

Modeling and Design of Real-Time Pricing Systems Based on Markov Decision Processes

Koichi Kobayashi^{1*}, Ichiro Maruta², Kazunori Sakurama³, Shun-ichi Azuma²

¹School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa, Japan

²Graduate School of Informatics, Kyoto University, Kyoto, Japan

³Graduate School of Engineering, Tottori University, Tottori, Japan

Email: *k-kobaya@jaist.ac.jp

Received 2 April 2014; revised 2 May 2014; accepted 9 May 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A real-time pricing system of electricity is a system that charges different electricity prices for different hours of the day and for different days, and is effective for reducing the peak and flattening the load curve. In this paper, using a Markov decision process (MDP), we propose a modeling method and an optimal control method for real-time pricing systems. First, the outline of real-time pricing systems is explained. Next, a model of a set of customers is derived as a multi-agent MDP. Furthermore, the optimal control problem is formulated, and is reduced to a quadratic programming problem. Finally, a numerical simulation is presented.

Keywords

Markov Decision Process, Optimal Control, Real-Time Pricing System

1. Introduction

In recent years, there has been growing interest in energy and the environment. For problems on energy and the environment such as energy saving, several approaches have been studied (see, e.g., [1] [2]). In this paper, we focus on real-time pricing systems of electricity. A real-time pricing system of electricity is a system that charges different electricity prices for different hours of the day and for different days, and is effective for reducing the peak and flattening the load curve (see, e.g., [3]-[6]). In general, a real-time pricing system consists of one controller deciding the price at each time and multiple electric customers such as commercial facilities

*Corresponding author.

and homes. If electricity conservation is needed, then the price is set to a high value. Since the economic load becomes high, customers conserve electricity. Thus, electricity conservation is achieved. In the existing methods, the price at each time is given by a simple function with respect to power consumptions and voltage deviations and so on (see, e.g., [6]). In order to realize more precisely pricing, it is necessary to use a mathematical model of customers.

In this paper, using a Markov decision process (MDP), we propose a mathematical model of real-time pricing systems. Since in many cases, the status of electricity conservation of customers is discrete and stochastic, it is appropriate to use an MDP. Then, a set of electricity customers is modeled by a multi-agent MDP. Furthermore, we consider the finite-time optimal control problem. By appropriately setting the cost function, it is achieved that customers conserve electricity actively. This problem can be used for the model predictive control method, which is a control method that the finite-time optimal control problem is solved at each time. In addition, the finite-time optimal control problem can be reduced to a quadratic programming problem. The proposed approach provides us with a basic of real-time pricing systems.

This paper is organized as follows. In Section 2, the outline of real-time pricing systems is explained. In Section 3, a model of electricity customers is derived. In Section 4, the optimal control problem is formulated, and its solution method is derived. In Section 5, a numerical simulation is shown. In Section 6, we conclude this paper.

Notation: Let \mathcal{R} denote the set of real numbers. Let I_n , $0_{m \times n}$ denote the $n \times n$ identity matrix, the $m \times n$ zero matrix, respectively. For simplicity, we sometimes use the symbol 0 instead of $0_{m \times n}$, and the symbol I instead of I_n . For two events A, B , let $E[A|B]$ denote the conditional expected value of A under the event B .

2. Outline of Real-Time Pricing Systems

In this section, we explain the outline of real-time pricing systems studied in this paper.

Figure 1 shows an illustration of real-time pricing systems studied in this paper. This system consists of one controller and multiple electric customers such as commercial facilities and homes. For an electric customer, we suppose that each customer can monitor the status of electricity conservation of other customers. In other words, the status of some customer affects that of other customers. For example, in commercial facilities, we suppose that the status of rival commercial facilities can be checked by lighting, Blog, Twitter, and so on. Depending on power consumption, *i.e.*, the status of electricity conservation, the controller determines the price at each time. If electricity conservation is needed, then the price is set to a high value. Since the economic load becomes high, customers conserve electricity. Thus, electricity conservation is achieved.

In this paper, the status of electricity conservation of each customer is modeled by a Markov decision process (MDP). Then a set of customers is modeled by a multi-agent MDP (MA-MDP). Furthermore, by using the obtained MA-MDP model, we consider the optimal control problem and its solution method.

