A System Marginal Price Forecasting Based on an Artificial Neural Network Adapted with Rough Set Theory

Jeong-Kyu Lee, Jong-Bae Park, Member, IEEE, Joong-Rin Shin, Kwang Y. Lee, Fellow, IEEE

Abstract—This paper presents a forecasting technique of the short-term system marginal price (SMP) using an Artificial Neural Network (ANN) adapted with Rough Set theory. The SMP forecasting is a very important element in an electricity market for the optimal biddings of market participants as well as for market stabilization of regulatory bodies. Input data is grouped into similar pattern using rough set theory, and the resulting patterns are used to train the ANN. In training of ANN, it is more efficient because some patterns are combined into one pattern. After training with the combined patterns adapted with rough set, the SMP is forecasted using the ANN.

The proposed method is applied to the historical real-world data from the Korea Power Exchange (KPX) to verify the effectiveness of the technique.

Index Terms—System marginal price, forecasting, artificial neural network, rough set theory

I. INTRODUCTION

The electric power industry in many countries all over the world has been undergoing deregulation, privatization and restructuring through the introduction of competition. Korea has been keeping pace with this situation. Restructuring of electric power industry in Korea started with the establishment of the Korea Power Exchange (KPX) in April 2001, and the breakup of the generation sector from the Korea Electric Power Corporation (KEPCO) into six generation companies (GENCOs). Currently, Cost-Based Pool (CBP) electricity market is in operation.

The goal of power system planning and operation for the previous vertically integrated industry was to minimize production and operation cost. However, the goal is now changed into maximizing the profit or return to the market participants since the introduction of competition in the electricity market. Indeed, it may become possible to use strategies in order to maximize the profit.

Every market participants have to invest in their facilities, such as generators, transmission lines and distribution networks, etc. to maximize their profit. Generation expansion, transmission expansion and distribution planning cannot be completed in short-term, such as in one or two years, and must be considered under long-term market conditions. That is to say, market participants should have appropriate facilities in accordance with the electricity price forecasting in the future. Market participants bid at a specific time period to trade the electric power. In this bidding, each participant can maximize their profit through a bidding strategy that is considered under several power system conditions, such as characteristics of electric power demand at each time period. Therefore, the energy trading levels between market participants is highly dependent on the short-term price forecast [1].

In general, hard or soft computing techniques could be used to estimate the spot prices. Hard computing techniques are built on an exact model of the system, and the solution is found using algorithms that consider the physical phenomena governing the system. This approach can be very accurate, but it requires a great deal of information, and the computational cost can be very high. Soft computing techniques, on the hand, are very simple in structure, and it can be as accurate as the hard computing technique if the correct inputs are considered [2]. The main stream of soft computing techniques is artificial intelligence, such as artificial neural network, fuzzy and neuro-fuzzy systems, etc.

Several techniques are proposed to forecast the electricity price accurately in response to the increasing importance of electricity price forecasting raised by market participants. Firstly, an artificial neural network (ANN) is used to forecast the SMP in Victorian Power System by Szku1ta et al. [1]. This work defined two factors that impacted the SMP, that is, the System Potential Demand (SPD) and System Power Reserves (SPR). Also, it showed the variation of current SMP due to the variation of each factor to verify the effectiveness of SPD and SPR. The SMP forecasting was implemented for a case from 14 May 1997 to 20 May 1997 using these two factors; however, this technique is impractical to apply in the real power market, because the forecasting error is relatively high. Rodriguez et al. developed the factors that impact energy price considering many power system conditions and then
applied to Ontario Competitive Power System Market [2]. In this work, traditional neural network and neuro-fuzzy system are applied to forecast the energy price and the forecasting error was compared between these two techniques. Moreover, various combinations of the factors that influenced the energy price were analyzed to determine the most accurate forecasting method. In addition, short-term energy price forecasting method was developed using time series modeling by Nogales et al. [3] and price forecasting method was developed using dynamic fuzzy system by Liu et al. [4].

In this paper, input data is classified and combined in each pattern adapted with the rough set theory in order to increase the efficiency and accuracy of the SMP forecasting using ANN.

II. FORMULATION

The SMP is the market price that is determined to consider the characteristics of generators in bidding, under a given power system condition and demand. It is determined by the highest bidding price among generators that is succeeded in the bidding. In general, factors that impact the energy price are power demand, temperature, operating reserves, predicted shortfalls, etc., and the main variable that drives the price is the demand [2]. The SMP has a specific characteristic for each time period, such as daily cycle and seasonal variation, which is the basis for forecasting the future SMP based on the past SMPS.

