

# Optimal Electricity Supply Bidding by Markov Decision Process

Haili Song, Chen-Ching Liu, *Fellow, IEEE*, Jacques Lawarrée, and Robert W. Dahlgren, *Member, IEEE*

**Abstract**—The bidding decision making problem is studied from a supplier's viewpoint in a spot market environment. The decision-making problem is formulated as a Markov Decision Process - a discrete stochastic optimization method. All other suppliers are modeled by their bidding parameters with corresponding probabilities. A systematic method is developed to calculate transition probabilities and rewards. A simplified market clearing system is also included in the implementation. A risk-neutral decision-maker is assumed, the optimal strategy is calculated to maximize the expected reward over a planning horizon. Simulation cases are used to illustrate the proposed method.

**Index Terms**—Electricity Market, Bidding Strategies, Decision-making, Markov Decision Process, Power System Economics.

## I. INTRODUCTION

THE power industry is evolving into an open-access, competitive environment. In this environment, economics and profitability are primary objectives of the market players. For each generation or distribution company, decisions have to be made on transactions, e.g., contract types and parameters. Electricity and services can be sold or purchased through bilateral contracts or the spot market [1]. The spot market usually operates as a pool, i.e., the market participants submit bids to a market that determines the transactions based on rules agreed upon by the participants.

To achieve efficiency in generation and consumption of electricity, an economic pricing scheme plays an important role. Properties of the prices of optional forward contracts are discussed in [2]. Alternative policies concerning access to and pricing of transmission are studied in [3]. Game theory models are used to estimate the possible effects of various policies upon productive efficiency and the distribution of gains among all market players.

In this paper, the problem of bidding decision-making is studied from the viewpoint of a generation company. Strategic bidding behavior has been studied in other fields such as commodity markets but less so in the electricity market. In [4], competitors are modeled by probability distributions of their bids. A method for updating probability distributions when

new data are observed is discussed. Some studies on bidding strategies have been conducted for the electricity market. In [5], a framework for an energy brokerage is proposed and a sub-optimal bidding strategy is developed according to the competitor's bidding probability density function. In [6], game theory is used to determine the suppliers' pricing strategy and it is assumed that suppliers bid with linear marginal price functions without capacity limits. Game theory is also used in [7] to simulate the decision making process for defining offered prices in a deregulated environment. A genetic algorithm is developed in [8] to select bidding strategies in the double auction electricity marketplace. Technical issues related to an auction and bidding market structure are analyzed in [9].

Besides bidding and pricing methods, other studies have been conducted on the subject of electricity transactions. For example, to handle the increasing number of electric energy transactions, an algorithm is proposed for identification of conflicting conditions between contracts [10].

In the literature, the optimal strategy is one that gives the decision-maker the maximum expected return for one bidding period. However, in the daily electricity market, the decision-maker's bid may influence the future market or the action may affect his/her own market position in the future. For example, a supplier's bid can affect the spot price and the other suppliers' bidding behavior may change according to the spot price. A supplier is also subject to resource constraints. The production limit is obvious for a hydro producer. A gas or coal producer may have fuel contracts that define the production limit over a time horizon. In this environment, the decision-maker should look into the future when a bidding decision is to be made. The bidding strategy that gives the best profit for one day may not be optimal when the expected profit over the planning horizon is desired.

The market has various uncertainties, e.g., price and load. Hence, the market model proposed here is stochastic. The decision-maker is assumed to be risk-neutral; the optimal strategy that maximizes the expected profit is desirable for the decision-maker. To develop a tractable model, a Markov process is assumed. The purpose of the proposed method is to optimize the expected reward over a planning horizon. This paper reports new results on the application of a Markov Decision Process (MDP) to optimize the bidding decisions. The MDP here is of the discrete-state and discrete-time type. MDP provides a systematic way to solve multiple stage probabilistic decision-making problems. In an MDP, the stochastic process evolves in a sequence of time stages. As shown in Fig. 1, at stage  $t$ , the market can be in any of a number of states numbered from 1 to  $N$ . In each state, the decision-maker can choose one decision  $a$  from a set

Manuscript received March 16, 1998; revised February 3, 1999. This research is sponsored by National Science Foundation under Grant no. ECS-9612636 with matching support from ALSTOM ESCA Corporation.

H. Song and C.-C. Liu are with the Department of Electrical Engineering, University of Washington, Seattle WA 98195.

J. Lawarrée is with the Department of Economics, University of Washington, Seattle WA 98195.

R. W. Dahlgren is with the ALSTOM ESCA Corp. Bellevue, WA.

Publisher Item Identifier S 0885-8950(00)03796-2.

