Social Game for Building Energy Efficiency: Incentive Design

Lillian J. Ratliff, Ming Jin, Ioannis C. Konstantakopoulos, Costas Spanos, S. Shankar Sastry

***Abstract*— We present analysis and results of a social game encouraging energy efficient behavior in occupants by distributing points which determine the likelihood of winning in a lottery. We estimate occupants utilities and formulate the interaction between the building manager and the occupants as a reversed Stackelberg game in which there are multiple followers that play in a non-cooperative game. The estimated utilities are used for determining the occupant behavior in the non-cooperative game. Due to nonconvexities and complexity of the problem, in particular the size of the joint distribution across the states of the occupants, we solve the resulting the bi-level optimization problem using a particle swarm optimization method. Drawing from the distribution across player states, we compute the Nash equilibrium of the game using the resulting leader choice. We show that the behavior of the agents under the leader choice results in greater utility for the leader.**

I. Introduction

Energy consumption of buildings, both residential and commercial, accounts for approximately 40% of all energy usage in the U.S. [1]. Lighting is a major consumer of energy in commercial buildings; one-fifth of all energy consumed in buildings is due to lighting [2].

There have been many approaches to improve energy efficiency of buildings through control and automation as well as incentives and pricing. From the meter to the consumer, many control methods, such as model predictive control, have been proposed as a means to improve the efficiency of building operations (see, e.g., [3]–[8]). From the meter to the energy utility, many economic solutions have been proposed, such as dynamic pricing and mechanisms including incentives, rebates, and recommendations, to reduce consumption (see, e.g., [9], [10]).

Many of the past approaches to building energy management only focus on heating and cooling of the building. We are advocating that due to new technological advances in building automation, incentives can be designed around more than just heating, ventilation and air conditioning (HVAC) systems. In particular, our experimental set-up allows us to design incentives based on lighting and individual plug-load in addition to HVAC. In the set-up, the building manager interacts with occupants through a social game.

Social games have been used to alleviate congestion in transportation systems [11] as well as in the healthcare domain for understanding the tradeoff between privacy and desire to win by expending calories [12].

There are many ways in which a building manager can be motivated to encourage energy efficient behavior. The most obvious is that they pay the bill or, due to some operational excellence measure, are required to maintain an energy effcient building. Beyond these motivations, recently demand response programs are being implemented by utility companies with the goal of correcting for improper load forecasting (see, e.g., [13], [14], [15]). In such a program, consumers enter into a contract with the utiltiy company in which they agree to change their demand in accordance with some agreed upon schedule. In this scenario, the building manager may now be required to keep this schedule.

Our approach to efficient building energy management focuses on office buildings and utilizes new builiding automation products such as the Lutron lighting system[1]. We design a social game aimed at incentivizing occupants to modify their behavior so that the overall energy consumption in the building is reduced. The social game consists of occupants logging their vote for the lighting setting in the office. They win points based on how energy efficient their vote is compared to other occupants. After each vote is logged, the average of the votes is implemented in the office. The points are used to determine an occupant's likelihood of winning in a lottery.

We designed an online platform so that occupants can log in and vote, view their points, and observe all occupants consumption patterns and points. This platform also stores all the past data allowing us to use it for estimating occupant behavior.

In a recent paper, we described the experimental setup and formulated the follower game [16]. At the core of our approach is the fact that we modeled the occupants as non-cooperative agents who play Nash. Under this assumption, we were able to use necessary and sufficient first- and second-order conditions [17] to cast the utility estimation problem as a convex optimization problem in the parameters of the occupants' utility functions. We showed that estimating agent utility functions via this method results in

*This work is supported by NSF CPS:Large:ActionWebs award number 0931843, TRUST (Team for Research in Ubiquitous Secure Technology) which receives support from NSF (award number CCF-0424422), FORCES (Foundations Of Resilient CybEr-physical Systems) which receives support from NSF (award number CNS-1239166), the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

**The authors are with the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley, Berkeley CA 94720. {ratliffl, jinming, ioanniskon}@eecs.berkeley.edu {spanos, sastry}@eecs.berkeley.edu

[1] http://www.lutron.com/en-US/Pages/default.aspx

a predictive model that out performs several other standard techniques.

In this paper, we are able to leverage the fact that we modeled the occupants as utility maximers in a game-theoretic framework in the formulation of the building manager's problem as a reversed Stackelberg game. In particular, we formulate the building manager's optimization problem as a bi-level optimization problem in which the inner optimization problem is a non-cooperative game between the occupants and the outer optimization problem is the maximization of the building manager's utility over the total points and default lighting setting.

Given the data from our social game experiment, we estimate the occupants' utility functions. We determine a distribution for each occupant over the set of events which include the occupant states *present and active*, *present and remaining at the default*, and *absent*. We refer to these as the player states and shorten them to *active*, *default*, and *absent*. Due to the number of events in the joint distribution across possible occupant states, we employ a particle swarm optimization method for solving the building manager's bi-level optimization problem for the total points and default lighting setting. This results in a suboptimal solution; however, we show that the solution leads to a occupant behavior that results in a larger utility for the building manager as compared to previously implemented schemes.

The rest of the paper is organized as follows. We begin in Section II by describing the experimental setup for our social game test-bed. In Section III, we present the game formulation. There are games at two levels; the inner non-cooperative continuous game between the occupants and the outer reversed Stackelberg game between the building manager and the followers. We describe the utility estimation and incentive design (solution to the building manager's optimization problem) in Section IV. We conclude with some discussion and proposal for future work in Section V.

