An Autonomous Demand Response Program in Smart Grid with Foresighted Users

Shahab Bahrami and Vincent W.S. Wong
Department of Electrical and Computer Engineering
The University of British Columbia, Vancouver, Canada
email: {bahramis, vincentw}@ece.ubc.ca

***Abstract*—In smart grid, demand response is a viable approach to motivate users towards shifting the demand during the peak load periods. Each user can also benefit by reducing its total cost. In most of the existing studies, the demand response program is modeled as a one-shot game among myopic users, who aim to minimize their cost in one period of time. In this paper, we show that the Nash equilibrium (NE) in the one-shot game can be inefficient in reducing the peak load demand and the users' cost. We address the inefficiency of the NE by modeling the demand response program as a repeated game. A grim-trigger strategy is proposed to determine the subgame perfect equilibrium. To address the issue of fairness, we partition the set of users into groups. In each time period, only one group of users are required to participate in the demand response program. Simulation results show that the proposed demand response repeated game can benefit both the users, by reducing their long-term cost, and the utility company, by reducing the peak-to-average ratio in the aggregate load demand.**

I. Introduction

Demand response programs aim to reduce the peak load by encouraging the users to either shift or reduce their demand voluntarily. In smart grid, each household is equipped with an energy management system (EMS) responsible for scheduling the energy usage of the users. The success of the demand response program depends in part on the users' active role in managing their energy usage.

Game theory is a viable approach to model the interaction among the participating users in the demand response program. In most of the existing studies (e.g., [1]–[4]), the demand response program is modeled as a one-shot game among *myopic* users. That is, the game models the interaction among the users in one period of time (e.g., in one day) and the users aim to minimize their cost in that period. In [1], the one-shot demand response game is used to jointly minimize the aggregate cost of the system and the peak-to-average ratio (PAR) in the aggregate load demand. However, the users' dissatisfaction from load shifting is not studied in this work. In [2], the dissatisfaction of the users is modeled by using a discomfort cost function for each user. However, the efficiency and fairness of the Nash equilibrium (NE) are not studied. In [3], a billing mechanism is proposed to minimize the aggregate cost of the system. The proposed billing mechanism is shown to be fair in charging the users.

In the aforementioned studies, the inefficiency of the NE in one-shot demand response game is not addressed. That is, there may exist an alternative strategy profile which has a lower cost for the users than in the NE. For example, in [4], it is shown that the cost of the users can be lower in the competitive market equilibrium as compared with the one-shot NE. One way to address the inefficiency of the NE is to play the game repeatedly [5]–[7]. In [5], the demand response program is formulated as a repeated game to minimize the aggregate cost of all users. In [6], the demand response program is modeled as a repeated game with critical peak pricing (CPP) scheme to minimize the aggregate cost of all users. In [7], the users' interaction in the demand response program is modeled as a repeated game with incomplete information. It is shown that there exists a unique Bayesian NE that minimizes the aggregate cost of all users. In the repeated games, the players are *foresighted*. That is, they have long-term plan to reduce their cost [8]. One advantage is that the set of Nash equilibria can include cooperative strategies based on predetermined agreement. Thus, it can lead to a lower cost for all users in long-term. The agreement among the users is a set of rules to modify the users' load pattern. A user does not follow the agreement can be punished according to some predetermined rules [9]. Another advantage is that, the repeated game model can capture the long-term interaction among the users as the utility companies typically design the demand response program for a long period of time, e.g., several months to several years.

In this paper, we design a repeated demand response game to motivate cooperation among the users. The contributions of this paper are as follows:

- We first show that the NE for a one-shot demand response game can be inefficient with real-time pricing (RTP) scheme. We formulate a repeated demand response game to address the inefficiency of the NE.
- We design a grim-trigger strategy to determine the subgame perfect equilibrium (SPE). To maintain fairness between users, the users are divided into groups in the proposed strategy. In each time period, only one group of users participate in the demand response program. Users in other groups consume electricity as they desire.
- Simulations are performed on a smart grid with 2100 users. When compared with the one-shot game, results show that the repeated demand response game can reduce the user's average cost up to 20% and the PAR up to 30%.

Our method can partly be compared with [6]. The problem addressed in this paper is different from [6] in two respects. First, the repeated demand response game in [6] is formulated for CPP scheme. The proposed approach cannot be used in an electricity market with RTP scheme. On the other hand, we formulate the demand response game with RTP scheme. Second, the demand response program in [6] is designed from the users' perspective to minimize the aggregate cost of all users. In contrast, our proposed demand response program is from the utility company's perspective to achieve a desired value for the PAR in the aggregate load demand.

The rest of this paper is organized as follows. In Section II, we determine the NE for the one-shot and repeated demand response games. In Section III, we propose a grim-trigger strategy to determine the SPE. In Section IV, we evaluate the performance of the proposed repeated demand response game through simulations. Conclusions are drawn in Section V.

