon foot-print of the organization. Various studies have proposed methods to reduce the energy consumption of computing devices at server farms, commercial offices and in consumer households. However, a system that impacts the habits of consumers requires rigorous inspection. From our previous experiences, we found that most of the techniques used in evaluating the response of devices to efficient energy management strategies fall short of representing the actual consumption. This has severe consequences for utilization and adaptation of efficiency measures. To this end in this paper we propose a unique simulation framework which accurately represents the consumption habits of the computing devices and, as our results show, provide a more realistic picture of the impact of efficiency measures both on consumer impact and on efficiency response.

© 2015 The Authors. Published by Elsevier Ltd.
Selection and/or peer-review under responsibility of ICAE

Smart Grids; simulation; Demand side management;

1. Introduction

It has been argued that conserving energy is the most important fuel for electricity in the future smart grid [8]. It is no surprise that in enterprises where electricity usage is a major part of cost of operations, efficient use of electricity becomes critical. There are various strategies to improve the efficiency of device usage ranging from more energy efficient equipment to device management policy to reduce the electricity cost based on the prices or availability of electricity. Another aspect of reducing cost and green house emissions is integration of renewable energy sources with grid electricity supply. To maximize the utilization of renewable sources, demand side management, or willfully managing the device consumption rates has been applied to great effect [8]. Many researchers have proposed solutions to reduce the cost of electricity usage for commercial entities and home consumers. Some of these studies have been conducted on actual machines [11] but the majority of the solutions have been validated through simulations.

However, for an office space to adapt a strategy to reduce energy consumption where its machines are controlled by an algorithm require very precise evaluations. This was observed in our attempt to apply the CDRS [11] and ColorPower [10] algorithm in our computing lab. The algorithm turned the power consumption of specific devices to backup power to increase the throughput of renewable energy source. In simulation studies the algorithm worked very well but the demand profiles, built following the standard state-of-the-art simulation strategies discussed of the US department of Energy's EnergyPlus [4] and [3] did not represent the actual demand profile. This resulted in the algorithm assuming a specific load profile for a device but the variation in the actual load resulted in computers runningout of their backup storage and causing disruption in the operations of the lab.

Our study of the literature for simulating loads profile resulted in two divergent groups. Some studies such as those by Burke and colleagues [3] and by Usman and colleagues [1] built complex and very accurate models of consumption. However, Burke did not incorporate human centric variables. The model proposed by Usman and colleagues catered for human specific actions as well but it was too abstract for proving the impact of load management for a specific office space. Secondly, the models are so complex that using them for traditional study is very cumbersome. The second simulation methods were mostly used by demand side management and demand response authors [2, 5, 9, 10, 12, 13, 14]. In almost all the cases observed, the researchers assumed the consumption to be off specific random variable distribution. A simulation would be randomly selecting a consumption value from the distribution. We found the first strategy as overkill and the second as underwhelming to convince the lab directors in our university to adapt our system.

To ameliorate this problem, we propose a simulation framework which builds its load profile and their response from previous experiences within the lab. Instead of assuming some random distribution to represent the work load, we build a machine learning model which learns the habits and patterns of the consumer and of devices. A simulation is a response of the learnt model against the natural and efficiency measure stimuli. As our results show, the resulting simulation is more accurate and correct. Accurate in the sense that the simulated output's mean squared error is less than the probability distribution models and correct in that the states that machine learnt simulation produces are valid.

The paper is organized in the following way; in the next section we describe our simulation methodology. This is followed by description of data collection and simulation execution. Next we discuss how we have used case based reasoning (CBR) as the machine learning engine. Though CBR is not a classical machine learning engine, but we use it to showcase the fact that learning can be effective in simulation in this specific scenario. This is followed with results and our conclusion and future works.

2. Simulation Strategy

Simulating consumption of a device in a natural setting is a non-trivial task. Within an office space the consumption of various electrical devices such as air-conditioners, computers, and lighting and ventilation systems is dependent upon the human, environmental, and structural aspects. Modeling such a complex interrelated system requires complex models[3] Various simulators exist which provide a very accurate simulation of an environment. US department of energy's EnergyPlus simulator and the model proposed by Burke achieve this goal very effectively [3,4]. However, these energy models are primarily made for static analysis to evaluate the impact of device installation. The models require detailed information such as heater to ambient conduction coefficient, indoor dimension etc. Collection and utilization of such information require additional cost. Secondly, the models to date do not consider human habit, computing, and device specific measures to create the simulation resulting in a complex yet incomplete model.

