Application of Rough Set Theory on Distribution System

ZENG Li hua¹, FENG Juan²
1 College of Mechanical and Electrical Engineering, 2 College of Information Science and Technology, Agricultural University of Hebei, P.R.China, 071001
lihuazeng@yahoo.com.cn

Abstract: This paper introduced the feature of distribution network and rough set theory. The application in power system was elaborated, such as on load forecasting, fault diagnosis, system-state analysis and data mining. Then given an illustration that using RS attribute reduction algorithm obtain the correlative factors in distribution system load forecasting. These factors were input vector of neural network. This simplified the network structure greatly, reduced forecasting time and improved forecasting efficiency.

Keywords: Rough Set, Distribution Network, Load Forecasting, Neural network

1 Introduction

There are lots of different between the distribution network and the transmission network. For example, the structure of the distribution network is a closed loop, while the working state is opened. The structure is radiating on its steady state, and short time working with closed loop is possible while the something is out of order or rotating load. Furthermore, the total length of distribution network is longer than transmission network, the branch line is more, and the wire size is smaller. These bring on the higher ratio of R/X (Resistance/ reactance), always more than one. So the capacitance of the line can be ignoring. There are more PQ nodes and less PV nodes than the transmission network. In addition, it is connected with the user directly. The reliability and quality of the power supply is referred directly to the national economy and peoples’ daily life. Consequently, the distribution system is a multidimensional dynamic system, which couldn’t establish accurate mathematical model.

2 Rough Set Theory

Rough set theory (RST), proposed by Zdzislaw Pawlak in the early 1980's, is a mathematics tool for dealing with uncertainty and vagueness aspects of knowledge model, discovering and revealing the potential rules. The notable feature is that it could use merely the information offered by the data itself and does not need any else prior knowledge. The main research content is based on the data processing of decision table. As follows: data preprocessing, making the entire decision table, which have default value completion and discretization, simplifying the decision table while keeping the relation between condition and decision, and so on. Rough sets can be used for feature selection, feature extraction, data reduction, decision rule generation, and pattern extraction (templates, association rules) etc.

However, the study on power system appears in recent years. Until 1997, Lambert, a Brazil scholar, published the first paper about the application to power system based on RS theory[1]. Nevertheless, lately the researches on power system come forth increasingly. The main fruits of RS are surveyed in equipment fault diagnosis, fault diagnosis of electric distribution network, transient stability classification, data mining etc. This paper mainly introduced the research status on local power system, and illustrated the application of load forecasting in distribution system.

3. The research status of Rough Set Theory in power system

Rough Set Theory is an effective mathematical tool for dealing with uncertain and incomplete data, especially adapting for complex system. Application of the theory in power system is very important for improving capacity and efficiency of data mining.[2]

3.1 Load forecasting
Load forecasting predicts the market demand of power. In the planning and running management, it decides to arrange reasonably in the aspects of power generation, power transmission and distribution of electric power. And it is all-important to ensure safe and economical operation and the development of national economy.

To find the potential rules in a great deal of historical data is useful for load forecasting in power system. At present, most of the forecasting models are based on neural network. But many factors influencing load forecasting in power system are uncertain. Therefore, prediction errors of many methods are larger. The data mining based on Rough Set Theory is applied to load forecasting in order to reduce the error of prediction brought by dirty data and incomplete data. Then using the reduction algorithm of rough set theory, the most factors to load forecasting would be found. This would reduce forecasting time and improve forecasting efficiency.\(^\text{[3]}\)

### 3.2 Fault diagnosis

#### 3.2.1 Equipment Fault diagnosis

The fault diagnosis of power equipment is a problem on pattern classification. It is difficult to establish exact mathematical model between fault phenomena and fault cause. Furthermore, the fault information obtained on the spot has often a great deal of noises. As far as the fault is concerned, one fault signature may embody in many measurement signals, i.e., there are redundancies at some degree among the fault symptoms.

As a new method of dealing with incomplete information, it makes use of reduction theory in fault diagnosis of power equipment. It is known that obtaining complete knowledge base is the bottleneck of fault diagnosis expert system. If the knowledge database established is incomplete, it may lead the confusion of expert reasoning and draw a wrong conclusion. Therefore, applying RS theory establishes and maintains knowledge base of fault diagnosis expert system of power transformer in reference \([4]\). If was a new thought of resolving the bottleneck problem on complete knowledge obtained difficulty by expert system. Moreover, the reference \([5]\) presented that combination between RS theory and fuzzy-neural network is used for transformer’s fault diagnosis. The thought is making use of RS method extracting simplified rules, structuring a fuzzy-neural network based on these rules in order to reduce network scope greatly, improve learning speed and overcome the problem of being difficult for convergence on neural network training at the time of numerous samples.