II. Experimental Setup

In this section we briefly describe the experimental setup. In [16], we provide a more detailed description of the setup.

The social game for energy savings that we have designed is such that occupants in an office builing vote according to their usage preferences of shared resources and are rewarded with points based on how *energy efficient* their strategy is in comparison with the other occupants. Having points increases the likelihood of the occupant winning in a lottery. The prizes in the lottery consist of three Amazon gift cards.

We have installed a Lutron system for the control of the lights in the office. This system allows us to precisely control the lighting level of each of the lights in the office. We use it to set the default lighting level as well as implement the average of the votes each time the occupants change their lighting preferences.

There are 22 occupants in the office which is divided into five lighting zones each with four occupants.

We have developed an online platform in which the occupants can login and participate in the game. In the platform the occupants can log their lighting setting votes, view point balances of all occupants, and observe all the behavior (voting) patterns of all occupants. Figure 1(a) shows a display of how an occupant can select their lighting preference and Figure 1(b) shows a sample of how occupants can see their point balance.

(a) (b)

Fig. 1. (a) Display of how occupants can log their lighting vote. (b) Display of an occupant's point balance.

An occupant's vote is for the lighting level in their zone as well as for neighboring zones. The lighting setting that is implemented is the average of all the votes.

There is a default lighting setting. An occupant can leave the lighting setting as the default after logging in or they can change it to some other value in the interval $[0, 100]$ depending on their preferences.

Each day when an occupant logs into the online platform the first time after they enter the office, they are considered present for the remainder of the day. If they actively change their vote from the default to some other value, then we consider them *active*. On the other hand, if they choose not to change their vote from the default setting, then they are considered *default* for the day. If they do not enter the office on a given day, then they are considered *absent*.

III. Game Formulation

We model the interaction between the building manager (leader) and the occupants (followers) as a leader-follower type game. We use the terms leader and building manager interchageably and, similarly, for follower and occupant.

In this model the followers are utility maximizers that play in a non-cooperative game for which we use the Nash equilibrium concept. The leader is also a utility maximizer with a utility that is dependent on the choices of the followers. The leader can influence the equilibrium of the game amongst the followers through the use of incentives which impact the utility and thereby the decisions of each follower.

The leader desires to reduce the energy consumption in the building as well as formulate a model of how the occupants make decisions about their energy usage. In order to achieve this goal, the leader implements a *social game* in which the followers are pitted against one another. The occupants win points based on their energy consumption choices. These points are then used to determine the individual follower's chance at winning in a lottery. In the particular social game we study in this paper, the occupants select a lighting setting to be implemented in the office.

A. Follower Game

We begin by describing the game-theoretic framework used for modeling the interaction between the occupants.

Let the number of occupants participating in the game be denoted by n. We model the occupants as utility maximizers having utility functions composed of two terms that capture the tradeoff between comfort and desire to win. We model their comfort level using a Taguchi loss function which is interpreted as modeling occupant dissatisfaction as increasing as variation increases from their desired lighting setting [18]. In particular, each occupant has the following Taguchi loss function as one component of their utility function:

$$\psi_i(x_i, x_{-i}) = -\left(\bar{x} - x_i\right)^2 \tag{1}$$

where $x_i \in \mathbb{R}$ is occupant i's lighting vote, $x_{-i} = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$, and

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

is the average of all the occupant votes and is the lighting setting which is implemented.

Each occupant's desire to win is modeled using the following function

$$\phi_i(x_i, x_{-i}) = -\rho\left(\frac{x_i}{100}\right)^2 \tag{3}$$

where ρ is the total number of points distributed by the building manager. The points are distributed by the leader using the relationship

$$\rho\frac{x_b - x_i}{nx_b - \sum_{j=1}^{n} x_j} \tag{4}$$

where $x_b = 90$ is the baseline setting for the lights, i.e. the lighting setting that occurred before the implementation of the social game in the office. In our previous work [16] we modeled the function ϕ_i, i.e. the desire to win, using the natural log of (4). We found that the form of ϕ_i as defined in (3) provides a better estimation and prediction of all the occupant's behavior. It appears that it captures the occupants' perceptions about how the points are distributed and the value of the points as determined by each of the occupants more accurately. We are currently exploring more general non-parametric and data-driven methods for estimating the occupants' utility functions.

Each occupant's utility function is then given by

$$f_i(x_i, x_{-i}) = \psi_i(x_i, x_{-i}) + \theta_i\phi_i(x_i, x_{-i}) \tag{5}$$

where θ_i is parameter unknown to the leader.

The i-th occupant faces the following optimization problem:

$$\max_{x_i \in S_i} f_i(x_i, x_{-i}) \tag{6}$$

where $S_i = [0, 100] \subset \mathbb{R}$ is the constraint set for x_i.

Note that each occupant's optimization problem is dependent on the other occupants' choice variables.

We can explicitly write out the constraint set as follows. Let $h_{i,j}(x_i, x_{-i})$ for $j \in \{1, 2\}$ denote the constraints on occupant i's optimization problem. In particular, following Rosen [19], for occupant i, the constraints are

$$h_{i,1}(x_i) = 100 - x_i \tag{7}$$

$$h_{i,2}(x_i) = x_i \tag{8}$$

so that we can define $\mathcal{C}_i = \{x_i \in \mathbb{R} \mid h_{i,j}(x_i) \geq 0,\ j \in \{1, 2\}\}$ and $\mathcal{C} = \mathcal{C}_1 \times \cdots \times \mathcal{C}_n$. Thus, the occupants are non-cooperative agents in a continuous game with convex constraints. We model their interaction using the Nash equilibrium concept.