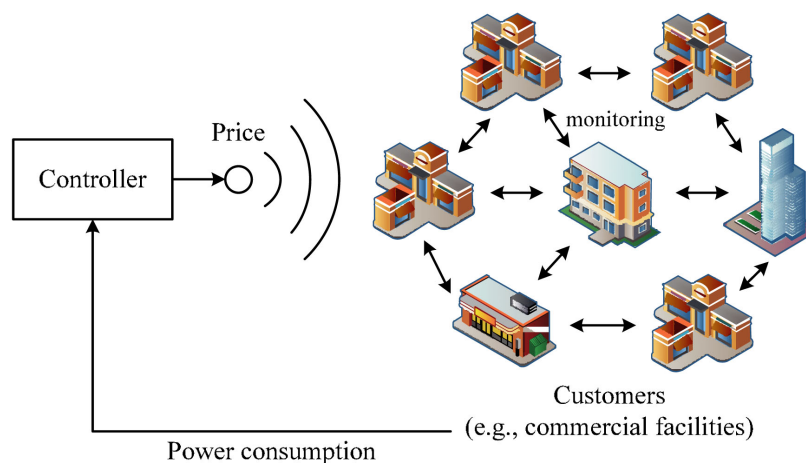


Figure 1. Illustration of real-time pricing systems.

3. Model of Customers

First, consider modeling the dynamics of each customer by a one-dimensional MDP. The value of the state x is randomly chosen among the finite set $\{0, 1, \dots, n-1\}$. The element of $\{0, 1, \dots, n-1\}$ expresses the status of electricity conservation, and “0” implies the status that a customer conserves electricity maximally, “ $n-1$ ” implies the status that a customer does not conserve electricity. Then the MDP of a customer is given by

$$\pi(t+1) = P(u(t))\pi(t) \quad (1)$$

where $u(t) \in \mathcal{R}$ is the control input, and corresponds to the price. The vector

$\pi(t) = [\pi_0(t) \ \pi_1(t) \ \dots \ \pi_{n-1}(t)]^T \in [0, 1]^n$ denotes the probability distribution, that is, $\pi_i(t) \in [0, 1]$ implies the probability that the state is i at time t . In addition, the initial probability distribution must satisfy the following condition:

$$\sum_{i=0}^{n-1} \pi_i(0) = 1.$$

The transition probability matrix $P(u(t))$ is given by

$$P(u(t)) = \begin{bmatrix} a_{11} + b_1 u(t) & \dots & a_{1n} + b_1 u(t) \\ a_{21} + b_2 u(t) & \dots & a_{2n} + b_2 u(t) \\ \vdots & \ddots & \vdots \\ a_{n1} + b_n u(t) & \dots & a_{nn} + b_n u(t) \end{bmatrix}.$$

The control input is determined under the condition for each element:

$$0 \leq a_{ij} + b_i u(t) \leq 1 \quad (2)$$

and the condition for each column:

$$\sum_{i=1}^n a_{ij} + b_i u(t) = 1. \quad (3)$$

Next, consider modeling the dynamics of a set of customers by an MA-MDP. The number of customers is given by q . For the customer i , the state is given by $x^i \in \{0, 1, \dots, n-1\}$, and from (1), the MDP model is given by

$$\pi^i(t+1) = P^i(u^i(t))\pi^i(t).$$

Then, we suppose that the MA-MDP model expressing the dynamics of a set of customers is given by

$$\begin{bmatrix} \pi^1(t+1) \\ \pi^2(t+1) \\ \vdots \\ \pi^q(t+1) \end{bmatrix} = \begin{bmatrix} \lambda_{11} P^1(u^1(t)) & \lambda_{12} I & \dots & \lambda_{1q} I \\ \lambda_{21} I & \lambda_{22} P^2(u^2(t)) & \dots & \lambda_{2q} I \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{q1} I & \lambda_{q2} I & \dots & \lambda_{qq} P^q(u^q(t)) \end{bmatrix} \begin{bmatrix} \pi^1(t) \\ \pi^2(t) \\ \vdots \\ \pi^q(t) \end{bmatrix}, \quad (4)$$

where λ_{ij} expresses the effect of couplings between customers, and is a constant satisfying the following condition:

$$\sum_{j=1}^q \lambda_{ij} = 1, i = 1, 2, \dots, q. \quad (5)$$

For simplicity of discussion, coupling terms are given by $\lambda_{ij} I$, but may be replaced with matrices satisfying some condition corresponding to (5).

