EnergyCoupon: A Case Study on Incentive-based Demand Response in Smart Grid*

Bainan Xia, Hao Ming, Ki-Yeob Lee, Yuanyuan Li, Yuqi Zhou, Shantanu Bansal
Srinivas Shakkottai and Le Xie

Dept. of ECE, Texas A&M University

Email: {xiabainan, mingming511, kiyeoblee, ylnff, zhouyuqi, shantanu88, sshakkot, le.xie}@tamu.edu

ABSTRACT

In competitive electricity market systems such as Texas, Load Serving Entities (LSEs) purchase energy in a wholesale market and then sell it in a retail market. In wholesale market, LSEs see dynamic real-time prices, whereas in retail market, LSEs typically provide flat rate contracts to end-consumers. An intuitive idea to induce savings for LSEs is to shift loads from high-price hours to low-price hours of the wholesale market by incentive-based demand response. We design and implement such a system, called EnergyCoupon, to provide coupon incentives to users and collect their responses. We run the experiment over household participants through the summer of 2016. The experimental results suggest behavioral changes in energy consumption can be achieved, which could be beneficial to both end users and LSEs. This paper presents the system set up, key algorithms, as well as experimental results analysis.

KEYWORDS

EnergyCoupon, incentive-based demand response, load serving entity (LSE), electricity market

ACM Reference format:

Bainan Xia, Hao Ming, Ki-Yeob Lee, Yuanyuan Li, Yuqi Zhou, Shantanu Bansal
Srinivas Shakkottai and Le Xie
Dept. of ECE, Texas A&M University
Email: {xiabainan, mingming511, kiyeoblee, ylnff, zhouyuqi, shantanu88, sshakkot, le.xie}@tamu.edu. 2017. EnergyCoupon: A Case Study on Incentive-based Demand Response in Smart Grid. In *Proceedings of e-Energy '17, Shatin, Hong Kong, May 16-19, 2017*, 11 pages.
DOI: <http://dx.doi.org/10.1145/3077839.3077862>

*Research was funded in part by NSF grants CNS-1149458, AST 1443891, EFRI-1440969, CCF-1331863, ECCS-1546682, and IIS-1636772. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding agencies

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. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

e-Energy '17, Shatin, Hong Kong

© 2017 ACM. 978-1-4503-5036-5/17/05...\$15.00

DOI: <http://dx.doi.org/10.1145/3077839.3077862>

1 INTRODUCTION

The integration of intermittent renewable energy sources such as wind and solar power into the electricity grid has resulted in much interest in demand side management that benefits the whole power system. For example, the State of Texas serves as high as 20% of its peak electricity demand using wind energy, whose total contribution is quite variable. Integration of heterogeneous energy sources in several states in the US, as well as many other countries in the world is achieved using a wholesale level electricity market, wherein aggregators (such as load serving entities, LSEs) purchase energy from generating companies. In the US, these markets are run by Independent System Operators (ISOs), who ensure the reliable availability of electricity. Here, the wholesale price of electricity seen by aggregators changes temporally as it follows diurnal variations in demand and supply. In a market run by the Electric Reliability Council of Texas (ERCOT), electricity can be purchased day-ahead using predictions of demand and supply, or in near-real-time using fine grain data available every fifteen minutes. Aggregators in this market typically use a combination of both approaches.

The aggregators in turn participate in retail level markets, wherein individual customers are provided with different options of electricity services. A good example is the retail market in several cities in Texas, in which over a hundred LSEs compete for residential customers via (slightly) different pricing plans. Retail customers typically pay a fixed or tiered fee, and such a pricing scheme does not allow real-time variations with respect to the wholesale price of electricity. Consequently, load serving entities (LSEs) would benefit from end users' shift of energy consumption from high-wholesale-price hours to low-wholesale-price hours. Achieving such load shifts requires the willing participation of end users, which is referred to Demand Response (DR)[19]. Several ISOs in the US actively engage in DR programs for the purpose of energy resource management, capacity planning and ancillary services [7–9, 13, 18], often at the level of large scale industrial consumers by offering fixed rates for demand reduction.

Retail demand response at the level of residential users requires incentive schemes at the level of individuals, who, through their collective actions, could potentially reduce demand quite significantly. For example, [12] suggests using a game-theoretic model that an LSE could reduce electricity purchase costs to the tune of about \$1 to \$2 per home per week during the summer months in Texas (depending on

the wholesale price of electricity) through a coupon-based incentive scheme. The implicit assumption is that the LSE is a price-taker and that the variation in its demand profile will not perturb the wholesale prices. However, actually implementing such a system requires many components, including predicting the price of electricity, identifying the likely baseline usage of residential customers, designing an incentive scheme to encourage reductions at the correct times, and providing incentive information to users.

The objective of this work is to design, develop and demonstrate EnergyCoupon, a prototypical Coupon Incentive-based Demand Response (CIDR) system, wherein users obtain coupons in return for desirable actions. These coupons are then used to participate in a weekly lottery with gift cards as the prizes. What is an effective architecture to prototype such a system? What kind of impact would it have on user behavior? Would the savings that result lead to a viable system?