Definition 1: A point $x \in \mathcal{C}$ is a **Nash equilibrium** for the game $(f_1, \dots, f_n)$ on $\mathcal{C}$ if

$$f_i(x_i, x_{-i}) \geq f_i(x_i', x_{-i}) \quad \forall\ x_i' \in S_i \tag{9}$$

for each $i \in \{1, \dots, n\}$.
The interpretation of the definition of Nash is as follows: no player can unilaterally deviate and increase their utility.

If the parameters $\theta_i \geq 0$, then the game is a concave n-person game on a convex set.

Theorem 1 ([19]): A Nash equilibrium exists for every concave n-person game.
Define the Lagrangian of each player's optimization problem as follows:

$$L_i(x_i, x_{-i}, \mu_i) = f_i(x_i, x_{-i}) + \sum_{j \in A_i(x_i)} \mu_{i,j} h_{i,j}(x_i) \tag{10}$$

where $A_i(x_i)$ is the active constraint set at x_i. We can define

$$\omega(x) = \begin{bmatrix} D_1 L_1(x, \mu_1) \\ \vdots \\ D_n L_n(x, \mu_n) \end{bmatrix} \tag{11}$$

where $D_i L_i$ denoets the derivative of L_i with respect to x_i.

It is the local representation of the differential game form [17] corresponding to the game between the occupants.

Definition 2 ([17]): A point $x^* \in \mathcal{C}$ is a **differential Nash equilibrium** for the game $(f_1, \dots, f_n)$ on $\mathcal{C}$ if $\omega(x^*, \mu^*) = 0$, $z^T D_{ii} L_i(x^*, \mu_i^*) z < 0$ for all $z \neq 0$ such that $D_i h_{i,j}(x_i^*)^T z = 0$, and $\mu_{i,j} > 0$ for $j \in A_i(x_i^*)$.

Proposition 1: A differential Nash equilibrium of the n-person concave game $(f_1, \dots, f_n)$ on $\mathcal{C}$ is a Nash equilibrium.

Proof: The proof is straightforward. Indeed, suppose the assumptions hold. The constraints for each player do not depend on other players' choice variables. We can hold x_{-i}^* fixed and apply Proposition 3.3.2 [20] to the i-th player's optimization problem

$$\max_{x_i \in \mathcal{C}_i} f_i(x_i, x_{-i}^*) \tag{12}$$

Since each f_i is concave and each C_i is a convex set, x_i^* is a global optimum of the i-th player's optimization problem under the assumptions. Since this is true for each of the

$i \in \{1,\ldots,n\}$ players, x^* is a Nash equilibrium. ■

A sufficient condition guaranteeing that a Nash equilibrium x is isolated is that the Jacobian of $\omega(x,\mu)$, denoted $D\omega(x,\mu)$, is invertible [17], [19]. We refer to such points as being *non-degenerate*.

B. Leader Optimization Problem – Incentive Design

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 [21].

Both the leader and the followers wish to maximize their pay-off determined by the functions $f_L(x,y)$ and $\{f_1(x,\gamma(x)),\ldots,f_n(x,\gamma(x))\}$ respectively where we now consider each of the follower's utility functions to be a function of the incentive mechanism $\gamma : x \mapsto y$ where leader's decision is $y = (d,\rho)$ with d being the default lighting setting and ρ the total number of points. The followers' decisions is denoted by x. The leader's strategy is γ.

The basic approach to solving the reversed Stackelberg game is as follows. Let y and x take values in $Y \subset \mathbb{R}^2$ and $S_i \subset \mathbb{R}$, respectively and let $f_L, f_i : \mathbb{R}^n \times \mathbb{R}^2 \to \mathbb{R}$ for each $i \in \{1,\ldots,n\}$. We define the desired choice for the leader as

$$(x^*, y^*) = \arg\max_{x,y} \{f_L(x,y) |\ y \in Y, x \in \mathcal{C}\}. \qquad (13)$$

The incentive problem can be stated as follows:

Problem 1: Find $\gamma : X \to Y$, $\gamma \in \Gamma$ such that x^* is a differential Nash equilibrium of the follower game $(f_1,\ldots,f_n)$ subject to constraints and $\gamma(x^*) = y^*$ where Γ is the set of admissible incentive mechanisms.

By insuring that the desired agent action x^* is a non-degenerate differential Nash equilibrium ensures structural stability of equilibrium helping to make the solution robust to measurement and environmental noise [22].

For the lighting social game, the leader's utility function is given as follows:

$$f_L(x,y) = \mathbb{E}\left[-g(y,x) - c_1 \sum_{i=1}^{n} \beta_i f_i(x_i, x_{-i}, y) - c_2 p(\rho)\right] \qquad (14)$$

where $g(y,x)$ is the energy cost in kilowatt-hours (kWh), $p(\cdot)$ is a cost function on the points ρ and $c_1, c_2 \in \mathbb{R}_+$ are scaling factors for the last two terms describing how much utility and total points respectively the leader is willing to exchange for 1 kWh. The second term is the *benevolence term* where the β_i's are the *benevolence factors*. This term captures the necessity for the leader to care about the followers' satisfaction which is related to their productivity level (see [23] for a similar formulation). The expectation is taken with respect to the joint distribution defined by distributions across the player states *absent*, *active*, *default* for each player.

Since the prize in the lottery is currently a fixed monetary value delivered to the winner through an Amazon gift card, varying the points does not cost the leader anything explicitly. However, we model the cost of giving points by a function $p(\cdot)$ which captures the fact that after some critical value of ρ the points no longer see as valuable to the followers. The followers' perceive the points that they receive has having some value towards winning the prize. The leader's goal is to choose ρ and d so they induce the followers to play the game and choose the desired lighting setting.