4. Optimal Control

4.1. Problem Formulation

Consider the following problem.

Problem 1. Suppose that for the MA-MDP model (4) expressing the dynamics of customers, the initial state $x^i(0) = x_0^i$, $i = 1, 2, \dots, q$, the desired state x_d , and the prediction horizon N are given. Then, find a control input sequence $u^i(t)$, $i = 1, 2, \dots, q$, $t = 0, 1, \dots, N-1$ minimizing the following cost function

$$J = \sum_{t=0}^N \sum_{i=1}^q E \left[Q_i \left(x^i(t) - x_d \right)^2 + R_i \left(u^i(t) \right)^2 \mid x^i(0) = x_0^i \right] \quad (6)$$

subject to the following constraint:

$$f \left(u^1(t), u^2(t), \dots, u^q(t) \right) \leq M, \quad (7)$$

where $f(\cdot)$ is a given linear function, M is a given vector. Q_i, R_i are given weights.

Hereafter, for simplicity of notation, the condition $x^i(0) = x_0^i$ in the cost function (6) is omitted.

By using the constraint (7), the input constraint such as $u_{\min} \leq u^i(t) \leq u_{\max}$ can be imposed. In addition, by adjusting Q_i, R_i , several specifications such that the state $x^i(t)$ must converges to the neighborhood of the desired state x_d can be considered.

4.2. Solution Method

We derive a solution method for Problem 1. First, consider the MDP model (1). The MDP model is a class of nonlinear systems. However, in this case, it can be transformed into a linear system. The MDP model (1) can be rewritten as

$$\pi(t+1) = A\pi(t) + \begin{bmatrix} b_1 u(t) (\pi_1(t) + \dots + \pi_n(t)) \\ b_2 u(t) (\pi_1(t) + \dots + \pi_n(t)) \\ \vdots \\ b_n u(t) (\pi_1(t) + \dots + \pi_n(t)) \end{bmatrix},$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

By the property of the probability distribution, the relation $\pi_1(t) + \dots + \pi_n(t) = 1$ holds. From this fact, the MDP model (1) can be equivalently transformed into the following linear system:

$$\pi(t+1) = A\pi(t) + Bu(t), \quad (8)$$

where $B = [b_1 \ b_2 \ \dots \ b_n]^T$.

Next, by using the linear system (8), consider representing the MA-MDP model (4) as a linear system. The linear system for the customer i is denoted by

$$\pi^i(t+1) = A^i \pi^i(t) + B^i u^i(t).$$

Then, the MA-MDP model (4) can be equivalently transformed into the following linear system:

$$\begin{bmatrix} \pi^1(t+1) \\ \vdots \\ \pi^q(t+1) \end{bmatrix} = \bar{A} \begin{bmatrix} \pi^1(t) \\ \vdots \\ \pi^q(t) \end{bmatrix} + \bar{B} \begin{bmatrix} u^1(t) \\ \vdots \\ u^q(t) \end{bmatrix}, \quad (9)$$

where

$$\bar{A} = \begin{bmatrix} \lambda_{q1}A^1 & \cdots & \lambda_{q1}I \\ \vdots & \ddots & \vdots \\ \lambda_{q1}I & \cdots & \lambda_{qq}A^q \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \lambda_{q1}B^1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{qq}B^q \end{bmatrix}.$$

Finally, consider the cost function (6). Define

$$L_1 := [0 \ 1 \ \cdots \ n-1],$$

$$L_2 := [0^2 \ 1^2 \ \cdots \ (n-1)^2].$$

Then we can obtain

$$E[x^i(t)] = L_1\pi^i(t),$$

$$E[(x^i(t))^2] = L_2\pi^i(t).$$

Therefore, the cost function (6) can be rewritten as

$$\begin{aligned} J &= \sum_{t=0}^N \sum_{i=1}^q E \left[Q_i (x^i(t) - x_d)^2 \right] + R_i (u^i(t))^2 \\ &= \sum_{t=0}^N \sum_{i=1}^q Q_i E \left[(x^i(t))^2 \right] - 2Q_i x_d E[x^i(t)] + Q_i x_d^2 + R_i (u^i(t))^2 \\ &= \sum_{t=0}^N \sum_{i=1}^q Q_i (L_2 - 2x_d L_1) \pi^i(t) + R_i (u^i(t))^2 + Q_i x_d^2. \end{aligned} \quad (10)$$

From the above discussion, Problem 1 is equivalent to the following problem.

