Incentive Design and Utility Learning via Energy Disaggregation

Lillian J. Ratliff [*] **Roy Dong** [*] **Henrik Ohlsson** [*]
S. Shankar Sastry [*]

[*] *University of California, Berkeley, Berkeley, CA 94720 USA*
{ratliffl, roydong, ohlsson, sastry}@eecs.berkeley.edu

Abstract: The utility company has many motivations for modifying energy consumption patterns of consumers such as revenue decoupling and demand response programs. We model the utility company–consumer interaction as a reverse Stackelberg game and present an iterative algorithm to design incentives for consumers while estimating their utility functions. Incentives are designed using the aggregated as well as the disaggregated (device level) consumption data. We simulate the iterative control (incentive design) and estimation (utility learning and disaggregation) process for examples.

Keywords: Game Theory, Economic Design, Energy Management Systems

1. INTRODUCTION

Currently, most electricity distribution systems only provide aggregate power consumption feedback to consumers, in the form of a energy bill. Studies have shown that providing device-level feedback on power consumption patterns to energy users can modify behavior and improve energy efficiency (Gardner and Stern, 2008; Laitner et al., 2009).

The current infrastructure only has sensors to measure the aggregated power consumption signal for a household. Even advanced metering infrastructures currently being deployed have the same restriction, albeit at high resolution and frequency (Armel et al., 2013). Additionally, deploying plug-level sensors would require entering households to install these devices. Methods requiring plug-level sensors are often referred to as *intrusive load monitoring*, and the network infrastructure required to transmit high resolution, high frequency data for several devices per household would be very costly.

A low cost alternative to the deployment of a large number of sensors is *non–intrusive load monitoring*. We consider the problem of nonintrusive load monitoring, which, in the scope of this paper, refers to recovering the power consumption signals of individual devices from the aggregate power consumption signal available to our sensors. This is also sometimes referred to as *energy disaggregation*, and we will use the two terms interchangeably. This problem has been an active topic of research lately. Some works include

Dong et al. (2013a); Froehlich et al. (2011); Johnson and Willsky (2012).

We propose that the utility company should use incentives to motivate a change in the energy consumption of consumers. We assume the utility company cares about the satisfaction of its consumers as well as altering consumption patterns, but it may not be able to directly observe the consumption patterns of individual devices or a consumer's satisfaction function.

In brief, the problem of behavior modification in energy consumption can be understood as follows. The utility company provides incentives to myopic energy consumers, who seek to maximize their own utility by selecting energy consumption patterns. This can be thought of as a control problem for the utility company. Additionally, the utility company does not directly observe the energy consumption patterns of individual devices, and seeks to recover it from an aggregate signal using energy disaggregation. This can be thought of as an estimation problem. Further, the consumer does not report any measure of its satisfaction directly to the utility. Thus, it must be estimated as well.

There are many motivations for changing energy consumption patterns of users. Many regions are beginning to implement revenue decoupling policies, whereby utility companies are economically motivated to decrease energy consumption (Eom, 2008). Additionally, the cost of producing energy depends on many variables, and being able to control demand can help alleviate the costs of inaccurate load forecasting. Demand response programs achieve this by controlling a portion of the demand at both peak and off-peak hours (Mathieu et al., 2012). We propose a model for how utility companies would design incentives to induce the desired consumer behavior.

In this paper, we consider three cases of utility learning and incentive design. In the first, the utility company designs an incentive based entirely on the aggregate power consumption signal. We propose an algorithm to estimate

the satisfaction function of the consumer based on the consumer's aggregated power consumption signals in Section 3. Then, in Section 4.1, we consider the case where the utility company knows the power consumption signal of individual devices and an unknown satisfaction function. Finally, in Section 4.2, we consider the case when the utility company only has access to the aggregated power consumption signal, and uses an energy disaggregation algorithm to recover the power consumption of individual devices. This disaggregated signal is used to allocate incentives, but the results will depend on the accuracy of our estimator, the energy disaggregation algorithm. We conclude the paper by showing the results from simulations of two examples of designing incentives while estimating the consumer's satisfaction function in Section 5. Finally, in Section 6 we make concluding remarks and discuss future research directions.
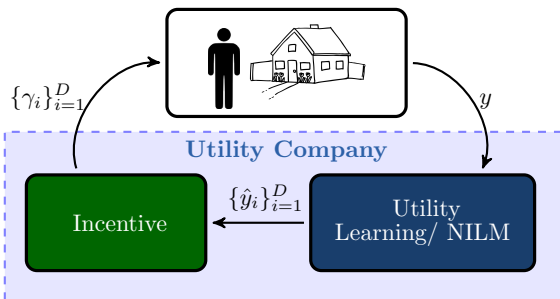

Fig. 1. Closing the Loop: Behavior modification via incentives γ_i is a *control* problem. The consumer decides when to use devices resulting in device level consumption y_i. Non–intrusive load monitoring (NILM) is used to *estimate* problem device level usage $\hat{y}_i$. Similarly, utility learning is an estimation problem

2. INCENTIVE DESIGN PRELIMINARIES

A reverse Stackelberg game is a hierarchical control problem in which sequential decision making occurs; in particular, there is a leader that announces a mapping of the follower's decision space into the leader's decision space, after which the follower determines his optimal decision variables (Groot et al., 2012).