Currently we do not add individual rationality constraints to the leader's optimization problem which would ensure that the players' utilities are at least as much as what they would get by selecting the default value. The impact being that this constraint would ensure players are active. With respect to economics literature, the default lighting setting compares to the outside option in contract theory. It is interesting that in the current situation the leader has control over the outside option. We leave exploring this for future work.

Due to the complexity of computing the expectation for the joint distribution across player states *absent*, *active*, *default* for $n = 22$ players, we currently restrict the set of admissible incentive mechanisms to be the map $\gamma(x) = (\gamma_d(x), \gamma_\rho(x))$ such that the i-th player's utility is

$$f_i(x,\gamma(x)) = \psi_i(x) - \theta_i \gamma_\rho(x) \left(\frac{x_i}{100}\right)^2 \qquad (15)$$

where $\gamma(x) \equiv \rho$ for all $i \in \{1,\ldots,20\}$. In addition, the nature of $\gamma_d(x)$ is that it is an option provided to the followers which they must actively vote in order for this value not to be taken as their current vote when they are present in the office. In sense, it is the outside option. Thus, the leader only selects the constants (d,ρ). This reduces the solution of the reversed Stackelberg game to a bi-level optimization problem that we solve with a particle swarm optimization (PSO) technique (see, e.g., [24]–[26]).

The particle swarm optimization method is a population based stochastic optimization technique in which the algorithm is initialized with a *population* of random solutions and searches for optima by updating *generations*. The potential solutions are called *particles*. Each particle stores its coordinates in the problems space which are associated with the best solution achieved up to the current time. The best over all particles is also stored and at each iteration the algorithm updates the particles' velocities.

At the inner level of the bi-level optimization problem, we replace the condition that the occupants play a Nash equilibrium with the dynamical system determined by the gradients of each player's utility with respect to their own choice variable, i.e.

$$\dot{x}_i = D_i f_i(x_i, x_{-i}, y),\ x_i \in \mathcal{C}_i,\ \forall\, i \in \{1,\ldots,n\}. \qquad (16)$$

It has been show that by using a projected gradient descent method for computing stationary points of the dynamical (16) derived from an n-person concave games on convex strategy spaces converges to Nash equilibria [27]. In our simulations, we add the constraint to the leader's optimization problem that at the stationary points of this dynamical system, i.e. the Nash equilibria, the matrix $-D\omega$ be positive definite thereby ensuring that each of the equilibria are isolated.

Denote the set of non-degenerate stationary points of the dynamical system $\dot{x}$ as defined in (16) as $\text{Stat}(\dot{x})$. The leader then solves the following problem: given the joint distribution across player states *active*, *default*, *absent*, find

$$\max_{y\in Y} f_L(y,x) \quad (17)$$
$$\text{s.t. } x \in \text{Stat}(\dot{x})$$

For each particle in the PSO algorithm, we sample from the distribution across player states and compute Nash for the resulting game via simulation of the dynamical system (16). We compute the mean of the votes at the Nash equilibrium to get the lighting setting. We repeat this process and use the mean of the lighting settings over all the simulations to compute the leader's utility for each of the particles.

We are currently exploring other techniques for solving bi-level optimization problems in which the degree of complexity of computing leader's utility is very high.

IV. Utility Estimation and Incentive Design

In this section, we present our results on both the utility estimation problem and the incentive design problem in which the leader optimizes their cost with respect to the total points to be distributed per day and the default lighting setting.

A. Utility Estimation – Results

We briefly describe the utility estimation problem in this section and refer the interested reader to [16] for a more detailed description including results on the efficacy of our estimations.

We formulate the utility estimation problem as a convex optimization problem by using first-order necessary conditions for Nash equilibria. In particular, the gradient of each occupant's Lagrangian should be identically zero at the observed Nash equilibrium.

For each observation $x^{(k)}$, we assume that it corresponds to occupants playing a strategy that is approximately a Nash equilibrium where the superscript notation $(\cdot)^{(k)}$ indicates the k-th observation. Thus, we can consider first-order optimality conditions for each occupants optimization problem and define a residual function capturing the amount of sub-optimality of each occupants choice $x_i^{(k)}$ [28], [23].

We consider the residual defined by the stationarity and complementary slackness conditions for each occupant's optimization problem:

$$r_{\text{s},i}^{(k)}(\theta_i,\mu_i) = D_i f_i(x_i^{(k)}, x_{-i}^{(k)}) + \sum_{j=1}^{n} \mu_i^j h_{i,j}(x_i^{(k)}) \quad (18)$$

$$r_{\text{c},i}^{j,(k)}(\mu) = \mu_i^j h_{i,j}(x_i^{(k)}) \;\; j\in\{1,2\} \quad (19)$$

Define $r_{\text{s}}^{(k)}(\theta) = [r_{\text{s},1}^{(k)}(\theta_1,\mu_1) \;\cdots\; r_{\text{s},n}^{(k)}(\theta_n,\mu_n)]^T$ and $r_{\text{c},i}^{(k)}(\mu_i) = [r_{\text{c},i}^{1,(k)}(\mu_i)\;\; r_{\text{c},i}^{2,(k)}(\mu_i)]$ so that we can define $r_{\text{c}}^{(k)} = [r_{\text{c},1}^{(k)}(\mu_1) \;\cdots\; r_{\text{c},n}^{(k)}(\mu_n)]^T$ where $\mu_i = (\mu_i^1, \mu_i^2)$.

