

Data Center Demand Response Pricing Using Inverse Optimization*

Athanasios Tsiligkaridis
Boston University
Boston, MA, USA
atsili@bu.edu

Ioannis Ch. Paschalidis
Boston University
Boston, MA, USA
yannisp@bu.edu

Ayse Coskun
Boston University
Boston, MA, USA
acoskun@bu.edu

ABSTRACT

In Demand Response (DR), consumers regulate their power based on requests from an energy supplier. Data Centers (DC) are among the promising candidates to perform DR to help stabilize the power grid due to their flexibility and controllability. In this work, we present a novel framework for offering incentives to DCs so they can dynamically adjust their electricity consumption and provide DR to the grid. Coordination between an Independent System Operator (ISO) and DCs is done through pricing where the ISO computes optimal prices which elicit desired responses from the DCs. We model DCs using realistic cost functions based on Quality of Service (QoS) requirements of the DC workloads and present an inverse optimization approach to cost function parameter estimation for precise and efficient pricing along with simulation results that highlight the strength of our approach.

CCS CONCEPTS

• **Hardware** → Enterprise level and data centers power issues;

KEYWORDS

Demand response, inverse optimization, pricing, data centers.

ACM Reference Format:

Athanasios Tsiligkaridis, Ioannis Ch. Paschalidis, and Ayse Coskun. 2019. Data Center Demand Response Pricing Using Inverse Optimization. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems (e-Energy '19)*, June 25–28, 2019, Phoenix, AZ, USA. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3307772.3330166>

1 INTRODUCTION

DR is an emerging solution to the issue of power instability; a DR program is utilized as a means of electricity demand management for service stability.

A DR program requires a market model that encourages user participation. One such model is a price-based approach [1, 2] where a set of users respond to a given price from an ISO by solving a cost minimization problem. Previous pricing approaches do not consider a means of customer cost function estimation which can provide

*Research partially supported by the NSF under grants DMS-1664644 and CNS-1645681, and by the ONR under MURI grant N00014-16-1-2832.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
e-Energy '19, June 25–28, 2019, Phoenix, AZ, USA
© 2019 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6671-7/19/06.
<https://doi.org/10.1145/3307772.3330166>

key information to the supplier; estimated cost functions can be used to make informed decisions in setting the best DR-inducing price by allowing anticipation of customer responses.

We consider DCs as a DR resource due to their flexibility and ease of monitoring and adjustment; if provided the necessary policies for smart pricing, DCs can deliver flexibility to the grid. DR programs can harness DC flexibility to allow for the decrease of energy usage costs by providing financial benefits from participation [3].

In this work, we focus on a real-time pricing market model where DCs are given price incentives to dynamically reduce their power consumption during times of power supply deficit. We present a novel model of a DR mechanism utilizing DCs and we construct realistic DC cost functions that capture the provisioning of QoS by the DC to the jobs it services. Finally, we present a means of reconstructing each cost function by using historical data observations of the interactions between an ISO and the DCs through an inverse optimization framework along with simulation results that highlight the strength of our approach.

2 PROBLEM FORMULATION

We focus on a DC DR setting which involves an ISO and a number S of DCs. The ISO provisions resources (energy) to each DC; but, it may experience an energy supply deficit and become unable to service all DCs. To handle this, the ISO aims to induce each DC to adjust its energy consumption such that the overall energy consumption by DCs matches available supply. For DC i , $i \in \{1, \dots, S\}$, to use less energy, it must reduce the number of servers that it has running from a nominal amount \bar{N}_i to a new smaller amount N_i .

We model a DC as a G/G/1 queuing system where jobs arrive and are carried out by its running servers. The queueing model assumes a single server representing a pooled resource consisting of all available servers in the DC; we scale the service rate linearly with the number of available servers N_i for DC i .

Next, we model the QoS of a given DC using the probability that the queue length L in the system is no less than some value U . For DC i , we use results from [4] to approximate this probability as: $P[L \geq U] \sim e^{-\theta_i^* U}$, where $\theta_i^* > 0$ depends on the arrival process A and service process B , respectively. Specifically, θ_i^* is the largest positive root to: $\Lambda_A(\theta_i^*) + \Lambda_B(-\theta_i^*) = 0$, where Λ_A and Λ_B are the corresponding limiting log-moment generating functions.