II. System Model

Consider a utility company serving N residential users. Let $\mathcal{N} = \{1, \ldots, N\}$ denote the set of users. Each household is equipped with an EMS connected to the utility company via a two-way communication network. A day is divided into H time slots with equal length. Let $\mathcal{H} = \{1, \ldots, H\}$ denote the set of time slots. For user $i \in \mathcal{N}$, we assume that the load demand at time slot $h \in \mathcal{H}$ consists of uncontrollable load (i.e., base load) and controllable load demand. Lighting and TV are examples of uncontrollable loads. The washing machine and dish washer are examples of controllable loads. Let $b_{i,h}$ and $x_{i,h}$ denote the uncontrollable load and controllable load for user i at time slot h, respectively. Let vector $\mathbf{x}_i = (x_{i,1}, \ldots, x_{i,H})$ denote the controllable load profile for user i. Let vector $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_N)$ denote the controllable load profile for all users. Let $L_i = \sum_{h\in\mathcal{H}} (x_{i,h} + b_{i,h})$ denote the aggregate load demand for user i in one day. The aggregate load at time h is denoted by $l_h = \sum_{i\in\mathcal{N}} (x_{i,h} + b_{i,h})$. Let $p_h(l_h)$ denote the electricity price at time slot h. Similar to [4], the price function $p_h(l_h)$ at time h is modeled as an increasing and convex function of the aggregate load l_h. In the following subsection, the one-shot demand response game is modeled.

A. One-shot Demand Response Game

A demand response game can be formulated to determine the optimal daily controllable load profile for price anticipating users. The price function $p_h(l_h)$ at time slot h is informed by the utility company to the users through a communication infrastructure. The users are assumed to be myopic. They aim to minimize their short-term (typically, one day) cost. The interaction among the myopic users can be modeled as a one-shot game denoted by tuple $\mathcal{G}_1 (\mathcal{N}, \{\Omega_i\}_{i\in\mathcal{N}}, \{c_i\}_{i\in\mathcal{N}})$. The users are the players. Ω_i is the strategy space for user i. It indicates the possible vectors that the controllable load profile $\mathbf{x}_i$ can take. Hence,

$$\Omega_i = \left\{ \mathbf{x}_i \,\middle|\, \sum_{h\in\mathcal{H}} (x_{i,h} + b_{i,h}) = L_i,\ x_{i,h} \geq 0, \forall\, h \in \mathcal{H} \right\}, \quad (1)$$

where the aggregate load L_i and base load $b_{i,h}$, $h \in \mathcal{H}$ for user i are known *a priori*. The users only shift their load demand from one time slot to some other time slots. Thus, the daily aggregate load L_i for user i is unchanged. Let $\mathbf{x}_{-i} = (\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_N)$ denote the vector of controllable load profiles for all users except user i. The function c_i in game $\mathcal{G}_1$ is the cost function for user i. The cost for user i includes the energy payment and the discomfort cost. It is defined as follows:

$$c_i(\mathbf{x}_i, \mathbf{x}_{-i}) = \sum_{h\in\mathcal{H}} p_h(l_h)(x_{i,h} + b_{i,h}) + d_i(\mathbf{x}_i),\ i \in \mathcal{N}. \quad (2)$$

In (2), c_i is a function of both $\mathbf{x}_i$ and $\mathbf{x}_{-i}$ since the price $p_h(l_h)$ at time h is a function of the aggregate load l_h. The function $d_i(\mathbf{x}_i)$ is the discomfort cost for user i. It is used as a metric in monetary unit to express the dissatisfaction of the user with changing its load profile from the desired pattern to the scheduled pattern. Let $a_{i,h}$ denote the desired controllable load demand for user i at time slot h, which is equal to the controllable load demand before participating in the demand response program. Thus, $a_{i,h}$ is known *a priori* by user i. Similar to [6], the discomfort cost function is modeled as a weighted Euclidean distance between the scheduled and the desired controllable load profiles. That is

$$d_i(\mathbf{x}_i) = \sum_{h\in\mathcal{H}} \omega_{i,h} \left| x_{i,h} - a_{i,h} \right|, \qquad i \in \mathcal{N}, \quad (3)$$

where $\omega_{i,h}$, $i \in \mathcal{N}$, $h \in \mathcal{H}$ are the weighting coefficients measured in cents/kW to reflect the user's discomfort caused by changing the demand from its desired amount. A larger $\omega_{i,h}$ indicates that user i has a lower preference to change the load level at time slot h.

Let $\mathbf{x}^{\text{NE}} = (\mathbf{x}_1^{\text{NE}}, \ldots, \mathbf{x}_N^{\text{NE}})$ denote the controllable load profile for all users in the NE. The vector $\mathbf{x}_i^{\text{NE}} = (x_{i,1}^{\text{NE}}, \ldots, x_{i,H}^{\text{NE}})$ for user $i \in \mathcal{N}$ is the solution of the following optimization problem when the load profiles of other users' are fixed.

$$\begin{aligned} &\underset{\mathbf{x}_i}{\text{minimize}} \sum_{h\in\mathcal{H}} \Big(p_h(l_h)(x_{i,h} + b_{i,h}) + \omega_{i,h} \left| x_{i,h} - a_{i,h} \right| \Big) \\ &\text{subject to}\ \ \mathbf{x}_i \in \Omega_i. \end{aligned} \quad (4)$$

