REST: A Reliable Estimation of Stopping Time Algorithm for Social Game Experiments*

Ming Jin, Lillian J. Ratliff, Ioannis Konstantakopoulos, Costas Spanos, and Shankar Sastry
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley, Berkeley, CA
{ jinming, ratliffl,ionniskon,spanos,sastry}@eecs.berkeley.edu

ABSTRACT

Through a social game, we integrate building occupants into the control and management of an office building that is instrumented with networked embedded systems for sensing and actuation. The goal of the social game is to both incentivize building occupants to be more energy efficient and learn behavioral models for occupants so that the building can be made sustainable through automation. Given a generative model for the occupants behavior in the competitive environment created by the social game, we develop a method for learning the parameters of the behavioral model as we conduct the experiment by adopting a *learning to learn* framework. Using tools from statistical learning, we provide bounds on the parameter inference error. In addition, we provide an algorithm for computing the stopping time required for a specified level of confidence in estimation. We show the performance of our algorithm in several examples.

Categories and Subject Descriptors

F.1.2 [**Theory of Computation**]: COMPUTATION BY ABSTRACT DEVICES—*Modes of Computation*; G.3 [**PROBABILITY AND STATISTICS**]: Experimental design; I.6.m [**SIMULATION AND MODELING**]: Miscellaneous

*This work is supported by TRUST (Team for Research in Ubiquitous Secure Technology) which receives support from NSF (CCF-0424422), FORCES (Foundations Of Resilient CybEr-physical Systems) which receives support from NSF (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.

ICCPS '15, April 14 - 16, 2015, Seattle, WA, USA
Copyright 2015 ACM 978-1-4503-3455-6/15/04 ...$15.00
http://dx.doi.org/10.1145/2735960.2735974 .

1. INTRODUCTION

Energy consumption of buildings, both residential and commercial, accounts for approximately 40% of all energy usage in the U.S. [8]. Lighting is a major consumer of energy in commercial buildings; one-fifth of all energy consumed in buildings is due to lighting [16].

Intelligent lighting control system, among other building automation and control (BAC) systems, has been designed and implemented to ensure both energy efficiency and user comfort. Striking the right balance between these two objectives is fundamental aspect of BAC systems, which requires thorough understanding of the preferences and behaviors of the occupants. BAC systems are composed of many embedded devices linked together for information sharing, sensing, and actuation. They are inherently cyber-physical systems (CPS).

Some tasks that a building manager might be interested in include predicting energy load, incentivizing users to change behavior, or—at a more basic level—learning user preferences for improving building automation. In all of these tasks the building manager needs a way to distinguish between occupants. One way to think about this problem is as a classification task.

We design a social game in which building occupants compete for points that determine their likelihood of winning a lottery by voting according to their lighting preferences via online platform—the more energy efficient an occupant's behavior, the more likely they are to win. Our initial experiments demonstrate desirable reduction of energy consumption while ensuring the lighting comfort of occupants [12].

The goal of the building manager in this setting is two-fold; the desire to induce the occupants to be more energy efficient as well as desire to have an accurate model of occupant behavior for improved automation. Given a desired classification accuracy, we apply results from statistical learning to provide bounds on parameter inference error as well as estimate stopping time that achieves the classification rate. In particular, strategies for stopping time estimation, as presented in this study, include lower bound based on Cramér-Rao bounds types of bounds that have been used in statistics to conduct feasibility analysis and judgment of proposed estimators, as well as upper bounds based on Delta method [5] and concentration inequalities [11]. The latter provide the theoretical foundation for the reliable estimation of stopping time (REST) algorithm (Section 6, Algorithm 1).

The remainder of the paper is organized as follows. We begin with a brief discussion of motivations and an overview of the social game experimental setup in Sections 2 and 3 respectively. In Section 4 we present the theoretical formulation including the game-theoretic, generative behavior model based on a game–theoretic view of player interactions and the *learning to learn* framework. We provide our results on parameter inference error bounds which lead to stopping time bounds in Section 5. In Section 6 we present the REST algorithm. Finally, in Section 7 we illustrate the applicability of the REST algorithm to stopping time estimation in the learning of behavior parameters, the prediction of lighting energy consumption, and classification of players into behavior–based categories. We conclude in Section 8 with some discussion on the results and comments on future work.

2. BACKGROUND AND MOTIVATION

The estimation of stopping times is useful for the proposal, experimental design, and evaluation of performance of various CPS with socioeconomic and environmental considerations. For instance, to evaluate the structure wherein nearby homes explicitly share energy with each other to balance local energy harvesting and demand in microgrids proposed in [18], we need to decide how long to run the experiment to arrive at a reliable conclusion. Another example is the evaluation of the control algorithm and embedded platform for HVAC systems proposed for energy savings [7]. Although the experimental periods in these studies are often chosen according to time and resource constraints, a recommendation of experimental period based on the complexity of the learning objectives is often informative in practice.

The social game based on lighting control systems is a pilot experiment to study the behavior of occupants in non-cooperative setting and to evaluate the potential of energy savings. We are interested in knowing how long we should run the experiment if we scale up from 22 people as is in the current setting to, for instance, 100 people in a much larger office to achieve more substantial impact on energy savings. This is one of the major motivations for the work presented in the present article.

The problem of determining the sample size has been long studied in theoretical statistics by an approach called power analysis that obtains the sample size required to correctly reject the null hypothesis (see, e.g., [3]). In addition, statistical learning theory provides generalization error bounds based on sampling complexity, such as the Rademacher complexity that is used for progressive sampling (see e.g. [1]). The probably approximately correct (PAC) property is similar to the *probably close enough* criteria (see e.g. [4]) where model–based dynamic sampling methods are evaluated. Another related line of work is called *virtual sample generation* that uses techniques, such as intervalized kernel methods and bootstrapping, to produce extra information for expediting learning (see, e.g., [6]). It was shown that virtual sample generation can be used to improve the small-data-set learning in manufacturing systems. Since only few examples can be obtained in the early stages in social experiments, we adopted a similar strategy to produce virtual samples based on our generative behavior model to reliably estimate the stopping time.

3. SOCIAL GAME

In this section, we provide a brief overview of the experimental setup. We have instrumented an office space with a heating, ventilation, and air conditioning (HVAC) system, automated lighting control (Lutron system[1]), plug-load metering and carbon dioxide sensors. The social game for energy savings that we have designed is such that occupants in an office building 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 experiments we have conducting thus far focus only on controlling the lighting dim level in the office.

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 dim level votes, view point balances of all occupants, and observe all the behavior (voting) patterns of all occupants.