Both the leader and the follower wish to maximize their pay–off determined by the functions $J_L(v, y)$ and $J_F(v, y)$ respectively. The leader's decision is denoted v; the follower's decision, y; and the incentive, $\gamma : y \mapsto v$. The basic approach to solving the reversed Stackelberg game is as follows. Let v and y take values in $V \subset \mathbb{R}$ and $Y \subset \mathbb{R}$, respectively; $J_L : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$; $J_F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. We define the desired choice for the leader as

$$(v^d, y^d) = \arg\max_{v,y} \left\{ J_L(v, y) | \ v \in V, y \in Y \right\}. \quad (1)$$

The incentive problem can be stated as follows:
Problem 1. Find $\gamma : Y \to V$, $\gamma \in \Gamma$ such that

$$\arg\max_{y \in Y} J_F(\gamma(y), y) = y^d \quad (2)$$

$$\gamma(y^d) = v^d \quad (3)$$

where Γ is the set of admissible incentive mechanisms.

3. INCENTIVE DESIGN USING AGGREGATE POWER SIGNAL

We cast the utility–consumer interaction in a reversed Stackelberg game framework in which the utility company is the leader and the consumer is the follower (see Figure 1). The leader's true utility is assumed to be given by

$$J_L(v, y) = g(y) - v + \beta f(y) \quad (4)$$

where $g(\cdot)$ is a concave function of the consumer's energy usage y over a billing period, v is the value of the incentive paid to the consumer, $f : Y \to \mathbb{R}$ is the consumer's satisfaction function for energy consumption which we assume is concave and β is a multiplying factor capturing the degree of benevolence of the utility company.

We assume that $v \in V = [0, \bar{v}]$ since the utility company should not take additional money away from the consumer on top of the cost of their usage and v should be less than some maximal amount the leader is willing to pay to the consumer $\bar{v}$. Similarly, let $y \in Y = [0, \bar{y}]$ where $\bar{y}$ is the upper bound on the allowed energy usage and let $\mathring{Y} = (0, \bar{y})$.

In a regulated market with revenue decoupling in place, a simplified model may consider

$$g(y) = -y \quad (5)$$

representing the fact that the utility wants the consumer to use less energy. Similarly, if the utility company has aspirations to institute a demand response program, a simplified model may consider

$$g(y) = - \left(y - y^{\text{ref}} \right)^2 \quad (6)$$

where y^{ref} is the reference signal prescribed by the demand response program.

The consumer's true utility is assumed to be

$$J_F(\gamma(y), y) = -py + \gamma(y) + f(y) \quad (7)$$

where p is the price of energy set and known to all and $\gamma : Y \to \mathbb{R}$ is the incentive mechanism. Thus, the consumer solves the optimization problem

$$\max_y \{ J_F(\gamma(y), y) | \ y \in Y \}. \quad (8)$$

We assume that the consumer is a *household* who is not strategic in the sense that they take the incentive γ and the price p and optimize their utility function without strategically choosing y. In particular, we assume that the consumer is myopic in that he does not consider past or future incentives in his optimization problem.

Incentives are designed by solving Problem 1 where we assume Γ to be the set of quadratic polynomials from Y to $\mathbb{R}$.

The leader does not know the follower's satisfaction function $f(\cdot)$, and hence, must estimate it as he solves the incentive design problem. We will use the notation $\hat{f}$ for the estimate of the satisfaction function and $\hat{J}_L$ and $\hat{J}_F$ for the player's cost functions using the estimate of f.

We propose an algorithm for iteratively estimating the agent's satisfaction function and choosing the incentive $\gamma(\cdot)$. We do so by using a polynomial estimate of the agent's satisfaction function at each iteration and applying

first-order optimality conditions. The use of more general sets of basis functions is left for future research.

Suppose that $\gamma^{(0)}$ and $\gamma^{(1)}$ are given *a priori*. At each iteration the leader issues an incentive and observes the follower's reaction. The leader then uses the observations up to the current time along with his knowledge of the incentives he issued to estimate the follower's utility function.