The price function $p_h(.)$ is increasing and convex by assumption. Thus, problem (4) is a convex optimization problem. An iterative algorithm can be used to determine the NE for the one-shot game [4], [10]. Let $l_h^{\text{NE}} = \sum_{i\in\mathcal{N}} \left(x_{i,h}^{\text{NE}} + b_{i,h} \right)$ denote the aggregate load demand at time slot h in the NE. Let $\overline{h}$ and $\underline{h}$ denote the time slots, in which l_h^{NE} becomes maximum and minimum, respectively. That is, $\overline{h} \in \underset{h\in\mathcal{H}}{\text{argmax}}\, l_h^{\text{NE}}$ and $\underline{h} \in \underset{h\in\mathcal{H}}{\text{argmin}}\, l_h^{\text{NE}}$. In Theorem 1, we show that there exists a load profile with lower cost for all users as compared with the cost in the NE.

Theorem 1: *If all users shift a sufficiently small amount of controllable load from time slot $\overline{h}$ to time slot $\underline{h}$, then all the users will have a lower daily cost than in the NE.*

Proof: Let ϵ denote a sufficiently small positive number. Let $\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})$ denote the change in the cost of user i, when all users shift ϵ amount of load from time slot $\overline{h}$ to time slot $\underline{h}$. We obtain

$$\begin{aligned}
&\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i}) = \\
&p_{\overline{h}}\big(l_{\overline{h}}^{\text{NE}} - N\epsilon\big)\big(x_{i,\overline{h}} - \epsilon + b_{i,\overline{h}}\big) - p_{\overline{h}}\left(l_{\overline{h}}^{\text{NE}}\right)\left(x_{i,\overline{h}} + b_{i,\overline{h}}\right) \\
&+ \omega_{i,\overline{h}}\Big(\big|x_{i,\overline{h}} - \epsilon - a_{i,\overline{h}}\big| - \big|x_{i,\overline{h}} - a_{i,\overline{h}}\big|\Big) \\
&+ p_{\underline{h}}\big(l_{\underline{h}}^{\text{NE}} + N\epsilon\big)\big(x_{i,\underline{h}} + \epsilon + b_{i,\underline{h}}\big) - p_{\underline{h}}\big(l_{\underline{h}}^{\text{NE}}\big)\left(x_{i,\underline{h}} + b_{i,\underline{h}}\right) \\
&+ \omega_{i,\underline{h}}\Big(\big|x_{i,\underline{h}} + \epsilon - a_{i,\underline{h}}\big| - \big|x_{i,\underline{h}} - a_{i,\underline{h}}\big|\Big). \qquad (5)
\end{aligned}$$

At time slot $\overline{h}$, we have $x_{i,\overline{h}} \leq a_{i,\overline{h}}$. At time slot $\underline{h}$, we have $x_{i,\underline{h}} \geq a_{i,\underline{h}}$. Thus, we can determine the sign of the absolute values in (5). By rearranging equation (5), we have

$$\begin{aligned}
\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i}) = &\Big(p_{\overline{h}}\left(l_{\overline{h}}^{\text{NE}} - N\epsilon\right) - p_{\overline{h}}\left(l_{\overline{h}}^{\text{NE}}\right)\Big)\big(x_{i,\overline{h}} + b_{i,\overline{h}}\big) \\
&- \epsilon\, p_{\overline{h}}\left(l_{\overline{h}}^{\text{NE}} - N\epsilon\right) + \epsilon\, \omega_{i,\overline{h}} \\
&+ \Big(p_{\underline{h}}\big(l_{\underline{h}}^{\text{NE}} + N\epsilon\big) - p_{\underline{h}}\big(l_{\underline{h}}^{\text{NE}}\big)\Big)\big(x_{i,\underline{h}} + b_{i,\underline{h}}\big) \\
&- \epsilon\, p_{\underline{h}}\big(l_{\underline{h}}^{\text{NE}} + N\epsilon\big) + \epsilon\, \omega_{i,\underline{h}}. \qquad (6)
\end{aligned}$$

By taking the derivative of $\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})$ with respect to ϵ, we obtain

$$\begin{aligned}
\frac{d\,\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})}{d\,\epsilon} = &-N \frac{d\,p_{\overline{h}}\left(l_h\right)}{d\,l_h}\Big|_{l_h = l_{\overline{h}}^{\text{NE}}} \left(x_{i,\overline{h}} + b_{i,\overline{h}}\right) - p_{\overline{h}}\left(l_{\overline{h}}^{\text{NE}}\right) \\
&+ N \frac{d\,p_{\underline{h}}\left(l_h\right)}{d\,l_h}\Big|_{l_h = l_{\underline{h}}^{\text{NE}}} \left(x_{i,\underline{h}} + b_{i,\underline{h}}\right) + p_{\underline{h}}\left(l_{\underline{h}}^{\text{NE}}\right) \\
&+ \omega_{i,\overline{h}} + \omega_{i,\underline{h}}. \qquad (7)
\end{aligned}$$

