

User-sensitive Scheduling of Home Appliances

Tanuja Bapat
Dept. of CSA, Indian Institute of Science
Bangalore, India
tanuja@csa.iisc.ernet.in

Neha Sengupta, Sunil K. Ghai, Vijay Arya,
Yedendra B. Shrinivasan, Deva Seetharam
IBM Research India
{nesengup, sunilkrghai, vijay.arya,
yshriniv, dseetharam}@in.ibm.com

ABSTRACT

Demand response (DR) programs encourage end-use customers to alter their power consumption in response to DR events such as change in real-time electricity prices. Facilitating household participation in DR programs is essential as the residential sector accounts for a sizable portion of the total energy consumed. However, manually tracking energy prices and deciding on how to schedule home appliances can be a challenge for residential consumers who are accustomed to fixed price electricity tariffs. In this work, we present *Yupik*, a system that helps users respond to real-time electricity prices while being sensitive to their context and lifestyle. *Yupik* combines sensing, analytics, and optimization to generate appliance usage schedules that may be used by households to minimize their energy bill as well as potential lifestyle disruptions. *Yupik* uses jPlugs, appliance level energy metering devices, to continuously monitor the power usage by various home appliances. The consumption patterns as well as data from external sources are analyzed using data mining algorithms to infer user's preferred usage profile. Using the preferred profile as a reference, *Yupik*'s optimization engine generates multiple usage plans that attempt to minimize energy and inconvenience costs. Some of *Yupik*'s capabilities are demonstrated with the help of preliminary data collected from a home that was instrumented with jPlugs to monitor the power usage of a few devices.

Categories and Subject Descriptors

I.m [Computing Methodologies]: Miscellaneous

General Terms

Measurement, Design, Human Factors

Keywords

Demand Response, Energy Consumption Scheduling, Home Automation, Smart Grids

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.

GreenNet'11, August 19, 2011, Toronto, Ontario, Canada.
Copyright 2011 ACM 978-1-4503-0799-4/11/08 ...\$10.00.

1. INTRODUCTION

Demand response (DR) programs encourage end-use customers to alter their power consumption in response to incentives or real-time electricity prices so that peak demand may be reduced [1]. Reducing peak demand preempts the need to invest in additional generation capacity that gets utilized only during narrow peak periods, thereby minimizing costs and added environmental pollution. Studies indicate that even minor shifts in peak demand have major implications in terms of savings for both consumers and utilities [2]. A Lawrence Berkeley National Lab report estimates that during 2006-08, the potential size of peak load reduction from existing DR resources in the US was as much as 5-5.8% of peak national demand [3]. As per another study, DR programs alone could achieve up to half of EU's 2020 targets concerning energy savings and CO₂ emissions [4].

Although traditionally manual DR programs have been used by industrial customers, facilitating household participation is likely to result in considerable savings given that this is a growing sector and accounts for a sizable portion of the total energy consumed. For instance in EU-27 nations, the domestic sector consumed 24.6% of total energy in 2007 while the sector grew by 8% from 1990 and 2007 [5].

Despite the savings possible through DR, the success of these programs essentially hinges upon user participation and their timely response to DR events such as increase or decrease of electricity prices. One of the main barriers in involving households to participate in DR is the lack of effective home automation systems [6, 7]. Manually tracking hourly prices and deciding on how to schedule home appliances can be a challenge for residential consumers who are accustomed to fixed price electricity tariffs. In fact households may not even have the time or knowledge to correctly respond to real-time prices [8]. Automating user's demand response is a challenge as home appliances need to be scheduled to maximize savings while respecting the comfort of consumers and minimizing any lifestyle disruptions.

In this work, we present *Yupik*, a system that helps users respond to real-time electricity prices while being sensitive to their context and lifestyle. Real-time prices are generally published by utilities in advance or can be predicted for the next couple of days as in [8]. *Yupik* is essentially a *planner* that uses variable hourly prices and computes optimal appliance usage schedules for the next *planning horizon* (e.g. a day or two, or a week). The generated schedules can then be used either by the households to plan their usage and/or by the automatic load control systems for scheduling appliances.