#### 3.2.2 Fault diagnosis of distribution network

Because the fault information used for distribution diagnosis, was from the terminal unit that out of door. The operation environment was bad. And the probability of equipment broken or information missing was high. In addition, the circuit breaker or protector misoperation, communication set failure etc, would bring about the imperfect alarm signal. Rough set theory is good at dealing with the imperfect data. So it could be using the RS theory and artificial intelligence technology to complete the fault diagnosis of complex system.

A method was proposed in reference \([6]\). Making use of improved discernible matrix, the fault samples are reduced and the reduced samples are layered according to the value of weighted mean roughness. Thereby a fault diagnosis layer model was established. With the proposed method the blindness and redundancy of layering can be avoided and the space for diagnosis model is evidently diminished. In addition, the weighted mean roughness is easy to be calculated and realized. In the fault diagnosis system of distribution system, it is very important to make unified reduction of continuous and discretizing signals with the rough sets at first in reference \([7]\). In this way, it can make the important analogues of distribution system such as voltage, current take part in the fault diagnosis, and improve the precision of fault identification.

#### 3.3 System-state analysis

Rough Set Theory can classify the sample data based on indistinguishable relation. And the classification process needn’t the domain knowledge or else extra information. Applying RS Theory to analyses and judgment of power system has a better objectivity.
3.4 Data mining

Data mining is a quite active research field in the world and is the outcome of combination between artificial intelligence and database technology. DM makes using of special algorithm to find potential rules in a great deal of data, withdraw useful knowledge and resolve deficient knowledge conflicting with great quantity of data.

Under the condition of being lack of prior knowledge of data, RS theory can only make a base on classified ability of observed data to resolve the analysis and management of fuzzy or uncertain data and reveal potential rules. It has been one of main methods in current DM. Moreover, the decision table is similar to the relational database on representation. This would bring great convenience on actual application. For example, using RS much key information was extracted from real-time database of power system control center, which helps operator to master system state quickly\[8\]. Li qiudan applied rough set theory and artificial neural network for load forecasting \[9\]. Moreover, it was given in the reference \[10\] that a fault diagnosis model based on generalized RS and artificial neural network, which applied in the distribution network.

4 Load forecasting of distribution system

Power load of Distribution system has an influence on all kinds of complex factors, which are also random variation. Therefore, the first step is to determine the type of forecast: the long-term load forecasting or the short-term load forecasting, and the qualitative forecast or quantitative forecast and so on. Then according to the specific condition of different area and different time, make sure the reasonable forecast content and forecast target.

4.1 Data source

After determining the forecast content, it is essential to collect systemic and all-around information, which is required. Then analyze and process the data according to the forecast target. Take the data (including load and weather information) of the August of a county in recent years as the basic data set. Thirty nine influencing factors were considered about the predicted day and the previous week, which include the date type, the hours of sunlight, maximum temperature, minimum temperature, mean temperature, maximum humidity, minimum humidity, weather condition and so on. Where the date type is divided into two types as follows: working day and holiday.

4.2 Data preprocessing

The data that we have collected unavoidably have exceptional value. For example, sometimes would black out. Furthermore, some data would lose or inaccurate. Therefore, it is very important to improve the load forecast precision by identifying the poor data accurately and completing the lack data rightly. According to RS theory, this paper completes the exceptional data with the most common or the mean value.

4.3 Attribute discretization

Some attribute that described with literal were discontinuous, such as date type, wind direction and weather condition and so on. Hence, discretization was necessary. For example, “1” represents working day and “2” represents holiday in the date type. Considering northern area, the wind direction was classified with historical experience summarized by expert and replaced literal with number. The classification was shown in Table1. The classification of weather condition was indicated in Table2.
Table 1: Classification of wind direction

<table>
<thead>
<tr>
<th>Class</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wind Direction</td>
<td>South</td>
<td>Southeast</td>
<td>Southwest</td>
<td>East</td>
<td>West</td>
<td>North</td>
<td>Northeast</td>
<td>Northwest</td>
<td>Windless</td>
</tr>
</tbody>
</table>