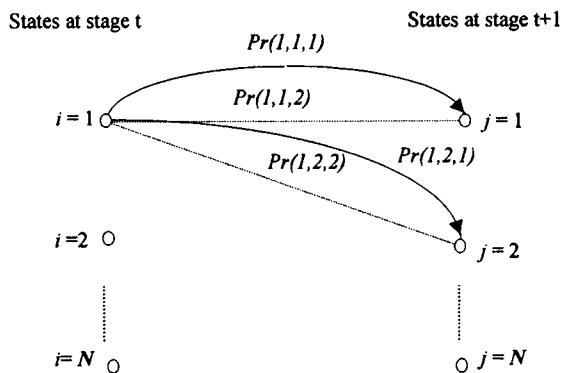


Fig. 1. States and state transition in MDP.

of feasible decision options. Corresponding to a decision  $a$ , the transition probability from a state  $i$  to another state  $j$  is given by  $Pr(i, j, a)$ . The decisions from the first stage to the end of the planning horizon form one strategy (policy). The decision-maker receives a reward  $r(i, j, a)$  from each transition. The MDP model is mathematically well established and its applications can be found in many areas [11]–[12].

The significance of the proposed method lies in the fact that the method calculates the optimal decision *over a planning horizon*. The supplier’s production limit is incorporated in the proposed bidding decision-making method. The information structure can incorporate detailed information about the market accumulated over time. Hence, it is believed that the proposed methodology is promising as a practical market decision support tool.

## II. PROBLEM FORMULATION

This study assumes a day-ahead auction system with no demand-side bidding; however, the proposed method can be adapted for other types of markets. It is assumed that the bids are cleared for each hour. The suppliers submit bid with the price and quantity at which they intend to sell during the next day. The market clears the bids. The spot price is determined according to the day-ahead load forecast and the bids from all suppliers. A single spot price covers all purchases and sales in each hour. Suppliers are paid according to this spot price and possibly other market payment rules. The proposed market model resembles the structures that have been established in England / Wales and New Zealand.

In Fig. 2, the environment in which the proposed electric energy Bidding Decision Support tool, called *BIDS*, is illustrated. The decision-maker’s bidding options and the competitors’ bids are specified depending on the state of the market. After the bids are cleared and transactions are completed, the market changes its state depending on the decision-maker’s decision and other players’ bids. The reward from the bid is calculated based on the spot price and the cost associated with the transaction.

The load profile within one day is divided into different load periods, i.e., peak load and off-peak load. The average load for each load period is judged to be sufficient to represent the overall load profile. Thus, in the proposed method, *the state is represented by the production limit, the average load and average*

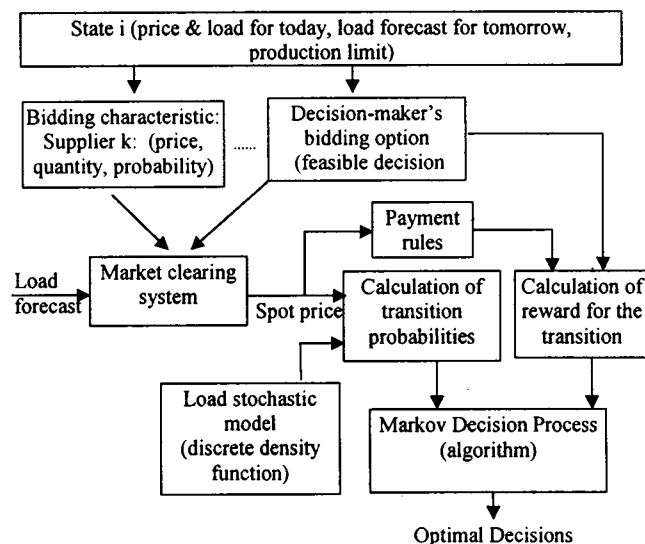


Fig. 2. Market and bidding decision making.

*spot price over each load period for today’s market, and the average load over each load period for tomorrow’s forecast.* There are 7 variables in the state definition: 1 to represent the production limit, 2 to represent the prices, 2 to represent today’s load, and 2 more for tomorrow’s load forecast. If 10 grid-numbers were used to represent different data levels, the number of states would be  $10^7$ . It is clear that the MDP formulation can suffer from the “curse of dimensionality” problem if the number of states is large, which is likely for a complex decision making problem. Aggregation can be used to reduce the number of states. For example, the prices and demand can be represented by ranges such as “high”, “medium”, or “low” to avoid a large number of states using specific numbers. Suboptimal decisions can be obtained for the reduced state space. Practical considerations can be used to further reduce the number of states. For example, unusually low demand in one load period and unusually high demand in another load period may not happen on the same day, so these combinations can be eliminated. Using similar techniques to further reduce the number of states, the state spaces of the study cases in Sec. VI contain 210 and 21 states, respectively.