Problem 2

$$\begin{aligned} &\text{find } u^i(t), i = 1, 2, \dots, q \\ &\quad t = 0, 1, \dots, N-1, \\ &\text{min Cost function (10),} \\ &\text{subject to System (9),} \\ &\quad \text{Inequality constraint (2),(7),} \\ &\quad \text{Equality constraint (3).} \end{aligned}$$

Problem 2 is reduced to a quadratic programming (QP) problem, and can be solved by a suitable solver such as MATLAB and IBM ILOG CPLEX. In addition, if $R_i = 0$, then Problem 2 is reduced to a linear programming (LP) problem (we remark that $Q_i x_d^2$ in the cost function (10) is a constant). See [7] for further details.

5. Numerical Example

Since it is difficult to use data in real systems, we present an artificial example. The state is chosen among the finite set $\{0, 1, 2, 3\}$. The number of consumers is given by $q = 5$. The coefficient matrices A^i, B^i in the linear system for the consumer i are given by

$$A_1 = \begin{bmatrix} 0.3 & 0.4 & 0.3 & 0.5 \\ 0.4 & 0.3 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 \\ 0.1 & 0.1 & 0.3 & 0.1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.5 & 0.6 & 0.5 & 0.3 \\ 0.2 & 0.2 & 0.1 & 0.4 \\ 0.2 & 0 & 0.2 & 0.3 \\ 0.1 & 0.2 & 0.2 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0.6 & 0.3 & 0.1 & 0.3 \\ 0.1 & 0.3 & 0.4 & 0.2 \\ 0.2 & 0.3 & 0.4 & 0.1 \\ 0.1 & 0.1 & 0.1 & 0.4 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} 0.7 & 0.5 & 0.4 & 0.5 \\ 0.1 & 0.2 & 0.1 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.2 \\ 0.1 & 0.2 & 0.4 & 0.1 \end{bmatrix}, A_5 = \begin{bmatrix} 0.3 & 0.5 & 0.2 & 0.4 \\ 0.1 & 0.2 & 0.3 & 0.4 \\ 0.3 & 0 & 0.3 & 0.1 \\ 0.3 & 0.3 & 0.2 & 0.1 \end{bmatrix}, B_i = \begin{bmatrix} 0.6 \\ -0.2 \\ -0.2 \\ -0.2 \end{bmatrix}.$$

The parameters λ_{ij} in (9) are given by

$$\begin{bmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} & \lambda_{14} & \lambda_{15} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} & \lambda_{24} & \lambda_{25} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & \lambda_{34} & \lambda_{35} \\ \lambda_{41} & \lambda_{42} & \lambda_{43} & \lambda_{44} & \lambda_{45} \\ \lambda_{51} & \lambda_{52} & \lambda_{53} & \lambda_{54} & \lambda_{55} \end{bmatrix} = \begin{bmatrix} 0.7 & 0.3 & 0 & 0 & 0 \\ 0.2 & 0.6 & 0.2 & 0 & 0 \\ 0 & 0.1 & 0.8 & 0.1 & 0 \\ 0 & 0 & 0.2 & 0.7 & 0.1 \\ 0 & 0 & 0 & 0.1 & 0.9 \end{bmatrix}.$$

The parameters N , x_d , Q_i , and R_i are given by $N=10$, $x_d=0$, $Q_i=1$, and $R_i=0$, respectively. From $R_i=0$, Problem 2 is reduced to an LP problem. The initial state is given by $x^i(0)=[0 \ 0 \ 0 \ 1]^T$. In addition, the input constraint $0 \leq u^i(t) \leq 1$ is imposed.

In this numerical example, we consider the following two cases:

- The price for each customer is the same (i.e., $u^1(t) = u^2(t) = \dots = u^q(t)$ holds).
- The price for each customer is different.