In comparison, the efficiency measures research, predominantly in the field of demand side management uses a simple way to evaluate their algorithms. This simplicity stems from the complexity in using the aforementioned models for evaluation. Majority of the systems in our literature survey [7] evaluated their strategies by assuming, based on data measurements, a probabilistic distribution for the consumption rate. There are three distributions commonly used. Normal distribution is used by. Average is assumed by. Uniform is assumed by. However, as our experiments show the resulting output due to these assumptions is incorrect.

Our approach to simulate the consumption behavior of the consumers is a median of the two strategies above. Instead of building an explicit model through analyzing the physical environment, we build the model implicitly by observing the consumption of the devices. We argue that since the goal of the simulation is the regular consumption of the device, if we can train a model to learn the behavior of the device then we can replay this model under different parameters to simulate the existing system.

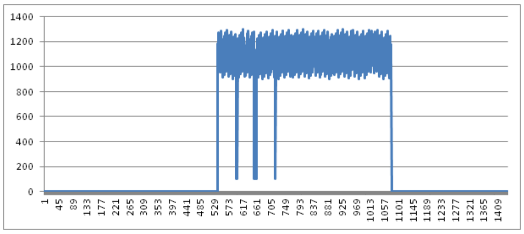

Figure 1 Consumption profile for an Air conditioner over a 24 hour period.

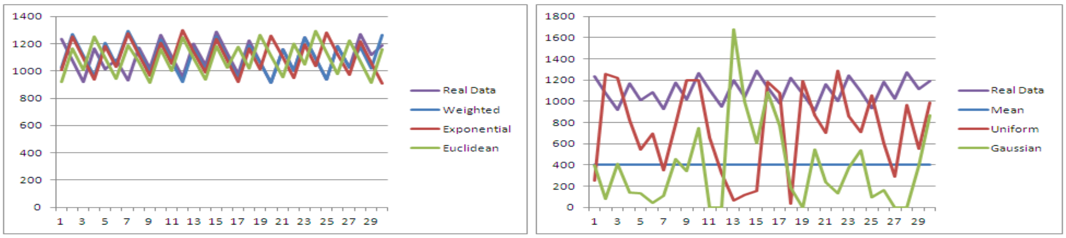

Figure 2 a) Comparison of real output and of CBR b) Comparison of real output and standard simulation methods

To explain this further, consider the consumption behavior of a desktop as shown in figure1. As a casual observation the consumption is between 0 Watts and 1300 Watts. However, if we look closely then we see three states of device. Off at 0 consumption, standby at 122 Watts and running ranging between 885 and 1300 Watts. The timing and duration of the consumption is driven by the consumption habits of the consumer. If we use the casually observed range as is the norm then the output will be similar to the displayed in figure 2(b). Here we show a 30 minute simulation using the standard simulation strategies. As you can see none of the patterns follow the real pattern (in purple). In comparison if we pick datum from the history of the device patterns matching the closest historical record then using three different measures the output will be different. This is shown in figure 2(a). As can be seen the output does not match exactly but is in the correct range. It maybe argued that if we chose the distribution bounds tightly then we get a better results but with the vast variation in device behaviors, finding such a bound is a non-trivial task and is not used in standard practice as observed by the authors. This strategy maybe applied in future but for now is beyond the scope of this work.

What we propose is that we observe the consumption behavior for a number of days. We then use a machine learning engine to learn the trends of consumption. Machine learning in essence is learning from experience. We can then provide the stimulus to this machine learning model to reproduce the behavior of the consumer-device combination to simulate a given scenario. For instance, we provide the model with the day of the week, the model builds the consumption profile which it learnt for that specific day of the week.

3. Case-based Simulation Framework

Our simulation strategy is to learn the device consumption behavior by observing its past history and then build a simulation based on the stimuli we require to simulate. To achieve this goal we propose Case based simulation framework (CBSF).

Case based reasoning is not a traditional machine learning technique such as artificial neural networks, support vector machines etc. However, it is a machine learning technique which in our case is more suited to the specific problem we are attempting to resolve.

We apply case based reasoning for simulation in a wholesomely different way than the traditional method. Our goal is to store previous experience and find the closest matching example from the history. We use the case base of CBR for this purpose. Data from previous days and response to specific stimulus is saved as cases. When the simulation is required to simulate the device behavior for a certain type of actions then case base is searched, the closest matches are retrieved and appropriate changes are made to the response before it is provided to the simulation operator.

Simulation device data for future smart grids is a very new and emerging concept. It is important that we use the right metrics to evaluate the results. For this reason, we will first discuss the evaluation criterions we used to assess our results. We will then discuss the experimental setup which includes the discussion on data collection followed by the results.

