Energy Coupon: A Mean Field Game Perspective on Demand Response in Smart Grids

Jian Li[‡], Bainan Xia[‡], Xinbo Geng[*], Hao Ming[‡],
Srinivas Shakkottai[‡], Vijay Subramanian[†], and Le Xie[*]
[‡*]Dept. of ECE, Texas A&M University, [†]Dept. of EECS, University of Michigan
Email: [‡]{lj0818, xiabainan, mingming511, sshakkot}@tamu.edu,
[*]{gengxbtamu, lxie}@ece.tamu.edu, [†]vgsubram@umich.edu

1. INTRODUCTION

There has recently been much interest in understanding *societal networks,* consisting of interconnected communication, transportation, energy and other networks that are important to the functioning of society. Research into these networks often takes the form of behavioral studies on decision making by the participants, and whether it is possible to provide incentives to modify their resource usage pattern in such a way that the society as a whole benefits [5].

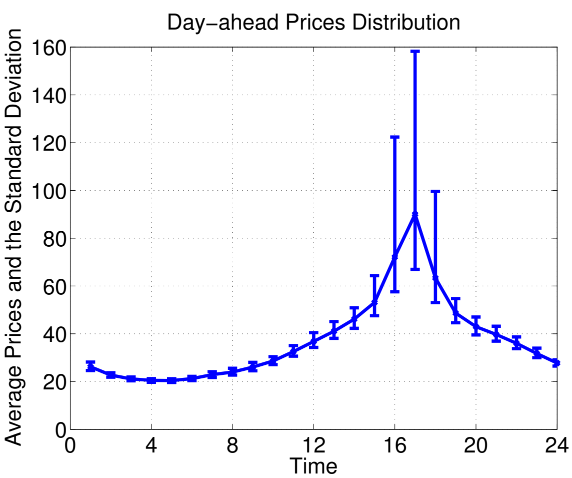

Figure 1: Day-ahead electricity market prices in dollars per MWh on an hourly basis between 12 AM to 12 PM, measured between June–August, 2013 in Austin, TX. Standard deviations above and below the mean are indicated separately.

Our candidate application in this paper is that of a Load Aggregator (LA) (*e.g.,* a utility company) trying to reduce its exposure to daily electricity market volatility by incentivizing demand response in a Smart Grid setting. For instance, consider Figure 1, which shows the (wholesale) price of electricity at different hours of day during the summer months in Texas. The data was obtained from the Electric Reliability Council of Texas (ERCOT) [1], an organization that manages the deregulated wholesale energy market in the state. The price shows considerable variation during the day, and peaks at about 5 PM, which is the time at which maximum demand occurs. A major source of this demand in Texas is air conditioning, which in each home is of the order of 30 kWh per day [2]. Incentivizing customers to move a few kWh of peak-time usage to the sides of the peak each day could lead to much reduced risks of peak price borne by the LA. Such demand shaping could also have a positive effect on environmental impact of power plant emissions.

Changing the energy usage pattern could potentially cause a small increase in the mean and deviation of the internal home temperature, which is a discomfort cost borne by the customer. In our system, the LA awards a number of "Energy Coupons" to the customer in proportion to his usage at the non-peak times, and these coupons are used as tickets at a lottery conducted by the LA. A higher number of coupons would be obtained by choosing an option that potentially entails more discomfort, and would also imply a higher probability of winning at the lottery.

2. ANALYTICAL MODEL

In our analytical model, each agent has a set of actions that it can take in each play of a repeated game, with each action having a corresponding cost. Higher cost actions yield a higher number of coupons. At the end of each play, the agents participate in a lottery in which they are randomly permuted into groups, and one or more prizes are given in each group. The state of each agent is measured using his surplus, which captures the history of plays experienced by the agent, and is a proxy to capture his interest in participating in the incentive system. Each win at the lottery increases the surplus, and each loss decreases it. Furthermore, we assume that the agent is risk neutral and the surplus is translated to value/utility using a concave and increasing function. This captures the idea that each successive win yields a reduced marginal happiness to the agent. Any agent could depart from the system with a fixed probability, and a departing agent is replaced by a new entrant with a randomly drawn surplus. In this setting, we would like to understand how agents would decide on their actions.

Our model is well suited to large scale systems in which any given subset of agents interact only rarely. This is the setting of a Mean Field Game (MFG) [3], which we will use as a framework to study equilibria in societal networks.

SIGMETRICS'15, June 15-19, 2015, Portland, OR, USA.
ACM 978-1-4503-3486-0/15/06.
http://dx.doi.org/10.1145/2745844.2745890 .

Here, the system is viewed from the perspective of a single agent, who assumes that each opponent's action would be drawn independently from an assumed distribution, and plays a best response action. We say that the system is at a Mean Field Equilibrium (MFE) if this best response action turns out to be a sample drawn from the assumed distribution.

Our first result is to characterize the best response policy of the mean field agent, using a dynamic programming formulation. We find that under our assumptions the value function of certain state of surplus is increasing, continuous and submodular. Further, we show using this result that given our ordering in which higher cost actions result in a higher probability of winning the lottery (due to more coupons being given), the choice of one action versus another depends on threshold values the surplus.