If we consider all possible factors that affect the SMP, forecasting will be very accurate, which, however, is very difficult to do in the real market. Therefore, this paper considers only the past power demand and SMP to forecast the future SMP. The data can be defined on a point \( (w,t) \) in the two-dimensional space, with the week \( w \) and time \( t \) as the coordinates. In the paper, the time \( t \) is defined as the chronological time for one week, starting from Monday, 12:00 AM and ending at Sunday midnight. The past demand and SMP data are then defined on an input domain \( \Omega = \Omega_x \times \Omega_t \), where \( \Omega_x = [w, w - 1, w - 2, \ldots, w - n] \) and \( \Omega_t = [t, t - 1, t - 2, \ldots, t - m] \). Here \( n \) and \( m \) represent the data length in the respective coordinates.

Then, the SMP forecasting problem can be formulated as follows:

\[
y(w,t) = f(Y(w,t), D(w,t)) \tag{1}
\]

where

\[
Y(w,t) = \{ y(\overline{w}, \overline{t}) : (\overline{w}, \overline{t}) \in \Omega_x, (\overline{w}, \overline{t}) \neq (w,t) \}
\]

\[
D(w,t) = \{ d(\overline{w}, \overline{t}) : (\overline{w}, \overline{t}) \in \Omega_t \}
\]

\[
y(w,t) : \text{SMP in week } w, \text{ (chronological) time } t
\]

\[
d(w,t) : \text{power demand in week } w, \text{ time } t
\]

Input data is defined on a two-dimensional space. However, when we train a neural network, data need to be fed sequentially as a one-dimensional sequence.

III. ROUGH SET THEORY

Rough Set Theory was developed by Pawlak in the early 1980’s [5],[6]. The concept of the Rough Set is a new mathematical approach to imprecision, vagueness and uncertainty in data analysis. The Rough Set philosophy is founded on the assumption that with every object of the universe of discourse we associate some information [7].

Briefly, the relevant Rough Set terminology is stated below. An information system is a pair \( S = (U, A) \), where \( U \) is a non-empty and finite set, called the universe, and \( A \) is a non-empty, finite set of attributes. Every \( a \in A \) defines a total function: \( a : u \rightarrow V_a \), \( V_a \) being the range of \( a \). With every subset \( B \subseteq A \) we associate a binary relation (the indiscernibility relation), as follows:

\[
\text{IND}(B) = \{ (x, y) \in U^2 / \forall a \in B, a(x) = a(y) \} \tag{2}
\]

Every subset \( B \subseteq A \) is called an attribute. If \( B \) has only one element, it is a primitive attribute. Otherwise, it is a compound one [8].

In this paper, past SMP and the demand data are divided into five regions because SMP and the demand data are real number. If we use the data as it is, it is impossible to find the similar data pattern. Details on Rough Set Theory can be found in [5]-[8].

In this paper, the past SMP and demand data are divided into five regions. If we use the data as it is, it is impossible to find the similar data pattern. Table I represents the divided regions on the past SMP and demand data

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMP (won/kWh)</td>
<td>0~20</td>
<td>20~40</td>
<td>40~50</td>
<td>50~55</td>
<td>55~</td>
</tr>
<tr>
<td>Demand (MW)</td>
<td>0~25,000</td>
<td>25,000~30,000</td>
<td>30,000~35,000</td>
<td>35,000~40,000</td>
<td>40,000~</td>
</tr>
</tbody>
</table>

IV. ARTIFICIAL NEURAL NETWORKS

Many researchers have been studying human brain in order to realize a machine that can compute, judge and recognize like a human being. Artificial neural network (ANN), which mimics the human brain, has drawn much attention recently, because its massive parallel structure can be utilized in computation, which is much more efficient than in the traditional serial-type computer. ANN is applied in many field of study such as pattern recognition, noise filtering, forecasting, etc. [1]. In power systems, ANNs have already been used to solve problems such as load forecasting [13], component and system fault diagnosis, security assessment, unit commitment, etc. [9].