Case (i) is the conventional case in real-time pricing systems. In Case (ii), we suppose that the difference in the price is covered by using local concurrencies such as the Eco-point point system [8]. The Eco-money system [9] in Japan were introduced to stimulate the economy and raise awareness of global warming. In the Eco-point system, many points, which correspond to money in a local concurrency, are given for the products that are effective from the viewpoints of electricity conservation and the environment. Such a system for energy management systems has been discussed in [10].

Next, we present the computational results. First, the computational result in Case (i) is explained. Figures 2-6 show the probability distribution for each customer. From these figures, we see that $\pi_0^i(t)$ increases and $\pi_3^i(t)$ decreases. Thus, the state converges to 0, which corresponds to the status that a customer conserves electricity maximally, with a certain probability. Furthermore, the optimal value of the cost function is 97.5902, and the optimal control input is derived as

$$u(0) = 0.5, u(1) = 0.65, u(2) = 0.6637, u(3) = 0.5664, u(4) = 0.5168, \\ u(5) = 0.5151, u(6) = 0.5184, u(7) = 0.5189, u(8) = 0.5192, u(9) = 0.5192.$$

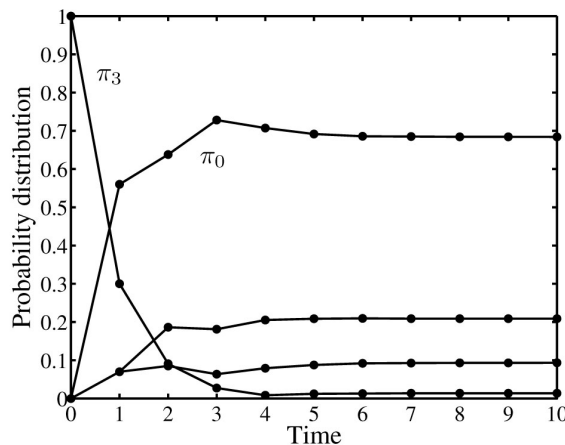


Figure 2. $\pi^1(t)$ in Case (i).

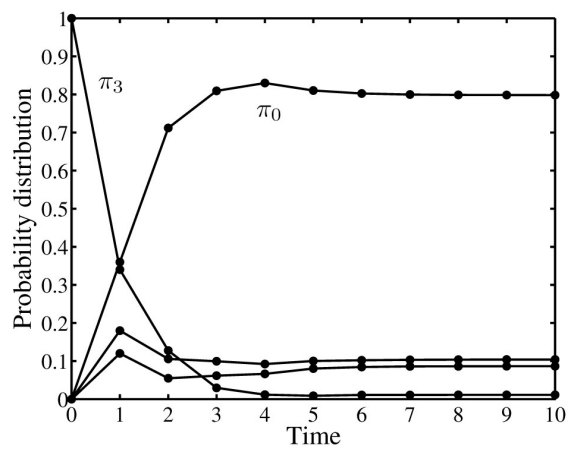


Figure 3. $\pi^2(t)$ in Case (i).

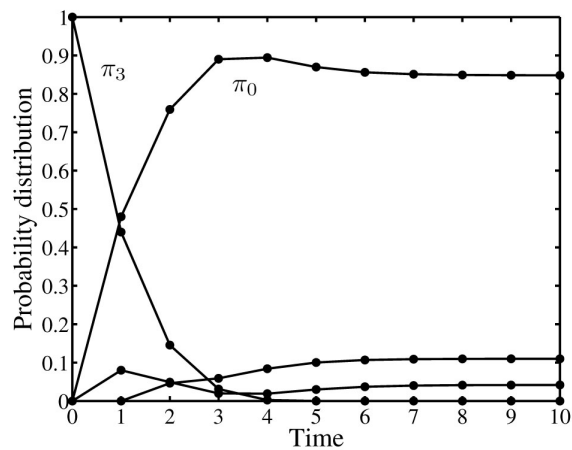


Figure 4. $\pi^3(t)$ in Case (i).