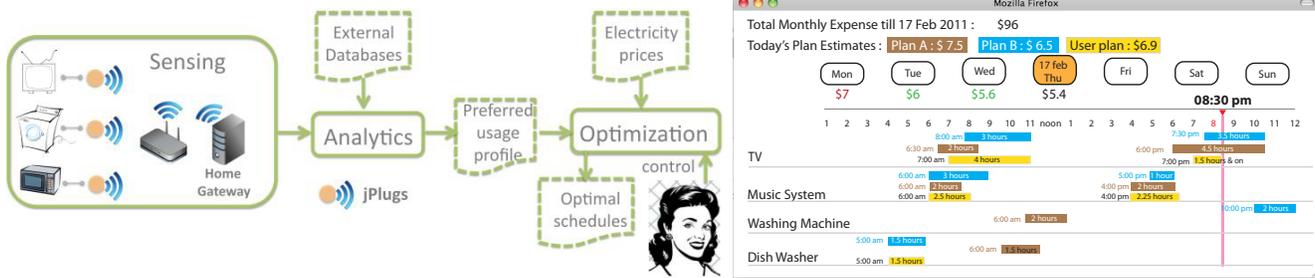


Figure 1: Architecture of Yupik that generates convenient low cost appliance usage schedules. A webpage on the right shows a Gantt chart with sample proposed schedules.

Yupik combines *sensing*, *analytics*, and *optimization* to generate convenient low cost appliance usage schedules. It uses jPlugs, appliance level energy metering devices, to sense household’s appliance usage patterns. The jPlugs continually monitor the power consumed by various appliances at home. The gathered data is processed and clustered by Yupik’s *analytics engine* to derive operating characteristics of appliances such as the number of operating levels, the average power consumed at each level, as well as the time and duration for which the appliance is operated at each of those levels. From historical usage patterns as well as historical and forecasted values of external factors that influence consumption (e.g. day of week, holiday, weather, etc), the engine infers user’s *preferred usage profile* i.e., the likely times, durations, and power levels of appliance usage for the next planning horizon. For instance, the engine considers the day of the week to estimate when the user may utilize her washing machine. It includes the local weather forecast while estimating the time of use (days and hours) and the power levels at which an AC may be used during the next planning horizon.

Taking user’s preferred usage profile as a reference, Yupik’s *optimization engine* schedules appliances by considering their time and load elasticities, operational characteristics, and the cost of inconvenience to the user. Savings come from the fact that usage is deferred to a time slot where the prices are low and/or usage levels (power levels) are reduced. The time elasticity determines whether a device can be deferred from its usual usage times. For e.g., it is possible to reschedule washing machines, but a television does not offer much flexibility. On the other hand, the load elasticity determines whether an appliance can be operated at reduced power levels. For e.g., appliances such as AC, refrigerators and washing machines can operate on multiple levels and the power consumption can change nonlinearly between those levels. For appliances with high time elasticity such as washing machine, Yupik attempts to schedule them on their preferred days that also have lower rates. For appliances that have low time elasticity and are usually non-deferrable, Yupik schedules these close to their preferred usage times. In order to minimize discomfort and increase flexibility, Yupik outputs multiple schedules with costs close to the optimal. An interface displays the schedules along with their projected electricity bill amounts and the costs for the past usages, thus allowing users to track their monthly expenses.

While there have been prior efforts [8, 9] to optimally schedule home appliances in the presence of real-time prices, most of them pay little attention to user’s comfort and lifestyle which is the focus of our work. Lu et al [10] introduce a parameter to consider temperature-based comfort zones, but

the approach is limited to ventilation and cooling devices. In [8], authors allow devices to be scheduled within a user-specified time window. On the other hand our work infers usage preferences based on the data collected from jPlugs which give much richer information about the usage patterns and operating characteristics than what can be provided by the consumers. In order to minimize lifestyle disruptions, devices are scheduled close to their normal operating routines. We also take into account the hold times of devices (i.e. the duration for which a device should remain ON after being switched on to do useful work) which is not considered in prior work. In order to provide flexibility to the user, we output multiple low cost schedules instead of only one optimal as opposed to prior work. Our work has some similarities with [11] where authors assume the availability of a preference matrix that lists user’s preferences by rank. However they do not specify how the matrix could be populated. In our work, historical usage patterns along with data from external sources is used to infer user’s preferred usage profile. While in [11], authors focus on scheduling multiple users to balance the total load, our focus is on scheduling appliances in a single household considering hourly energy prices.