In the NE, no user has an incentive to unilaterally deviate from its strategy. Hence, the change in the cost of user i is zero when it shifts a sufficiently small amount of load from time $\overline{h}$ to $\underline{h}$. Similar to the aforementioned approach, we obtain

$$\begin{aligned}
&- \frac{d\,p_{\overline{h}}\big(l_h\big)}{d\,l_h}\Big|_{l_h = l_{\overline{h}}^{\text{NE}}} \big(x_{i,\overline{h}} + b_{i,\overline{h}}\big) - p_{\overline{h}}\big(l_{\overline{h}}^{\text{NE}}\big) + \omega_{i,\overline{h}} \\
&+ \frac{d\,p_{\underline{h}}\big(l_h\big)}{d\,l_h}\Big|_{l_h = l_{\underline{h}}^{\text{NE}}} \big(x_{i,\underline{h}} + b_{i,\underline{h}}\big) + p_{\underline{h}}\big(l_{\underline{h}}^{\text{NE}}\big) + \omega_{i,\underline{h}} = 0. \qquad (8)
\end{aligned}$$

Substituting (8) into (7), we obtain

$$\begin{aligned}
\frac{d\,\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})}{d\,\epsilon} =&(N-1)\Big(-\frac{d\,p_{\overline{h}}\big(l_h\big)}{d\,l_h}\Big|_{l_h = l_{\overline{h}}^{\text{NE}}} \big(x_{i,\overline{h}} + b_{i,\overline{h}}\big) \\
&+ \frac{d\,p_{\underline{h}}\big(l_h\big)}{d\,l_h}\Big|_{l_h = l_{\underline{h}}^{\text{NE}}} \big(x_{i,\underline{h}} + b_{i,\underline{h}}\big)\Big). \qquad (9)
\end{aligned}$$

Since the price function $p_h(.)$ is increasing and convex, its derivative is increasing. For $l_{\overline{h}}^{\text{NE}} \geq l_{\underline{h}}^{\text{NE}}$, we have $\frac{d\,p_{\overline{h}}(l_h)}{d\,l_h}\Big|_{l_h = l_{\overline{h}}^{\text{NE}}} \geq \frac{d\,p_{\underline{h}}(l_h)}{d\,l_h}\Big|_{l_h = l_{\underline{h}}^{\text{NE}}}$. Besides, the peak load is greater than the off-peak load. Thus, $\big(x_{i,\overline{h}} + b_{i,\overline{h}}\big) \geq \big(x_{i,\underline{h}} + b_{i,\underline{h}}\big)$. Hence, $\frac{d\,\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})}{d\,\epsilon}$ is non-positive for $N \geq 2$. Moreover, $\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})$ approaches zero, when ϵ approaches zero. Thus, $\Delta c_i(\mathbf{x}_i, \mathbf{x}_{-i})$ is non-positive for $N \geq 2$ and sufficiently small ϵ. The proof is completed. ■

Theorem 1 states that if users cooperate with each other, then they can modify their load pattern to incur a lower cost than in the one-shot NE. If we model the demand response as a repeated game, then we can determine the corresponding Nash equilibria with cooperative strategy. In the next subsection, we present the repeated demand response game.

B. Repeated Demand Response Game

In the repeated demand response game, the users are foresighted and aim to minimize their cost in long-term. To determine the cost, we first introduce the stage game, the game history and the strategy of a user. Let $\mathcal{T} = \{1, 2, \ldots, T\}$ denote the set of T time periods that the repeated game is being played. A time period can represent one day. The stage game is the game played at each time period [8]. In our model, the one-shot demand response game is the stage game. The superscript t in an arbitrary parameter y^t indicates the value of parameter y in time period t. Let Ω_i^t denote the strategy space for user i in time period $t \in \mathcal{T}$. For $t = 1$, Ω_i^t corresponds to the strategy space in the one-shot demand response game defined in (1). Ω_i^t for time period $t > 1$ can be defined similarly. Let Ω^{t-1}, $t > 1$, denote all possible $t-1$ histories of the strategy profiles. Thus, we can express Ω^{t-1} as

$$\Omega^{t-1} = \prod_{i \in \mathcal{N}} \Omega_i^1 \times \cdots \times \prod_{i \in \mathcal{N}} \Omega_i^{t-1}. \qquad (10)$$

Let the sequence $\mathbf{X}_i \equiv \{\mathbf{x}_i^t\}_{t=1}^T$ denote the strategy profile for user i in the repeated game with T time periods, where $\mathbf{x}_i^t : \Omega^{t-1} \to \Omega_i^t$ is the controllable load profile in time period t. We assume that the game has perfect information. That is, the entire past history is known to all users. The repeated game with imperfect information is beyond the scope of this paper. The interested reader is referred to [11] for more details.

In general, the users have no common knowledge on the ending time of the game. This game is called an infinitely repeated game. Each user considers a probability $q \in [0, 1]$ for the game to be continued in the next period. This probability can be reflected in a discount factor δ. Besides, the users may value a dollar received today more than a dollar received later. Let s denote the interest rate. The present value of the money is inversely proportional to the interest rate. That is, if s is high, then the value of the money will depreciate quickly [12]. Thus, δ can be expressed as $\delta = \frac{q}{1+s}$.