An occupant chooses a value in the interval $[0, 100]$ representing their vote for the dim level in their zone as well as neighboring zones. The lighting setting that is implemented in each zone is the average of all the votes weighted according to proximity to that zone. For analysis, the occupants can be in one of three states. If an occupant actively logs in and votes, we say the occupant is in the state *active*. If an occupant is present

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

in the office but leaves their vote at the default dim level, we say their state is *default*. Finally, if the occupant is not present, we say their state is *absent*.

4. THEORETICAL FORMULATION

In this section we provide the theoretical formulation of our approach to estimation of stopping time. The problem is formulated as an inference problem in a transfer learning setting [2,10]. A fundamental assumption is that the learner's uncertainty concerning the true model (or equivalently, the true prior) is large, while the dimensionality of the true model is in fact quite low.

In our experimental setup we aim to learn agents preferences with respect to usage of shared resources. We assume that consistency in the office setting in terms of schedules and behaviors will lead to a low dimensional model for occupant preferences.

4.1 Generative Behavioral Model

In previous work we took a game-theoretic approach for designing incentives and estimating utility functions of the occupants [12]. We were able to estimate the parameters of each occupant's utility function via solving a convex optimization problem. In this section we present the essential model; for the interested reader, a more detailed description can be found in [12].

Let n denote the number of players. Each player optimizes their utility function $\Psi_i(x_i, x_{-i})$ over some set $\mathcal{C}_i = \{x_i \in \mathbb{R} |\ c_{i,j}(x_i) \geq 0, j \in \{1, \ldots, m_i\}\}$. Let $\mathcal{C} = \prod_{i=1}^n \mathcal{C}_i$ be the joint strategy space. We parameterized each occupant's utility functions as follows:

$$\begin{aligned}\Psi_i(x_i, x_{-i}, \gamma; \xi) =& \psi_0(x_i, x_{-i}, \gamma(x_i, x_{-i})) \\ &+ \xi_i \psi_1(x_i, x_{-i}, \gamma(x_i, x_{-i})) \end{aligned} \quad (1)$$

where γ is the announced incentive, ψ_0 and ψ_1 are concave basis functions representing the occupants *comfort* and *desire to win* respectively, and ξ_i is the parameter. We remark that this framework easily extends to any finite number of basis functions and therefore, parameters; we choose only two basis functions as this is a reasonable model for the energy management social game we conducted.

Then we relaxed the first-order conditions for (differential) Nash equilibria. In particular, we defined residuals using the stationarity and complementary slackness conditions of each player's individual optimization problem at each iteration $\{x^k, \gamma^k\}_{k=1}^K$ where x^k is the observed Nash equilibrium of the γ^k induced game at iteration k. The stationarity residual is given by

$$r_{s,i}^k(\xi_i, \lambda_i) = D_i \Psi_i(x^k, \gamma^k(x^k)) + \sum_{j=1}^{m_i} \lambda_{i,j} D_i c_{i,j}(x_i^k) \quad (2)$$

where $m_i = |A_i(x_i^k)|$ is the cardinality of the active constraint set at x_i^k and $\lambda_i = (\lambda_{1,1}, \ldots, \lambda_{i,m_i})$. The complementary slackness residual is similarly defined by

$$r_{c,i}^{j,k}(\mu_i) = \lambda_{i,j} c_{i,j}(x_i^k), \quad \forall\ j \in \{1, \ldots, m_i\}. \quad (3)$$

Define $r_s^k(\xi, \lambda) = [r_{s,1}^k(\xi_1, \lambda_1)\ \cdots\ r_{s,n}^k(\xi_n, \lambda_n)]^T$ and $r_{c,i}^k(\lambda_i) = [r_{c,i}^{1,k}(\lambda_i)\ \cdots\ r_{c,i}^{m_i,k}(\lambda_i)]$ so that we can define $r_c^k(\lambda) = [r_{c,1}^k(\lambda_1)\ \cdots\ r_{c,n}^k(\lambda_n)]^T$. We took a non-negative, convex penalty function, $\chi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ where $m = \sum_{i=1}^n m_i$, and defined the following optimization problem in the parameters $\xi = (\xi_1, \ldots, \xi_n)$ and the Lagrange multipliers $\lambda = (\lambda_1, \ldots, \lambda_n)$:

$$\min_{\lambda \geq 0, \xi \geq 0} \sum_{k=1}^K \chi(r_s^k(\xi, \lambda), r_c^k(\lambda)) \quad (4)$$

Note that since the basis functions are concave and we force the parameters to be positive, the resulting game is a n–person concave game and thus we know a Nash equilibrium exists [14]. If the observations are non-degenerate differential Nash equilibria in the games determined by the estimated parameters, then we know the observed Nash are *unique* [13]. The condition required to verify this is checking the derivative of the differential game form, i.e. a matrix of the second-order partial derivatives of each player's utility function, is negative definite. This condition may be added as a constraint in the above optimization problem.

4.2 Learning to Learn

Having an *objective* prior used in a classification or other learning task is extremely advantageous. Part of our goal in this paper is to utilize the *learning to learn* framework to learn an objective prior by starting with a *subjective* prior [2].

We model the human behavior as a hierarchical Bayesian model with two levels. The first level represents individuality, such as personality and psychology, which is responsible for decision making, denoted as $\pi \in \Pi$. Given the individuality parameter, the parameter ξ of the generative behavior model proposed in Section 4.1 is distributed according to the objective prior, $p(\xi|\pi)$, and we assume by realizability that $\pi^* \in \Pi$ corresponds to the true prior.

The data generated by the behavior model, $x \in \mathcal{C}$, allows direct inference to the parameters ξ. Our framework is general enough to incorporate other types of data, e.g. survey data, which allows for direct inference to the individuality π of each of the players. In addition, we remark that this framework does not preclude the inclusion of labels in the data set. We do not observe all the factors in the environment that influence agent behavior and individuality and we treat these unobservable environmental factors as sources of randomness.

In practice, since the social game will have a large number of players, we can model the parameters of each players utility functions as following a normal distribution. We thus establish the hierarchical Bayesian model for the group as follows:

Group–level variation (first layer): $\mu_j \sim \mathcal{N}(\nu, \sigma)$, $\theta_j \sim \Gamma(\alpha, \beta)$ *where* μ_j *is the mean response sampled from the normal distribution and* θ_j *is the variance of the response sampled from the gamma distribution for player* j.

Individual randomness (second layer): $\xi_i^k \sim \mathcal{N}(\mu_i, \theta_i)$ *is the random parameter for individual* i *sampled at time* k *according to the normal distribution whose parameters* μ_i *and* θ_i *are sampled from the first layer.*

Observation (third layer): *We will denote the data collected up to time* t *by* $\boldsymbol{x}^t = \{x^1, \ldots, x^t\}$ *where* x^k *are the responses generated from the model in Section 4.1 where* $\xi_i^k \sim \mathcal{N}(\mu_i, \theta_i)$ *is the parameter used in the model. Since the game is a concave game, there exists a unique Nash equilibrium so that there is a one-to-one mapping from the parameters* $\xi_1^k, \ldots, \xi_n^k$ *to the Nash equilibrium* $x^k = (x_1^k, \ldots, x_n^k)$.