Formally, at the k-th iterate the leader will observe the follower's reaction $y^{(k)}$ to a delivered incentive $\gamma^{(k)}$ where we suppress the dependence of the incentive on y. The follower's reaction $y^{(k)}$ is optimal with respect to $J_F(\gamma^{(k)}(y), y)$ subject to $y^{(k)} \in Y$.

We use the observations $y^{(0)}, \ldots, y^{(k)}$ to estimate the parameters in the follower's satisfaction function given by

$$\hat{f}^{(k)}(y) = \sum_{i=0}^{j} \alpha_i y^{i+1} \tag{9}$$

where j is the order of the polynomial estimate to be determined in the algorithm and we restrict $\alpha = (\alpha_1, \ldots, \alpha_j) \in \mathcal{A}$ a convex set, e.g. $\mathcal{A} = \mathbb{R}_+^{j+1}$.

As in Keshavarz et al. (2011), we assume that an appropriate constraint qualification holds and we use Kharush-Khun-Tucker (KKT) conditions to define a notion of *approximate optimality*. Thus, we can allow for some error in the estimation problem either from measurement noise or suboptimal consumer choice. In particular, for each $i = 0, \ldots, k$ let $\gamma_{y^{(i)}}^{(i)} = \gamma^{(i)}(y^{(i)})$ and define

$$r_{\text{ineq}}^{(i)} = (g_\ell(\gamma_{y^{(i)}}^{(i)}, y^{(i)}))_+, \ \ell = 1, 2 \tag{10}$$

$$r_{\text{stat}}^{(i)}(\alpha, \lambda^{(i)}) = \nabla J_F(\gamma_{y^{(i)}}^{(i)}, y^{(i)}) + \sum_{\ell=1}^{2} \lambda_\ell^{(i)} \nabla g_i(\gamma_{y^{(i)}}^{(i)}, y^{(i)}) \tag{11}$$

$$r_{\text{comp}}^{(i)}(\lambda^{(i)}) = \lambda_\ell^{(i)} g_\ell(\gamma_{y^{(i)}}^{(i)}, y^{(i)}), \ \ell = 1, 2 \tag{12}$$

where $(\cdot)_+ = \max\{\cdot, 0\}$,

$$g_1(\gamma, y) = -y \le 0 \ \text{ and } \ g_2(\gamma, y) = y - \bar{y} \le 0 \tag{13}$$

with Lagrange multipliers $\lambda^{(i)} = (\lambda_1^{(i)}, \lambda_2^{(i)})$.

Then, for $\{(\gamma^{(i)}, y^{(i)})\}_{i=0}^{k}$, we can solve

$$\min_{\alpha, \lambda} \left\{ \sum_{i=0}^{k} \phi(r_{\text{stat}}^{(i)}, r_{\text{comp}}^{(i)}) | \ \alpha \in \mathcal{A}, \lambda^{(i)} \ge 0, \ i = 0, \ldots, k \right\} \tag{P-2}$$

where the inequality for $\lambda^{(i)}$ is element-wise and $\phi : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}_+$ is a nonnegative convex penalty function (e.g. any norm on $\mathbb{R} \times \mathbb{R}^2$) satisfying

$$\phi(x_1, x_2) = 0 \iff \{x_1 = 0, \ x_2 = 0\}. \tag{14}$$

The optimization problem (P-2) is convex since $r_{\text{stat}}^{(i)}$ and $r_{\text{comp}}^{(i)}$ are linear in α and $\lambda^{(i)}$ and the constraints are convex. If we solve (P-2) and ϕ is zero at the optimal solution, $r_{\text{stat}}^{(i)}$ and $r_{\text{comp}}^{(i)}$ are zero at the optimal solution for each i. If, in addition, $r_{\text{ineq}}^{(i)}$ is zero at the optimal solution for each i, then the estimate $\hat{f}^{(k)}$ at iteration k is exactly consistent with the data.

If $y^{(i)} \in \mathring{Y}$, the problem simplifies to a checking a linear algebra condition. Indeed, consider

$$\hat{J}_F(\gamma(y), y) = -py + \gamma(y) + \hat{f}(y). \tag{15}$$

In the case that $\hat{f}$ is concave and under our assumption that the follower is rational and hence plays optimally, the observation $y^{(i)}$ is a global optimum at iteration i. Otherwise, the observation $y^{(i)}$ is a local optimum; the follower plays myopically. In both cases, we use the necessary condition (Bertsekas, 1999)

$$\nabla \hat{J}_F^{(i)}(\gamma(y^{(i)}), y^{(i)}) = 0 \tag{16}$$

for each of the past iterates $i \in \{0, \ldots, k\}$ to determine estimates of the coefficients in $\hat{f}^{(k)}$.