1.1 Related Work

The CIDR system that we aim to develop is a form of “nudge engine” that is designed to encourage beneficial behavior. One such experimental scheme uses a lottery-based-incentive scheme to encourage uniform load on public transportation systems [14]. Dynamic rewards for reduction of congestion in roadways have also been applied [1]. Another experiment attempts to design algorithms to coordinate demand flexibility to enable the full utilization of variable renewable generation [11]. Similarly, [21] conducts an experiment on the use of social games with monetary rewards to promote reduction in office electricity usage (such as turning off lights). In the context of DR, [27] present a simulation study of a system that illustrates the voluntary, self-controlled, penalty-free, and low equipment requirements of CIDR. The social value of energy savings is determined in [3], which provides bi-monthly energy saving information for a year for free, and then asks how much users will pay for the information in the following year. Different from these works, our focus is on demand estimation of individual homes, accurate price prediction, active incentivization of users to respond to DR events and system behavior with repeated decision making of end-users.

On the modeling and analysis side, [15] model the system as a Stackelberg game with two stages: setting the coupon values followed by the consumer’s choice. [23] study demand-response as trading off the cost taking beneficial actions in terms of changing electricity consumption as compared to the likelihood of winning a lottery using the concept of a mean field game. In both the above the game is played in one shot, and there is no evolution over time. The dynamic version of the DR problem with lottery-based incentives is studied in [12], also using the mean field game model. Here, the LSE awards different numbers of coupons in proportion to the customer energy consumption in different periods of time during the day. The customers maintain a belief about the number of coupons available with the others, and optimize

their actions based on this belief. The main result is to show existence of a so-called mean field equilibrium in this setting, and to numerically evaluate the sustainability of the system. Our system design is inspired by this work, and we aim to show how such a system might be constructed in practice.

1.2 System Overview

We design and implement the EnergyCoupon system that provides electricity usage targets to customers in a real-time fashion, measures their responses, awards coupons accordingly, and conducts a periodic lottery to reward users. Figure 1 illustrates the system architecture, which consists of five parts that includes three classes of functionalities (shown using different colors/shades), an SQL Database and an Android/iOS App that forms the user interface. In what follows, we provide a brief overview of the different parts, and postpone detailed descriptions to later sections.

Figure 1: System Architecture

SQL Database: The database stores all the historical and real-time information of the EnergyCoupon system, and is hosted on a core server that runs 24 hours a day during the experiment. The database interacts with all parts of the system via SQL APIs.

Android/iOS App: Apps are developed for different mobile OS. End-users are requested to install the App on their mobile devices. DR at the user-end is handled by the App. Figure 2 illustrates the activities on the Android version of the App. On the left is the home page that provides personalized tips on electricity usage, displays the past day’s usage (solid blue line) and real-time usage targets (green and yellow) for the current day. Hitting the targets results in some number of coupons being awarded. The middle activity provides the user with their electricity usage history over the past two weeks, and also allows them to compare the coupons received against the user that received the most coupons. The activity on the right is the lottery interface, which allows the user to specify the number of coupons that he wishes to enter in a particular lottery.

Figure 2: EnergyCoupon App Activities

The functionality at the backend of the system is separated into three parts: data acquisition (blue), online algorithms (green) and incentives and responses (pink).

Smartmeter Texas: This is the source of electricity consumption data of each customer, obtained at a 15 minute resolution. Details are presented in Section 3.1.

ERCOT Data: This is the source of real time and day ahead electricity pricing data. Details are presented in Section 3.1.

Weather Data: This is the source of weather information needed by our price prediction and usage estimation algorithms. Details are presented in Section 3.1.

Peak Time Estimate: This is a price prediction algorithm, whose purpose is to determine whether the realtime electricity price is going to be high or low two hours in advance of the event. Details are presented in Section 2.1.

Baseline Estimate: A good estimate of the customer's likely electricity usage is required in order to determine what kind of incentives to provide, and to calculate the impact that such incentives have on behavior. We develop an algorithm to calculate this baseline usage that involves historical usage and weather data. Details are presented in Section 2.2.

Tips and Usage Statistics: We provide suggestions to our users on when coupons are likely to be offered, and how their weekly electricity usage compares with those of their neighbors. These personalized tips are generated based on specific usage statistics of all homes. Details are presented in Section 3.3.

Coupon Generation: The system generates personalized coupon targets based on the results of the price prediction and baseline estimation routines. Users are awarded coupons based on the percentage reduction in their consumption. Details are presented in Section 3.2.

Lottery: We run a weekly lottery, in which the user has a chance of converting coupons into real gift cards. Users participate in a lottery by deciding how many coupons to enter in that particular lottery (or wait until they accumulate more coupons). Details are presented in Section 3.4.

1.3 Main Results

This work is perhaps first-of-its-kind in deploying actual systems to solicit end users' electricity flexibility via innovative incentive mechanisms such as lottery coupons. The main results of this paper consist of

- (1) a description of the entire experimental system, as well as the key supporting algorithms;
- (2) empirical analysis of end users' energy consumption behavior subject to the lottery coupons; and
- (3) analysis of the load serving entities' cost and benefit, as well as projection of the impact on society with larger scale deployment.

First, we implemented a real system to conduct CIDR and ran a small scale practical experiment on it. Though, limited by the small number of participants, we believe it represents a valid prototype of building a complete practical DR system. Details of system implementation are available in Section 3. Second, it is shown from the experimental results that the energy consumption profile of the active user in our system can be changed by DR in a quite effective manner. Details are presented in Section 5.2. Finally, based on the posterior experiment analysis, the LSEs reveal clear savings in purchasing energy, roughly \$1 per active user per week, via the CIDR in our proposed system. More discussion on this result is available in Section 5.3.