A decision-maker has to set the bid parameters, i.e., the price and quantity have to be specified for each hour period of the next day. In this study, a staircase supply function is used to describe the bid prices for different MW ranges over a day. One bid price is used for each MW range. Also, the decision-maker specifies one MW quantity for each load period.

An MDP moves from one state  $i$  to another state  $j$  according to a transition probability  $p_{ij}$ . A transition occurs as a result of a change in price and / or load demand. A major difference between a Markov Process and an MDP is that the latter incorporates decision options for each state, which affect the transition probabilities and rewards. These probabilities represent the uncertainty of the market. In Sec. III, an algorithm is proposed for calculation of transition probabilities and rewards based on load forecast, decisions for each state, and bidding characteristics of other suppliers. The reward from a transition is the total revenue

minus the cost. It is assumed that the decision-maker knows its own cost.

### III. CALCULATION OF TRANSITION PROBABILITIES AND REWARDS

It should be kept in mind that the proposed decision support tool, *BIDS*, is developed from the viewpoint of a supplier, i.e., the decision-maker using the MDP model. The decision-maker observes and predicts the trend of the load demand, predicts the other suppliers' bidding parameters, and optimizes his or her own profits accordingly.

To simplify the notation, the calculation is described in such a way that there is only one unit for each supplier and one bid price is set for each unit. Also, the time variable is not explicitly shown. However, the proposed method does allow suppliers to have multiple units and more than one bid price can be set for any unit. The algorithm is identical when multiple units and multiple prices are used. For time-varying data, the transition probabilities and rewards are calculated accordingly, and a non-stationary MDP is needed. The terminology is described in the following.

$N$ :	# of suppliers (supplier $m$ is the decision-maker)
$K(i, n)$ :	# of bidding options for supplier $n$ in state $i$
$T$ :	# of load periods ( $T = 2$ in this study, peak/off-peak load period)
$hours(t)$ :	# of hours in load period $t$
$p(i, n, k)$ :	Bid price of supplier $n$ , given bidding option $k$
$q(i, n, k)$ :	Bid quantity of supplier $n$ for load period $t$ , given $k$
$Pr(i, n, k)$ :	Probability that supplier $n$ ( $n \neq m$ ) chooses option $k$ in state $i$
$PL(i)$ :	Decision-maker's production limit in state $i$
$SP(i, t)$ :	Spot price at load period $t$ in state $i$
$LD(i, t)$ :	Load demand at load period $t$ in state $i$
$LF(i, t)$ :	Load forecast at load period $t$ in state $i$
$Pr(i, LF(j, t))$ :	Probability that the load forecast for the day after tomorrow is $LF(j, t)$ for load period $t$ , $t = 1, \dots, T$ , given the present state $i$
$Pr(i, j, a)$ :	Transition probability from state $i$ to $j$ if the decision-maker makes decision $a$
$Pr(i, s)$ :	Probability of scenario $s$
$SP(i, s, t)$ :	Spot price at load period $t$ for scenario $s$
$\Omega(i, j)$ :	Subset of scenarios that results in the spot price, cleared quantity, and production limit of state $j$ from those of state $i$
$Cost(q)$ :	Cost of the decision-maker to produce $q$ MWh
$r(i, s)$ :	Reward for scenario $s$
$r(i, j, a)$ :	Reward corresponding to the transition from state $i$ to state $j$ , given decision $a$
$q_s(i, j, t)$ :	The decision-maker's quantity that is called into operation in load period $t$ for scenario $s$
$PL(i, s)$ :	The decision-maker's updated production limit if scenario $s$ takes place, starting from state $i$

The objective is to calculate the transition probability that the system transfers from state  $i$  to a state in the next day, say state  $j$ . Since the spot price for the next day should match the supply with tomorrow's load forecast, as the system moves to state  $j$ , the production limit is denoted by  $PL(j)$ , the spot price is  $SP(j, t)$ , the cleared load demand will be the tomorrow's load forecast that is available in state  $i$ . The load forecast for the day after tomorrow is  $LF(j, t)$ . It is assumed that different suppliers' bid options are independent. Therefore, the probability that supplier  $n1$  chooses option  $k1$  and supplier  $n2$  chooses option  $k2$  will be  $Pr(i, n1, k1) * Pr(i, n2, k2)$ . These probabilities should be determined by market data observations.