2. SYSTEM

Fig 1 shows the high level architecture of Yupik. Yupik combines sensing, analytics, and optimization to generate usage plans that minimize both cost and inconvenience.

2.1 Sensing

Yupik monitors the energy consumed by individual appliances in a household through smart plugs known as jPlugs. A jPlug looks like a normal power strip. The appliance that needs to be monitored is plugged into jPlug’s socket which in turn plugs into the wall socket. jPlugs sense the voltage, frequency and current drawn by the connected appliance and compute parameters such as as power factor, active, reactive, and apparent power. All these quantities are time stamped, affixed with appliance specific labels, and transmitted over a secure Wi-Fi link (802.11b) to a home gateway periodically every ten seconds (jPlugs implement the entire IP stack). The gateway accumulates data from multiple jPlugs which is then processed by Yupik to generate optimal appliance usage schedules.

Although Yupik is a planner that suggests optimal usage schedules, these schedules can also be used to control home appliances via jPlugs. A jPlug is equipped with a solid-state switching relay and the appliance connected to it can be switched on/off over the network. Unlike most commer-

cially available smart plugs, jPlugs maintain accurate timing information and use the Simple Network Time Protocol (SNTP) to synchronize their clocks over the network. This is valuable for DR systems.

2.2 Analytics

Yupik’s analytics engine analyzes consumption data collected from jPlugs as well as the historical and forecasted values of external parameters and applies data mining methods to infer user’s preferred usage profile for the next planning horizon.

For days that have passed, the diurnal usage pattern and the operational duration corresponding to each home appliance are extracted from jPlug data. For each device, the timestamped data entries of power usages are processed using the *Two-step clustering approach* [12] to determine the number of operating levels and the mean power consumed at each level. The day is divided into fixed size time slots (e.g. hourly) and the usage pattern of a device is obtained by marking its operating level in each slot. For example, if the device was used in time slots 2, 3, and 6 at power levels 10W, 30W, and 20W respectively, the usage pattern is $\{(2, 10W), (3, 30W), (6, 20W)\}$. The operational duration is simply the number of slots where the device was in ON state (in the above example, the duration is 3 time slots).

The user’s *preferred usage profile* for a device on a certain day of the planning horizon is essentially a frequently occurring usage pattern i.e. the set of time slots when the user prefers to use the device along with the power level for each time slot. The novelty of the analytics engine is that it estimates the preferred usage profile by relating variations in appliance usage with the user’s *Context Vector*. The context vector includes parameters that influence user’s consumption and whose historical as well as true/forecasted values are available. Examples of such parameters include day of week, weather (temperature), summer or winter, holiday or not, TV schedule, etc. For e.g., a context vector could hold the following 3 parameters: day of week $\in \{1, \dots, 7\}$, holiday $\in \{0, 1\}$, Cricket match on TV $\in \{0, 1\}$. Thus the consumption data of a device d for the previous n days can be viewed as a list of context vectors along with the associated usage patterns $\{(\mathbf{C}_i, \mathbf{U}_i^d)\}_{i=1}^n$. Since the parameters in the context vector are such that their values are available for days in the planning horizon, the problem of estimating user’s preferred usage profile reduces to determining the likely usage pattern *given* the context vector. The estimation is performed in two steps. Firstly the operational duration for the device is predicted. Then, the likely usage pattern with the predicted operational duration is mined.

Prediction of operational duration (\hat{y}^d): The prediction is modeled as a regression problem, where the input is the context vector and the output is the predicted operational duration. Given the past data $\{(\mathbf{C}_i, y_i^d)\}_{i=1}^n$, where \mathbf{C}_i is the context vector $(\theta_i^{(1)}, \dots, \theta_i^{(k)})$ of k parameter values and y_i^d is the number of time slots used by device d , the regression problem is to predict the value of y^d for a new context vector \mathbf{C} such that $y^d = f(\mathbf{C})$. We use the linear regression model that takes the form as:

$$f(\mathbf{C}) = w_0 + \sum_{j=1}^m w_j \phi_j(\mathbf{C})$$

where $\phi_j(\cdot)$ are basis functions operated on \mathbf{C} . The prediction model learns the weights w_j by solving the regularized

least squares formulation on the past data of the device and the learnt parameters are used for the prediction of future operational duration of the same device with a confidence interval ω_d . Since the estimate may be a real number, it is rounded to an integer.