3.1. *Evaluation*

The goal of the simulation is to produce a correct simulation from which we can quantify the savings. To evaluate the response of the simulation we use two metrics. First is the accuracy and the second is the correctness. Accuracy is the closeness of the output of the simulation from the actual result. We measure this by calculating the mean squared error (MSE) of the simulated output. Lower MSE value translates to better accuracy. The correctness is defined as the number of simulation outputs which are valid. Valid technically are states which a system can take. For instance in figure 1, the valid ranges are 131 and 179 for active mode and 95 and 105 in sleep mode. Any value which is not in this range is a consumption which is not possible for this device due to the circuitry involved.

To validate our results and see the impact of the size of data we setup 3 experiments with different data sizes. We collected 60,120 and 180 sample events from our system. These observations were used to setup or train the simulation engine. In the case of CBR the observations and the response were placed in the case base. We argue that correctness is a much better measure to evaluate a simulation. A simulation attempts to build a scenario in future under certain constraints. The goal is not to produce an output which exactly matches the actual output but is to provide a possible outcome. If multiple simulations are conducted then the probability distribution of the actual should be followed but not in every case. This is measured more accurately by correctness than by MSE. For instance, a student worker works late on three days out of 5. If we are not concerned on which days she works late then a simulation which simulates three late sittings on any of the 5 days will accurately represent the scenario. However, if we compare the actual results with the simulation and the days of late work don't match then MSE can be high. Nevertheless we report the MSE numbers from our experiments for comparison and completeness.

We used a single case based reasoning engine which stored the data for all the 8 experimental setups. The 120 data points prior to the shutdown point were taken as the input. The CBR engine compared these data points with its case base and returned a 30 datapoint output simulating the result of a shutdown on the actual system. We have used Euclidean, Weighted and Exponential as our distance measures to find the best possible simulated output.

Table 1. Results of proposed framework and of competing strategies.

Technique	Accuracy (MSE)	Correctness (%age)	Accuracy (MSE)	Correctness (%age)	Accuracy (MSE)	Correctness (%age)	Accuracy (MSE)	Correctness (%age)	Accuracy (MSE)	Correctness (%age)	Accuracy (MSE)	Correctness (%age)
Device	**Desktop**		**AC**		**Desktop +CRT**		**Desktop +LCD**		**Solar Power (complete)**		**Solar Power (1stFloor)**	
CBSF												
Exponential	3.13	100	71.17	100	3.84	100	1.929	100	11.41	100	0.081	100
Weighted	3.17	100	47.99	100	2.97	100	15.80	0	11.41	100	26.21	100
Euclidean	2.8	100	75.47	100	2.97	100	15.66	0	11.41	100	26.21	100
Uniform Dist	15.3	60	215.93	40	25.67	36.66	33.22	13.33	245.04	73.33	24.17	23.33
Mean	27	0	305.57	0	31.53	0	40.06	0	132.82	100	10.45	0
Gaussian	12.26	25	426.79	13.33	40.04	3.33	47.28	3.33	183.29	73.33	41.43	33.33

4. Results

We compared our results with the simulations used to validate various DR systems such as those proposed in [5,10,12,13,14]. As discussed earlier, we observed that there are three prevalent methods to simulate household loads. Researchers such as Danabeh and colleagues [5] and Zhang and colleagues [14] used uniform or weighted uniform distribution. Ranade and Beal used both gaussian and uniform. where as Tang and colleagues [1] used mean. In comparison we used the CBSF with the three distance measures discussed above. Table 1shows the mean and correctness measures of the three standard simulation methods and the results of the CBSF.

5. Discussion

Although the result seem very extreme but follow the intuitive feel that the authors observed previous to the use of CBSF for forecasting. As can be observed, correctness of mean value is either 100% or 0%. This makes sense as the mean value (which is the same through out the experiment) is either a valid and correct value or is not. We observed that the mean was a valid value in only 1 out of 6 devices (16%) which is very low. In comparison correctness for Gaussian and normal distribution varies between 0% and 60%. This also makes intuitive sense as we can see from figure 2, only few values of Gaussian and normal fall in the range which a device produces.

In comparison, the correctness of the CBSF is 100% in almost all the cases except for the distance measures of weighted and Euclidean for Desktop+LCD. The 100% result is easy to infer since as we use the actual load profiles of devices, it is impossible for the device profile to be non-existent in real data. The two incidences of lower correctness is due to the closeness in consumption profile of Desktop+CRT and Desktop+LCD. The distance measures of weighted and Euclidean picked the wrong profiles in eachexperiment resulting in the results. Though this points to a better performance for exponential but we at this time do not have sufficient data to claim exponential smoothing as a better distance measure.