For the  $N$  suppliers and  $T$  load periods, a scenario  $s$  is a combination of the price and quantity from each supplier's bidding parameters, i.e.,

$$s = \{(p(i, n, k), q(i, n, t, k))\} \quad n = 1, \dots, N; t = 1, \dots, T$$

The probability of this scenario,  $Pr(i, s)$ , is:

$$Pr(i, s) = \prod_{n=1, \dots, N}^{n \neq m} Pr(i, n, k) \quad (1)$$

Supplier  $n1$  has  $K(i, n1)$  options and supplier  $n2$  has  $K(i, n2)$  options. All suppliers' options are independent. Combining all the suppliers' options, the possible scenarios can be enumerated and the total number of scenarios is

$$W = \prod_{n=1, \dots, N} K(i, n) \quad (2)$$

To calculate the spot price for each scenario, a market clearing system has to be modeled. Unit Commitment based [13] or Optimal Power Flow based [14] market clearing systems have been proposed. Without considering the security constraints and other market characteristics, a simplified market clearing system is adopted in this study. The suppliers' bids are ranked from the cheapest to the most expensive for each load period. Suppliers' bids are chosen from the cheapest until the load in that period is met. For all units that are called into operation, the most expensive bid price defines the spot price in that load period. It should be pointed out that the average load does not necessarily lead to the average spot price, due to the nonlinear relation between the load demand and the resulting spot price. For tractability, this study uses the spot price corresponding to the average load as an approximate average price for each load period.

The spot price for a scenario  $s$ ,  $SP(i, s, t)$ , is calculated by matching the suppliers' bids with the load forecast at state  $i$ ,  $LF(i, t)$ . The spot price and the decision-maker's bid decision determine the amount to be produced by the decision-maker. The decision-maker's production for scenario  $s$  is calculated by summing up the production in all load periods. The decision-maker's production limit for the next stage is updated by

$$PL(i, s) = PL(i) - \sum_{t=1}^T [q_s(i, s, t) * hours(r)] \quad (3)$$

For each scenario  $s$  in  $\Omega(i, j)$ , the system moves from state  $i$  to state  $j$  if the load forecast for the day after tomorrow is

$LF(j, t)$  and scenario  $s$  takes place. Since all the scenarios are exclusive, the probability that results in spot price  $SP(j, t)$ ,  $t = 1, \dots, T$ , and production limit  $PL(j)$  is the sum of the probabilities of all the scenarios in  $\Omega(i, j)$ . In our notation, this means  $SP(i, s, t) = SP(j, t)$ ,  $LF(i, t) = LF(j, t)$ ,  $t = 1, \dots, T$ , and  $PL(i, s) = PL(j)$  for scenario  $s$ . For all scenarios in  $\Omega(i, j)$ , the probability can be evaluated by

$$\begin{aligned} & \Pr(SP(j, t) \& PL(j) \& LF(i, t) = LF(j, t)) \\ &= \sum_{s \in \Omega(i, j)} \Pr(i, s) \end{aligned} \quad (4)$$

The above probability represents all scenarios that match the spot price, the production limit, and tomorrow's load forecast with the definition of state  $j$ . State transition is also caused by the change of load forecast between tomorrow's load and the day after tomorrow's load. The probability that the system moves from state  $i$  to state  $j$  for the decision-maker's decision option  $a$  can now be calculated by

$$\begin{aligned} \Pr(i, j, a) &= \Pr(SP(j, t) \& PL(j) \& LF(i, t) \\ &= LF(j, t)) * \Pr(i, LF(j, t)) \\ &= \sum_{s \in \Omega(i, j)} \Pr(i, s) * \Pr(i, LF(j, t)) \end{aligned} \quad (5)$$

The above equations lead to transition probabilities that depend only on decision-maker's decision and the two states  $i$  and  $j$ . As mentioned previously, if transition probabilities are time-dependent, the MDP is a nonstationary stochastic process.