We want $P[L \geq U]$ to be small; thus, θ_i should be large. We use this parameter to design DC i 's cost function such that it yields a small cost for a large θ_i and vice versa. We model DC i 's cost function, $C_i(\theta_i)$, as a *convex, non-increasing* function.

2.1 Forward problem

DC i solves a cost minimization problem when responding to a price p . We form the general forward optimization problem as:

$$\begin{aligned}
& \underset{N_i, \theta_i}{\text{minimize}} && -p\alpha_i(\bar{N}_i - N_i) + C_i(\theta_i) \\
& \text{subject to} && \Lambda_A(\theta_i) + \Lambda_{D_i}(-\theta_i) = 0, \\
& && \bar{N}_i(1 - \zeta) \leq N_i \leq \bar{N}_i,
\end{aligned} \tag{1}$$

where α_i is the amount of power use per server, $\zeta \in [0, 1]$ is used to form bounds on the server counts, and A and $D_i = N_i B$ are arrival and scaled service processes, respectively. The intuition behind this formulation is that DC i wants to minimize its cost when responding to the price p . DC i will be paid $p\alpha_i(\bar{N}_i - N_i)$ by the ISO for reducing its running server count from \bar{N}_i to N_i ; it will also be subject to a QoS cost $C_i(\theta_i)$ for running N_i servers.

For our simulations, we assume that A and D_i are i.i.d. Gaussian processes. This causes the first constraint (QoS) in (1) to be linear in N_i and θ_i ; since it will be satisfied with equality, the forward problem can be expressed in terms of one variable, N_i , where the cost function becomes: $C_i(\theta_i) = C_i(gN_i + h)$, $g \in \mathbb{R}$, $h \in \mathbb{R}$.

2.2 Inverse problem and finding the best price

To find the best DR-inducing price p^* , the ISO minimizes a social cost metric $G(p)$:

$$G(p) = q(T - \sum_{i=1}^S \alpha_i(\bar{N}_i - N_i^*(p))) + \sum_{i=1}^S C_i(gN_i^*(p) + h). \tag{2}$$

The first term represents the ISO's penalty for not achieving a target load reduction T over all DCs; each DC's individual reduction is $\alpha_i(\bar{N}_i - N_i^*(p))$ where $N_i^*(p)$ is the optimal solution of (1). The second term represents social welfare since the ISO wants to encourage participation in the DR program.

Finding p^* requires knowledge of DC i 's *convex, non-increasing* cost function, $C_i(\theta_i) \forall i$. We express this cost function as the inner product of a vector of scaling parameters \mathbf{k}_i and a vector of *convex, non-increasing* constituent functions $\mathbf{c}(\theta_i) = (c_1(\theta_i), \dots, c_V(\theta_i))$: $C_i(\theta_i) = \mathbf{k}_i^T \mathbf{c}(\theta_i)$. We assume that the ISO knows the constituent functions but does not know the coefficients \mathbf{k}_i , $\forall i$.

Through the use of the inverse variational inequality framework [5], the ISO can use data observations to learn the scaling coefficients. For a specific price p_j , each DC i provides a response N_i^j ; we use M observations of $\langle p_j, N_i^j \rangle$, $i \in \{1, \dots, S\}$, $j \in \{1, \dots, M\}$ as data observations of the DR system interactions.

For observed price p_j , define the objective function of the forward problem (1) as: $\phi_i(N_i^j, \mathbf{k}_i, \boldsymbol{\gamma}_{i,j}) = -p_j \alpha_i(\bar{N}_i - N_i^j) + \mathbf{k}_i^T \mathbf{c}(gN_i^j + h)$, where $\boldsymbol{\gamma}_{i,j} = (p_j, \alpha_i, \bar{N}_i, g, h)$. With this, we set up the inverse problem by applying Theorem 3 from [5] as:

$$\begin{aligned}
& \min_{\mathbf{k}_1, \dots, \mathbf{k}_S, \mathbf{y}, \boldsymbol{\epsilon}} \|\boldsymbol{\epsilon}\|_\infty && (3) \\
& \text{s.t. } -\mathbf{y}_i^j \leq 0, \forall i, \forall j, \\
& -\mathbf{y}_i^j - \frac{\partial}{\partial N_i^j} \phi_i(N_i^j, \mathbf{k}_i, \boldsymbol{\gamma}_{i,j}) + \bar{N}_i(1 - \zeta) \leq 0, \forall i, \forall j, \\
& \sum_{i=1}^S \left[\frac{\partial}{\partial N_i^j} \phi_i(N_i^j, \mathbf{k}_i, \boldsymbol{\gamma}_{i,j}) N_i^j + \bar{N}_i y_i^j \right] - \epsilon_j \leq 0, \forall j, \\
& \frac{\partial}{\partial N_i^{j_m}} \phi_i(N_i^{j_m}, \mathbf{k}_i, \boldsymbol{\gamma}_{i,j_m}) = 0, \forall i,
\end{aligned}$$