In this paper, a multilayer feed-forward neural network (or back-propagation neural network) is used to forecast SMP. Fig. 1 represents a basic structure of a three-layer back-propagation neural network with one input layer, one hidden
layer, and one output layer [9,12]. In the back-propagation neural network, each input unit receives an input signal and broadcasts this signal to neurons in the hidden layer. Each hidden unit then computes its activation and sends its signal to neurons in the output layer. Each output unit computes its activation to form the response of the ANN for the given input pattern. This activation is compared with a target value and then weights in the network are adjusted to reduce the difference between the activation and the target value.

V. FORECASTING OF SMP WITH NEURAL NETWORK

The neural network for forecasting SMP is affected by the activation function, the number of hidden layers, and the number of units in each layer. Moreover, even if the same input data and activation functions are used, it may lead to a very different result according to the manner in which the input data pattern is organized. Therefore, it is very important to find an appropriate manner of organizing the input data pattern in order to improve the efficiency of the algorithm. In this section, the Rough Set theory is applied to the input data pattern in order to group and combine the similar data pattern, and then the result data is used in training of the ANN to forecast SMP. Fig. 2 illustrates the procedure of forecasting the SMP with neural network adapted with the Rough Set theory.

The SMP forecasting is based on the observation that the hourly power demand in a given week exhibits similar pattern from week to week. For example, the load cycle in the second week in April, from Monday to Sunday, is repeated in the third week in April without much deviation. Another observation is that the hourly changes can be predicted from the pattern learned for a given week.

In this paper, Rough Set theory is applied to the given four weeks of data. The result is the reduced three, two or one pattern. This reduced pattern is used to train the ANN. Therefore, the efficiency is higher than the use of the patterns before applying the Rough Set theory.

The concept of this approach is depicted in Fig. 3, where 4 patterns of hourly SMP are given for 4 consecutive weeks, and an SMP for the following week is forecasted. The first 4 patterns are used to train the neural network, and the SMP for the current week for the next time instant is forecasted based on the SMPs in the previous time instants. This method reflects the characteristics of hourly, daily, and seasonal variation of power demand.
The following nonlinear model is proposed for the SMP forecasting:

\[ y(w,t) = F(W, y(w,t-1), y(w,t-2),..., y(w,t-m),
\]
\[ d(w,t), d(w,t-1), d(w,t-2),..., d(w,t-m)) \]

where

\( y(w,t) \): SMP in week \( w \), at (chronological) time \( t \)
\( d(w,t) \): power demand in week \( w \), time \( t \)
\( W \): weight matrix of ANN
\( m \): index for input data length
\( F(.) \): nonlinear function representing ANN

The weight matrix in (4) is adjusted in the learning mode of ANN by applying a number of SMP patterns. For example, in Fig. 3, four SMP patterns corresponding to weeks \( \{w_i: i=1,2,3,4\} \) are used to train the ANN, where each pattern has the input data length of four \( (m=4) \).

Once the weight matrix at time \( t \) is estimated, the SMP is forecasted with the SMPs of previous hours as well as the power demand as follows:

\[ \hat{y}(w,t) = F(\hat{W}, y(w,t-1), y(w,t-2),..., y(w,t-m),
\]
\[ d(w,t), d(w,t-1), d(w,t-2),..., d(w,t-m)) \]

where \( \hat{W} \) denotes the weight estimate and \( \hat{y}(w,t) \) indicates the SMP forecast for week \( w \), time \( t \).

VI. CASE STUDY

The proposed forecasting procedure was implemented using the past SMP and demand data obtained from the Korea Power Exchange (KPX) for the month of April 2002.

Table II presents the SMP forecasting error for a month. RS means the Rough Set theory in table II. Although there are some exceptions, the SMP forecasting adapted with the Rough Set theory shows better results in general. The average forecasting error for the case with RS and without RS are 6.04% and 5.92%, respectively. If we exclude Monday, holiday, and the day after holiday, the average forecasting error for the case with RS and without RS are 4.70% and 4.81%, respectively. Therefore, it can be concluded that the accuracy of SMP forecasting can be improved by adapting the Rough Set theory.

In holiday, demand is between 25,000MW and 32,000MW for a whole day in a region that have generators using different fuel types. Therefore, although the demand change is small, SMP changes sharply in this region, depending on which fuel type a generator uses.

Fig. 4 shows the actual SMP and the forecasted SMP that is predicted by the ANN adapted with the Rough Set theory for one month, April 2002. In the figure, pattern of the forecasted SMP is very similar to, or coincide with the pattern of the actual SMP; but when the actual SMP rises or fall sharply, forecasted SMP cannot follow the actual pattern.