Inference of preferred usage profile: In order to determine the frequently used usage pattern of a device, firstly the frequently used operational time slots of the device are determined. As the usage of appliances varies on a day-to-day basis and depends on parameters in the context vector, the usage history is filtered based on the given context vector. Within the filtered data, the problem of locating frequently used operational time slots is modeled as the *frequent item-set mining* problem from data mining. Using standard notation [13], let \mathcal{D} be the database of diurnal "transactions", each containing the set of operational time slots of the device over the set of all possible time slots \mathcal{I} in a day. The collection of frequent item-sets of different sizes in \mathcal{D} is defined as $\mathcal{F}(\mathcal{D}, \sigma) := \{X \subseteq \mathcal{I} \mid \text{support}(X, \mathcal{D}) \geq \sigma\}$ where $\text{support}(X, \mathcal{D})$ denotes the proportion of transactions in \mathcal{D} that contain the item-set X and $\sigma \in [0, 1]$ is a threshold. We are interested in picking the most likely frequent item-set of time slots whose size matches the operational duration predicted by regression $(\hat{y}^d \pm \omega_d)$. For this, we start with a large value of support threshold σ and decrement it step-wise until we obtain item-sets of required size. At each step the Apriori algorithm [13] is used to determine the collection of frequent item-sets $\mathcal{F}(\mathcal{D}, \sigma)$.

Having located the most frequent item-set of operational time slots, each time slot in the item-set is marked by the average operating power level considering the subset of data records that contain this frequent item-set. This results in a frequent usage pattern that is regarded as the user’s preferred usage profile for the specified day. If the planning horizon spans multiple days, then the preferred usage profile of each day is used by the optimization module as a base to schedule appliances.

2.3 Optimization

The optimization engine proposes plans to schedule home appliances based on given hourly energy prices and the preferred usage profile. The scheduling problem is modeled as an integer linear program (ILP) wherein the objective is to minimize both the energy and inconvenience costs for the consumer over a planning horizon. This is subject to consumer and appliance constraints such as the operational duration and hold time per device, number of appliances that can be run together, need to run appliances at specific times, etc. Both deferrable and reducible devices are considered. If a device is deferrable, its usage can be shifted in time, and if reducible, its power level can be changed. The planning horizon is divided into fixed sized time slots and the scheduling problem is to determine the slots where each device is used and at what power level.

Notation Let the planning horizon $T = \{1, \dots, \kappa\}$ consist of κ time slots and R_t be the energy price at time slot t . Let \mathcal{D} be the set of all devices to be scheduled. Let L_i denote the set of permissible power levels¹ for device i including 0 for OFF state. Let H_i denote the *hold time* for i i.e. once the device is ON, it should be ON for at least H_i slots. Let $y_{it} \in L_i$ denote the device usage indicator to be determined

¹We only consider power levels that do useful work and exclude modes such as *stand by* that need not be scheduled.

i.e., $y_{it} = 0$ if i is OFF in time slot t and $y_{it} = p > 0$ if i is ON state at power level p .

Elasticity Given the user's preferred usage profile for a device, that is, the time slots and power levels at which the user regularly uses it, the optimization engine schedules the device close to these time slots. For every device i , the user provides a *tolerance* of δ_i time slots. If in the preferred usage profile, the device i is scheduled at time τ for a duration of s slots, then it is rescheduled somewhere in the interval $[\tau - \delta_i, \tau + s + \delta_i]$ in the optimized schedule. δ can be set large or small depending on the time elasticity of the device, for e.g. a high value for washing machine and a low value for a water heater or TV.

Multiple Chunks If a device is used multiple times in the usage profile, then each continuous usage chunk is considered as a separate device for scheduling over T . Let $D(i)$ denote the set of chunks corresponding to device i . Let the set of all such time chunks including all devices be D' . Let S_i and M_i denote the start time and length of chunk i respectively. Let P_{it} denote the given power level in the t th slot of chunk i . Let P_i^{avg} denote the average power level of chunk i considering all its slots.