2 KEY ALGORITHMS IN THE SYSTEM

2.1 Price Prediction

The EnergyCoupon system aims to incentivize end-consumers to shift loads from high-price hours to low-price hours in wholesale market. The market clearing prices in RT market are only published 15 minutes prior to the time to which they apply, which is not sufficient for the users in our system to react. We believe a 2-hour time window is reasonable for participants to plan their actions. Hence, the system needs to predict the electricity price 2-hour ahead accurately. There is much earlier research on electricity price prediction [5, 6, 10, 17, 26]. However, most of the literature propose models based on the Time Series Analysis, which could achieve a higher accuracy but only works well within half an hour of the target time.

Different from the classic price prediction, our system has certain specific requirements that requires new algorithm design. We are typically interested in identifying the electricity price as being either "high" or "low" 2-hours in advance, rather than producing an exact value prediction. Based on our proposed scheme, whenever the user shifts load from one time period to another, if the price in original time period is higher, it will yield extra savings for LSE in purchasing certain amount of energy in wholesale market. Inspired by these needs, we formulate the price prediction in our system into a classification problem. A classifier is designed to catch the time periods at which the price is higher than most of the rest of the day. Since the algorithm needs to run online, low

computing complexity is essential. One of the simplest and widely used classifiers that fits our situation is the decision tree. We first go through the classic procedure of building the corresponding tree structure, and then summarize this part with a performance analysis posterior to the experiment.

2.1.1 Data Preparation. As a data-driven approach, the training data set needs to be carefully prepared. The experiment is conducted in Houston, TX. Because of the hot weather in Texas summer, the energy consumption is dominated by air conditioning usage. Since weather has an impact on renewable energy availability, such as wind farms, we construct our data sets with five fundamental feature classes: Price (π), Demand (P), Temperature (T), Humidity (H) and WindSpeed (W). Considering the time-periodicity of the features, we believe constructing the training set using the data from the whole year of 2015 is sufficient. For details on data acquisition, please see section 3.1. The next step is to label the sample set. According to our requirements, the price label is binary, i.e. either “high” or “low”, which implies a proper threshold of high price is needed. Table 1 shows the average number of high price appearances per day by different thresholds in the chosen data set.

Table 1: Daily Average High Price Appearances by Thresholds

Price Threshold (\$/MWh)	10	20	30	40	50	60	70	80
Average High Price #	94	57	13	4	3	2	2	1

Note that we pick time resolution to be 15 minutes which gives 96 samples per feature class per day. Table 1 indicates the high price appears several times a day with a threshold of 40 and becomes very rare with thresholds above 50. Meanwhile, based on the observation in the training stage, a lower threshold gives higher fitting error rate. Thus, we choose the high price threshold to be 50 in the following study.

2.1.2 Feature Selection. Given the labeled data set, feature selection is conducted based on the self/cross-correlations between different features and the label. In a particular time period, the price label is related to the current feature values as well as earlier ones. For example, we divide the time space into 96 15-minute slots in a day. The electricity price in time slot t , π_t , corresponds not only to π_t, P_t, H_t, W_t , but also to the earlier values, say T_{t-1} or W_{t-10} . We calculate the self/cross-correlations between each feature class with different time-slot lags and the label, and then select the top four features in each of the five feature classes. Figure 3 shows the self-correlation of electricity price with different lags. Since we are predicting the price label 2-hour in advance, the top four price features are the latest published ones, i.e. $\pi_{t-8}, \pi_{t-9}, \pi_{t-10}, \pi_{t-11}$. Figure 4 shows the cross-correlation between price and temperature. Choosing the top four ones gives us $T_t, T_{t-1}, T_{t-2}, T_{t-3}$. Similarly, the selected features in the rest three feature classes are $P_t, P_{t-1}, P_{t-2}, P_{t-3}, H_{t-47}, H_{t-48}, H_{t-49}, H_{t-50}, W_{t-88}, W_{t-89}, W_{t-90}, W_{t-91}$. Note that for the demand and weather related features, accurate

predictions are published a couple of hours ahead, so we could use them directly.

2.1.3 Training and Validation. Given the selected feature set, we construct a training data set containing around 30000 data samples based on the data in 2015. Each sample is a 21-dimension vector including 20 features and 1 label. Before we train the decision tree, two parameters, penalty ratio and minimal leaf size, need to be determined.

Penalty ratio defines as the weight ratio of two kinds of errors: False Positive (FP) and False Negative (FN). In our EnergyCoupon system, FP error means the actual price is low but our prediction is high and FN error defines in the opposite way. Similarly, we define the terms True Positive (TP) and True Negative (TN) to capture the correct predictions respectively. The relationship between the four terms are as following:

$$FP + FN = \text{Total Prediction Errors}$$

$$TP + TN = \text{Total Correct Predictions}$$

$$FP + TN = \text{Total Negative (Low Price) Samples}$$

$$TP + FN = \text{Total Positive (High Price) Samples}$$

FP errors cause more coupons being issued and FN errors imply misses in catching the high price events, which may cause a loss for LSEs. Due to the natural property of lottery scheme, issuing more coupons to customers hardly hurts the system. Thus, it implies that we should put more weight on FN errors to catch more high price events. However, as we increase the weight of FN errors, the total error rate given by the cross validation, which is a classic fitting performance check, increases as well. Thus, the determination of penalty ratio meets a trade-off between overall fitting performance and sensitivity, which is defined by the ratio, $TP/(TP+FN)$.

The minimal leaf size is a specific parameter when building decision trees. It defines the minimum number of samples in each leaf of the final tree. The larger minimal leaf size yields higher fitting errors. However, small minimal leaf size gives higher risky of overfitting. One extreme case is to build a tree per training data sample per leaf, which gives a zero fitting error but only fits the training data. Thus, the determination of minimal leaf size meets a trade-off between fitting errors and risk of overfitting.