The reward is the difference between the revenue and the cost of the decision-maker as a supplier. The cost function of the decision-maker is assumed to be deterministic and known to the decision-maker. Each generating unit has a cost function, which is assumed to be a piece-wise linear function of MW. As the market clears, the amount to be produced by each unit is allocated. Therefore, the cost of each unit can be calculated accordingly. The revenue at the load period  $t$  is the product of the spot price and the quantity that is called into operation. For each scenario, the reward for the decision-maker is calculated as

$$\begin{aligned} r(i, s) &= \sum_{t=1}^T [SP(i, s, t) * q_s(i, s, t) * hours(t)] \\ &\quad - \sum_{t=1}^T Cost(q_s(i, s, t) * hours(t)) \end{aligned} \quad (6)$$

The reward has to be calculated for all scenarios associated with the transition from  $i$  to  $j$ . Also, the reward for each transition from  $i$  to  $j$  has to be weighted by the conditional probability that the system moves from state  $i$  to state  $j$ . Finally, the reward for the transition from state  $i$  to state  $j$  and decision option  $a$  is the sum of the rewards for individual scenarios weighted by the conditional probabilities of these scenarios.

$$r(i, j, a) = \sum_{s \in \Omega(i, j)} \{r(i, s) * [\Pr(i, s) / \sum_{s \in \Omega(i, j)} \Pr(i, s)]\} \quad (7)$$

#### IV. VALUE ITERATION

After the transition probabilities and rewards are calculated, the value iteration method is applied for this MDP model [11]. Value iteration is similar to backward dynamic programming. Let  $V(i, \tau + 1)$  be the total expected reward in  $\tau + 1$  remaining stages starting from state  $i$  if an optimal policy is followed. At the last stage,  $\tau = 0$ . Value iteration searches for the decision that results in the maximum total expected reward in the remaining  $\tau + 1$  stages, i.e.,

$$V(i, \tau + 1) = \max_{a=1}^{K(i, m)} \sum_{j=1}^N \{\Pr(i, j, a) * [r(i, j, a) + V(j, \tau)]\} \quad (8)$$

Equation (8) is the recursive relation that is used to calculate the optimal decision from the last stage ( $\tau = 0$ ) to the first stage. These rewards from a transition to  $j$  are weighted by the probability of the transition,  $\Pr(i, j, a)$ , to obtain the total expected earnings.

#### V. MODEL VALIDATION

For validation purpose, it is necessary to accumulate the actual data and observations from the market over a reasonable period of time, say, 1 year. The market data set provides the actual scenarios. For each scenario, according to the *BIDS* representation, the decision-maker's reward is  $r(i, j, a)$  for a transition from state  $i$  to state  $j$  and decision option  $a$ . The *actual* reward is known from the market data. This actual reward is denoted by  $w(i, j, a)$ . The relationship between the estimated and the actual rewards can be analyzed by linear regression. If a linear relationship is assumed between the estimated and the actual rewards, then

$$w(i, j, a) = A * r(i, j, a) + B \quad (9)$$

The coefficients  $A$  and  $B$  are found by minimizing the average absolute error over the market data set,

$$R = \frac{1}{M} \sum_{m=1}^M |w_m(i, j, a) - A * r_m(i, j, a) - B| \quad (10)$$

where  $M$  is the total number of observations for a transition from state  $i$  to state  $j$ . Ideally,  $A$  should be close to 1 and  $B$  should be close to 0. A significant deviation from these values implies the necessity to tune the reward calculation method.

#### VI. IMPLEMENTATION AND CASE STUDIES

A 3-supplier-5-generator system is used for illustration of the *BIDS* method. The suppliers are *GenCoA*, *GenCoB*, and *GenCoC*. *GenCoA* is the decision-maker using our MDP technique. The decision-maker owns one generating unit. *GenCoB* owns two units and *GenCoC* owns two other units. All three suppliers bid in the spot market. The planning horizon is 7 days, i.e., the bid decision for the next day considers the entire week ahead. In each state, *GenCoA* makes its decision from a set of pre-specified decision options. In each state, *GenCoA* does not know exactly how *GenCoB* and *GenCoC* are going to bid. However, their individual bidding behavior is modeled by bid prices,

TABLE I  
SUPPLIER CAPACITIES AND POSSIBLE LOAD LEVELS

Possible system load levels	Peak	220, 240, 260 MW
	Off-peak	160, 180, 200 MW
GenCoA capacity		120MW
GenCoB capacity		2*50MW
GenCoC capacity		2*50MW

TABLE II  
GENCOA'S COST, PRODUCTION LIMIT AND DECISION OPTIONS

Marginal cost (\$/MWh)	0-40MW	40-80MW	80-120MW
	16	17	18
Production limit: 0 – 9,600MWh			
No. of decision options : 6 (for every state)			
Option 1: bidding price 18\$/MWh; intend to sell 40MW for both peak and off-peak load.			