Given observations $\{x^{(k)}\}_{k=1}^K$ where each $x^{(k)} \in \mathcal{C}$, we can solve the following convex optimization problem:

$$\min_{\mu,\theta} \sum_{k=1}^{K} \chi(r_s^{(k)}(\theta,\mu), r_c^{(k)}(\mu)) \quad (20)$$

$$\text{s.t. } \;\theta_i \geq 0, \mu_i \geq 0 \;\; \forall\, i\in\{1,\dots,n\} \quad (21)$$

where $\chi : \mathbb{R}^n \times \mathbb{R}^{2n} \to \mathbb{R}_+$ is a nonnegative, convex penalty function satisfying $\chi(z_1,z_2)=0$ if and only if $z_1=0$ and $z_2=0$, i.e. any norm on $\mathbb{R}^n\times\mathbb{R}^{2n}$, and the inequality $\mu_i \geq 0$ is elementwise.

Note the constraint that the θ_i's be non-negative. This is to ensure that the estimated utility functions are concave. We add this restriction so that we can employ techniques from simulation of dynamical systems to the computation of the Nash equilibrium in the resulting n-person concave game with convex constraints. In particular, define a gradient-like system using the local representation of the differential game form [17] and using the estimated θ_i's

$$\dot{x}_i = D_i f_i(x_i, x_{-i}; \theta_i) \;\; \forall\, i\in\{1,\dots,n\}, \quad (22)$$

and consider the feasible set defined by the constraints

$$\left.\begin{array}{ll} h_{i,1}(x_i) & = 100 - x_i \geq 0 \\ h_{i,2}(x_i) & = x_i \geq 0 \end{array}\right\} \;\; \forall\, i \in\{1,\dots,20\} \quad (23)$$

Then, the subgradient projection method applied to the dynamics (22) and the constraint set defined by (23) is known to converge to the unique Nash equilibrium of the constrained n-person concave game [27].

By drawing from the joint distribution across player states (*active*, *default*, *absent*), we simulate the game using the estimated utility functions. In figure 2, we can see that our model captures most of the variation in the true votes.

B. Incentive Design – Results

We collected data on the energy consumption of the lights for different lighting settings (see Figure 3) and created a piecewise affine map from the setting to energy consumption in kilowatt–hours (kWh). Using this map, we formulate a utility for the leader which takes the average lighting votes as the input and returns the difference between the maximum consumption in kWh, i.e. 25 kWh, and the piecewise affine map for energy comsuption of the lights.

We include a second term in the leader's utility which is a *benevolence* term. The leader must care about the occupants' satisfaction for various reasons including productivity and safety. We include in the leader's utility the sum of the occupant utilities multiplied by a *benevolence factor* as is described in Section III-B in (14).

We use the function $p(\cdot) = \rho^2$ for the last term in the leader's utility so that it has the form

$$f_L(y,x) = \mathbb{E}\left[K - g(y,x) - c_1 \sum_{i=1}^{n} \beta_i f_i(x_i, x_{-i}, y) - c_2\rho^2\right] \quad (24)$$

where K is the maximum consumption of the Lutron lighting system in kWh's and $g(y,x)$ is the energy consumption at

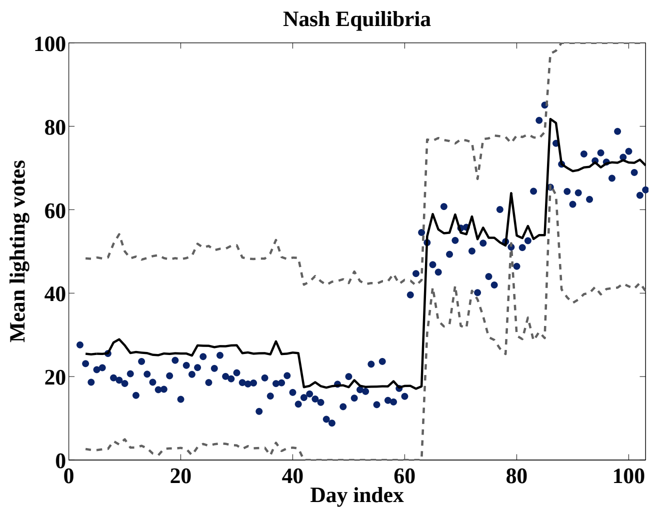

Fig. 2. Blue dots are the true mean of the lighting votes for each day over the druation of the experiment. The solid black line is the mean of the mean of the Nash equilibria for each day obtained via simulating the game with the estimated utilities. The two dashed black lines are one standard deviation of the mean for the simulations. Notice that the mean of the Nash equilibria for the simulated games is very near the true votes and all the variation is captured with in one standard deviation of the mean.

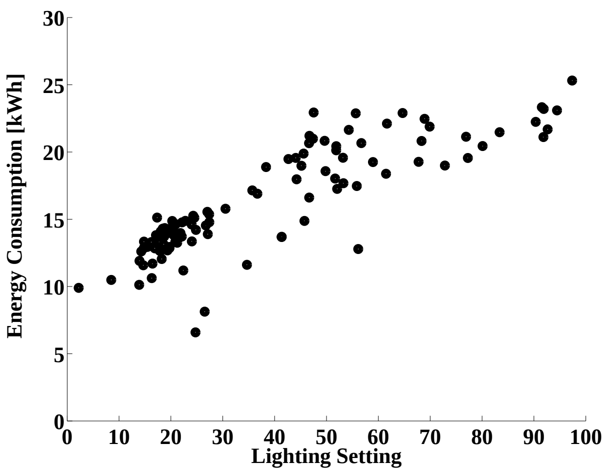

Fig. 3. Energy conumption data for the Lutron lighting system in kWh as a function of the lighting setting.

a given (y, x). c_1, c_2 are constants converting the last two terms to the correct units.