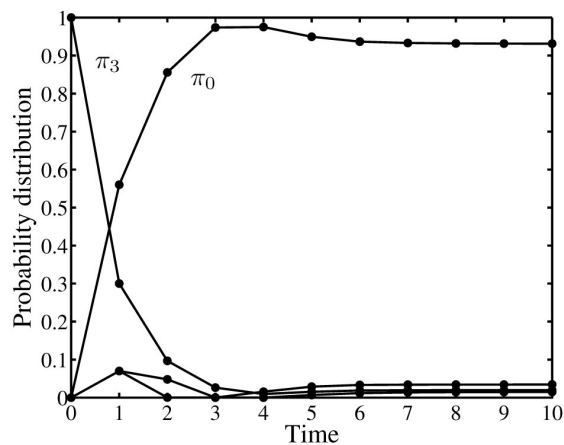


Figure 5. $\pi^4(t)$ in Case (i).

Next, the computational result in Case (ii) is explained. **Figures 7-11** show the probability distribution for each customer. Comparing **Figures 2-6** with **Figures 7-11**, we see that transient responses of $\pi_0^i(t)$ are improved in Case (ii). In particular, for the customer 5, the steady state of $\pi_0^5(t)$ is also improved (see

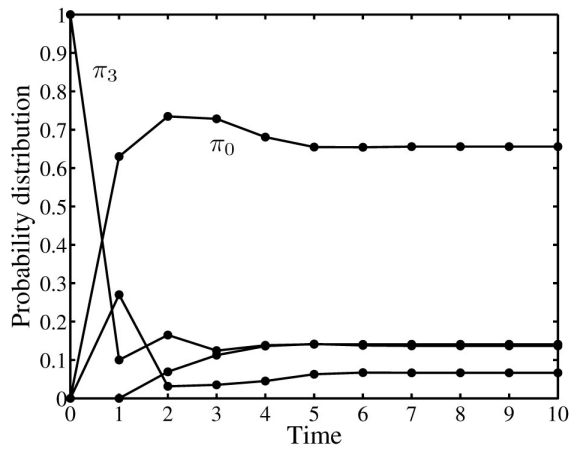


Figure 6. $\pi^5(t)$ in Case (i).

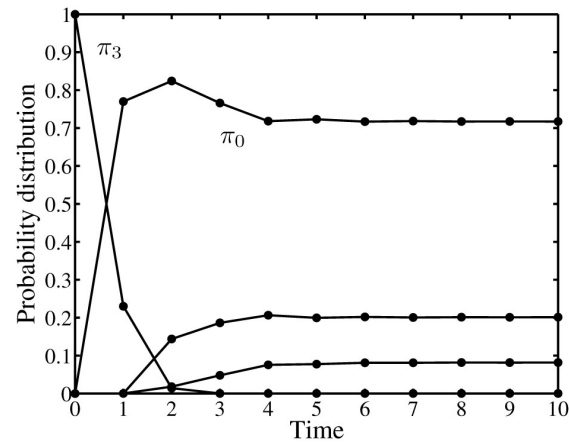


Figure 7. $\pi^1(t)$ in Case (ii).

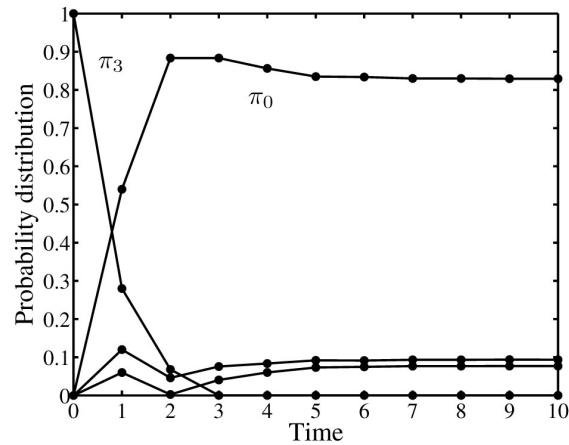


Figure 8. $\pi^2(t)$ in Case (ii).