We want to choose a proper pair of the two parameters such that the sensitivity is high and the cross-validation error is low. Figure 5 illustrates the scatter plot of the parameter pairs on the plane of sensitivity and cross validation error. Each marker is a pair of the two parameters. We apply the filters in two dimensions: sensitivity ≥ 0.7 and cross validation error ≤ 0.12 , and then choose the largest minimal leaf size over the resulting set. The final value of the two parameters are minimal leaf size = 70 and penalty ratio = 1 : 8.

Table 2 shows the overall performance statistics for price prediction in both fitting stage and experiment/testing stage. Note that in the training stage, we have around 30000 samples from the whole year of 2015, while in the testing stage, we

Figure 3: Self-correlation of Electricity Price

Figure 4: Cross-correlation between Temperature and Price

Figure 5: Parameter Pairs on Sensitivity vs. Cross Validation Error Plane

only have around 5000 samples from the 3 months of 2016 summer during the experiment. We observe that the error rate doubles in the testing stage and sensitivity decreases about 8%. Various reasons could cause a performance drop of a classifier in testing. In our situation, one critical issue is the electricity price could be affected by other reasons rather than the features in our feature set, such as network congestion and natural gas price changes. Furthermore, we use the historical values of the features in the training stage, but take the predicted values which are updated every 15 minutes in online testing, where the realizations are only revealed posterior to the time.

Table 2: Price Prediction Performance

	TP	TN	FP	FN	Sensitivity	Error
Fitting	605	32838	1273	234	72.11%	4.31%
Testing	126	4809	472	71	63.96%	9.91%

2.2 Baseline Prediction

Baseline is defined as the regular electricity consumption without DR events. We should point out that only posterior energy consumption can be observed in our experiment. The actual baseline is invisible due to the DR. However, a good estimation of the baseline is essential for demand reduction measurement and load shifting analysis, which are critical in our study. Equation (1) provides the definition of demand reduction on interval k of day D .

$$P_{DR}^D(k) = P_{base}^D(k) - P_{real}^D(k) \quad (1)$$

Different from the traditional load forecast, when considering baseline prediction in DR, there are some phenomena that we need to consider specifically. The first one called “**user’s dilemma**”: if an active user keeps reducing his energy consumption from the predicted baseline, the baseline estimation, which only takes the most recent history into consideration, will keep reducing as well. As a result, this user will have to reduce the energy consumption dramatically to meet the targets and eventually quit the DR program.

The second is the “**baseline cheating**” effect [4]. By intentionally increasing electricity consumption over a few days, one could create a high baseline for the following day, and get high rewards without making costly energy reductions before the algorithm recovers. A good baseline prediction method should eliminate both phenomenon. In the rest of this section, we will first introduce two traditional methods, then discuss the Hybrid Method that we used in the experiment and lastly propose the Similar-day Method for post analysis.

2.2.1 Conventional Method. Conventional load forecast by averaging consumptions in past few days is widely used by the ISO in DR programs [16]. In equation (2), N is the number of past days considered (usually $N = 10$).

$$P_{base}^D(k) = \frac{1}{N} \sum_{i=1}^N P_{base}^{D-i}(k) \quad (2)$$

It gives relatively precise predictions via simple calculations. However, it is obvious that since the calculation depends on consumption over the past few days, an active user faces lower and lower baseline (“**user’s dilemma**”); In addition, “**baseline cheating**” is also possible. Thus, the conventional method is not quite suitable for DR programs.

2.2.2 Regression Method. An alternative option for baseline prediction is using regression, such as Gaussian process [25] and Artificial Neural Network (ANN). These give more accurate predictions, but still take into account the most recent histories, even last-minute energy consumption, as inputs. In this way, “**user’s dilemma**” and “**baseline cheating**” cannot be avoided. Moreover, many regression methods have bad performance in rear future predictions, such as day ahead/hour-ahead.

2.2.3 Hybrid Method. During the experiment, a hybrid method is designed by using the weighted average of the output of conventional method and the control group consumptions.

$$P_{base}^D(k) = \beta \frac{1}{N} \sum_{i=1}^N P_{base}^{D-i}(k) + (1 - \beta) \frac{1}{M} \sum_{j=1}^M P_{base,j}(k) \quad (3)$$

$\beta \in [0, 1]$ is the weight parameter. We apply $N = 10$ on weekdays and $N = 4$ at weekends. $P_{base,j}(k)$ represents the typical energy usage during interval k for control group member j . The control group is pre-selected from Pecan Street Dataport, which is a good data source for non-DR consumers [2]. $M = 40$, which is the number of selected anonymous residential consumers with similar house size and house age. The baseline of a control group member on Monday for a particular month is the average consumption of all Mondays in that month during time interval k . So we have specific baselines for all 7 days in a week in control group.

The Hybrid Method eliminates both the “user’s dilemma” and “baseline cheating” effects. In addition, by introducing a control group, the influences of ambient variables (weather, humidity) are naturally involved. However, in the post-experiment analysis, it reveals significant behavior differences between control group and the users in the experiment as well as large bias on the total energy consumption. With these concerns, we propose the following Similar Day Method.

2.2.4 “Similar Day” Method. Historical energy consumption *before DR programs* and temperature data are required in this method. First we divide 24 hours in a day into several sections. For each section, historical days that have similar temperature pattern (measured by mean square error (5)) with the target section are selected and simple average is calculated across those days. N_s indicates the number of similar days selected and N_t is the number of intervals in a particular section.

$$P_{base}^D(k) = \frac{1}{N_s} \sum_{i=1}^{N_s} P_{base}^i(k) \quad (4)$$

$$T_{MSE}^{ij} = \frac{1}{N_t} \sum_{k=1}^{N_t} (T^i(k) - T^j(k)) \quad (5)$$

In order to eliminate the impact of other factors, which are independent of ambient variables such as temperature, on the energy usage and only consider pure energy shifting, we further normalize the total energy consumption with the realizations posterior to the experiment as a post refinement. In other words, we assume our users only shift energy usage from DR intervals to the other intervals without increments or decrements in total energy consumption.