References

[1] Naveed Arshad, Usman Ali, and Fahad Javed. A highly configurable simulator for assessing energy usage. Energy Procedia, 42:308{317, 2013.
[2] Jacob Beal, Je_rey Berliner, and Kevin Hunter. Fast precise distributed control for energy demand management. In Self-Adaptive and Self-Organizing Systems (SASO), 2012 IEEE Sixth International Conference on, pages 187{192. IEEE, 2012.
[3] WJ Burke and DM Auslander. Modular and extensible systemic simulation of demand response networks. In CIGRE Canada Conference on Power System, Toronto, Canada, 2008.
[4] Drury B Crawley, Linda K Lawrie, Frederick C Winkelmann, Walter F Buhl, Y Joe Huang, Curtis O Pedersen, Rishard K Strand, Richard J Liesen, Daniel E FIsher, Michael J Witte, et al. Energyplus: creating a new-generation building energy simulation program. Energy and Building, 33(4):319{331, 2001.
[5] Anna Danabeh, Long Zhao, and Bo Zeng. Job scheduling with uncertain local genertion in smart buildings: Two state robust approach. IEEE Transactions on Smart Grids, 2014.
[6] C.W. Gellings. The concept of demand-side management for electric utilities. Proceedings of the IEEE, 73(10):1468 { 1470, oct. 1985.
[7] Fahad Javed. Towards Self-Managing Demand Side Management. PhD thesis, Lahore, Pakistan, 2013.
[8] Fahad Javed and Naveed Arshad. A penny saved is a penny earned: Applying optimization techniques to power management. In 16th IEEE International Conference on the Engineering of Computer-Based Systems (ECBS 2009), 13-16 April 2009, San Francisco, CA, USA, 2009.
[9] Bingnan Jiang and Yunsi Fei. Dynamic residential demand response and distributed generation management in smart microgrid with hierarchical agents. Energy Procedia, 12(0):76 { 90, 2011.
[10] Vinayak V. Ranade and Jacob Beal. Distributed control for small customer energy demand management. In SASO'10, pages 11{20, 2010.
[11] M. Rehman, F. Javed, N. Arshad, M. Ikram, A. Babri, and U. Masood. A cooperative demand responsesystem for small scale commercial consumers. In 5th International Conference on Applied Energy (ICAE 2013), 1 - 4 July 2013, Pretoria, South Africa, 2013.
[12] Bruce Stephen, Stuart Galloway, and Graeme Burt. Self-learning load characteristics model for smart appliances. IEEE Transactions on Smart Grids, 2014.
[13] Shaojie Tang, Quiyuan Huang, Ziang-Yang LI, and Dapeng Wu. In INFOCOM 2013, Proceeding IEEE, 2013.
[14] Wei Zhang, Wenxin Liu, Xin Wang, Liming Liu, and Frank Ferrese. Distributed multiple agent system based online optimal reactive power control for smart grids. IEEE Transactions on Smart Grids, 2014

Biography

Maria Zaffar is currently working as the lead research associate at GIFT University's smart grids technology lab. She completed her bachelors from Kinnaird college.

Dr. Fahad Javed completed his Ph.D. from LUMS university in computer science in 2013. He is currently working as assistant professor at GIFT university where he is leading the research in smart grids technologies and innovations.

Dr Naveed Arshad completed his Ph.D. from University of Colorado at Boulder, USA. Before joining LUMS, Dr Naveed Arshad has worked with ABN AMRO Global IT Systems, Pakistan International Airline. He is part of the Software Engineering Research Group (SERG) at LUMS. This group is undertaking research in various areas of software engineering such as engineering of autonomic systems, conceptual modeling, large scale systems development, etc.

ADDITIONAL INFORMATION FOR ADMINSTRATION USE BY THE ORGANIZATION COMMITTEE

(* ICAE2015 organization committee reserves the right to decide the paper type in the final program)

PLEASE SELECT (CLICK) ONE OF THE FOLLOWING OPTIONS:

I PREFER TO PRESENTING MY PAPER AS

√ ORAL ~~☐ POSTER~~

ALL PAPERS PRESENTED AT ICAE2015 WILL BE PUBLISHED IN ENERGY PROCEDIA, IF YOU DO NOT WANT TO INCLUDE YOUR PAPER IN THE ENERGY PROCEDIA, PLEASE CLICK

☐ NO, I DO **NOT** WANT TO INCLUDE MY PAPER IN THE ENERGY PROCEDIA.