Figure 6 and Figure 11). Furthermore, the optimal value of the cost function is 84.5057, and we see that the optimal value of the cost function is improved. The optimal control input

$$u(t) = [u^1(t) \ u^2(t) \ u^3(t) \ u^4(t) \ u^5(t)]^T$$

is derived as

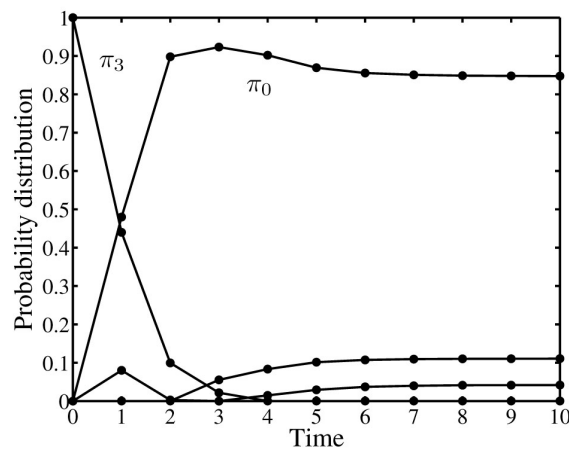


Figure 9. $\pi^3(t)$ in Case (ii).

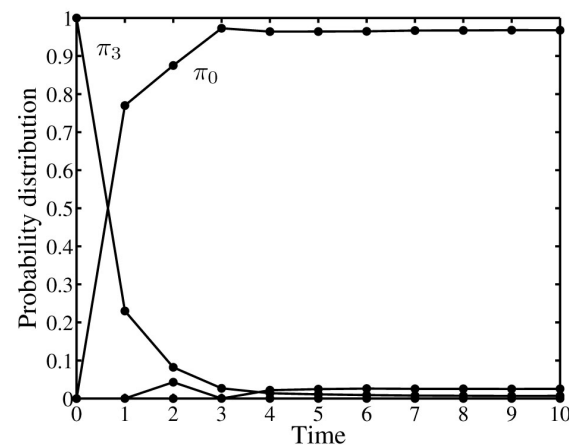


Figure 10. $\pi^4(t)$ in Case (ii).

$$\begin{aligned}
 u(0) &= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0.5 \end{bmatrix}, u(1) = \begin{bmatrix} 1 \\ 1 \\ 0.615 \\ 0.785 \end{bmatrix}, u(2) = \begin{bmatrix} 0.6637 \\ 0.6792 \\ 0.6081 \\ 0.5664 \\ 0.7622 \end{bmatrix}, u(3) = \begin{bmatrix} 0.5478 \\ 0.5942 \\ 0.5492 \\ 0.5135 \\ 0.7790 \end{bmatrix}, u(4) = \begin{bmatrix} 0.5755 \\ 0.5719 \\ 0.5086 \\ 0.5279 \\ 0.7627 \end{bmatrix}, \\
 u(5) &= \begin{bmatrix} 0.5755 \\ 0.5825 \\ 0.5068 \\ 0.5473 \\ 0.7752 \end{bmatrix}, u(6) = \begin{bmatrix} 0.5811 \\ 0.5831 \\ 0.5057 \\ 0.5576 \\ 0.7669 \end{bmatrix}, u(7) = \begin{bmatrix} 0.5811 \\ 0.5850 \\ 0.5048 \\ 0.5608 \\ 0.7726 \end{bmatrix}, u(8) = \begin{bmatrix} 0.5817 \\ 0.5850 \\ 0.5045 \\ 0.5625 \\ 0.7684 \end{bmatrix}, u(9) = \begin{bmatrix} 0.5816 \\ 0.5853 \\ 0.5043 \\ 0.5629 \\ 0.7713 \end{bmatrix}.
 \end{aligned}$$

From these values, we see that in the steady state, $u^5(t)$ is widely different to $u^i(t)$, $i = 1, 2, 3, 4$. Thus, in the system considered here, it is appropriate to use a local concurrency.