The cost function in time period t is denoted by $c_i^t\big(\mathbf{x}_i^t, \mathbf{x}_{-i}^t\big)$. For time period $t = 1$, $c_i^t(.)$ is defined in (2). Let vector $\mathbf{X}_{-i} = (\mathbf{X}_1, \ldots, \mathbf{X}_{i-1}, \mathbf{X}_{i+1}, \ldots, \mathbf{X}_N)$ denote the strategy profile for all users except user i. The discounted cost for user i is denoted by $c_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i})$. It is defined as

$$c_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i}) = \sum_{t \in \mathcal{T}} \delta^{t-1} c_i^t\big(\mathbf{x}_i^t, \mathbf{x}_{-i}^t\big), \qquad i \in \mathcal{N}. \qquad (11)$$

After T periods, we define an average discounted cost for user i as

$$\overline{c}_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i}) = \frac{1-\delta}{1-\delta^T} c_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i}). \qquad (12)$$

That is, if user i incurs cost $\overline{c}_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i})$ at all time periods $1 \leq t \leq T$, then its discounted cost will be equal to $c_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i})$. For an infinitely repeated game, T approaches ∞. Thus, $\overline{c}_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i}) = (1-\delta) c_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i})$. Let $\mathbf{X}^{\text{NE}} = (\mathbf{X}_1^{\text{NE}}, \ldots, \mathbf{X}_N^{\text{NE}})$ denote the strategy profile for all users in the NE. In the infinitely repeated game, user i aims to determine the strategy profile $\mathbf{X}_i^{\text{NE}}$ that minimizes its discounted cost (11). Since $\overline{c}_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i}) = (1-\delta) c_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i})$ and δ is given, minimizing the discounted cost is equivalent to minimizing the average discounted cost. The optimal strategy profile $\mathbf{X}_i^{\text{NE}}$, $i \in \mathcal{N}$ is the solution of the following optimization problem when the strategy profiles of other users are fixed.

$$\begin{aligned} &\underset{\mathbf{X}_i}{\text{minimize}}\ \overline{c}_i^{\text{disc}}(\mathbf{X}_i, \mathbf{X}_{-i}) \\ &\text{subject to}\ \ \mathbf{x}_i^t \in \Omega_i^t, \qquad\qquad t \in \mathcal{T}. \end{aligned} \tag{13}$$

We use the SPE solution concept in the proposed repeated game. A subgame is a subset of the game that contains all possible actions starting from time period $t \in \mathcal{T}$. A strategy profile is an SPE if and only if it is an NE in every subgame [9]. For an infinitely repeated game, we can construct a particular strategy, namely the grim-trigger strategy, that can lead to the SPE [9]. In the grim-trigger strategy, the players start by cooperating and continue to cooperate as long as everyone has cooperated in the past. If one player does not cooperate, then the others will punish this player either forever or for a specific period of time. In the next section, we propose a grim-trigger strategy for the repeated demand response game.

III. Equilibrium Strategy Design

We propose a demand response program based on the grim-trigger strategy that can lead to the SPE. In the proposed grim-trigger strategy, the users modify their controllable load pattern to reduce the peak load demand in a cooperative manner. To address the issue of fairness between users, the utility company divides the users into groups. One group of users participate in the demand response program in each time period and cooperate to reduce the peak load. Moreover, each group of users participate in the program periodically, e.g., once a week or every other day. The non-participating groups of users can consume electricity according to their desired pattern.

The proposed grim-trigger strategy is summarized in Algorithm 1. In Line 1, the time period t is set to 1. The utility company selects an integer number $M > 1$ and partitions the set of the users into M non-overlapping and non-empty groups denoted by $S_1, \ldots, S_M$. To maintain fairness between users, the utility company can divide the users to M equal-sized groups randomly. In Line 2, the utility company informs each EMS about the user's group number and the number of users' groups M. Users in group S_r, $1 \leq r \leq M$, participate in the demand response program periodically in time periods $t = kM + r$, where k is a non-negative integer. For example, for $M = 7$, the users are divided into seven groups. If $t = 1$ is Monday, then the users in group S_2 participate in the demand response program every Tuesday. The loop in Lines 3 to 7 describes the action of the users in each time period. In Line 4, the participating user i sets $\omega_{i,h}^t = 0$ for all time slots $h \in \mathcal{H}$. Otherwise, it sets $\omega_{i,h}^t$ to a sufficiently large number (e.g., $\omega_{i,h}^t = 10^6$) for all time slots $h \in \mathcal{H}$. In Line 5, all users play the one-shot demand response game with updated $\omega_{i,h}^t$ in time period t. An iterative algorithm similar to one given in [4, Theorem 7] can be used to determine the load profile of user i in the one-shot NE in time period t. When participating user i sets $\omega_{i,h}^t = 0$, then the user modifies its controllable load profile without considering its discomfort cost. Therefore, the participating user shifts its controllable load demand from peak time slots to reduce the peak load demand in a cooperative manner. On the other hand, a non-participating user consumes electricity according to its desired pattern to avoid a high discomfort cost. In Line 6, the time period is updated at the end of time period t.

Algorithm 1 Executed by EMS $i \in \mathcal{N}$.