Theoretically “similar day” method should eliminate both “user’s dilemma” and “baseline cheating”. Since baseline only depends on the external temperature of the target section, as well as the historical temperature and consumption *before the DR program begins*, it is not influenced by user’s behavior during the program. Therefore, the baseline is purely objective, and *robust to manipulation* (“baseline cheating”). “User’s dilemma” is eliminated by the same reason.

We conclude in Table 3 that “similar day” is the most suitable baseline prediction method to apply in our system.

Table 3: Comparison of Baseline Estimation Methods

Method	Conventional	Regression	Hybrid	Similar Day
User’s dilemma	Yes	Yes	Partial	No
Baseline cheating	Yes	Yes	Partial	No
Complexity	Low	High	Low	Low
Suitability	No	No	Partial	Yes

3 SYSTEM IMPLEMENTATION

3.1 Data Acquisition

Data sources are essential in applying the algorithms described in Section 2 to our EnergyCoupon mechanism. As shown in Figure 1, there are three main sources of data:

- SmartMeterTexas (SMT)** [22] is a public website supported by Public Utilities Commission of Texas and several large utilities in Texas. The aim of this website is to secure communications between customers and utilities, and enable each smart-meter user to monitor their energy consumption. By downloading high resolution electricity consumption data (15-min) from SMT, EnergyCoupon is able to calculate baseline for each consumer, and also calculate electricity usage shifts after DR events.
- Electric Reliability Council of Texas (ERCOT)** [20] manages the 90% of Texas’s electric load. Much market information is on its website. Our EnergyCoupon app automatically downloads day-ahead and real-time prices, and uses it as part of the input to the prediction of the real-time price for upcoming intervals.
- Weather Underground (WU)** [24] provides real-time and historical weather data including temperature and humidity. This data is used in both baseline prediction and real-time price estimation.

3.2 Target Setting and Coupon Generation

As we discussed in Section 2, the high price threshold is set to be 50\$/MWh since it is greater than the electricity price at most times of the day. Thus, if the user cuts down the energy consumption as we suggest in a high price hour, they are potentially using more energy in the rest of the time periods, which implies a load shift from high price hours to low price hours. Hence, the coupon target should be set whenever a high price occurrence, i.e. a DR event, is predicted. The number of coupons awarded depends on the percentage of the baseline-energy-consumption reduction during the DR event. Obviously, the baseline-energy-consumption should be personalized, and is determined by the baseline prediction algorithm in Section 2.

The multi-layer structure of coupon targets is shown in Figure 6. During each DR event, 2 energy reduction targets and 3 status markers (white, yellow and green) are shown in the App. For example, if a consumer could save 30% to 70% energy during DR event, he/she will be awarded 2 coupons and ends up in the yellow region. Similarly a reduction of more than 70% results in 5 coupons being awarded, and the user is said to have hit a green target. If the consumer saves

less than 30% in a certain DR event, his/her response will be located in white region and no coupon is awarded.

Figure 6: Setting Targets

An important supplement to the scheme of dynamic coupon target setting is that of “fixed coupon” intervals, which are manually set according to the historical DA market analysis. “Fixed coupon” is set to the time periods that high price often occurs in the DA market and has the same appearance with normal targets. Since DA price has a much more consistent pattern compared with RT price, the coupon targets set accordingly are “fixed” in terms of time. Those “fixed coupons” are also created for research purpose: we would have more DR events during the experiment, and more human behavioral changes could be observed.

Figure 7 summarizes how coupon targets is generated by using algorithms from Section 2 and 3.1, 3.2.

Figure 7: *EnergyCoupon* Algorithm Flow Chart

3.3 Tip Generation

Tip generation mechanism provides information so that consumers can learn from their historical behavior, and adjust their consumption during DR events. Tip generation consists of four parts: “received coupons”, “peak-price prediction”, “behavior evaluation” and “fact-based tips”. Among these four parts, “behavior evaluation” gives medals (gold, silver

and bronze) to consumers according to their energy saving for the previous week. By this means consumers can get brief idea about how their neighbors are doing, and further incentivize themselves to engage further in *EnergyCoupon*. “Fact-based tips” provides information such as weekly energy consumption and expected savings in electricity bill. In addition to the partial information shown in the App (Figure 2), complete information is available in weekly email sent to each participant (Figure 8).

Weekly Overview

Figure 8: *EnergyCoupon* Weekly Report

3.4 Lottery

Lottery is widely used in DR to aggregate the small impacts of individual actions in order to provide acceptable incentives to large group of people [11, 14]. Participants can take part in a lottery each Friday by submitting a certain a desired number of coupons, no higher than the number that they possess in their account 2(c). Remaining coupons can be retained for use in future lotteries. Three prizes \$20, \$15 and \$5 of gift cards are issued every week, and users who have larger bids would have higher chance to win these prizes. As the last part of DR period in our system, the lottery completes a DR cycle by collecting the bids from end-consumers and giving out dollar value gifts to the winners.

Some information about the competitors are provided via App and weekly reports as shown in Figure 8. The purpose of this information is to promote strategic behavior in the

lottery as well as competition to obtain coupons. Further, in order to ensure larger participation in the lottery, we allow the customer to collect coupons over several weeks if they choose to do so.