However, we can observe that the forecasted SMP with the proposed technique is very close to the actual SMP in general. Also, judging from the results shown in Table II, and Fig. 4, the technique proposed in this paper can provide useful information to market participants.

VII. CONCLUSIONS

In this paper, the system marginal price (SMP) is forecasted using a back-propagation neural network (NN) adapted with the Rough Set theory.

<table>
<thead>
<tr>
<th>Day</th>
<th>Monday</th>
<th>Tuesday</th>
<th>Wednesday</th>
<th>Thursday</th>
<th>Friday</th>
<th>Saturday</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
</tr>
<tr>
<td>Method</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
</tr>
<tr>
<td>Err(%)</td>
<td>4.84</td>
<td>3.06</td>
<td>2.68</td>
<td>2.71</td>
<td>2.53</td>
<td>3.42</td>
</tr>
<tr>
<td>Method</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
</tr>
<tr>
<td>Err(%)</td>
<td>11.38</td>
<td>11.41</td>
<td>4.29</td>
<td>4.15</td>
<td>2.45</td>
<td>2.38</td>
</tr>
<tr>
<td>Method</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
</tr>
<tr>
<td>Err(%)</td>
<td>5.23</td>
<td>5.17</td>
<td>2.42</td>
<td>2.48</td>
<td>1.95</td>
<td>1.95</td>
</tr>
<tr>
<td>Method</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
<td>with RS</td>
<td>without RS</td>
</tr>
<tr>
<td>Err(%)</td>
<td>3.71</td>
<td>3.75</td>
<td>6.12</td>
<td>5.86</td>
<td>5.03</td>
<td>4.82</td>
</tr>
</tbody>
</table>

TABLE II

SMP FORECASTING ERROR FOR APRIL 2002
The power demand and the past SMP data obtained from the Korea Power Exchange (KPX) are used as input data for the ANN.

Input data are classified into similar patterns, and the resulting patterns are used to train the ANN. In the result, the SMP forecasting adapted Rough Set theory shows better forecasting error.

Except on holiday, the day after holiday and weekends, the forecasting error is very small. Therefore, the proposed technique can be applied to a real power market for short-term price forecasting, and provide a useful information to market participants in establishing optimal strategies.

As a future work, a more accurate method should be developed to forecast for Monday, weekend and holiday.
Also, a technique that can search or eliminate the bad data needs to be developed.

VIII. ACKNOWLEDGEMENT
This work was financially supported by MOCIE through IERC program.

IX. REFERENCES

BIOGRAPHIES

Jeong-Kyu Lee received his B.S. and M.S. degrees in Electrical Engineering from Konkuk University in 2000 and 2003, respectively. He was at the Pennsylvania State University as a visiting intern in 2004. Currently, he is enrolled in a doctoral program in Konkuk University. His major research topics include power system operation and economics, and application of intelligent system techniques in power systems.

Joong-Bae Park received his B.S., M.S., and Ph.D. degrees in Electrical Engineering from Seoul National University in 1987, 1989, and 1998, respectively. For 1989-1998, he was with Korea Electric Power Corporation, and for 1998-2001 he was an Assistant Professor at Anyang University, Korea. Currently, he is with the Department of Electrical Engineering at Konkuk University, Seoul, Korea. His major research topics include power system operation, planning, economics, and energy markets.

Kwang Y. Lee received the B.S. degree in Electrical Engineering from Seoul National University, Seoul, Korea, in 1964, the M.S. degree in Electrical Engineering from North Dakota State University, Fargo, ND, in 1967, and the Ph. D. degree in Systems Science from Michigan State University, East Lansing, in 1971. He is currently a Professor of Electrical Engineering at the Pennsylvania state University, University Park, PA. His research interests include control theory, computational intelligence and their application to power systems. Dr. Lee is currently the Director of Power Systems Control Laboratory at Penn State. He is a Fellow of IEEE, an Associate Editor of IEEE Transactions on Neural Networks, and Editor of IEEE Transactions on Energy Conversion.

Joong-Rin Shin received his B.S., M.S., and Ph.D. degrees in Electrical Engineering from Seoul National University in 1977, 1984, and 1989, respectively. For 1977-1990, he was with Korea Electric Power Corporation as a research staff. Since 1990, he has been with Konkuk University, Seoul, Korea, where he is currently a Professor of Electrical Engineering. His major research field is in power system operation and planning.