TABLE III  
PARAMETERS OF STATE 1

Production limit (MWh)	Spot price (\$/MWh)		System load (MW)		Tomorrow's load forecast	
	Peak	Off-peak	Peak	Off-peak	Peak	Off-peak
960	23	20	260	200	260	200

quantities and the associated probabilities based on *GenCoA*'s knowledge and information. The transition probabilities and rewards are calculated according to the algorithm described in Sec. III. The *BIDS* value iteration algorithm is implemented in Matlab.

**Example 1:** The decision-maker has a production limit over the planning horizon. In this simulation, the number of states is 210 and there are 10 grid numbers between 0 and the production limit. The production limit for one week is set at the level such that if the unit runs at its capacity, the production limit will be reached in 4 days.

Due to space limitation, only selected data is shown in Tables I–III.

The user specifies the state parameters. The parameters for one of the states, state 1, are listed in Table III for illustration.

The other suppliers' bid parameters and probabilities are specified depending on the state (from state 1 to state 210). For example, one of the bid options of *GenCoB*'s in state 1 is

*Unit1: bid price 15\$/MWh for the first 20MW block, 16\$/MWh for the next 30MW, maximum output 50MW for both peak & off-peak load period;*

*Unit2: bid price 17\$/MWh for the whole capacity (50MW), maximum output 50MW for both peak & off-peak load period*

*The probability of this bidding option: 0.25*

Using the data above, the transition probabilities and rewards are calculated by the proposed algorithm. For example, for the decision-maker's 6<sup>th</sup> decision option: selling nothing, the transition probability from state 2 to state 1,  $Pr(2, 1, 6) = 0.17188$  which corresponds to a reward  $r(2, 1, 6) = 0$ .

The value iteration algorithm is used to calculate the optimal decision for each state and each stage. The decisions at some of the states on the first day, day 1, are illustrated in Table IV. One

TABLE IV  
SELECTED RESULTS FROM EXAMPLE 1

On day 1	Results from <i>BIDS</i>		Peak-load strategy
	Decision No.	Expected reward (\$)	Expected reward (\$)
...			
State 190	No. 3*	43,973	39,138
State 191	No. 1	42,061	38,511
State 192	No. 1	42,912	38,699
State 193	No. 4*	42,646	37,134
...			

**\*Decision Option 3:**

*Block 1: 18\$/MWh; 40MW for peak & off-peak load;*

*Block 2: 19\$/MWh; 40MW for peak & off-peak load;*

**\*Decision Option 4:**

*Block 1: 19\$/MWh; 40MW for peak & off-peak load;*

*Block 2: 20\$/MWh; 40MW for peak & off-peak load;*

*Block 3: 21\$/MWh; 40MW for peak; 0MW for off-peak;*

rational strategy, *Peak-Load Strategy*, is chosen for comparison with the *BIDS* results. By taking *Peak-Load Strategy*, the decision-maker bids at the cost, divides the production limit evenly over the week, and sells only at the peak load period on each day. According to the production limit in this case study, *GenCoA* can sell 97MW each day for the peak-load period. The result shows that the expected rewards from the *BIDS* bidding strategy are consistently higher than that of the *Peak-Load Strategy*.

The resulting optimal strategy is *time dependent*. For example, *the optimal strategy for state 190 is no. 3 for day 1 to day 4 and it is no. 4 for day 5 to day 7*. The optimal decision for day 1 to day 4 is to sell less (option 3) and save the capacity for day 5 to day 7. The production limit is 9,600MWh in state 190. If the system happens to be in state 190 on day 5, the production limit will not be violated even if the unit runs at its capacity for the remaining 3 days. The optimal decision (option 4) is to sell more electricity, i.e., the 3 blocks in option at increasing prices.

From day 1 to day 7, the optimal decision changes according to the production limit even if the price and demand are identical. In some of the states, the resulting optimal decision is not to sell (option 6); the resources are saved for more profitable days.

**Example 2:** The decision-maker has market power and no production limit. The decision-maker can manipulate the bid to influence the spot price. Other suppliers will bid according to the spot price and load demand information. The decision-maker makes the bidding decision to maximize the expected reward over the planning horizon.

Without production limit, there are 21 states in the MDP model. The feasible decision set is taken to be different from that in Example 1. For this example, the *Daily Maximum Strategy* is chosen for comparison with the *BIDS* results. The *Daily Maximum Strategy* maximizes the daily reward without considering how the bid affects the market trend. The results are shown in Table V.