We use the additional notation $\boldsymbol{x}_i^t = \{x_i^1, \ldots, x_i^t\}$ to denote the data for player i up to time t where $x^k = (x_1^k, \ldots, x_n^k)$ is the Nash equilibrium at time k.

5. BOUNDS ON INFERENCE ERROR

In this section we utilize existing statistical theory to derive lower bounds on the inference error of the hyper–parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_4) = (\alpha, \beta, \nu, \sigma)$. Since the hyper–parameters are deterministic, we can apply the maximum likelihood estimator (MLE) given by

$$\hat{\boldsymbol{\theta}}^{MLE} = \arg\max_{\boldsymbol{\theta}} \ln p(\boldsymbol{x}^t | \boldsymbol{\theta}) \tag{5}$$

which maximizes the log–likelihood of the data $\boldsymbol{x}^t$ collected up to time t. For any estimators which are unbiased, the Cramér–Rao bound [17] provides a lower bound on the mean square error (MSE) of the estimation if certain regularity conditions are met (see, e.g., [9, Def. 7.21]). According to the transfer learning framework, each occupant $i \in \{1, \ldots, n\}$ will provide one instance of the subjective prior as they are responsible for contributing $\boldsymbol{x}_i^t$ to the data set. Intuitively, simultaneously learning a group of participants will give us a better inference of the hyper–parameters, which is demonstrated in the following propositions.

Proposition 1. *Suppose* $\boldsymbol{x}_1^t, \ldots, \boldsymbol{x}_n^t$ *are i.i.d. from an unknown distribution* $p_{\boldsymbol{\theta}}(\cdot)$ *parameterized by the hyper–parameters* $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_4)$. *For any estimators that meet the regularity conditions, the MSE of estimation of* $\boldsymbol{\theta}_i$ *is lower bounded by*

$$\boldsymbol{E}_{\boldsymbol{x}^t}\left[(\boldsymbol{\theta}_i - \hat{\boldsymbol{\theta}}_i)^2\right] \geq \frac{1}{n\zeta_i} \tag{6}$$

where $\zeta_i = -\boldsymbol{E}_{\boldsymbol{x}^t}\left[\frac{\partial^2 \ln p(\boldsymbol{x}^t|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}_i^2}\right]$ *is the curvature whose expectation is taken over the random variable* $\boldsymbol{x}^t$.

Now we consider estimating the random parameters μ_i, θ_i at the second–level given the observations $\boldsymbol{\xi}_i^t = (\xi_i^1, \ldots, \xi_i^t)$ for each $i \in \{1, \ldots, n\}$. Note that the optimization problem to find the admissible set of parameters $\{\xi_1^k, \ldots, \xi_n^k\}$ that induce the observed Nash $(x_1^k, \ldots, x_n^k)$ is convex. Hence, given an observed Nash equilibrium, through solving the convex optimization problem, we have a unique set of parameters $\{\xi_1^k, \ldots, \xi_n^k\}$. Thus we can consider our data to be $\boldsymbol{\xi}_i^t$ for each $i \in \{1, \ldots, n\}$.

Define the MSE matrix for estimator as $\hat{\boldsymbol{\theta}}(\boldsymbol{\xi}^t)$ as

$$\boldsymbol{R}(\hat{\boldsymbol{\theta}}) = \boldsymbol{E}_{\boldsymbol{\xi}^t, \boldsymbol{\theta}}\left[\left(\hat{\boldsymbol{\theta}}(\boldsymbol{\xi}^t) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\boldsymbol{\xi}^t) - \boldsymbol{\theta}\right)^T\right]. \tag{7}$$

One possible estimator that could be used is the maximum a posteriori estimator (MAP) given by

$$\hat{\boldsymbol{\theta}}^{MAP} = \arg\max_{\boldsymbol{\theta}} \left[\ln p(\boldsymbol{\theta}|\boldsymbol{\xi}^t)\right]. \tag{8}$$

In Bayesian estimation, the performance of any estimator $\hat{\boldsymbol{\theta}}(\boldsymbol{\xi}^t)$ can be bounded by the Bayesian Cramér–Rao bound (BCRB), under suitable regularity conditions [15, Def. 2.78].

Proposition 2. *Suppose we have a social game that has been carried out for time* T *with* n *players whose behavior model parameters are distributed as* $\xi_i^k \sim \mathcal{N}(\mu_i, \theta_i)$ *where* $\mu_i \sim \mathcal{N}(\nu, \sigma)$ *and* $\theta_i \sim \Gamma(\alpha, \beta)$ *for* $k \in \{1, \ldots, T\}$ *and* $i \in \{1, \ldots, n\}$. *Suppose that* μ_i *and* θ_i *are independent. Then, under suitable regularity conditions, the MSE matrix is lower bounded:*

$$\boldsymbol{R}\left(\hat{\mu}_i, \hat{\theta}_i\right) \geq \begin{bmatrix} \frac{(\alpha-1)\beta\sigma}{T\sigma - (\alpha-1)\beta} & 0 \\ 0 & \frac{2(\alpha-1)(\alpha-2)\beta^2}{T+2\alpha-2} \end{bmatrix}. \tag{9}$$

We remark that the importance of the lower-bound in the above proposition is that it bounds the MSE achievable by the optimal estimator and reveals the relationship between the MSE and the number of rounds, T, required in the experiment to achieve that bound. The error bound also decreases as a function of $\frac{1}{T}$.

As we see the BCRB involves the hyper–parameters of the subjective priors, which are non–random variables and often unknown. The hybrid Cramér–Rao bound (HCRB) [17] is applicable to the joint estimation of random and non–random parameters. The following proposition provides a HCRB for the parameter estimation for an individual.