At the k-th iteration, we have data $\{(\gamma^{(i)}, y^{(i)})\}_{i=0}^{k}$. Since we require $\gamma^{(i)} \in \Gamma$, we can express each $\gamma^{(i)}$ as

$$\gamma^{(i)}(y) = \xi_1^{(i)} y + \xi_2^{(i)} y^2. \tag{17}$$

Then, using Equation (16), we define

$$b^{(i)} = p - (\xi_1^{(i)} + 2\xi_2^{(i)} y^{(i),d}) \tag{18}$$

and

$$\tilde{y}_j^{(i)} = \begin{bmatrix} 1 & 2y^{(i)} & \cdots & (j+1)(y^{(i)})^j \end{bmatrix} \tag{19}$$

for $i \in \{0, \ldots, k\}$.

We want to find the lowest order polynomial estimate of f given the data. We do so by checking if $b^{(k)} \in \text{range}(Y^{(k)})$ where

$$Y^{(k)} = \begin{bmatrix} - \ \tilde{y}_j^{(0)} \ - \\ \vdots \\ - \ \tilde{y}_j^{(k)} \ - \end{bmatrix}, \ b^{(k)} = \begin{bmatrix} b^{(0)} \\ \vdots \\ b^{(k)} \end{bmatrix} \tag{20}$$

starting with $j = 2$ and increasing it until (20) is satisfied or we reach $j = k$. Suppose that it is satisfied at $j = N$, $2 \le N \le k$. Then, we estimate $\hat{f}^{(k)}$ to be an $(N+1)$-th order polynomial. We determine α_i for $i \in \{0, \ldots, N\}$ by solving

$$b^{(k)} - Y^{(k)}\alpha = 0, \ \text{ where } \ \alpha = [\alpha_0 \ldots \alpha_N]^T. \tag{21}$$

If $b^{(k)} \notin \text{range}(Y^{(k)})$ for any $j \in [2, k]$, we terminate.

Our algorithm prescribes that the leader check if $y^{(i)} \in \mathring{Y}$ for each i. If this is the case, then he shall find the minimum order polynomial given the data as described above. On the other hand, he shall solve the convex problem (P-2).

Using $\{\alpha_i\}_{i=0}^{j}$, $\hat{J}_L^{(k)}$, and $\hat{J}_F^{(k)}$, the leader solves the incentive design problem. That is, the leader first solves

$$(v^{(k+1),d}, y^{(k+1),d}) = \arg \min_{v \in V, y \in Y} \hat{J}_L^{(k)}(v, y) \tag{22}$$

$$= \arg \min_{v \in V, y \in Y} \left\{ g(y) - v + \beta \hat{f}^{(k)}(y) \right\} \tag{23}$$

Then, the leader finds $\gamma^{(k+1)} \in \Gamma$ such that

$$\arg \max_{y \in Y} \hat{J}_F^{(k)}(\gamma^{(k+1)}(y), y) = y^{(k+1),d} \tag{24}$$

$$\gamma^{(k+1)}(y^{(k+1),d}) = v^{(k+1),d} \tag{25}$$

If $y^{(k+1),d} \in \mathring{Y}$, then since we restrict $\gamma^{(k+1)}$ to be of the form (17) the above problem reduces to solving $A^{(k+1)}\xi^{(k+1)} = \tilde{b}^{(k+1)}$ where

$$A^{(k+1)} = \begin{bmatrix} 1 & 2y^{(k+1),d} \\ y^{(k+1),d} & (y^{(k+1),d})^2 \end{bmatrix}, \xi^{(k+1)} = \begin{bmatrix} \xi_1^{(k+1)} \\ \xi_2^{(k+1)} \end{bmatrix}, \quad (26)$$

and

$$\tilde{\mathrm{b}}^{(k+1)} = \begin{bmatrix} p - \alpha_0 - 2\alpha_1 y^{(k+1),d} \\ v^{(k+1),d} \end{bmatrix}. \quad (27)$$

If $\tilde{\mathrm{b}}^{(k+1)} \in \mathrm{range}(A^{(k+1)})$ then a solution $\xi^{(k+1)}$ exists and if $A^{(k+1)}$ is full rank, the solution is unique. Otherwise, if $y^{(k+1),d} \notin \mathring{Y}$, we terminate the algorithm.

Remark 1. The algorithm is motivated by the case when the consumer's satisfaction function is a polynomial of order k and the utility company does not know k, by following the algorithm past even $k+1$ iterations, the utility company will be playing optimally. Alternatively, if incentives $\gamma^{(i)}$ were chosen randomly, the utility company would not know when to stop choosing random $\gamma^{(i)}$'s; thus, after $k+1$ iterations would begin playing suboptimally.