Using the past data, i.e. data collected for default settings $\{10, 20, 60, 90\}$, and θ_i estimates for each occupant, we create a piecewise affine map for interpolating the parameters of the occupants utility functions for different default settings. Similarly, we interpolated the joint distribution across player states (*absent*, *active*, *default*) as a function of the default setting. This allows use to optimize the leaders utility function over both the total points ρ and the default setting d. Due to the complexity of the expectation and the nature of the bi-level optimization problem, we solve the leader's problem by employing a particle swarm optimization method.

Example 1 (Solution Leader Optimization Problem): The following example is a sample solution to the leader's optimization problem under some selection of the parameters c_1, c_2 and the benevolence factor $\beta = (\beta_1, \ldots, \beta_n)$.

In the implementation of the leader's optimization problem in this example we make the following choices for the parameters and scaling of the leader's utility function. For each particle in the PSO algorithm, we map each follower's true utility f_i to $\hat{f}_i$ taking a value in the range $[0, 100]$ by finding the global maximum and minimum of their utility under the current particle to determine an appropriate affine scaling of their original utility. We use $\hat{f}_i$ in place of f_i in the leader's utilty.

We use $c_1 = 1/2$ which represents the fact that the leader is willing to exchange 1 kWh savings for a utility value of 2 in the total sum of the followers' utilities $\sum_i \beta_i \hat{f}_i$ under the current particle value for $y = (d, \rho)$. Similarly, we use $c_2 = 1/500$ which represents the fact that the leader is willing to exchange 500 points in return for 1 kWh of savings.

At present the choice of these parameters is just for the purpose of creating an example with interesting behavior and we leave full exploration of these parameters to future work in which we implement various solutions in practice and obtain feedback from the occupants' via survey about their satisfaction.

Examining each of the occupant's estimated utility functions has given us a sense of which occupants are the most sensitive to changes in ρ and d. Occupant 2 is quite inflexible to changes in the points ρ and appears to care less about winning and more about his comfort level (see Figure 4). This fact is also reflected in the very low parameter estimate for θ_2. It is also the case that occupant 2's behavior is largely affected by others' votes.

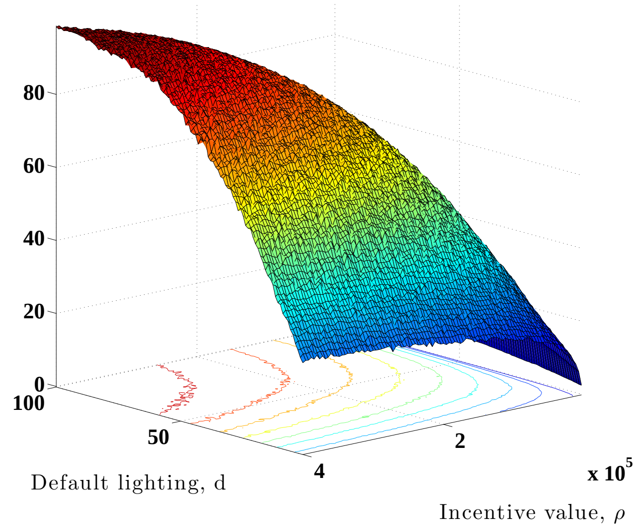

Fig. 4. Utility of occupant 2 as a function of (d, ρ) at the mean Nash equilibrium after running 1000 simulations. Notice that for fixed values of d the utility value is near constant in ρ. Also, occupant 2 has very large utility when the default setting is around 70.

In addition, occupants in the set $C = \{2, 6, 8, 14, 20\}$ are

(d,ρ) \ β	(0.9,0.1)	(0.75,0.25)	(0.6,0.4)	(0.45,0.55)	(0.3,0.7)	(0.2,0.8)
$(10,7000)$	\$2.01	\$2.10	\$2.19	\$2.28	\$2.37	\$2.42
$(20,7000)$	\$1.98	\$2.01	\$2.06	\$2.08	\$2.10	\$2.13
$(60,7000)$	\$1.70	\$1.67	\$1.66	\$1.65	\$1.65	\$1.64
$(90,7000)$	\$1.35	\$1.33	\$1.32	\$1.31	\$1.31	\$1.30

TABLE I

LEADER'S UTILITY IN DOLLARS FOR THE PREVIOUSLY IMPLEMENTED (d,ρ) FOR VARIOUS BENEVOLENCE FACTORS $\beta = (\beta_2, \sum_{j\in A}\beta_j)$ WHERE $A = \{6,8,14,20\}$. THE VALUE IS INTERPRETED AS THE ENERGY SAVED IN DOLLARS BY THE LEADER PLUS THE UTILITY AS MEASURED IN DOLLARS. WE USE A RATE OF \$0.12 PER KWH AS THIS IS THE APPROXIMATE RATE CHARGED BY THE BUILDINGS ON THE UC BERKELEY CAMPUS. COMPARE TO TABLE II

$(d,\rho,\beta_2,\sum_{j\in A}\beta_j)$	utility
$(63, 200\times 10^3, 0.9, 0.1)$	\$4.56
$(56, 169.6\times 10^3, 0.75, 0.25)$	\$4.73
$(55.5, 175.2\times 10^3, 0.6, 0.4)$	\$4.67
$(48, 142.2\times 10^3, 0.45, 0.55)$	\$4.69
$(10.47, 173\times 10^3, 0.3, 0.7)$	\$5.07
$(7.23, 194.6\times 10^3, 0.2, 0.8)$	\$5.43