1: Initialization: set $t := 1$. The utility company selects M and divides the users to M non-overlapping and non-empty subsets $S_1, \ldots, S_M$.
2: Receive the the number M and the group number from the utility company.
3: **Repeat**
4: Set $\omega_{i,h}^t := 0$ for all $h \in \mathcal{H}$ if user i is participating. Otherwise, set $\omega_{i,h}^t := 10^6$ for all $h \in \mathcal{H}$.
5: Use an iterative algorithm to determine the NE strategy for updated $\omega_{i,h}^t$.
6: $t := t + 1$ at the end of time period t.
7: **Until** at least one user deviates from the agreement in Lines 4 and 5.
8: Play the one-shot demand response game at remaining time periods.

The EMS of each participating user's household can collect data about the controllable load demand of the user at the beginning of each time period and communicate to the utility company by using a two-way communication network. Then, the utility company can communicate the aggregate controllable load demand of the participating users to all users. The participating users cooperate to shift their load from peak time slots as much as possible. Thus, their aggregate controllable load pattern can be predicted by all users when no participating user deviates from the proposed cooperative strategy. In Line 7, the users can monitor the aggregate controllable load pattern of the participating users at the end of each time period to detect any deviation. It is possible that some participating users deviate out of necessity, e.g., they need to increase the load for some periods not anticipated before. Hence, the users tolerate deviation from the cooperative strategy to some degree. For example, they can consider a threshold level for the aggregate controllable load in peak time slots. At the end of each time period, if the aggregate controllable load level is greater than the threshold value, then deviation is detected. In Line 8, if a deviation is detected, then user i selects the one-shot NE strategy. This punishment strategy can motivate the participating users to cooperate since the cost of the deviating user is higher in the one-shot game than the cost in the repeated game.

From the Folk theorem [13], if the discounted factor δ is sufficiently close to 1, then the proposed grim-trigger strategy leads to the SPE. We denote the threshold value of the discounting factor by δ_{th}. That is, for $\delta \geq \delta_{\text{th}}$, no user has incentive to deviate unilaterally. In Algorithm 1, the users will punish the deviating user forever. However, it is also possible for the users to forgive after passing some periods of time and return to the proposed cooperative strategy. We will further demonstrate the effect of forgiveness in the next section.

IV. Performance Evaluation

In this section, we evaluate the performance of the proposed repeated demand response game with RTP scheme. Consider a smart grid with a utility company serving $N = 2100$ users. A time period represents one day. The day is divided to $H = 24$ one-hour time slots. For RTP scheme, we use the electricity market database [14] for Ontario, Canada, to approximate the price as a linear function $p_h^t(l_h^t) = \lambda_h^t l_h^t$, where $\lambda_h^t = 4 \times 10^{-4}\,\text{cents/kW}$ and l_h^t is measured in kilowatts. Similar to [6], the weighting coefficients $\omega_{i,h}^t$ for the users' discomfort function are randomly selected from interval $[10\,\text{cents/kW}, 20\,\text{cents/kW}]$. Unless stated otherwise, the users consider probability $q = 0.95$ for the stage game to be continued in the next period. The Canada daily interest rate $s = 0.0029\%$ (1% for one year) is used to calculate δ [15]. Thus, $\delta = 0.949$.

For the sake of simplicity, we model the load profile of the users for one day, and we assume that it is repeated periodically for the other days. The model can be generalized by considering the load profile for any periodic time horizon such as one week. To model the load pattern for a user, we use a load pattern for about 5 million households from Ontario, Canada power grid database in March 25, 2015 [14]. We compute the average load for each household. Then, the load demand of a user at each time slot is selected from a normal distribution with the mean value of the computed average load for each household and the standard deviation of 0.3 kW. The amount of controllable load $x_{i,h}$ at time slot h is chosen randomly between 20% to 40% of its total load at each time slot [6]. The generation capacity of the utility company is sufficient to meet the aggregate load demand. The desired controllable load demand is the load without demand response. Fig. $1(a)$ shows the aggregate and the base load profiles without demand response program.

First, we consider the one-shot demand response game to model the interaction among the myopic users. All the users modify their load profiles in a non-cooperative manner. We have determined the NE using an iterative algorithm proposed in [4, Theorem 7]. The aggregate load profile in the NE is shown in Fig. $1(b)$. In the one-shot NE, the peak load is only reduced by about 4%. In fact, the myopic users prefer not to modify their load pattern to avoid a high discomfort cost.

In the second scenario, we consider a repeated demand response game with the proposed grim-trigger strategy for $M = 2$, 5, and 7. Fig. $1(b)$ shows that, for $M = 2$, 5, and 7, the peak load demand is reduced by about 20%, 15%, and 10%, respectively. Hence, comparing with the one-shot demand response game, the repeated game approach has better performance in reducing the peak load demand, although only $\frac{1}{M}$ of the users participate in each day. In fact, the participating users cooperate with each other and modify their load pattern

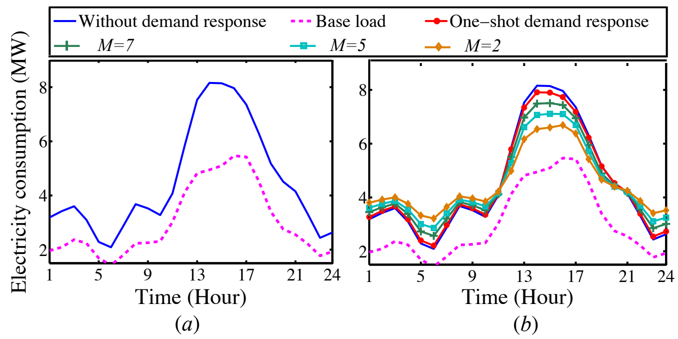

Fig. 1. (a) The aggregate and base load profiles without demand response. (b) The aggregate load profile for the one-shot game and repeated demand response game with number of groups $M = 2$, 5, and 7.