Energy Cost A device j at timeslot t consumes y_{jt} units of electricity priced at R_t . Therefore the energy cost is

$$E = \sum_{t=1}^k \left(\sum_{j=1}^{d'} y_{jt} \right) \times R_t$$

Inconvenience Cost The inconvenience cost is composed of two parts. If a device j is scheduled outside of the region it was being used earlier, then the distance of the new schedule from the current use is given by $I_1^j =$

$$\sum_{t=1}^{\lceil S_j \rceil - 1} (S_j - t) |y_{jt} - P_j^{avg}| + \sum_{t=\lceil S_j + M_j \rceil}^k (t - (S_j + M_j)) |y_{jt} - P_j^{avg}|$$

where the first term sums the distances times difference in power levels for slots scheduled before the region of regular usage. The second term does the same for slots after the region. If a device is scheduled within its current usage region, then the difference in power levels of use is given by the following cost

$$I_2^j = \sum_{t=\lceil S_j \rceil}^{\lceil S_j + M_j \rceil - 1} |P_j - y_{jt}|$$

Objective function Thus the objective function with both the energy and inconvenience costs is

$$\min_{y_{jt}, j \in D', t \in T} g = E + \sum_{j=1}^{d'} [\epsilon_1 * I_1^j + \epsilon_2 * I_2^j]$$

where ϵ_1 and ϵ_2 are control parameters.

Device level Constraint A device i at time slot t must be scheduled at a power level within set L_i . Besides as different chunks of the same device are scheduled independently and may overlap, the total power level used for a slot must also lie in L_i . To ensure both these, we introduce binary variables $b_{\lambda it} \in \{0, 1\}$, $\lambda \in L_i$ and the following constraints:

$$\sum_{j \in D(i)} y_{jt} = \sum_{\lambda \in L_i} \lambda * b_{\lambda it}, \quad \sum_{\lambda \in L_i} b_{\lambda it} = 1 \quad \forall t \in T, \forall i \in D$$

Conservation of Operational Hours Each device must be able to complete its task by running for the essential number of time slots, i.e., the sum of durations of all chunks of a device in the preferred profile must equal the number of assigned slots in the optimized schedule:

$$\sum_{j \in D(i)} L_j = \sum_{t=1}^k \sum_{\lambda \in L_i \setminus 0} b_{\lambda it} \quad \forall i \in D$$

Tolerance Constraint The tolerance value δ_j constrains the slots where the device j can be scheduled. Since it was previously scheduled at time S_j in the preferred profile, it can only be scheduled in slots $[S_j - \delta_j, S_j + M_j + \delta_j]$ in the optimized schedule. This constraint is written as:

$$\sum_{t=\lceil S_j \rceil - \delta_j}^{\lceil S_j + M_j \rceil + \delta_j - 1} y_{jt} = M_j \quad \forall j \in D'$$

Hold Time Constraint If the device j is switched ON at time slot t , then the device should be ON for at least next H_j time slots. To introduce this constraint, we define the change of state indicator variables $x_{jt} = y_{jt} - y_{jt-1}$, $\forall j \in D', t \in T$. Hence we wish to define the following constraint:

$$x_{jt} > 0 \quad \& \quad x_{jt} = y_{jt} \quad \Rightarrow \quad y_{jt+r-1} > 0 \quad \forall r \in \{2, \dots, H_j\}$$

The above if-condition can be linearized using two sets of binary variables a_{jt}, c_{jt} , $j \in D', t \in T$ and number of constraints of the order of $|D'| |T|$. This is done by assuming lower and upper bounds on x_{jt}, y_{jt} , and $x_{jt} - y_{jt}$. In particular, if u is the maximum power level of any device, then these bounds could be $-u, u$.

Fuse Capacity The sum of maximum power levels of all devices scheduled in the same time slot should not exceed the fuse capacity of the household Fuse. Let $L_i^{max} = \max L_i$.

$$\sum_{i=1}^d L_i^{max} \left(\sum_{\lambda \in L_i \setminus 0} b_{\lambda it} \right) \leq \text{Fuse} \quad \forall t \in T$$

Hard Constraints If the user wishes to explicitly provide a preference to switch ON/OFF devices during specific time periods, then hard constraints can be introduced on the decision variables y_{jt} as per the requirement.