TABLE II

LEADER'S UTILITY IN DOLLARS FOR THE VALUES $(d^*,\rho^*,\beta_2,\sum_{j\in A}\beta_j)$ WHERE β_2 IS THE BENEVOLENCE FOACTOR FOR USER 2 AND $1-\beta_2 = \sum_{j\in A}\beta_j$ IS THE SUM OF THE BENEVOLENCE FACTORS FOR THE OCCUPANTS $A = \{6,8,14,20\}$. THE UTILITY VALUE IS DETERMINED BY SOLVING THE LEADER'S OPTIMIZATION PROBLEM USING THE PSO METHOD AND IS INTERPRETED AS THE ENERGY SAVED IN DOLLARS BY THE LEADER PLUS THE UTILITY AS MEASURED IN DOLLARS. WE USE A RATE OF \$0.12 PER KWH.

the most active players in a probabilistic sense. As a result, in this example we give non-zero benevolence terms to players in this set. We refer to this set as the leader's *care-set*. For all $i \in \{1,\dots,20\}\backslash C$, we set $\beta_i = 0$. Further, we force $\sum_{j\in C}\beta_j = 1$. Since occupant 2 has particularly interesting behavior, we vary β_2, and let $\beta_j = (1-\beta_2)\frac{1}{|C|}$ for all $j \in C$ and where $|C|$ is the cardinality of C.

Tables I and II contain the energy savings in dollars for the leader per day given the energy cost of the lights and how much of the occupants' utility and the total points distributed per day that the leader is willing to exchange for 1 kWh in dollars using a cost per kWh of \$0.12. Table I has the leader's utility in dollars for previous values of (d,ρ) after the start of the social game. In Table II we report the values after optimizing over (d,ρ) for some given benevolence factor $\beta = (\beta_1,\dots,\beta_n)$. We can see that computing even the suboptimal (d,ρ) by solving the leader's bi-level optimization problem via PSO, the leader has a much higher utility.

We have not yet factored in the cost of the prize in the lottery. Currently it is at a value of \$100 per week. The values we report in Tables I and II are per day savings on weekdays. Hence, with a prize cost of \$20 per day for our particular experimental set-up the leader does not save. Using this case-study as proof-of-concept, we are in the process of implementing a social game in an entire building in Singapore with more than $1,000$ occupants. This social game will include options for the consumer to choose lighting setting, HVAC and personal cubicle plug-load consumption. In addition, we plan to implement a social game of this nature in Sutarja Dai Hall on the UC Berkeley campus. At this scale, with a week-day lottery cost of \$100 the building manager stands to save a considerable amount.

In Figure 5 we show the results of simulating the game under the (d,ρ)'s that we found for various benevolence factors. We show the mean of the lighting votes averaged over 1000 simulations. It is interesting to see that the average Nash equilibrium under the various default settings is actually less than the default setting itself except in the case when the default setting is below a threshold below which occupants actually log votes above the default setting. For example, with a default setting of 10.74, the mean of the Nash equilibria is ~ 15. The case when the default setting is above this threshold of basic operation, the most aggressive players' desire to win pushes the Nash equilibrium below the default. On the other hand, when the default is below this threshold, all the players' comfort comes into play and shifts the Nash equilibrium above the default setting.

V. DISCUSSION AND FUTURE WORK

We presented the results of a social game for encouraging energy efficient behavior in building occupants and modeling of occupant behavior patterns. We briefly discussed the utility estimation problem. Using the estimated utilities, we formulated and solved the building manager's bi-level optimization problem for the total points and default setting. Due to the large number of events underlying the joint distribution across player states and non-convexities, we utilized a particle swarm optimization method. We are exploring more efficient methods for solving for the optimal

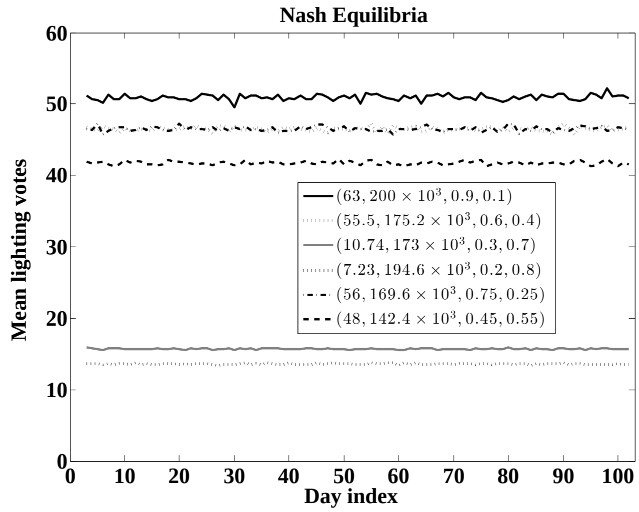

Fig. 5. Mean of the Nash equilibria of the simulated games over 103 days under estimated occupant utilities with leader incentives found via PSO and parameters given by $(d, \rho, \beta_2, \sum_{j\in A} \beta_j)$ where $A = \{6, 8, 14, 20\}$. Note that the mean Nash equilibrium in each case is slightly below the default setting.

points and default setting as well as implementing the current (d, ρ) that we found through PSO in our test bed.

The leader's utility function contains a number of parameters such as c_1, c_2 and the benevolence factor which represent how much utility or *happiness* the leader is willing to exchange for savings. We are in the process of examing the impact of these factors on the leader savings as well as the occupant satisfaction in practice. We are implementing surveys to collect additional data about the occupants' satisfaction which we plan to incorporate into our solution.