Proposition 3. *Suppose we have a social game that has been carried out for time* T *with* n *players whose behavior model parameters are distributed as* $\xi_i^k \sim \mathcal{N}(\mu_i, \theta_i)$

where $\mu_i \sim \mathcal{N}(\nu, \sigma)$ *and* $\theta_i \sim \Gamma(\alpha, \beta)$ *for* $k \in \{1, \dots, T\}$ *and* $i \in \{1, \dots, n\}$. *Let* $\boldsymbol{\theta}_{nr} = (\nu, \sigma, \alpha, \beta)$ *and* $\boldsymbol{\theta}_r = (\mu_i, \theta_i)$ *in the parameter vector* $\boldsymbol{\theta} = (\boldsymbol{\theta}_{nr}, \boldsymbol{\theta}_r)$. *Then, under suitable regularity conditions, the MSE matrix is lower bounded by*

$$\boldsymbol{R}(\hat{\boldsymbol{\theta}}) \geq \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}^{-1} \tag{10}$$

where

$$A = \begin{bmatrix} \frac{1}{\sigma} & 0 & 0 & 0 \\ 0 & \frac{1}{2\sigma^2} & 0 & 0 \\ 0 & 0 & \frac{\Gamma''(\alpha)\Gamma(\alpha) - \Gamma'(\alpha)^2}{\Gamma(\alpha)^2} & \frac{1}{\beta} \\ 0 & 0 & \frac{1}{\beta} & \frac{\alpha}{\beta^2} \end{bmatrix}, \tag{11}$$

$$B = \begin{bmatrix} -\frac{1}{\sigma} & 0 \\ 0 & 0 \\ 0 & -\frac{1}{(\alpha-1)\beta} \\ 0 & -\frac{1}{\beta^2} \end{bmatrix}, \tag{12}$$

and

$$C = \begin{bmatrix} \frac{T\sigma + (\alpha-1)\beta}{(\alpha-1)\beta\sigma} & 0 \\ 0 & \frac{T+2\alpha-2}{2(\alpha-1)(\alpha-2)\beta^2} \end{bmatrix}. \tag{13}$$

In particular, the MSE for the individual parameters $\boldsymbol{\theta}_r = (\mu_i, \theta_i)$ *is lower bounded by*

$$\boldsymbol{R}(\hat{\mu}_i, \hat{\theta}_i) \geq \begin{bmatrix} \frac{(\alpha-1)\beta}{T} & 0 \\ 0 & \frac{2(\alpha-1)(\alpha-2)\beta^2}{T + \frac{4c(\alpha-1)^2 - 4\alpha + 6}{\alpha^2 c - \alpha c - \alpha + 1}} \end{bmatrix} \tag{14}$$

where $c = \frac{\Gamma(\alpha)\Gamma''(\alpha) - \Gamma'(\alpha)^2}{\Gamma(\alpha)^2}$ *and* $\Gamma(\cdot)$ *is the gamma function.*

We remark that the above bounds are theoretical and often difficult to compute in practice. In the next section we present an approach to computing stopping time based on these bounds.

6. REST ALGORITHMS

In this section we present the reliable estimation of stopping time (REST) algorithm. Let us first describe the inputs to the REST algorithm. As before, let n be the number of players in the game. We use the notation $\boldsymbol{\theta}$ to denote the set of model parameters $\{\mu_i, \theta_i\}$ for each $i \in \{1, \dots, n\}$. Let $\mathbf{Y}^t \in \mathbb{R}^{t \times n}$ be a random matrix where $[\boldsymbol{Y}^t]_{i,j} \in \mathcal{C}_i$ is the random variable representing player i's vote at time $j \in \{1, \dots, t\}$ and $Y^k \in \mathcal{C}$ is the vector valued random variable representing the Nash equilibrium at time k. For the McDiarmid method, we are given a target function $f : \mathcal{C}^m \to \mathbf{R}$ that is a function of $Y^1, \dots, Y^t$ satisfying for all j and all $x^1, \dots, x^t, \hat{x}^j \in \mathcal{C}$ that

$$|f(x^1, \dots, x^j, \dots, x^t) - f(x^1, \dots, \hat{x}^j, \dots, x^t)| \leq c_j. \tag{15}$$

The target function represents the objective; for instance, it may be the average of the lighting votes or the amount of energy consumed by the system at the Nash equilibrium. Note that $\boldsymbol{Y}^t$ is the random variable and we again use the notation $\boldsymbol{x}^t$ to denote the data up to time t.

For the Delta method, we are given real-valued, differentiable function g that is a function of the sample mean. We use the notation $g'(\boldsymbol{x}^t)$ for the first derivative of g.

We define the following functions: $M_{est}(\mathbf{x}^t)$ returns the estimated model parameters $\boldsymbol{\theta}$, $M_{sim}(\boldsymbol{\theta}, t)$ is the generative model which returns the simulated data $\boldsymbol{x}^k_{sim}$ for time k, $\Phi(z; \mu, \theta)$ is the cumulative density function of $\mathcal{N}(\mu, \theta)$ evaluated at point $z \in \mathbb{R}$. In addition, let $\varepsilon > 0$ be the precision and $1 - \delta > 0$ probability bounds (see either (16) or (20)) and let t be the start time of the experiment.

The REST algorithm is given in Algorithm 1. The two methods at the core of the REST algorithm are the *McDiarmid* method [11] and the *Delta* method [5]. We have the following results showing that for each of the methods a probably approximately correct (PAC) property holds.

Algorithm 1 Reliable Estimation of Stopping Time

function REST($\boldsymbol{x}^t$,ϵ,δ,M_{est},f',M_{sim},L,M)
  Initialization:
    $\hat{\boldsymbol{\theta}} \leftarrow M_{est}(\boldsymbol{x}^t)$ ▷ estimation of model parameters
    $\tau \leftarrow L$ ▷ initialization of stopping time
    $t_{\text{start}} \leftarrow t$ ▷ initial time
    $k \leftarrow t_{\text{start}}$ ▷ iteration number
    $\hat{\sigma} \leftarrow std(\boldsymbol{x}^t)$ ▷ estimation of standard deviation
    $stop \leftarrow false$ ▷ stopping condition
  Main program:
    while $\neg stop \wedge (k < L)$ **do**
      $k \leftarrow k + 1$
      switch M **do**
        case Delta
          if $\left(\Phi\left(\epsilon; 0, \frac{f'(\boldsymbol{x}^t)\hat{\sigma}}{\sqrt{k}}\right) - \frac{1}{2}\right) \geq \frac{\delta}{2}$ **then**
            $stop \leftarrow true$
          end if
        case McDiarmid
          $\boldsymbol{x}^{k-t_{start}}_{sim} \leftarrow M_{sim}(\hat{\boldsymbol{\theta}}, k - t_{\text{start}})$
          $\boldsymbol{x}_{tot} \leftarrow [\boldsymbol{x}^t, \boldsymbol{x}^{k-t_{start}}_{sim}]$
          if $\exp\left(-\frac{2\varepsilon^2}{c^2_{ub}(\boldsymbol{x}_{tot})}\right) \leq \frac{1-\delta}{2}$ **then**
            $stop \leftarrow true$
          end if
    end while
  Outputs: $\tau \leftarrow k - t_{\text{start}}$
    ▷ Estimated stopping time from now
end function

Theorem 1. ***(McDiarmid's stopping time)** The output of stopping time, τ_{MD}^{EST}, according to the McDiarmid method of the REST Algorithm satisfies the following PAC property,*

$$P\left(|f(\boldsymbol{Y}^m) - \mathbf{E}\left[f(\boldsymbol{Y}^m)\right]| \geq \varepsilon \right) \leq 1-\delta, \forall m \geq \tau_{MD}^{EST} \tag{16}$$

where $f(\boldsymbol{Y}^m) \in \mathbb{R}$ is the random variable of investigation, ε is the desired precision, and $1-\delta$ is the specified probability bound.