Multiple Solutions The above ILP can output multiple solutions. Besides the difference between the best and the second best solutions may sometimes be minimal in terms of cost in dollars, however the difference in inconvenience may be significant. Therefore to provide flexibility to the users, the optimization engine outputs multiple solutions within a specified optimality gap. Let $g(Y)$ denote the objective function value for a solution $Y = \{y_{jt}\}, j \in D', t \in T$. Then the engine outputs all solutions within a relative optimality gap of $\alpha = |g(Y) - g(Y^*)|/g(Y)$, where Y^* is the optimal. α is set close to 0, for e.g. 5%. The solutions are aggregated and multiple usage plans are displayed to the user.

3. INTERFACE

The aggregated schedules output by the optimization engine are displayed using a Gantt chart on a webpage. For instance, these could be viewed by the user via a browser on a desktop or a smartphone. The interface provides the following information: (i) proposed times of use for each appliance as per the Yupik schedules (ii) Cost estimates of proposed schedules (iii) actual time of use for each appliance by the user (iv) Cost of a actual usage by the user (v) Actual daily and monthly expense incurred thus far.

Fig. 1 (right) shows the Gantt chart that has 24 time slots, one for each hour, with the current time shown using a red

line (08:30 pm). The chart displays two proposed plans A and B for the current day (17 Feb 2011). Each plan is identified by a unique color and shows the proposed times and durations of use for each appliance. A user may choose to use an appliance during proposed time slots or deviate from these. The yellow timelines show the actual usages by the user. The corresponding costs are shown under ‘Actual Use’. The current day is shown highlighted (orange) among days of the week and the actual daily expenses incurred are displayed below each day. When the expense for a day exceeds the optimal estimate, it is highlighted in red (e.g. Mon), else green (e.g. Tue). Thus the interface informs the user of possible optimal usage plans and their actual usages and helps them keep track of daily and monthly expenses.

4. EXPERIMENTS

We use preliminary data from an ongoing project where jPlugs are being tested in a few homes in Bangalore, India. For analytics, along with some of our code, we used the clustering and frequent item-set mining subroutines from data mining open-source WEKA [14]. CPLEX [15] was used to solve the ILPs for optimization (although solving ILPs can be NP-hard, household appliance scheduling does not result in a large-scale problem and is easily handled by CPLEX). Since the local utility does not provide variable electricity prices for residential customers, the day-ahead hourly prices published by ComED [16] were used.

We use four weeks of consumption data from a home where jPlugs monitored the usage of a TV, a music system, and a power strip. We did not have data from devices which consume more power since the instrumented jPlugs had a 5 Amps limit. Therefore in reality scheduling these devices may not result in significant savings in dollars. However, we use this data to validate the analytics module in terms of searching frequent usage patterns and the optimization module by considering normalized relative savings with respect to no scheduling. We also treat the power strip as a virtual device as we do not know which devices were connected to it.

Analytics Fig. 2(a) shows the histogram of power consumed by the TV. We clearly see two clusters with significant difference in their means. This shows that the TV operates at two power levels. The left cluster has its mean at 15.77W and represents the standby mode, while the right cluster has its mean at 89.15W and represents the normal operating mode. The two-step clustering approach essentially retrieves these two means. As jPlugs log the power consumed even when the device is switched off, zero power entries were filtered out before clustering. Fig. 2(b) shows the same for the power strip where it is hard to determine if it operates at either two or three different power levels (the rightmost power level is used only for a small fraction of time). This is perhaps because different devices may have been connected to the power strip at different times. The two-step clustering approach retrieves two power levels with means 20.50W and 55.66W.

Next, to retrieve user’s preferred profile for TV usage during the four weeks, the days were divided into 24 hourly time slots and each slot was marked on or off according to the usage (standby was considered off). Fig. 2(c) shows the frequency of usage for each time slot during the 4 weeks. We divided the days into weekdays and weekends and used regression to infer the number of operational hours for a

weekday, which resulted in 5.8 slots. Fig. 2(d) shows three frequently used item-sets of time slots A, B, and C of sizes 5, 4 retrieved by the apriori algorithm. The item-set A = 19-23h of size 5 was the closest in size to the predicted duration of 5.8 and was also the most frequent item-set with a support threshold of $\sigma = 0.4$. This means that for 40% of the days, the TV was on from 19-23h. Thus this is regarded as the regular/preferred usage profile for TV.