In addition, we did not include individual rationality constraints in the leader's optimization problem. It would be interesting to explore incorporating such a constraint in the optimization problem where we consider the outside good to be the default setting. This problem is slightly different than what is seen in the economics literature because the leader here has control over the default setting, and thus, the outside good.

Another interesting direction for future research that we are exploring is understanding the type (parameter) space of the occupants and how the Nash equilibria of the follower game depend on these parameters. Specifically it is interesting to take a dynamical systems perspective and understand under which parameter configurations the desired Nash equilibrium from the leader perspective is structurally stable.

References

[1] J. McQuade, "A system approach to high performance buildings," *United Technologies Corporation, Tech. Rep*, 2009.

[2] A. Swenson, "Lighting in commercial buildings," U.S. Energy Infomation Administration, Tech. Rep., 2009.

[3] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, "Identifying models of HVAC systems using semiparametric regression," in *Proceedings of the American Control Conference*, 2012.

[4] M. Boman, P. Davidsson, N. Skarmeas, K. Clark, and R. Gustavsson, "Energy saving and added customer value in intelligent buildings," in *Third International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology*, 1998, pp. 505–517.

[5] D. Bourgeois, C. Reinhart, and I. Macdonald, "Adding advanced behavioural models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control," *Energy and Buildings*, vol. 38, no. 7, pp. 814 – 823, 2006, special Issue on Daylighting Buildings.

[6] Y. Ma, G. Anderson, and F. Borrelli, "A distributed predictive control approach to building temperature regulation," in *Proceedings of the American Control Conference*. IEEE, 2011, pp. 2089–2094.

[7] F. Oldewurtel, A. Parisio, C. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and K. Wirth, "Energy efficient building climate control using stochastic model predictive control and weather predictions," in *Proceedings of the American Control Conference*, 2010, pp. 5100–5105.

[8] T. Lovett, E. Gabe-Thomas, S. Natarajan, E. O'Neill, and J. Padget, "'just enough' sensing to enliten: A preliminary demonstration of sensing strategy for the 'energy literacy through an intelligent home energy advisor' (enliten) project," in *Proceedings of the Fourth International Conference on Future Energy Systems*. New York, NY, USA: ACM, 2013, pp. 279–280.

[9] J. Mathieu, M. Dyson, D. Callaway, and A. Rosenfeld, "Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.

[10] M. Roozbehani, M. Dahleh, and S. Mitter, "Dynamic pricing and stabilization of supply and demand in modern electric power grids," in *First IEEE International Conference on Smart Grid Communications (SmartGridComm),*, oct. 2010, pp. 543 –548.

[11] D. Merugu, B. S. Prabhakar, and N. Rama, "An incentive mechanism for decongesting the roads: A pilot program in bangalore," in *Proceedings of ACM NetEcon Workshop*, 2009.

[12] A. Bestick, L. J. Ratliff, P. Yan, R. Bajcsy, and S. S. Sastry, "An inverse correlated equilibrium framework for utility learning in multiplayer, noncooperative settings," in *Proceedings of the 2nd ACM international conference on High confidence networked systems*. ACM, 2013, pp. 9–16.

[13] M. Albadi and E. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989 – 1996, 2008.

[14] J. L. Mathieu, P. N. Price, S. Kiliccote, and M. A. Piette, "Quantifying changes in building electricity use, with application to demand response," *IEEE Transactions on Smart Grid,*, vol. 2, no. 3, pp. 507–518, 2011.

[15] M. Lee, O.Aslam, B. Foster, D. Kathan, J. Kwok, L. Medearis, R. Palmer, P. Sporborg, and M. Tita, "Assessment of demand response and advanced metering," Federal Energy Regulatory Commission, Tech. Rep., 2013.

[16] I. C. Konstantakopoulos, L. J. Ratliff, M. Jin, S. S. Sastry, , and C. Spanos, "Social game for building energy efficiency: Utility learning, simulation, and analysis," arXiv, Tech. Rep., 2014.

[17] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "Characterization and computation of local nash equilibria in continuous games," in *Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing*, 2013.

[18] G. Taguchi, E. A. Elsayed, and T. C. Hsiang, *Quality engineering in production systems*. McGraw-Hill College, 1989.

[19] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, p. 520, 1965.

[20] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[21] N. Groot, B. De Schutter, and H. Hellendoorn, "Reverse stackelberg games, part i: Basic framework," in *IEEE International Conference on Control Applications,*, Oct 2012, pp. 421–426.

[22] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "Generictiy and structural stability of non–degenerate differential nash equilibria," in *Proceedings of the 2014 American Controls Conference*, 2014.

[23] L. J. Ratliff, R. Dong, H. Ohlsson, and S. S. Sastry, "Incentive design and utility learning via energy disaggregation," in *Proceedings of the 19th World Congress of the International Federation of Automatic Control*, 2014.

[24] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, Nov 1995, pp. 1942–1948.

[25] J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*. Springer, 2010, pp. 760–766.

[26] G. Venter and J. Sobieszczanski-Sobieski, "Particle swarm optimization," *AIAA journal*, vol. 41, no. 8, pp. 1583–1589, 2003.

[27] S. D. Flåm, "Solving non-cooperative games by continuous subgradient projection methods," in *System Modelling and Optimization*. Springer, 1990, pp. 115–123.

[28] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *IEEE International Symposium on Intelligent Control*. IEEE, 2011, pp. 613–619.