Due to the limited number of participants during our trial, we do not really observe the bidding strategies of the users posterior to the experiment. However, from surveys and discussion with participants, we did find that they were aware of the information provided on coupons possessed by their competitors, and were taking it into account while choosing their entries into the lottery. Participants also experimented with their home thermostat settings to see how best to obtain coupons. Analyzing the user decision problem at a larger scale would be an interesting future study.

4 EXPERIMENT DESIGN

We conducted our small scale trial with 10 residential participants in Houston, Texas. These participants were recruited through door-to-door canvassing, and had no relation to the authors. Participants were asked to create an account on SmartMeterTexas.com, and install the EnergyCoupon App. The purpose of the trial and the usage details of the App were explained to them under an IRB agreement. Demand response was totally voluntary and there were no penalties for not responding. The primary means of responding was for the participants to reduce their air conditioning usage via modifying their thermostat settings. As described earlier, participants were awarded EnergyCoupons for desired behavior. We held a weekly lottery each Friday in which they could use their EnergyCoupons, with the understanding that a the probability of winning was proportional to the number of coupons entered. Consumers who won the prizes received a gift card code through email.

From June 1, 2016 to June 10, 2016, participants were asked to get familiar with the functions of the EnergyCoupon App, and try to respond to DR events. The electricity usage pattern during this period was neither considered as experimental data, nor used as historical data in baseline prediction. From June 10 to August 26, 30-min DR events were generated between 1 PM to 7 PM, whenever the predicted real-time price exceeded the high price threshold of 50\$/MWh, as well in predetermined “fixed coupon intervals,” in which we had calculated that the average price of electricity to be high. These DR events corresponded to offering EnergyCoupons for reduction of electricity usage based on set targets. Normally, there were at least three 30-min events on weekdays and two on weekends. No DR events were generated outside the time interval of 1 PM to 7 PM, even if we predicted high prices.

As we explained in Section 2.2, electricity usage targets were set based on the baseline prediction calculated by the hybrid method (equation (3)). A control group was selected from 40 anonymous houses in the Pecan Street [2] data set, with parameter β is set to be 0.3 for the first week in experiment and 0.8 in the rest of the experiment. Further, the “similar day” method of baseline calculation was used post fact to analyze the data collected.

5 RESULTS AND ANALYSIS

5.1 Overview of Experimental Results

An overview of electricity consumption for all users during the three months of our trial (June, July and August 2016) is shown in Table 4. These were summer months, and air conditioning usage during this period was quite intense. From this table, we see that the total energy consumption P_{total} is highest in July and lowest in June. Within the total usage, the consumption during the period of interest to us (1-7 PM everyday, during which most high prices occur) occupies nearly $\eta_{aft} \approx 34\%$ of total consumption, which is quite consistent across the three months. In terms of consumption during DR events (30-min interval with RT price higher than triggering price when we offer EnergyCoupons for usage reduction), the proportions are different across the months. For instance, in June the consumption during DR events is only $\eta_{DR} = 7.97\%$ of daily energy consumption, whereas in August it is as high as 21.7%. This increase can be attributed the progression of summer with increasing air conditioning usage during the late afternoon period. As mentioned earlier, a concentrated load due to air conditioning is a major target of our DR scheme.

Table 4: Electricity Consumption by Month

Month	$P_{total}(kWh)$	$P_{aft}(kWh)$	$\eta_{aft}(\%)$	$P_{DR}(kWh)$	$\eta_{DR}(\%)$
June	11381	3935	34.6	907	7.97
July	20795	7196	34.6	2889	13.89
August	17138	5738	33.5	3710	21.65

Figure 9 indicates a week-by-week breakdown of total coupons earned by participants and their lottery participation. In terms of coupons earned, a gradual increase in first 7 weeks as participants became familiar with the system occurred, followed by fluctuation in the rest of the weeks can be observed from this figure. In general, participants received most coupons in August. This was likely because August had the more DR events June and July (Table 4). The fluctuation in August could have also be caused by other factors including holiday travel and extremely hot weather.

Figure 10 shows the participation statistics differently by showing the number of coupons and number of lotteries in which each user took part. Note that the User IDs shown are simply arbitrary numbers assigned to them for anonymity and all 10 have been shown in the figure.

An interesting finding is that there seems no obvious correlation between coupons and participation, which could be thought of as two measurement of participant enthusiasm. Note that all participants were aware of the highest coupons received each week (see App description in the Introduction), so they could calibrate their success probabilities. Most played conservatively, and ensured that they had a reasonable number of coupons before trying out the lottery. However, an outlier is user 44, who participated in the lottery every week, although he/she had very few coupons available. This user could be thought of as risk seeking, and playing each lottery although the odds were clearly against him/her. Of course, lottery participation only took a few minutes every

Figure 9: Weekly Coupons Issued & Lottery Participation

Figure 10: Per-User Coupons Earned & Lottery Participation

Figure 11: Load Shifting Effect to User 47 (Jul 9-Jul 13)

week, whereas coupons could only be accumulated through energy saving efforts. The latter might thus be a better indicator of enthusiasm than just lottery participation.

5.2 Energy Profile Analysis

One of our basic objectives of our system is to incentivize behavioral changes in terms of shifting electricity consumption from peak to off-peak times. We describe our results in achieving this objective in this subsection.