Proposition 1. Let f be polynomial of order $k+1$, $y^{(0)} \in \mathring{Y}$ and $\gamma^{(0)}, \gamma^{(1)}$ be given *a priori*. Suppose that at each iteration of the algorithm $\mathrm{b}^{(\ell)} \in \mathrm{range}(Y^{(\ell)})$, $\mathrm{rank}(Y^{(\ell)}) = \ell + 1$, $y^{(i)} \in \mathring{Y}$ and $\tilde{\mathrm{b}}^{(\ell+1)} \in \mathrm{range}(A^{(\ell+1)})$. Then, after k iterations the satisfaction function is known exactly and the incentive $\gamma^{(k+1)}$ induces the consumer to use the desired control.

Proposition 2. Suppose that f is polynomial up to order $k+1$ and that the leader has $k+1$ historical measurements

$$\gamma^{(-k)}, \dots, \gamma^{(1)}, y^{(-k)}, \dots, y^{(1)} \quad (28)$$

such that $Y^{(k)}$ is full rank where $y^{(i)} \in \mathring{Y}$ for $i = 0, \dots, k$, then the leader can estimate the follower's satisfaction function exactly and if there exists an incentive $\gamma^{(k+1)}$, then it induces the desired equilibrium.

We conclude this section by providing an example of the iterative process when f is a concave function.

Example 1. First, we suppose that $\gamma^{(0)}, \gamma^{(1)} \in \Gamma$ are chosen *a priori* and are parameterized as follows:

$$\gamma^{(i)}(y) = \xi_1^{(i)}y + \xi_2^{(i)}y^2 \quad (29)$$

for $i \in \{0, 1\}$. Then, the procedure goes as follows. The leader issues $\gamma^{(0)}$ and observes $y^{(0)}$. Subsequently, he issues $\gamma^{(1)}$ and observes $y^{(1)}$. Suppose $y^{(0)}, y^{(1)} \in \mathring{Y}$. The leader determines α_1, α_0 in the estimation of $\hat{f}(y) = \alpha_1 y^2 + \alpha_0 y$ by computing the derivative of $\hat{J}_F^{(0)}(\gamma^{(0)}(y), y)$ and $\hat{J}_F^{(1)}(\gamma^{(1)}(y), y)$ with respect to y, evaluating at $y^{(0)}$ and $y^{(1)}$ and equating to zero, i.e. he solves

$$-p - 2y^{(0)} + 2(\alpha_1 + \xi_2^{(0)})y^{(0)} + \alpha_0 + \xi_1^{(0)} = 0 \quad (30)$$

$$-p - 2y^{(1)} + 2(\alpha_1 + \xi_2^{(1)})y^{(1)} + \alpha_0 + \xi_1^{(1)} = 0 \quad (31)$$

for α_0 and α_1. If either $y^{(0)}$ or $y^{(1)}$ are on the boundary of Y, then the leader solves (P-2) for $\alpha = (\alpha_1, \alpha_0)$.

Using α_0, α_1, the leader solves the following incentive design problem for $\gamma^{(2)}$. First, find $(v^{(2),d}, y^{(2),d}) \in V \times Y$ such that

$$\hat{J}_L^{(2)}(v, y) = -y - v + \alpha_1 y^2 + \alpha_0 y \quad (32)$$

is maximized. Since we restrict to quadratic incentives, we parameterize $\gamma^{(2)}$ as in Equation (29) with $i = 2$. Now, given the utility $\hat{J}_F^{(2)}(\gamma^{(2)}(y), y)$, we find $\xi_1^{(2)}, \xi_2^{(2)}$ such that

$$\arg\max_{y \in Y} \hat{J}_F^{(2)}(y; \xi_1^{(2)}, \xi_2^{(2)}) = y^{(2),d} \quad (33)$$

$$\xi_1^{(2)} y^{(2),d} + \xi_2^{(2)} (y^{(2),d})^2 = v^{(2),d} \quad (34)$$

Assuming that $y^{(2),d} \in \mathring{Y}$, it will be an induced local maxima under the incentive $\gamma^{(2)}$. Hence, Equation (33) can be reformulated using the necessary condition

$$\nabla_y \hat{J}_F^{(2)}(y^{(2),d}; \xi_1^{(2)}, \xi_2^{(2)}) = 0. \quad (35)$$

Now, Equations (34) and (35) give us two equations in the two unknowns $\xi_1^{(2)}, \xi_2^{(2)}$ that can be solved; indeed,

$$-p + \xi_1^{(2)} + \alpha_0 + 2(\xi_2^{(2)} + \alpha_1)y^{(2),d} = 0 \quad (36)$$

$$\xi_1^{(2)} y^{(2),d} + \xi_2^{(2)} (y^{(2),d})^2 = v^{(2),d} \quad (37)$$

Solving these equations gives us the parameters for $\gamma^{(2)}$. Now, the leader can issue $\gamma^{(2)}$ to the follower and observe his reaction $y^{(2)}$. The leader can then continue in the iterative process as described above.