Proof. Consider $\boldsymbol{Y}^m = (Y^1, ..., Y^m)$ where the Y^k are independent random variables taking values in the set $\mathcal{C}$, and a real–valued function $f(\boldsymbol{Y}^m)$. Define the quantity based on the upper bounds

$$c_{ub}(\boldsymbol{Y}^m) = \sum_{j=1}^{m} c_j^2. \tag{17}$$

The algorithm produces bounds on the maximum difference between two realizations of function values:

$$\left| f(x^1, ..., x^j, ..., x^m) - f(x^1, ..., \hat{x}^j, ..., x^m) \right| \leq c_j \tag{18}$$

The stopping time estimator τ_{MD}^{EST} satisfies:

$$\exp\left(\frac{-2\varepsilon^2}{c_{ub}(\boldsymbol{Y}^{\tau_{MD}^{EST}})} \right) \leq \frac{1-\delta}{2} \tag{19}$$

Therefore by the McDiarmid's inequality [11] the PAC property is satisfied. □

Theorem 2. ***(Delta method stopping time)** The output of stopping time, τ_{Delta}^{EST}, by the Delta method of the REST Algorithm satisfies the following PAC property whose parameters are specified a priori:*

$$\boldsymbol{P}\left(|g(\boldsymbol{Y}^m) - \mathbf{E}\left[g(\boldsymbol{Y}^m)\right]| \geq \varepsilon \right) \leq 1-\delta, \forall m \geq \tau_{Delta}^{EST} \tag{20}$$

where the notation is consistent with Theorem 1.

Proof. Define the random variable $Z = g(\mathbf{Y}^m) - \mathbf{E}\left[g(\mathbf{Y}^m)\right]$ and let $\sigma_{\boldsymbol{Y}^m}$ be the variance. Then by the Delta method [5, Prop. 8.14], the asymptotic distribution of Z converges to $\mathcal{N}(0, |g'(\mathbf{E}\left[\boldsymbol{Y}^m\right])|\, \sigma_{\boldsymbol{Y}^m})$. Since the stopping time estimate $\tau_{Delta}^{\mathrm{EST}}$ satisfies

$$\boldsymbol{P}(|Z| \leq \varepsilon) = 2\Phi\left(\frac{\varepsilon}{|g'(\mathbf{E}\left[\boldsymbol{Y}^m\right])|\, \sigma_{\boldsymbol{Y}^m}} \right) - 1 \geq \delta \tag{21}$$

where $\Phi(\cdot)$ is the cumulative density function of $\mathcal{N}(0,1)$, the proof is completed. □

The basic idea of REST is to find the shortest span of experiment to satisfy the criteria shown in (19) and (21) so that the PAC property holds. The parameters that appear in the bounds, such as c_i in (18) as well as $\sigma_{\boldsymbol{Y}^m}$ and $g'(\mathbf{E}\left[\boldsymbol{Y}^m\right])$ in (21), are estimated from a set of real data and virtual samples that are generated to be consistent with the real data according to $M_{sim}(\boldsymbol{\theta}, t)$, which accounts for the fluctuations in the stopping time estimates. In particular, in (18) and (21) we use $c_{ub}(\tau_{MD}^{EST})$, $\hat{\sigma}_{\boldsymbol{x}^m}$, and $\mathbf{E}[\boldsymbol{x}^m]$ estimated from the data set $\boldsymbol{x}^m$.

7. APPLICATION OF REST

To illustrate the application of REST algorithms, we use it to estimate stopping time for tasks including learning behavior parameters, predicting energy consumption, and classifying users into categories.

7.1 Behavior Parameter Learning

In the building energy management social game focused on lighting control, we have shown that our game-theoretic model can capture the dynamics and behavior of the agents in the non–cooperative game [12]. The parameter we are interested in learning is ξ_i, which captures the *tradeoff* between winning as a function of the incentive and lighting comfort.

Denote the lighting vote matrix $\boldsymbol{x}^t \in \mathbb{R}^{n \times t}$, where t is the span of the experiment. At time k, user i's vote is x_i^k. We synthetically generate the votes by drawing the parameters ξ_i for $i \in \{1, \ldots, n\}$ from $\mathcal{N}(\mu_i, \theta_i)$. In addition, user i votes the default lighting level with probability p_i^{def}, votes according to the Nash equilibrium strategy with probability p_i^{act}, and is absent with probability p_i^{abs} in which case the user does not vote.

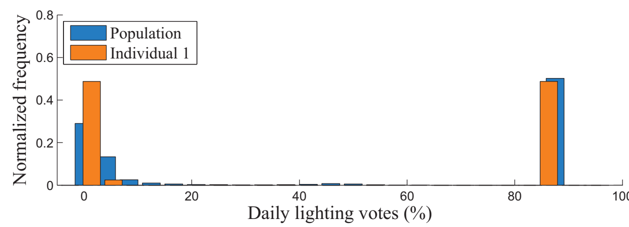

Figure 1: Distribution of lighting votes for the simulated population and individual user.

Figure 2 illustrates the lighting votes for a population of size $n = 100$ and user $i = 1$ and utility function parameter is ξ_1 which is drawn from the generative model (Section 4). Our task is to estimate the mean of the behavior parameter, μ_1, as appears in the individual randomness layer in the hierarchical Bayesian model in Section 4.2. The unbiased MLE estimator for μ_1 is the sample mean, $\hat{\mu}_1 = \frac{1}{t}\sum_{k=1}^{t} \xi_1^k$, where ξ_1^k is estimated by the method in Section 4.1 from the lighting votes up to time k.

The Cramér-Rao lower bounds derived in Section 5 directly apply to the inference of stopping time to satisfy the lower bounds on error. For the *Delta* method of REST, the sample standard deviation is applied to ξ_1^k for $k \in \{1, ..., t\}$. We define $g(\hat{\mu}_1) = \hat{\mu}_1$, and its derivative is unity. For the *McDiarmid* method, the bound

c_i in (18) is taken as $\frac{1}{t}(\max_k(\xi_1^k) - \min_k(\xi_1^k))$ so that $c_{ub}(\boldsymbol{x}^t) = \frac{1}{t}\big(\max_k(\xi_1^k) - \min_k(\xi_1^k)\big)^2$.

The Cramér-Rao lower bounds and the REST estimated stopping time are shown in Figure 3. The *Delta* method produces tighter bounds than the *McDiarmid* method. Both are lower bounded by the Bayesian and Hybrid Cramér-Rao bounds as expected.