The smartmeter data enables us to obtain the electricity consumed by the participant at a 15 minute resolution. Once we have the consumption on a day in which DR events take place (i.e., we offer incentives), the question arises as to what the consumption would have been, had we not offered any incentives. As discussed Section 2.2.4, the “similar day” method is based on generating a day with a similar temperature profile to a day of interest, and using historical data to accurately determine what the baseline usage would have been on such a day. The normalized version of the same method apportions the measured total usage on a given day in proportion to the “similar” day’s usage profile.

As an example, consider Figure 11. Here, the average electricity consumption of a particular participant is calculated over July 9 to July 13, 2016, shown by a solid red line. The behavior during this period is influenced by the DR events, and appears to show a shift towards the later hours. We take the same total electricity consumption for each day, and apportion it across the day based on the “similar day” method. Averaging over these baseline days yields the baseline shown by the dashed blue line. Comparing the two graphs, we see that on average 6.75 kWh of electricity is shifted each day from the interval 11:00 AM to 6:00 PM to the later interval of 6:00 PM to 10:00 PM, with a total shift over the five days of 33.7 kWh.

The above approach appears to yield a powerful method of understanding behavioral changes, since we calculate the baseline temporal usage based on the actual total ground truth usage, so shifts in energy usage times are clearly revealed. We seek to understand the total energy usage shifts away from the peak period (when we have DR events) on a

per participant basis. These shifts would then provide a metric whereby we can measure the average effect of providing EnergyCoupon.

Figure 12 summarizes the monthly energy shifts calculated per participant. We observe that on average, the effect of EnergyCoupon is to shift about 5% of the total electricity consumption from peak hours to off-peak hours. However, we do see some active participants who achieve monthly average shifts of 20% or even 30%. In addition, most users shift a lower percentage of energy in hot weather (August) than in relatively cooler conditions (June and July). This suggests that users have lower flexibility in terms of air conditioning as the weather gets hotter. Some of the results can be identified with reasons unique to the participant. For instance, according to our exit survey user 38 could not change his/her usage much due to the birth of a child. Similarly, user 47 who was very active went on vacation in August, causing the responsiveness to DR events to drop.

In theory, the number of coupons awarded to participants and the percentage energy shift during DR events should be closely correlated. Comparing Figure 12 with 10, we do indeed see that this expectation is largely correct. However, we do see outliers such as user 44, who moved into the residence relatively recently, which meant that the historical baseline was for a family with a different usage pattern. This reveals a limitation in the baseline calculations in that the presumed usage pattern of the participant varied widely with the historical data. This also reveals the fact that careful personalization of the baseline is needed to accurately calculate responsiveness to DR at the small scale that residential users can achieve.

Another interesting finding is that a positive linear correlation exists between energy shifts in DR events and the peak period between 1 PM to 7 PM, as visible in Figure 13. Here, the percentage energy shift for every user is plotted as a scatter plot and is distinguished into weekdays and weekends. The graph suggests that energy shift events during the peak period largely coincide with DR events, suggesting responsiveness of users to DR events. This finding is consistent with our observation in Figure 11: users tend to shift the

Figure 12: Per-User Monthly Energy Shift

Figure 13: Comparison of Energy Saving (1-7pm) and Total Energy Saving

Figure 14: Comparison of Energy Saving (1-7pm) and Total Money Saving

energy due to the DR event out of the peak period to later in the evening. This is a particularly valuable discovery since the peak usage interval often has high real-time prices, and shifting usage to lower-price intervals reduces LSE costs.

In summary, the EnergyCoupon system appears to perform reasonably well at load shaping in the experiment: more than half of the users shift 5% or more energy during the afternoon, while some active users can shift over 20%. Consumers tend to shift a large proportion of load from DR period to off-peak period in the morning or late evening. This helps the LSE to further reduce the cost due to the real-time price difference between peak and off-peak hours.

5.3 LSE's Cost Benefit Analysis

Thus far, we showed how the temporal energy usage profile is shaped during the experiment. We also expect a reduction to the wholesale electricity purchase cost to the LSE, since loads were shifted from peak to off-peak intervals.

As listed in Table 5, the LSE is expected to save totally \$48.8 from the whole experiment, with weekly saving of \$0.44 per participant. In addition, most of the saving comes from the month of August (74%), when many high price events occur. The conclusion is that even small shifts during the hot weather can yield substantial savings to the LSE due to the large variability in wholesale prices. Moreover, if the contribution of the top three users towards reducing LSE costs are summed up, the total value is \$37.8. If such users are defined as active participants, each active participant achieves a cost reduction of \$1.15 per week. This number agrees closely with numerical studies in other work [12].

Table 5: Money Saving for the LSE by Month

Month	All Users (\$)	Top 3 (\$)
June	6.4	10.4
July	6.2	6.5
August	36.2	20.9
Total	48.8	37.8
Average	0.44	1.15

Now, the cost reductions indicated in Table 5 do not include the cost of prizes to participants, which was three gift cards worth \$35 on most weeks. While a cluster of 50 active users would likely be a self sustaining pool with savings of a over a dollar per week per participant, our small scale trial with just three active users, who respond to coupon incentives most positively, and other less active users does not provide enough savings to sustain the system. However, it does validate the idea that behavioral shifts of the kind indicated in [12] are feasible in practice.

Another question that arises is whether there is correlation between energy shift and cost reduction. One would indeed expect positive correlation, given that the energy shift is incentivized based on predicted high prices. Figure 14 confirms this hypothesis. The larger the energy shift during DR event, the more rewards the participant will get, and the more savings the LSE will achieve. This further illustrates the win-win situation between the LSE and electricity users under our EnergyCoupon mechanism.