4. DEVICE LEVEL INCENTIVE DESIGN USING DISAGGREGATION ALGORITHM

In a manner similar to the previous section, we consider that the consumer's satisfaction function is unknown. However, we now consider that the utility company desires to design device level incentives. We remark that the utility company may not want to incentivize every device; the process we present can be used to target devices with the highest consumption or potential to offset inaccuracies in load forecasting.

4.1 Exact Disaggregation Algorithm

We first describe the process of designing device level incentives assuming the utility company has a disaggregation algorithm in place which produces no error. That is, they observe the aggregate signal and then applies their disaggregation algorithm to get exact estimates of the device level usage y_ℓ for $\ell \in \{1, \dots, D\}$ where D is the number of devices.

The utility company has the true utility function

$$J_L(v, y) = \sum_{\ell=1}^{D} g_\ell(y_\ell) - v_\ell + \beta_\ell f_\ell(y_\ell) \quad (38)$$

and the consumer has the true utility function

$$J_F(\gamma(y), y) = \sum_{\ell=1}^{D} -py_\ell + \gamma_\ell(y_\ell) + f_\ell(y_\ell). \quad (39)$$

The utility company could choose only to incentivize specific devices such as high consumption devices. This fits easily into our framework; however, for simplicity we just present the model in which incentives are designed for each device.

The implicit assumption that the player utilities are separable in the devices allows us to generalize the algorithm presented in the previous section. Let us be more precise. We again assume that $\gamma_\ell^{(0)}, \gamma_\ell^{(1)}$ for $\ell \in \{1, \dots, D\}$ are given *a priori*.

At the k-th iteration the utility company issues an incentive $\gamma_\ell^{(k)}$ for each device $\ell \in \{1, \dots, D\}$ and observes the

aggregate signal $y^{(k)}$. Then they apply a disaggregation algorithm to determine the device level usage $y_\ell^{(k)}$ for $\ell \in \{1, \dots, D\}$.

The utility company forms an estimate of the consumer's device level satisfaction function

$$\hat{f}_\ell^{(k)}(y_\ell) = \sum_{i=0}^{j} \alpha_{i,\ell} y_\ell^{i+1} \qquad (40)$$

and then solves the problem of finding the $\alpha_{i,\ell}$'s by solving for $\alpha_\ell = (\alpha_{0,\ell}, \dots, \alpha_{j,\ell})$ as in the previous section for each device $\ell \in \{1, \dots, D\}$.

Proposition 3. For $\ell \in \{1, \dots, D\}$, let f_ℓ be polynomial up to order $k_\ell + 1$, $\gamma_\ell^{(0)}, \gamma_\ell^{(1)}$ be given *a priori*, and $y_\ell^{(0)}, y_\ell^{(1)} \in \mathring{Y}$. Suppose that at each iteration of the algorithm $\mathrm{b}_\ell^{(m)} \in \mathrm{range}(Y_\ell^{(m)})$, $\mathrm{rank}(Y_\ell^{(m)}) = m_\ell + 1$, $y_\ell^{(m)} \in \mathring{Y}$, and $\tilde{\mathrm{b}}_\ell^{(m+1)} \in A_\ell^{(m+1)}$ for each $\ell \in \{1, \dots, D\}$. Then, after

$$k^* = \max_{\ell \in \{1,\dots,D\}} k_\ell \qquad (41)$$

iterations, the satisfaction function is known exactly and the incentives $\gamma_\ell^{(k^*+1)}$ induce the desired equilibrium.

Note that the notation $(\cdot)_\ell$ indicates the object defined in Section 3 for the ℓ-th device.

4.2 Disaggregation Algorithm with Some Error

Now, we consider that the leader has some error in his estimate of the device level usage due to inaccuracies in the disaggregation algorithm, i.e. the leader determines $\hat{y}_\ell$ such that $\|y_\ell - \hat{y}_\ell\| \le \varepsilon$ where $\varepsilon > 0$ is the resulting error from the estimation in the disaggregation algorithm. Bounds on ε can be determined by examining the fundamental limits of non–intrusive load monitoring algorithms (Dong et al., 2013b).

We again assume that $\gamma_\ell^{(0)}, \gamma_\ell^{(1)}$ for $\ell \in \{1, \dots, D\}$ are given *a priori*. Following the same procedure as before, at the k-th iterate the leader will issue $\gamma_\ell^{(k)}$ for $\ell \in \{1, \dots, D\}$ and observe $y^{(k)}$. Then apply a disaggregation algorithm to determine $\hat{y}_\ell^{(0)}$ where

$$\|y_\ell - \hat{y}_\ell\| \le \varepsilon \qquad (42)$$

for $\ell \in \{1, \dots, D\}$. The incentive design problem follows the same steps as provided in the previous section with the exception that the y_ℓ's are replaced with the estimated $\hat{y}_\ell$'s and we tolerate an error in solving for the minimal polynomial estimate of f_ℓ.