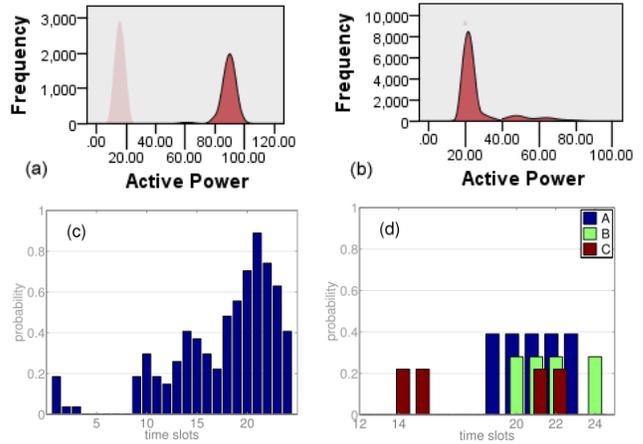


Figure 2: Clusters and frequent usage patterns

Optimization For optimization, we have a total of 3 devices i.e. $D = \{\text{TV, music system, power strip}\}$. The frequent usage pattern of each device was retrieved by the analytics engine and used as the user’s preferred usage profile. Scheduling was performed for a planning horizon of one day with 24 time slots i.e. $T = \{1, \dots, 24\}$. The optimization

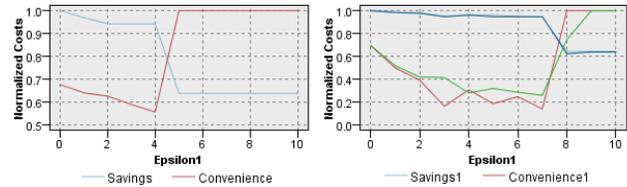


Figure 3: Variation with ϵ_1

ILP has 3 control parameters: tolerance $\delta_i \forall i \in D$, and ϵ_1, ϵ_2 that weigh the inconvenience costs. We show how variations in these parameters help users attain the desired level of balance between optimal energy price and comfort level.

Fig. 3(left) plots the normalized savings versus convenience costs when optimization is performed with a different value of ϵ_1 each time, keeping ϵ_2 and $\delta_i \forall i \in D$ fixed. Savings are computed with respect to the cost incurred by user’s preferred usage schedule retrieved from analytics, without any optimization. When ϵ_1 is low, the inconvenience cost of scheduling devices outside their regular/preferred usage regions remains low and the objective function pays more attention to reducing energy costs. This results in high savings but low convenience. As ϵ_1 increases, devices are scheduled closer to their regular usage regions resulting in more convenience and low savings. Fig. 3(right) shows the case when optimization outputs multiple schedules close to the optimal. We see that even though savings are more or less the same in two schedules, convenience costs differ, providing users more flexibility in scheduling devices. The

control parameter ϵ_2 weighs the inconvenience caused as a consequence of change in power levels. Fig. 4(left) shows the variation in savings and inconvenience as a function of ϵ_2 . We see that relative to ϵ_1 , higher values of ϵ_2 are needed to bring about a change in costs. This is because ϵ_2 only weighs the difference in power levels of slots which are scheduled at the preferred usage times whereas ϵ_1 considers both difference in time as well as difference in power levels for slots scheduled outside the preferred usage regions. The tolerance

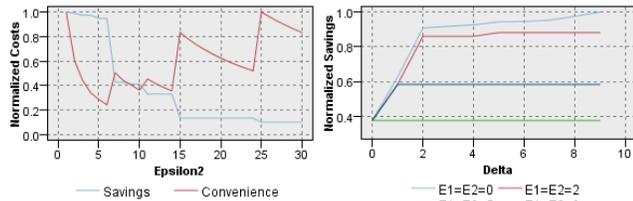


Figure 4: Variation with ϵ_2 , tolerance δ

parameter δ_i determines the time elasticity of each appliance i.e. devices are allowed to be scheduled only around a window of their frequently used time slots. Fig. 4(right) shows the normalized savings as a function of $\delta = \delta_i \forall i \in D$ for different values of ϵ_1, ϵ_2 . As ϵ_1, ϵ_2 move from low to high, the effect of δ reduces as the inconvenience costs become too high, resulting in devices being scheduled at their preferred times, thus lowering the savings. In practice, graphs of costs versus control parameters for average-case homes with standard appliances could be used to guide users to set appropriate values of control parameters for their homes.