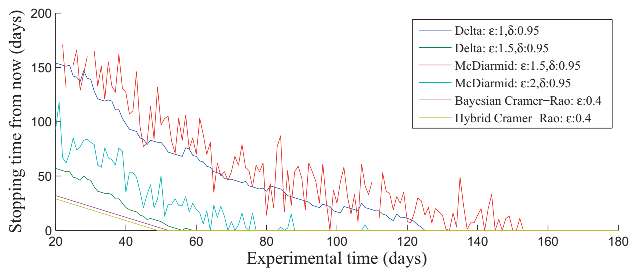

Figure 2: Estimated stopping time by REST for inference of μ_1 for user 1, and lower bounds given by Bayesian and Hybrid Cramér-Rao. For each day in the experiment (x-axis) we estimate the stopping time from that day forward (y-axis).

To evaluate the stopping time produced by REST, for a given period of experiment, we simulate the data 100 times and obtain the empirical distribution of the estimated $\hat{\mu}_1$ (see Figure 4). The estimation is concentrated within the bounds starting from day 100, which agrees with the *Delta* method. The *McDiarmid* method gives an overestimated result since the bounds are less tight.

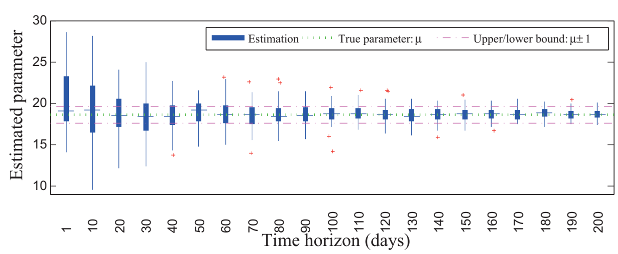

Figure 3: Distribution of $\hat{\mu}_1$ for 100 simulations of possible realizations when the size of dataset varies (x-axis). The edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points, and outliers are plotted individually. The true parameter and ± 1 upper/lower bounds are also indicated.

7.2 Energy Consumption Prediction

One objective of the social game is to estimate the effects of game dynamics on energy savings of the system. The energy consumption, E^{light}, is a monotone increasing function of the mean lighting vote and is given by

$$E^{\text{light}} = g\left(\frac{1}{|S \setminus S_{\text{absent}}|} \sum_{i \notin S_{\text{absent}}} x_i\right) \tag{22}$$

where S_{absent} and $S \setminus S_{\text{absent}}$ are the sets of absent and present (active and default) users respectively, and x_i is the individual lighting vote. The relationship between energy consumption and lighting level can be described by a piece-wise linear function, as demonstrated in Figure 5. There are two quantities that can be used to characterize the effect of the social game on energy savings, namely, the mean energy consumption, $\mu_E = \mathbf{E}\left[E^{\text{light}}\right]$ and the percentage of energy consumption below a certain threshold, $p_{E,\lambda} = \boldsymbol{P}\big(E^{\text{light}} < \lambda\big)$ where λ can be chosen, for instance, as a target level set by the building manager. The latter quantity is particularly useful for demand response programs; a building manager can determine the likelihood that in the next time period (e.g. 15 minutes or 1 hour) that the energy consumption will be below a certain threshold.

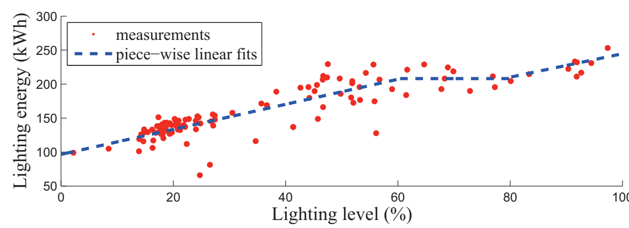

Figure 4: Lighting energy consumption (kWh) is fitted with a piece-wise linear function.

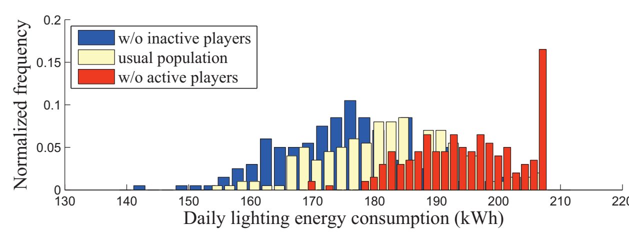

Figure 5: Distribution of daily lighting energy consumption (kWh) for different populations.

Compared to the fixed set point under the traditional control scheme, energy consumption in social games is random and depends on the competitive environment; the value depends on the set of players and also their strategies. The effects of the game are illustrated in Figure 6 where we show the normalized frequency of the energy consumption for the original population, the population with 10% of the active players (ranked by the active probability p^{act}) removed, and the population with 10% of the inactive players (ranked by the default probability p^{def}) removed. An interesting observation is that even though the average of the lighting votes is implemented, the influence of the introduction/removal of one active user, is larger than the weighting of its vote, $\frac{1}{|S \backslash S_{\text{absent}}|}$, since the dynamics of the game suggest that other users will change their votes to match this situation.

Given the lighting votes $\boldsymbol{x}^t$ up to time t and estimators $\hat{\mu}_E$ for the mean, the task is to estimate the stopping time of the experiment to satisfy the PAC property (15) with precision ε and probability bound $1-\delta$. For the *Delta* method, the target function is $g(\mathbf{E}[\boldsymbol{x}^t])$ as in (22). Its first derivative is given by the slope of the piecewise function. For the *McDiarmid* method, the upper bound c_i is obtained by $\frac{1}{t}(\max_i E_i^{\text{light}} - \min_i E_i^{\text{light}})$, and $c_{ub}(\boldsymbol{x}^t) = \frac{1}{t}\big(\max_i E_i^{\text{light}} - \min_i E_i^{\text{light}}\big)^2$.

Figure 7 illustrates REST estimated stopping time. For evaluation purpose, the empirical distribution of the estimates of mean energy are shown in Figure 8 thereby showing that the estimated stopping time achieves the desired result.

7.3 Classification of Player Category

Another potential goal of social game experiments is to classify agents into different categories based on their behavior. Classification of players is extremely important for incentive design because it provides a data–driven method for understanding the preference space of the competitive agents. Further, in the context of energy management, it is important for customer segmentation which can be used for targeting in demand response programs.

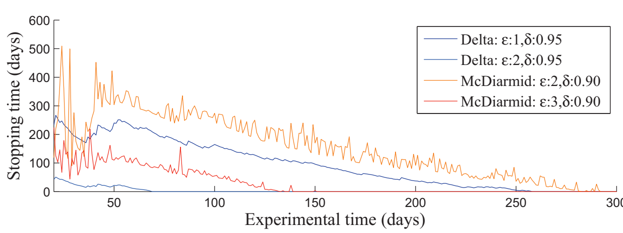

Figure 6: Estimated stopping time by REST for energy consumption prediction.