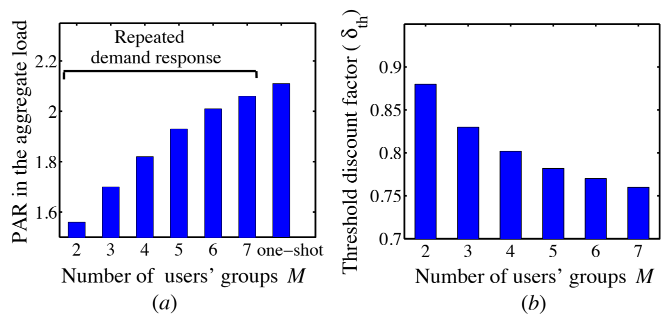

Fig. 2. (a) The PAR for the one-shot and repeated demand response games with different values of M. (b) The δ_{th} for different values of M.

to reduce the peak load demand. Fig. $2(a)$ shows the PAR in the aggregate load demand for the one-shot game and the repeated game with $M = 2$ to $M = 7$. For the higher values of M, the number of participating users in each time period is lower. Thus, the PAR increases gradually when M increases. The PAR is maximum for the one-shot game since the users modify their load pattern in a non-cooperative manner. Fig. $2(b)$ shows δ_{th} for $M = 2$ to 7. When M increases, each user participates more frequently and incurs a higher discomfort cost, thus δ_{th} increases gradually. When the utility company selects the number of groups M, it has to consider a trade-off between the value of the PAR and the value of the δ_{th}. For example, the utility company can select $M = 2$ when $\delta = 0.949$, since for $M = 2$, we have $\delta_{\text{th}} = 0.88$ and $\delta > \delta_{th}$.

Fig. 3 shows the convergence of the average discounted cost for user 1 in group S_1 for the one-shot game and the repeated demand response game with $M = 2$, 5, and 7. It can be observed that in the repeated demand response game, the user's average discounted cost converges to a value from 8% (for $M = 7$) to 20% (for $M = 2$) lower than the user's average discounted cost in the one-shot game. In fact, in the repeated demand response game, the participating users ignore their discomfort cost and shift all their controllable load demands from the peak hours to the off-peak hours. On the other hand, in the one-shot demand response game, the users consider a trade-off between their payment and their discomfort cost. Therefore, they modify their load pattern as long as their discomfort cost is not high.

Finally, we show that selecting the NE strategy as a

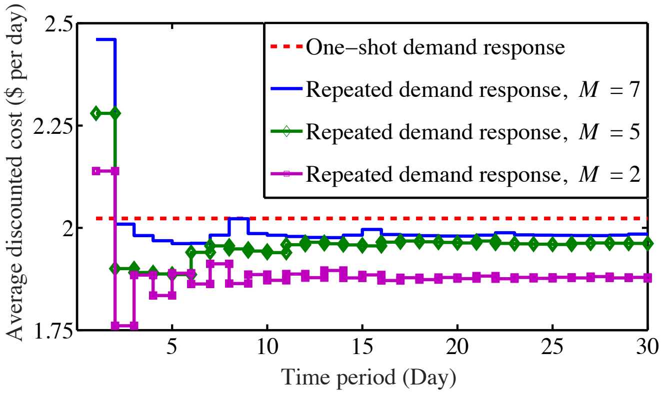

Fig. 3. The average discounted cost for user 1 in the one-shot game and the repeated demand response game with number of groups $M = 2$, 5, and 7.

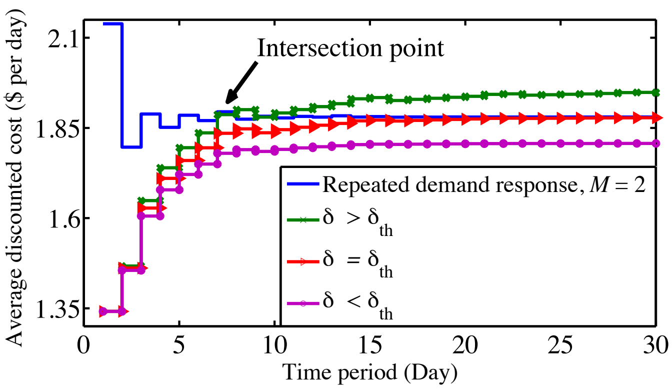

Fig. 4. The average discounted cost for user 1 in group S_1 in the repeated demand response game with number of groups $M = 2$ when user 1 deviates from the equilibrium strategy.