5. DISCUSSIONS AND FUTURE WORK

In this work, we presented Yupik, a planner that helps users respond to variable hourly electricity prices. Yupik combines sensing, analytics, and optimization to generate flexible appliance usage schedules that may be used by the users to minimize their energy bill as well as potential lifestyle disruptions. The consumption data collected via jPlugs is correlated with external data such as the day of week, holiday, etc to determine the regular or preferred operating routines of devices. Devices are scheduled close to their regular usage times by taking into account their time and load elasticities so that user inconvenience is minimized. We presented preliminary experimental results based on jPlug data collected from a few low power devices in a single home. In future work, we plan to collect and analyze data from more homes with devices consuming more power as well as experiment with data from external sources such as weather database.

Understanding user behavior and inferring their preferences using data collected from a number of sensor devices and external sources is a challenging problem. While our work estimates a unique preferred usage profile, each usage profile has an associated likelihood. Future work will consider the problem of scheduling appliances given multiple usage profiles and their likelihoods. To infer preferred usage profiles, future work will also consider the more general approach of association rule mining where we wish to mine rules of the form *context vector* \Rightarrow *usage pattern*. Another scheduling challenge is how to update schedules based on differences between predicted prices and real prices that become available before each hour.

6. REFERENCES

- [1] U.S. Department of Energy. Benefits of demand response in electricity markets and recommendations for achieving them, 2006. <http://eetd.lbl.gov/ea/emp/reports/congress-1252d.pdf>.
- [2] K. Spees and L. Lave. Impacts of Responsive Load in PJM: Load Shifting and Real Time Pricing. *The Energy Journal*, 29(2):101–122, 2008.
- [3] P. Cappers, C. Goldman, and D. Kathan. Demand Response in U.S. Electricity Markets: Empirical Evidence. Technical Report LBNL-2124E, Lawrence Berkeley National Lab, June 2009.
- [4] A. Chardon, O. Almen, P. E. Lewis, J. Stromback, and B. Chateau. Demand Response: a decisive breakthrough for Europe. How Europe could save Gigawatts, Billions of Euros and Millions of tons of CO2. Capgemini Report, June 2008.
- [5] European Environment Agency. Final energy consumption by sector (CSI 027/ENER 016) - Assessment published Sep 2010.
- [6] Demand response program evaluation—Final report Quantum Consulting Inc. and Summit Blue Consulting, LLC Working Group 2 Measurement and Evaluation Committee, and California Edison Company, Apr 2005.
- [7] M. Ann-Piette, G. Ghatikar, S. Kiliccote, D. Watson, E. Koch, and D. Hennage. Design and operation of an open, interoperable automated demand response infrastructure for commercial buildings. *Journal of Computing and Information Science in Engineering*, 9:1–9, June 2009.
- [8] A.-H. Mohsenian-Rad and A. Leon-Garcia. Optimal residential load control with price prediction in real-time electricity pricing environments. *Smart Grid, IEEE Transactions on*, 1(2):120–133, sept 2010.
- [9] A.J. Conejo, J.M. Morales, and L. Baringo. Real-time demand response model. *Smart Grid, IEEE Transactions on*, 1(3):236–242, dec 2010.
- [10] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse. The smart thermostat: using occupancy sensors to save energy in homes. In *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems, SenSys '10*, pages 211–224, 2010.
- [11] C. Wang and M. de Groot. Managing end-user preferences in the smart grid. In *Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking, e-Energy '10*, pages 105–114, 2010.
- [12] J. Bacher, K. Wenzig, and M. Vogler. SPSS Twostep Cluster - A First Evaluation. In *RC33 Sixth International Conference on Social Science Methodology*, 2004.
- [13] R. Agrawal and R. Srikant. Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases VLDB*, pages 487–499, September 1994.
- [14] Waikato Machine Learning Group. Weka, 2010. <http://mloss.org/software/view/16/>.
- [15] IBM ILOG CPLEX. <http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/>.
- [16] ComED. <http://www.thewattspot.com/>.