From the social game we previously conducted [12], we identified four categories that cover the profiles of all the users: those who care most about lighting comfort, those who care most about winning the lottery, those who are less extreme than the previous two, and those who are ambivalent about the game. The ambivalent type is easy to identify based on their votes or the number of times they log on to the website; therefore, we are most interested in classification of the first three

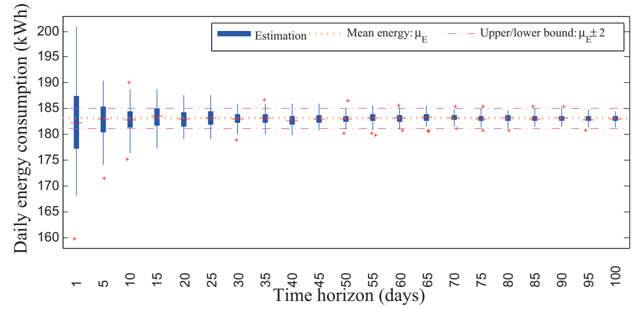

Figure 7: Distribution of estimated mean of energy consumption when the sample size varies, obtained from 100 simulations.

types of users into their categories.

Following the notation from Section 6, we denote the data for user i as $\boldsymbol{x}_i^t \in \mathcal{C}^t \subset \mathbb{R}^{1\times t}$ where t is the time. We denote the category $y_i \in \mathcal{Y}$ and the function that maps lighting votes to the parameter ξ_i^t as $\Upsilon : \boldsymbol{x}_i^t \mapsto \xi_i^t$. For each user we compare the empirical distribution $\boldsymbol{P}_i$ of $\xi_i^t = \Upsilon(\boldsymbol{x}_i^t)$ to the distribution $\boldsymbol{Q}_j$ of ξ_j for category j according to the Jensen-Shannon divergence, $JSD(\boldsymbol{P}||\boldsymbol{Q})$, given by

$$JSD(\boldsymbol{P}||\boldsymbol{Q}) = \frac{1}{2}D(\boldsymbol{P}||\boldsymbol{M}) + \frac{1}{2}D(\boldsymbol{Q}||\boldsymbol{M}) \qquad (23)$$

where $\boldsymbol{M} = \frac{1}{2}(\boldsymbol{P}+\boldsymbol{Q})$, and $D(\boldsymbol{P}||\boldsymbol{M}) = \sum_k \boldsymbol{P}(k) \ln \frac{\boldsymbol{P}(k)}{\boldsymbol{Q}(k)}$ is the Kullback-Leibler divergence and the summation is over the number of bins used in the computation of the empirical distribution.

Our classification rule is thus

$$h(\boldsymbol{x}_i^t) = \arg \min_{j\in\{1,\dots,4\}} JSD(\boldsymbol{P}_i||\boldsymbol{Q}_j). \qquad (24)$$

We apply the classification rule to the data of each user i in the population. The cost function $L(h;\boldsymbol{x}^t,\boldsymbol{y})$, where $\boldsymbol{x}^t = [\boldsymbol{x}_1^t, \dots, \boldsymbol{x}_n^t] \in \mathbf{R}^{t\times n}$ is the data matrix and $\boldsymbol{y}$ is the vector of categories for all the players, is given by

$$L(h;\boldsymbol{x}^t,\boldsymbol{y}) = \frac{1}{n}\sum_{i=1}^{n} 1(h(\boldsymbol{x}_i^t) \neq y_i); \qquad (25)$$

it is the proportion of misclassified users. Define the random function

$$f(h;\boldsymbol{x}^t,\boldsymbol{y}) = L(h;\boldsymbol{x}^t,\boldsymbol{y}) - \inf_{s\in(0,\infty)} \mathbf{E}\left[L(h;\boldsymbol{x}^s,\boldsymbol{y})\right]; \qquad (26)$$

it represents the deviation of the observed misclassification error from the best performance. If we assume that the task is κ_t*-learnable*, i.e $\mathbf{E}\left[f(h;\boldsymbol{x}^t,\boldsymbol{y})\right] \leq \kappa_t$ and the misclassification $\inf_{x\in(0,\infty)} \mathbf{E}\left[L(h;\boldsymbol{x}^s,\boldsymbol{y})\right]$ of the best performer is upper bounded by ϱ, then the task of estimating stopping time reduces to finding a constant c_i for all $i \in \{1,\dots,t\}$ such that when $\boldsymbol{x}^t$ and $\hat{\boldsymbol{x}}^t$ differ in only the i-th column, we have

$$\left|f(h;\boldsymbol{x}^t,\boldsymbol{y}) - f(h;\hat{\boldsymbol{x}}^t,\boldsymbol{y})\right| \le c_i. \tag{27}$$

By Algorithm 1 and Theorem 1 , the following property is satisfied:

$$\begin{aligned} P\big(L_t \ge \epsilon + \kappa_t + \varrho\big) &\le P\big(L_t \ge \epsilon + \kappa_t + \inf_t \mathbf{E}\left[L_t\right]\big) \\ &= P\big(f_t \ge \epsilon + \kappa_t\big) \\ &= P\big(f_t - \mathbf{E}\left[f_t\right] \ge \epsilon + \kappa_t - \mathbf{E}\left[f_t\right]\big) \\ &\le P\big(f_t - \mathbf{E}\left[f_t\right] \ge \epsilon\big) \le 1-\delta \end{aligned} \tag{28}$$

where we use the notation $L_t = L(h;\boldsymbol{x}^t,\boldsymbol{y})$ and $f_t = L_t - \inf_s \mathbf{E}\left[L_s\right]$. This property directly relates the misclassification error in practice to the bound obtained in Theorem 1, which implies that an algorithm with lower values for ϱ and κ_t will perform well in practice with high probability.

Using strategy profiles learned from experiments reported in [12], we simulate a large population interacting in a lighting control social game by making 33 replicates of each of the three prototypical users which correspond to three categories of interest. People in the same group share the same strategy profile, which is the empirical distribution of ξ_i, but on each day their specific parameter value, ξ_i, is random, independent and is revealed by the lighting votes through the mapping from the Nash equilibrium to parameter value (see Section 4.2).