5. NUMERICAL EXAMPLES

We simulate two examples of designing incentives while estimating the consumer's satisfaction function. In both examples we assume a unit price per unit of energy, i.e. $p = 1$.

5.1 Aggregate Signal and Log Satisfaction Function

We simulate a system in which the consumer has the true utility given by

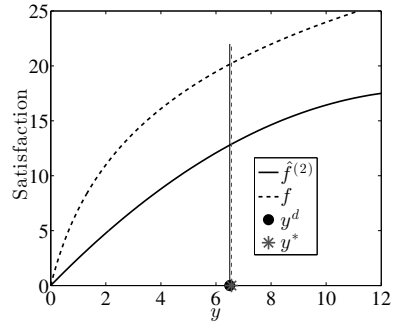$$J_F(\gamma(y), y) = -py + \gamma(y) + f(y) \qquad (43)$$

Fig. 2. Estimated satisfaction function $\hat{f}^{(2)}$ and true satisfaction function f. The true response $y^* = 6.56$ and the desired response $y^d = 6.5$. Notice that the slope of the estimated satisfaction function and the slope of the true satisfaction function are roughly equal at y^d and y^*.

where the satisfaction function is

$$f(y) = 10 \log(y + 1). \qquad (44)$$

We assume the utility company is in a regulated market; hence wants the consumer to consume less. Thus, the utility company has utility function

$$J_L(v, y) = -y - v + \beta f(y) \qquad (45)$$

where the benevolence factor is $\beta = 0.75$. We let $\bar{y} = \bar{v} = 100$. We choose two concave incentive function $\gamma^{(0)}(y)$ and $\gamma^{(1)}(y)$ defined as follows: $\gamma^{(0)}(y) = -y^2 + 10y$ and $\gamma^{(1)}(y) = -y^2 + 15y$.

We use the algorithm presented in Section 3 to design incentives while estimating α_0 and α_1. We simulate the utility company issuing $\gamma^{(0)}$ and then $\gamma^{(1)}$ where the consumer chooses his optimal response to each of the incentives. The responses are $y^{(0)} = 5.29$ and $y^{(1)} = 7.58$. After two iterations, we get a reasonable approximation of the true f and a quadratic incentive $\gamma^{(2)}$;

$$\hat{f}^{(2)}(y) = 2.57y - 0.093y^2, \ \gamma^{(2)}(y) = 0.33y - 0.05y^2. \ (46)$$

The optimal power usage under the incentive $\gamma^{(2)}$ is $y^* = 6.56$ and the desired power usage is $y^d = 6.5$. It is clear that the utility company could do better if he new the true satisfaction function. Figure 2 shows $\hat{f}^{(2)}(y)$ and $f(y)$. It is important to notice that y^* is nearly equal to y^d and at these two points the slope of $\hat{f}^{(2)}$ is approximately equal to that of the true f. This indicates that $\hat{f}^{(2)}$ is a good estimate of f.

5.2 Disaggregated Signal

We simulate a system in which the consumer's the true utility

$$J_F(\gamma(y), y) = \sum_{\ell=1}^{10} -py_\ell + \gamma_\ell(y_\ell) + f_\ell(y_\ell) \qquad (47)$$

where the satisfaction functions $f_\ell(y_\ell)$ are exactly quadratic for each device $\ell \in \{1, \dots, 10\}$;

$$f_\ell(y_\ell) = \alpha_{1,\ell} y_\ell^2 + \alpha_{0,\ell} y_\ell \qquad (48)$$

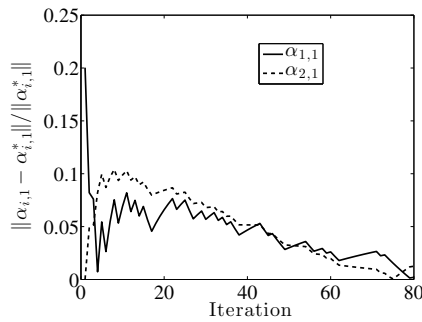The utility company's utility is given by

Fig. 3. Relative error in estimate of $\alpha_{i,1}$'s for device 1 with disaggregation error bound $\varepsilon = 0.15$. α_i^* is the true value. The relative error eventually decreases below the noise bound $\varepsilon = 0.15$.

$$J_L(v,y) = \sum_{\ell=1}^{10} -y_\ell - v_\ell + \beta_\ell f_\ell(y_\ell) \qquad (49)$$

where the benevolence factor (i.e. a representation of how much the utility company cares about the satisfaction of the consumer) is $\beta_\ell = 1$ for each ℓ. The utility company must disaggregate the aggregated energy signal y giving rise to estimates $\hat{y}_\ell$. If $\hat{y}_\ell = y_\ell$, i.e. there is no error in the disaggregation algorithm, then after two iterations the utility company would know the satisfaction function of each device exactly. Let's explore the case when the disaggregation algorithm has ε–error. In our examples we randomly generate noise within a given ε bound and add that to the true y_i's to simulate the error in the disaggregation step resulting from the disaggregation algorithm.