The expected reward is the accumulated reward from day 1 to day 7. It depends on the strategy from day 1 to day 7. In Table V, only the optimal decisions on day 1 are illustrated. Note that the

TABLE V  
SELECTED RESULTS FROM EXAMPLE 2

On day 1	Results from <i>BIDS</i>		Daily maximum strategy	
	Decision No.	Expected reward (\$)	Decision No.	Expected reward (\$)
...				
State 1	No. 4*	65,342	No. 2*	58,824
State 2	No. 5	62,802	No. 4	56,254
State 3	No. 4	60,592	No. 4	56,594
State 4	No. 4	63,823	No. 4	57,858
...				

\*Decision Option 4: (see Decision Option 4 in Table 4)

\*Decision Option 2:

Block 1: 21\$/MWh; 40MW for peak & off-peak load;

Block 2: 22\$/MWh; 40MW for peak & off-peak load;

Block 3: 23\$/MWh; 40MW for peak & off-peak load;

TABLE VI  
TRANSITION PROBABILITIES FOR DECISION No. 2 AND No. 4 IN STATE 1

From state $i=1$	To $j=1$	To $j=2$	To $j=3$	...	To $j=7$	To $j=8$	To $j=9$	...
Decision No. 2	0.125	0.25	0.125		0.125	0.25	0.125	...
Decision No. 4	0	0	0		0.25	0.5	0.25	

Daily Maximum Strategy is *time independent*; the decision-maker chooses the same decision option as long as the system is at the same state. The *BIDS* results show the same decisions on day 7 of the planning horizon. This is expected since the decision is also to maximize the expected profit for that day only. The results from *BIDS* are *time dependent*. The optimal decisions on day 7 are different from those of the previous days. For example, in state 1 (same as state 1 in example 1 but without production limit), decision option 4 is to bid at low prices and decision option 2 is to bid high prices. Since the decision-maker has market power, bidding at high prices gives the decision-maker a high immediate reward and the system transfers to high price states with higher probabilities. Similarly, bidding at low prices gives the decision-maker relatively low immediate reward and the system transfers to low price states with higher probabilities. The transition probabilities of decision no. 2 and no. 4 in state 1 are illustrated in Table VI. By making decision no. 2, the system possibly moves to states 1 to 3 and states 7 to 9 (with nonzero transition probabilities). By making decision no. 4, however, the system possibly moves only to states 7 to 9. The spot prices in states 1 to 3 are high, the spot prices in states 7 to 9 are lower. The peak spot price for the former is 23\$/MWh and 21\$/MWh for the latter. The optimal decision for state 1 is no. 4 on day 1 based on the *BIDS* method. The decision-maker actually earns more when the spot price remains to be relatively low.

On the last day, the optimal decision is the decision that gives the best daily return; on other days the decisions are chosen to maximize the *total* expected profits over the planning horizon. This demonstrates the importance of the planning horizon in

the *BIDS* algorithm. The decision-maker should not rely on the daily profit in making decisions. This optimal profit can not be achieved when the bidding strategy is derived from only one day's information.

## VII. CONCLUSION

The spot market bidding decision-making problem is studied in this paper. An algorithm is developed to calculate the transition probabilities and rewards for the MDP model. Simulation cases have been studied to validate the method. Compared to other analysis methods, MDP is able to optimize the decision over a planning horizon. Even though the modeling is fairly detailed for the market, the power system operational constraints are ignored for tractability of the model. These issues should be considered in the future.

A risk-neutral decision-maker is assumed in this paper. If the decision-maker is risk-averse, the variance of his/her profit should be incorporated in the bidding decisions. The optimal decisions will then be determined by maximizing the total benefit. There is no provision for incorporating risk attitude in an ordinary MDP. A slightly different formulation, risk-sensitive MDP [15], can be used to solve this problem.

Probabilistic bidding information of competitors is required to calculate the state transition probabilities and rewards. For some of the markets, the market participants' bidding data is public information, e.g., the United Kingdom Market. If the competitors' bidding data is not available, the historical market prices and load demand information can be used to identify state transition probabilities by statistical data analysis. The number of players and the number of generators affect the computational effort in calculating the state transition probabilities and rewards. However, it does not influence the size of the state space.

Since the competitor's strategy may change, it may be interesting to calculate the sensitivity of the policy / reward with respect to the competitor's bids. To analyze the sensitivities, one can vary the price, amount and the probability of this competitor's bids and re-calculate the transition probabilities, optimal policies, etc. using the proposed MDP algorithm.

## ACKNOWLEDGMENT

The authors would like to thank Roger Bjorgan for the many helpful discussions.