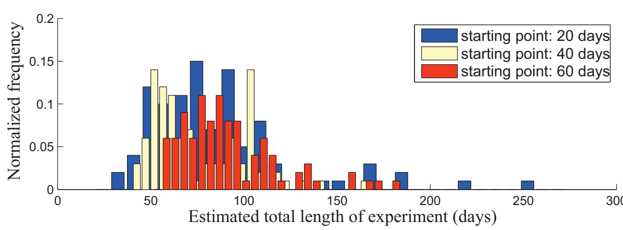

Figure 8: Distribution of estimated total time of experiment determined by REST for initial phases of length 20, 40, and 60 days. For example, for an initial phase of length 40**, we have collected data** $\boldsymbol{x}^{40}$ **and we apply REST to each of the 100 bootstrapped sample sets,** $\boldsymbol{x}^{40}_{boot,j}, j = 1,...,100$ **drawn from** $\boldsymbol{x}^{40}$ **to obtain 100 estimates of stopping time. This is added to the length of the initial phase to get the total experiment time. The three distributions are indistinguishable, which indicates that the total experiment time is stable over time.**

Starting from an initial phase, i.e. some period over which the experiment has been running, we provide estimation of the additional number of days required in order to correctly classify the users. The estimated total length of experiment is thus the sum of the length of the initial phase and additional days estimated by

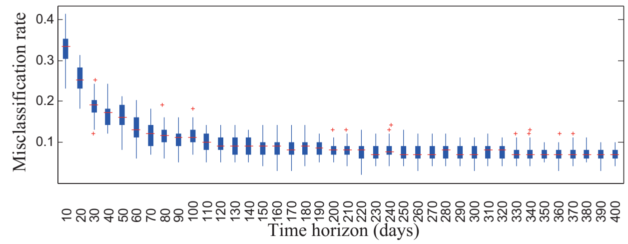

Figure 9: Empirical distribution of the misclassification error when the sample size varies (x-axis) for the user type classification problem.

REST. In Figure 9, we show the distribution of the estimated total experimentation time determined by REST for initial phases with 20, 40, and 60 days, and with parameters $\varepsilon = 0.1$, $\delta = 0.9$. The distributions are indistinguishable from each other, with mean of 89, 78, 94 days respectively, which indicate that sample size of around 90 days is sufficient to give us reliable estimates of each user's category.

To evaluate the usefulness of the stopping time recommended by REST, we obtain the distribution of misclassification error for a given sample size, as shown in Figure 10. The misclassification error generally decreases as we increase the size of sample set, and reaches a *plateau* around day 80. Although the stopping time provided by REST appears slightly overestimated, it captures the complexity of the problem and is useful in practice.

7.4 Summary of Results

In summary, if a problem can be formulated as a hierarchical Bayesian network and the target function and its derivative can be written out in an analytic form, as in the case of parameter inference and energy estimation, REST with the Delta method can be applied to provide tighter asymptotic estimation, and Cramér-Rao types of bounds can be derived for lower bounds. REST with *McDiarmid* method, nevertheless, can be applied to more general problems such as the classification of player category, as long as the upper bounds in (18) can be estimated. As in many CPS applications, the required length of experiment is a reflection of the complexity of the problem, and reliable estimation can be useful for project planning, experimental design, and budget management.

8. DISCUSSION

The length of social game type experiments based on shared resource usage in CPS with socioeconomic and sustainability concerns, such as intelligent office buildings, depends on resource constraints as well as system and environment complexity. We provided a method

(REST Algorithm 1) for estimating bounds on stopping time based on Cramér-Rao types of bounds, the Delta method, and concentration inequalities. We apply REST to problems of parameter inference, energy consumption estimation, and user type classification all of which are of interest in the operations of intelligent buildings. In particular, these problems and results can be used for increasing sustainability by improving automation and control, providing guarantees for demand response programs, and for developing game-theoretic behavioral models that can be used in the design of incentives used to induce energy efficient behavior.

There are a number of directions for future research including employing the REST in practice on our experimental platform. We are currently in the process of implementing such an experiment. Further, in previous work we designed incentives based on a game–theoretic behavioral model estimated from the data [12]. We are currently working on incorporating the design of incentives into the framework presented in this paper. One method of improving the behavioral model of competitive agents is to *explore* the input space by issuing incentives early in the game that maximize the information provided to the learner (or building manager). Another direction for future research is to apply the transfer learning framework to improve automation; we can transfer the behavioral model learned in the competitive environment (social game) to one in which the lighting settings are determined automatically.

9. REFERENCES

[1] P. L. Bartlett and S. Mendelson. Rademacher and gaussian complexities: Risk bounds and structural results. *The Journal of Machine Learning Research*, 3:463–482, 2003.

[2] J. Baxter. A Bayesian/Information Theoretic Model of Learning to Learn via Multiple Task Sampling. *Machine Learning*, 28(1):7–39, 1997.

[3] M. Desu. *Sample size methodology*. Elsevier, 2012.

[4] G. John and P. Langley. Static Versus Dynamic Sampling for Data Mining. In *Proc. of the 2nd Inter. Conf. on Knowledge Discovery and Data Mining*, pages 367–370. AAAI Press, 1996.

[5] R. Keener. *Theoretical Statistics: Topics for a Core Course*. Springer Tests in Statistics, 2010.

[6] D.-C. Li and Y.-S. Lin. Learning management knowledge for manufacturing systems in the early stages using time series data. *European Journal of Operational Research*, 184(1):169–184, 2008.

[7] M. Maasoumy, Q. Zhu, C. Li, F. Meggers, and A. Sangiovanni-Vincentelli. Co-design of control algorithm and embedded platform for HVAC systems. In *Proc. of the 4th Inter. Conf. on Cyber-Physical Systems*, 2012.

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

[9] R. C. Mittelhammer. *Mathematical statistics for economics and business*. Springer, 1996.

[10] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Trans. on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.

[11] M. Raginsky and I. Sason. Concentration of measure inequalities in information theory, communications and coding. *arXiv*, 1212.4663, 2012.

[12] L. Ratliff, M. Jin, I. C. Konstantakopoulos, C. Spanos, and S. S. Sastry. Social Game for Building Energy Efficiency: Incentive Design. In *Proc. of the 52nd Allerton Conf. on Communication, Control, and Computing*, 2014.

[13] L. J. Ratliff, S. A. Burden, and S. S. Sastry. On the Characterization of Local Nash Equilibria in Continuous Games. *arXiv*, 1411.2168, 2014.

[14] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520, 1965.

[15] M. J. Schervish. *Theory of statistics*. Springer, 1995.

[16] A. Swenson. Lighting in commercial buildings. Technical report, U.S. Energy Infomation Administration, 2009.

[17] H. L. Van Trees and K. L. Bell. *Bayesian bounds for parameter estimation and nonlinear filtering/tracking*. John Wiley & Sons, 2007.

[18] T. Zhu, Z. Huang, A. Sharma, J. Su, D. Irwin, A. Mishra, D. Menasche, and P. Shenoy. Sharing renewable energy in smart microgrids. In *Proc. of the 4th Inter. Conf. on Cyber-Physical Systems*, pages 219–228, 2013.