Figure 3 shows the relative error on the estimates of $\alpha_{i,1}$ for $i \in \{1,2\}$ as a function of the iteration. The relative error for other devices are similar. We used the error bound $\varepsilon = 0.15$ for the disaggregation error. The relative error decreases as the number of iterations increase.

As we iterate the noise introduced via disaggregation has minimal impact on the estimate of $\alpha_{i,\ell}$ for $i = \{1,2\}$ and $\ell \in \{1, \ldots, D\}$. We note that the designed incentive for this problem converges to zero as we increase the iterations and the impact of the noise is minimized. It becomes zero since the benevolence factor is $\beta_\ell = 1$ and the price $p = 1$; hence, the agent and the leader have the same utility functions after the leader learns the agent's satisfaction function. As we increase the noise threshold ε, the estimation of $\alpha_{i,\ell}$ degrades.

6. DISCUSSION AND FUTURE WORK

We modeled the utility company–consumer interaction as a reversed Stackelberg game. We defined a process by which the utility company can jointly estimate the consumer's utility function and design incentives for behavior modification. Whether the utility company is interested in inducing energy efficient behavior or creating an incentive compatible demand response program, the procedure we present applies. We are studying fundamental limits of *non–intrusive load monitoring* in order to determine precise bounds on the payoff to the utility company when

a disaggregation algorithm is in place and incentives are being designed. We seek to understand how these fundamental limits impact the quality of the incentive design problem as well as how they can be integrated into a stochastic contorl framework for incentive design when faced with non-strategic agents with unknown preferences.

The electrical grid is a social cyber-physical system (S-CPS) with human actors influencing the trajectory of the system. Inherent to the study of S-CPS's are privacy and security considerations. We remark that consumers may consider their satisfaction function to be private information. We are currently exploring the design of privacy–aware mechanisms for ε–incentive compatible problems (Nissim et al., 2012).

REFERENCES

Armel, K.C., Gupta, A., Shrimali, G., and Albert, A. (2013). Is disaggregation the holy grail of energy efficiency? The case of electricity. *Energy Policy*, 52, 213–234.

Bertsekas, D.P. (1999). *Nonlinear programming*. Athena Scientific.

Dong, R., Ratliff, L., Ohlsson, H., and Sastry, S.S. (2013a). Energy disaggregation via adaptive filter. In *Proceedings of the 50th Allerton Conference on Communication, Control, and Computing*.

Dong, R., Ratliff, L., Ohlsson, H., and Sastry, S.S. (2013b). Fundamental limits of non–intrusive load monitoring. *arXiv:1310.7850v1*.

Eom, J. (2008). Shareholder incentives for utility-delivered energy efficiency programs in California. In *Proceedings of the 28th USAEE/IAEE North American Conference*.

Froehlich, J., Larson, E., Gupta, S., Cohn, G., Reynolds, M., and Patel, S. (2011). Disaggregated end-use energy sensing for the smart grid. *IEEE Pervasive Computing*, 10(1), 28–39.

Gardner, G.T. and Stern, P.C. (2008). The short list: The most effective actions U.S. households can take to curb climate change. In *Environment: Science and Policy for Sustainable Development*.

Groot, N., De Schutter, B., and Hellendoorn, H. (2012). Reverse stackelberg games, part i: Basic framework. In *IEEE International Conference on Control Applications*, 421–426. doi:10.1109/CCA.2012.6402334.

Johnson, M.J. and Willsky, A.S. (2012). Bayesian nonparametric hidden semi-Markov models. *arXiv:1203.1365*.

Keshavarz, A., Wang, Y., and Boyd, S. (2011). Imputing a convex objective function. In *IEEE International Symposium on Intelligent Control*, 613–619. IEEE.

Laitner, J.A., Ehrhardt-Martinez, K., and McKinney, V. (2009). Examining the scale of the behaviour energy efficiency continuum. In *European Council for an Energy Efficient Economy*.

Mathieu, J.L., Haring, T., and Andersson, G. (2012). Harnessing residential loads for demand response: Engineering and economic considerations. In *Interdisciplinary Workshop on Smart Grid Design and Implementation*.

Nissim, K., Orlandi, C., and Smorodinsky, R. (2012). Privacy–aware mechanism design. *arXiv:1111.3350v2*.