6. Conclusions

In this paper, we have proposed a modeling method and an optimal control method of real-time pricing systems using the MDP-based approach. In many cases, the status of electricity conservation of customers is discrete and

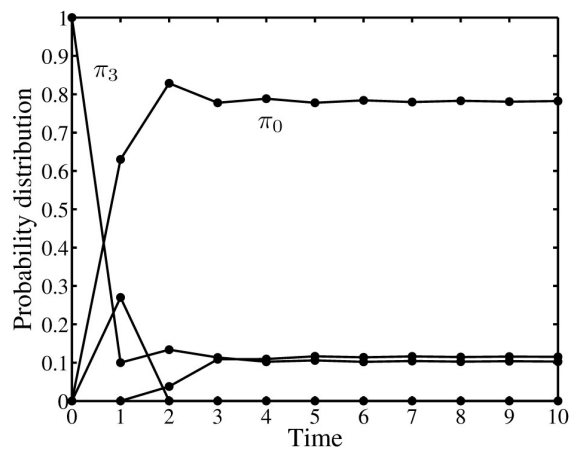


Figure 11. $\pi^5(t)$ in Case (ii).

stochastic, and the use of the MDP model is effective. A real-time pricing system is modeled by multi-agent MDPs, and the optimal control problem is reduced to a QP problem. Furthermore, a numerical simulation has been shown. The proposed method provides us with a new method in real-time pricing of electricity.

There are several open problems. First, it is important to develop the identification method of the MA-MDP model based on the existing result (see, e.g., [11]) for MDPs. Since the effect of couplings between customers was simplified, it is also important to consider modeling it more precisely. Next, the optimal control problem is reduced to a QP problem or an LP problem. These problems can be solved faster than a combinatorial optimization problem such as a mixed integer programming problem. However, for large-scale systems, the computation time for solving the optimal control problem will be long. Then, it is important to develop a distributed algorithm.

Acknowledgements

This research was partly supported by JST, CREST.

References

- [1] Camacho, E.F., Samad, T., Garcia-Sanz, M. and Hiskens, I. (2011) Control for Renewable Energy and Smart Grids. In: Samad, T. and Annaswamy, A.M., Eds., *The Impacy of Control Technology*, IEEE Control Systems Society, New York.
- [2] Ruihua, Z., Yumei, D. and Yuhong, L. (2010) New Challenges to Power System Planning and Operation of Smart Grid Development in China. *Proceedings of the 2010 International Conference on Power System Technology*, Hangzhou, 24-28 October 2010, 1-8.
- [3] Borenstein, S., Jaske, M. and Rosenfeld, A. (2002) Dynamic Pricing, Advanced Metering, and Demand Response in Electricity Markets. Center for the Study of Energy Markets, University of California, Berkeley.
- [4] Roozbehani, M., Dahleh, M. and Mitter, S. (2010) On the Stability of Wholesale Electricity Markets under Real-Time Pricing. *Proceedings of the 49th IEEE Conference on Decision and Control*, Atlanta, 15-17 December 2010, 1911-1918.
- [5] Samadi, P., Mohsenian-Rad, A.-H., Schober, R., Wong, V.W.S. and Jatskevich, J. (2010) Optimal Real-Time Pricing Algorithm Based on Utility Maximization for Smart Grid. *Proceedings of the 1st IEEE International Conference on Smart Grid Communications*, Gaithersburg, 4-6 October 2010, 415-420.
- [6] Vivekananthan, C., Mishra, Y. and Ledwich, G. (2013) A Novel Real Time Pricing Scheme for Demand Response in Residential Distribution Systems. *Proceedings of the 38th Annual Conference of the IEEE Industrial Electronics Society*, Monteral, 25-28 October 2012, 1954-1959.
- [7] Bello, D. and Riano, G. (2006) Linear Programming Solvers for Markov Decision Processes. *Proceedings of the 2006 IEEE Systems and Information Engineering Design Symposium*, Charlottesville, 28 April 2006, 90-95. <http://dx.doi.org/10.1109/SIEDS.2006.278719>
- [8] Eco-Point System for Housing. http://www.vec.gr.jp/english/topics/100217_1.htm

- [9] Kyoto Eco Money. http://www.city.kyoto.jp/koho/eng/topics/2012_8/index.html
- [10] Sawashima, K., Kubota, Y., Lu, H., Takemae, T., Yoshida, K. and Wan, Y. (2011) Socio-Personal Energy Management System. Keio ALPS2011 Group K Final Report. <http://lab.sdm.keio.ac.jp/alps2011k/FinalReport-ALPS2011-K.pdf>
- [11] Rust, J. (1994) Structural Estimation of Markov Decision Processes. In: *Handbook of Econometrics*, Elsevier, Amsterdam, Vol. IV, Chapter 51, 3081-3143.