punishment can motivate the participating users to cooperate if $\delta \geq \delta_{\text{th}}$. Fig. 4 shows the convergence of the average discounted cost for user 1 in group S_1 for $M = 2$ with $\delta_{\text{th}} = 0.88$ and $\delta = 0.95, 0.88$, and 0.8. We assume that any deviation can be detected by all users. If $\delta \geq \delta_{\text{th}}$ and user 1 deviates in the first day, then it has a lower cost in that day since its discomfort cost becomes zero. However, other users can punish user 1 by choosing their one-shot NE strategy in the following days. The discounted cost of user 1 increases gradually and converges to a value higher than or equal to the average discounted cost when it cooperates. Hence, deviation is not profitable for user 1. Similarly, we can show that deviation for all users is not profitable when $\delta \geq \delta_{\text{th}}$. Therefore, the proposed grim-trigger strategy is the SPE. Moreover, other users do not need to punish user 1 forever. The intersection point of the curves for the discounted cost with and without deviation shows that after one week, the average discounted cost of the deviating user 1 is equal to the average discounted cost if user 1 cooperates. Hence, in the eighth day, other users can forgive and cooperate again since deviation is not profitable for user 1. Fig. 4 also shows that if $\delta < \delta_{\text{th}}$ and user 1 deviates, then its average discounted cost converges to a lower value than its cost without deviation. Hence, user 1 can benefit from deviation. Since there exist at least one user that benefits from deviation, the users are not in the SPE. In summary, for $\delta < \delta_{\text{th}}$, at least one user has incentive to deviate. For $\delta \geq \delta_{\text{th}}$, the proposed grim-trigger strategy with or without forgiveness can lead to the SPE.

V. Conclusion

In this paper, we addressed the inefficiency of the NE in the one-shot demand response game by modeling the demand response program as an infinitely repeated game. A grim-trigger strategy is proposed to determine the SPE. To maintain fairness between the users in the proposed strategy, the set of users is partitioned into groups. In each time period, only one group of users participate in the program. Besides, the threat of being punished for deviation motivated users to reduce the peak load demand in a cooperative manner. Simulation results showed that the proposed repeated demand response game can benefit both the users, by reducing their discounted cost, and the utility company, by reducing the PAR in the aggregate load. We also showed that the proposed grim-trigger strategy can lead to the SPE even if other users forgive the deviating user after passing some periods of time. For future work, we plan to extend the model by considering the load uncertainty and time of use for the household's electrical appliances.

References

[1] A.H. Mohsenian-Rad, V.W.S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia, "Autonomous demand side management based on game-theoretic energy consumption scheduling for the future smart grid," *IEEE Trans. on Smart Grid*, vol. 1, no. 3, pp. 320–331, Dec. 2010.

[2] R. Arai, K. Yamamoto, and M. Morikura, "Differential game-theoretic framework for a demand-side energy management system," in *Proc. of IEEE SmartGridComm*, Vancouver, Canada, Oct. 2013.

[3] H. Chen, Y. Li, R. Louie, and B. Vucetic, "Autonomous demand side management based on energy consumption scheduling and instantaneous load billing: An aggregative game approach," *IEEE Trans. on Smart Grid*, vol. 5, no. 4, pp. 1744–1754, July 2014.

[4] L. Chen, N. Li, S.H. Low, and J.C. Doyle, "Two market models for demand response in power networks," in *Proc. of IEEE SmartGridComm*, Gaithersburg, MD, Oct. 2010.

[5] K. Ma, G. Hu, and C.J. Spanos, "Cooperative demand response using repeated game for price anticipating buildings in smart grid," in *Proc. of Int'l Conf. on Control, Automation, Robotics and Vision*, Marina Bay Sands, Singapore, Dec. 2014.

[6] L. Song, Y. Xiao, and M. van der Schaar, "Demand side management in smart grids using a repeated game framework," *IEEE Selected Areas in Comm.*, vol. 32, no. 7, pp. 1412–1424, July 2014.

[7] C. Eksin, H. Delic, and A. Ribeiro, "Distributed demand side management of heterogeneous rational consumers in smart grids with renewable sources," in *Proc. of IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Florence, Italy, May 2014.

[8] D. Fudenberg and J. Tirole, *Game Theroy*. MA: The MIT Press, 1991.

[9] G.J. Mailath and L. Samuelson, *Repeated Games and Reputations Long-Run Relationships*. NY: Oxford University Press, 2006.

[10] P. Chakraborty and P.P. Khargonekar, "A demand response game and its robust price of anarchy," in *in Proc. of IEEE SmartGridComm*, Venice, Italy, Nov. 2014.

[11] R. Umann and M. Maschler, *Repeated Games with Incomplete Information*. MA: The MIT Press, 1995.

[12] J.M. Perloff, *Microeconomics*, 7th ed. NJ: Prentice Hall, 2014.

[13] Q. Wen, "The Folk theorem for repeated games with complete information," *Econometrica*, vol. 62, no. 4, pp. 949–954, July 1994.

[14] Independent electricty system operator (IESO). [Online]. Available: http://www.ieso.ca

[15] Trading economics website. [Online]. Available: http://www.tradingeconomics.com/canada/interest-rate