## REFERENCES

- [1] H. Rudnick, R. Varela, and W. Hogan, "Evaluation of alternatives for power system coordination and pooling in a competitive environment," *IEEE Trans. Power Systems*, vol. 12, no. 2, pp. 605–613, May 1997.
- [2] T. W. Gedra, "Optional forward contracts for electric power markets," *IEEE Trans. Power Systems*, vol. 9, no. 4, pp. 1766–1773, Nov. 1994.
- [3] B. F. Hobbs and K. A. Kelly, "Using game theory to analyze electric transmission pricing policies in the United States," *European Journal of Operational Research*, vol. 56, no. 2, pp. 154–171, Jan. 1992.
- [4] F. H. Griffs, "Winning over key competitors," *J. Construction Engineering and Management*, vol. 118, no. 1, pp. 151–165, 1992.
- [5] J. W. Lamont and S. Rajan, "Strategic bidding in an energy brokerage," *IEEE Trans. Power Systems*, vol. 12, no. 4, pp. 1729–1733, Nov. 1997.
- [6] R. W. Ferrero, J. F. Rivera, and S. M. Shahidepour, "Application of games with incomplete information for pricing electricity in deregulated power pools," PES Paper no. PE-012-PWRS-1-05-1997.

- [7] R. W. Ferrero, S. M. Shahidehpour, and V. C. Ramesh, "Transaction analysis in deregulated power systems using game theory," *IEEE Trans. Power Systems*, vol. 12, no. 3, pp. 1340–1347, Aug. 1997.
- [8] C. W. Richter, Jr and G. B. Sheblé, "Genetic algorithm evolution of utility bidding strategies for the competitive marketplace," *IEEE Trans. Power Systems*, vol. 13, no. 1, pp. 256–261, Feb. 1998.
- [9] G. B. Sheblé, "Price based operation in an auction market structure," *IEEE Trans. Power Systems*, vol. 11, no. 4, pp. 1770–1777, Nov. 1996.
- [10] G. Rosenwald and C. C. Liu, "Consistency evaluation in an operational Environment involving many transactions," *IEEE Trans. Power Systems*, vol. 11, no. 4, pp. 1757–1762, Nov. 1996.
- [11] R. A. Howard, *Dynamic programming and Markov processes*, New York: The Technology Press of the MIT and John Wiley & Sons, Inc., 1960.
- [12] D. J. White, "A survey of applications of Markov Decision Processes," *J. Operational Research Society*, vol. 44, no. 11, pp. 1073–1096, 1993.
- [13] J. Ancona, "A bids solicitation and selection method for developing a competitive spot priced electric market," *IEEE Trans. Power Systems*, vol. 12, no. 2, pp. 743–748, May 1997.
- [14] T. Alvey, D. Goodwin, X. Ma, D. Streiffert, and D. Sun, "A security-constrained bid-clearing system for the New Zealand wholesale electricity market," PES Paper no. PE-923-PWRS-2-06-1997.
- [15] R. A. Howard and J. E. Matheson, "Risk-sensitive Markov decision processes," *Management science*, vol. 18, no. 7, pp. 356–369, March 1972.

**Haili Song** received her BSEE and MSEE degrees from Tsinghua University, Beijing, in 1991 and 1993, respectively. She was a faculty member at the department of Electrical Engineering in Tsinghua University from 1993 to 1996. She has been a Ph.D. student at the University of Washington since 1996. Her areas of interest are power system economics, optimization, game theory, and intelligent system applications.

**Chen-Ching Liu** received his Ph.D. from University of California, Berkeley. He is currently Professor of EE and Director of Electric Energy Industrial Consortium at the University of Washington. Dr. Liu serves as Chair of the Subcommittee on Intelligent System Applications of the Technical Committee on Power System Analysis, Computing and Economics. He is a Fellow of the IEEE.

**Jacques Lawarrée** received his Ph.D. in economics from the University of California at Berkeley in 1990. He is currently an associate professor of economics at the University of Washington and is a member of the European Centre for Advanced Research in Economics in Brussels. Dr. Lawarrée specializes in game theory, contract theory and industrial organization.

**Robert W. Dahlgren** received his BSEE and MSEE from the University of Washington, in 1994 and 1995, respectively. He is currently a Ph.D. student at the University of Washington and a full-time employee of ALSTOM ESCA in Bellevue, Washington. His research interests include energy interchange, scheduling and power markets under deregulation. He is a member of Eta Kappa Nu, Tau Beta Pi, and the IEEE.