The probability of winning the lottery defines the transition kernel (along with the regeneration distribution) of the Markov process of the surplus, and hence maps an assumed distribution across competitors states to a resultant stationary distribution. We show the existence of a fixed point of this kernel, which is the MFE. Since we have a discrete action and state space, showing a fixed point in the space of stationary distributions is quite intricate, which follows Nash's original argument for the existence of mixed equilibria. The proofs of main results are available in [4].

3. NUMERICAL MODEL

We conduct a data-based simulation in the context of electricity usage for home air conditioning to illustrate the performance our system.

Home model: We model the usage of electricity for air conditioning by an average home in Texas over the course of a day. Given the ambient temperature, from [2] we calculate the ON-OFF pattern of our typical air conditioner.

Actions: We pick a reasonable discrete subset of the action space for our study. We assume that the actions available to the customer involve transfering energy from maximum price period 3 ($5-6$ PM) to the cheaper periods indexed by $1, 2, 4$, and 5. Transferring energy is equivalent to increasing/decreasing the ON time of the AC. We choose a transfer unit of 5 minutes, and candidate actions are shown in Table 2.

Costs: We measure the state of the home under action a by the tuple consisting of the mean tempurature and the standard deviation, denoted by $[\bar{\tau}_a, \sigma_a]$. The baseline state is under action 0, denoted by $[\bar{\tau}_0, \sigma_0]$. The cost of taking action a is $\theta_a = |\bar{\tau}_0 - \bar{\tau}_a| + \lambda|\sigma_0 - \sigma_a|$, where we choose $\lambda = 10$ to make the numerical values of the mean and standard deviation comparable to each other.

Hazard: We measure the hazard experienced by the LA per unit load as the weighted sum of the mean and standard deviation of the day-ahead price of electricity. Denoting the mean price at time period j by π_j and the standard deviation by ϕ_j, we define the *hazard* of peroid j measured in dollars/MWh as $h_j(\pi_j, \phi_j) = \pi_j + \alpha\phi_j$, where α models the risk that the LA perceives. The hazard values for each period are shown in Table 1.

Coupons: Coupons are assigned heuristically to promote consumption in off-peak time. We identified 6 actions that appeared to have the most promise of being used, and these are shown in Table 2 with their attendant costs and number of coupons received.

Table 1: Hazards

Index	Period	Hazard/MWh
1	$3-4$ PM	\$61.2
2	$4-5$ PM	\$120.2
3	$5-6$ PM	\$154.2
4	$6-7$ PM	\$101.3
5	$7-8$ PM	\$54.05

The LA conducts an auction each week across clusters of M homes (we take $M = 50$ in our study) in each auction. For each cluster, there are K prizes (we take $K = 1$ here) for winning the lottery with a value of \$40 (we will see below that this choice is viable). We assume that customers are likely to remain in the system for an average of 50 lotteries.

Table 2: Actions, Costs and Energy Coupons

Index	Action Vector	Cost	Coupons
0	$(0, 0, 6, 0, 0)$	0	13.92
1	$(1, 0, 0, 0, 5)$	2.495	243.92
2	$(2, 0, 0, 0, 4)$	2.328	233.92
3	$(3, 0, 0, 1, 2)$	1.950	187.92
4	$(4, 0, 0, 1, 1)$	1.829	177.92
5	$(3, 0, 0, 2, 1)$	1.747	151.92

MFE: Given the above parameters, we numerically determine the properties of the MFE generated by our system. The mean field distribution of surplus indicates that customers win at a lottery between 1 and 2 times over an average lifetime of 50 time intervals, as is to be expected with a cluster size of 50 customers at each lottery. From the mean field action distribution, we find that the best action from the LA's perspective is action 1, which is chosen with probability 0.6. We use the mean field action distribution to find that the net reduction in hazard over 50 homes is \$82 per week. Thus, incentivizing customers by offering a prize of \$40 each week is certainly feasible. The MFE illustrates that even as small a shift as 30 minutes of AC usage each day over several homes can yield significant benefits.

4. ACKNOWLEDGEMENTS

This work was funded in part by NSF grants CNS-1149458, EFRI-1440969, AST-1443891, III-1219071, AST-1443972, AST-1343381, ECCS-1150944 and CCF-1331863.

5. REFERENCES

[1] Electric Reliability Council of Texas (ERCOT). Data set available at `http://www.ercot.com/`.

[2] Pecan Street. Data set available at `https://dataport.pecanstreet.org/`.

[3] J. M. Lasry and P. Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[4] J. Li, B. Xia, X. Geng, H. Ming, S. Shakkottai, V. Subramanian, and L. Xie. Energy coupon: A mean field game perspective on demand response in smart grids. *Arxiv preprint arXiv:1503.02951*, 2015.

[5] B. Prabhakar. Designing large-scale nudge engines. In *Proceedings of the ACM SIGMETRICS*, 2013.