6 CONCLUSION AND FUTURE WORK

According to the electricity market mechanism, a load shifting from peak hours to off-peak hours yields substantial savings to an LSE in purchasing energy. This paper exposit the empirical study of coupon incentive-based demand response for retail electricity market in the summer of 2016. The experiment results indicate that our system indeed changes the energy consumption profile of active users towards the desired direction. Furthermore, the posterior experiment analysis reveals potential savings for the LSEs.

From the theoretical perspective, an interesting decision making problem may be formulated from our practical lottery system in which participants have an estimate of others' coupons. From the engineering perspective, while the current scale of experiment was small, it did serve to test out all the major systems and validated the main hypothesis that we designed the system around. We are now preparing a much larger scale of experiment in 2017 summer in collaboration with a residential electricity provider.

REFERENCES

- [1] NuRide, Inc. <https://www.nuride.com>.
- [2] Pecan Street Dataport. [online] Available: <https://dataport.pecanstreet.org/>.
- [3] Hunt Allcott and Judd B Kessler. 2015. *The Welfare Effects of Nudges: A Case Study of Energy Use Social Comparisons*. Technical Report. National Bureau of Economic Research.
- [4] Hung-po Chao. 2011. Demand response in wholesale electricity markets: the choice of customer baseline. *Journal of Regulatory Economics* 39, 1 (2011), 68–88.
- [5] Antonio J Conejo, Miguel A Plazas, Rosa Espinola, and Ana B Molina. 2005. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE transactions on power systems* 20, 2 (2005), 1035–1042.
- [6] Javier Contreras, Rosario Espinola, Francisco J Nogales, and Antonio J Conejo. 2003. ARIMA models to predict next-day electricity prices. *IEEE transactions on power systems* 18, 3 (2003), 1014–1020.
- [7] Doug Hurley, Paul Peterson, and Melissa Whited. 2013. Demand response as a power system resource. *Synapse Energy Economics Inc* (2013).
- [8] California ISO. 2011. *Demand response & proxy demand resource-frequently asked questions*.
- [9] California ISO. July 2003. *Initial comments of the California Independent System Operator Corporation on the proposed decision adopting demand response activities and budgets for 2009 through 2011*.
- [10] Tryggvi Jónsson, Pierre Pinson, Henrik Aalborg Nielsen, Henrik Madsen, and Torben Skov Nielsen. 2013. Forecasting electricity spot prices accounting for wind power predictions. *IEEE Transactions on Sustainable Energy* 4, 1 (2013), 210–218.
- [11] A Ju. January 2017. *Experiment makes energy savings a game*. [online] Available: <http://news.cornell.edu/stories/2014/09/experiment-makes-energy-savings-game>.
- [12] Jian Li, Bainan Xia, Xinbo Geng, Hao Ming, Srinivas Shakkottai, Vijay Subramanian, and Le Xie. 2015. Energy coupon: A mean field game perspective on demand response in smart grids. *ACM SIGMETRICS Performance Evaluation Review* 43, 1 (2015), 455–456.
- [13] James McAnany. 2015. *2014 demand response operations–market activity report*.
- [14] Deepak Merugu, Balaji S Prabhakar, and N Rama. 2009. An incentive mechanism for decongesting the roads: A pilot program in bangalore. In *Proc. of ACM NetEcon Workshop*. Citeseer.
- [15] Hao Ming and Le Xie. 2014. Analysis of coupon incentive-based demand response with bounded consumer rationality. In *North American Power Symposium (NAPS), 2014*. IEEE, 1–6.
- [16] Saeed Mohajeryami. 2016. *Customer Baseline (CBL) Calculation for New York ISO*. [online] Available: <http://webpages.uncc.edu/smohajer/Papers/WTR1.pdf>.
- [17] Francisco Javier Nogales, Javier Contreras, Antonio J Conejo, and Rosario Espinola. 2002. Forecasting next-day electricity prices by time series models. *IEEE Transactions on power systems* 17, 2 (2002), 342–348.
- [18] New York Independent System Operator (NYISO). July 2009. *Day-ahead demand response program manual*.
- [19] U.S.Department of Energy. 2006. *Benefit of Demand Response in Electricity Market and Recommendations for Achieving Them*.
- [20] Electric Reliability Council of Texas. *ERCOT.com*. [online] Available: <http://www.ercot.com/>.
- [21] Shijia Pan, Yulai Shen, Zheng Sun, Priya Mahajan, Lin Zhang, and Pei Zhang. 2013. Demo abstract: saving energy in smart commercial buildings through social gaming. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 43–46.
- [22] Public Utilities Commission of Texas. *SmartMeterTexas.com*. Public Utilities Commission of Texas. <https://www.smartmetertexas.com>.
- [23] G. A. Schwartz, H. Tembine, S. Amin, and S. S. Sastry. 2012. Electricity Demand Shaping via Randomized Rewards: A Mean Field Game Approach. *Allerton Conference on Communication, Control, and Computing* (2012).
- [24] Weather Underground. *wunderground.com*. [online] Available: <https://www.wunderground.com>.
- [25] Yang Weng and Ram Rajagopal. 2015. Probabilistic baseline estimation via gaussian process. In *2015 IEEE Power & Energy Society General Meeting*. IEEE, 1–5.
- [26] Lei Wu and Mohammad Shahidehpour. 2010. A hybrid model for day-ahead price forecasting. *IEEE Transactions on Power Systems* 25, 3 (2010), 1519–1530.
- [27] Haiwang Zhong, Le Xie, and Qing Xia. 2013. Coupon incentive-based demand response: Theory and case study. *IEEE Transactions on Power Systems* 28, 2 (2013), 1266–1276.