where \mathbf{y} is a dual variable, $\boldsymbol{\epsilon}$ is a vector, j_m represents a median price index, and $\|\boldsymbol{\epsilon}\|_\infty = \max_j |\epsilon_j|$ is the infinity norm of $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_M)$.

The inverse is a minimization of an error norm such that we have an approximate solution to a variational inequality problem; this is a Linear Program and can be solved efficiently.

3 SIMULATION RESULTS

We aim to show that the ISO can determine the best DR-inducing price by first using data observations to learn all DCs' cost functions and then minimizing (2). We consider a system with 5 DCs along with a quadratic penalty function: $q(\cdot) = \eta(\cdot)^2/2$. We let $C_i(\theta_i) = k_i e^{-\beta_i \theta_i}$ where $k_i, \forall i$ are unknowns to the ISO.

We consider three cases: 1) Optimal (ground truth), where the ISO has full prior knowledge of each DC's cost function, 2) Inverse (our approach), where the ISO uses data to estimate all cost functions by solving (3), and 3) Averaging [1], where random realizations of cost coefficients are obtained for different prices to estimate social cost values over which we average and search for the price that yields the minimum average social cost.

Figure 1 shows derived prices for all cases. For our approach, we consider noiseless and noisy data. Perfect observations allow us to find the correct DR-inducing prices (those in case one) while noisy data yields prices close to the optimum. For the averaging approach, we display multiple curves that differ based on how we obtain realizations of the cost coefficients; we sample from Gaussian distributions that are either centered on the correct cost parameter or not with different variances. When we sample coefficients from the distributions centered on the true parameter value, we get close to obtaining the optimum DR price as we observe more data (purple and green curves); but, this situation is not practical since we initially have no knowledge about the true parameters. With wrong centering, correct prices cannot be obtained.

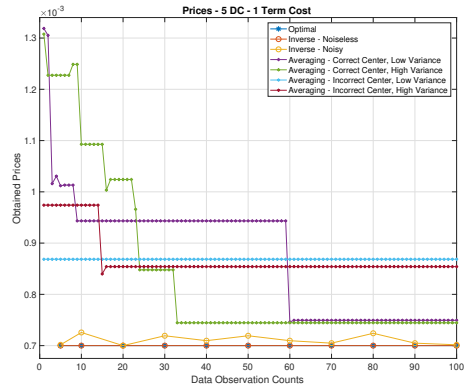


Figure 1: Optimal prices as a function of data observation counts. Our inverse approach successfully finds the best DR-inducing price.

4 CONCLUSION

In this work, we proposed a novel framework for DC DR pricing, modeled DC cost functions using QoS requirements, presented a data-driven inverse optimization approach for parametric estimation, and displayed simulation results showcasing the success of our inverse approach in yielding correct prices.

As future work, we can: 1) incorporate power network constraints for added realism to our ISO-DC system, and 2) consider ISO broadcasting of different prices.

REFERENCES

- [1] Z. Liu, I. Liu, S. Low, and A. Wierman. Pricing data center demand response. In *Proceedings of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, Austin, TX, USA, June 2014.
- [2] L. Chen, N. Li, S. H. Low, and J. C. Doyle. Two market models for demand response in power networks. In *IEEE International Conference on Smart Grid Communications*, Gaithersburg, MD, USA, October 2010.
- [3] A. Wierman, Z. Liu, I. Liu, and H. Mohsenian-Rad. Opportunities and challenges for data center demand response. In *International Green Computing Conference*, Dallas, TX, USA, November 2014.
- [4] Ioannis Ch. Paschalidis. *Large Deviations in High Speed Communication Networks*. PhD thesis, Massachusetts Institute of Technology, 1996.
- [5] D. Bertsimas, V. Gupta, and I. Ch. Paschalidis. Data-Driven Estimation in Equilibrium Using Inverse Optimization. *Mathematical Programming, Series A*, 153(2):595–633, 2015.