Table 2: Classification of weather condition

<table>
<thead>
<tr>
<th>Serial Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather Case</td>
<td>Clear (cloud cover &lt;50%)</td>
<td>Cloudy (cloud cover 50%-90%)</td>
<td>Overcast (cloud cover &gt;90%)</td>
<td>Light rain (daily precipitation &lt;10mm)</td>
<td>Moderate rain (daily precipitation 10-25mm)</td>
<td>Heavy rain (daily precipitation &gt;25mm)</td>
<td>Thunderstorm</td>
</tr>
</tbody>
</table>

4.4 Attribute reduction

After discretization, analysis of attribute reduction can be done for condition attribute with Rough Set theory. Then find the factors, which have direct correlation with load. Following the above steps, five influencing factors were screened out. These were weather condition of predicted day, data type of predicted day, average temperature of predicted day, average load of previous day and precipitation of previous week.

4.5 Neural network training and forecast

The attribute reduction set was the input of neural network. The value of corresponding load was output for training. Then using the trained neural network model predicts the load. In this way, the input of neural network was reduced highly, the structure of network was simplified greatly, and the training sample deleted redundant information was typical one. The whole create condition to improve the efficiency of network training. The result of load forecasting and the relative error of some distribution network on these dates from 1st to 15th of August were shown in Table 3. The relative error is in 5 percent and predicted effect was preferable.

Table 3: An example of load forecasting

<table>
<thead>
<tr>
<th>Serial Number</th>
<th>Date</th>
<th>Factual peak load (kw.)</th>
<th>Predicted peak load (kw.)</th>
<th>Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>August 1</td>
<td>57855</td>
<td>58365</td>
<td>0.88%</td>
</tr>
<tr>
<td>2</td>
<td>August 2</td>
<td>57640</td>
<td>58480</td>
<td>1.46%</td>
</tr>
<tr>
<td>3</td>
<td>August 3</td>
<td>58645</td>
<td>59650</td>
<td>1.71%</td>
</tr>
<tr>
<td>4</td>
<td>August 4</td>
<td>59275</td>
<td>59650</td>
<td>0.63%</td>
</tr>
<tr>
<td>5</td>
<td>August 5</td>
<td>58730</td>
<td>60250</td>
<td>2.59%</td>
</tr>
<tr>
<td>6</td>
<td>August 6</td>
<td>57130</td>
<td>58360</td>
<td>2.15%</td>
</tr>
<tr>
<td>7</td>
<td>August 7</td>
<td>58955</td>
<td>57605</td>
<td>2.29%</td>
</tr>
<tr>
<td>8</td>
<td>August 8</td>
<td>58680</td>
<td>59220</td>
<td>0.92%</td>
</tr>
<tr>
<td>9</td>
<td>August 9</td>
<td>58635</td>
<td>59890</td>
<td>2.14%</td>
</tr>
<tr>
<td>10</td>
<td>August 10</td>
<td>58685</td>
<td>59875</td>
<td>2.03%</td>
</tr>
<tr>
<td>11</td>
<td>August 11</td>
<td>58420</td>
<td>59655</td>
<td>2.11%</td>
</tr>
<tr>
<td>12</td>
<td>August 12</td>
<td>57110</td>
<td>58410</td>
<td>2.28%</td>
</tr>
<tr>
<td>13</td>
<td>August 13</td>
<td>56000</td>
<td>56550</td>
<td>0.98%</td>
</tr>
<tr>
<td>14</td>
<td>August 14</td>
<td>58405</td>
<td>59165</td>
<td>1.30%</td>
</tr>
<tr>
<td>15</td>
<td>August 15</td>
<td>58650</td>
<td>56720</td>
<td>3.29%</td>
</tr>
</tbody>
</table>

5 Conclusions

This paper introduced the feather of local distribution network and RS theory, elaborated the application of RS in power system and put forward a example of distribution network load forecasting based on RS (i.e. using RS theory attribute reduction algorithm obtains the most correlative factor with load and applies to load forecasting). From the simulation result, you could see that useful information
was discovered in the historical sample data by the rough set theory attribute reduction algorithm, avoided selecting input vector blindly and improved the accuracy and velocity of load forecasting. In a word, RS theory is a tool dealing with complicated system data; it makes use of information offered by the data itself and doesn’t need else prior knowledge, on the premise of keeping the dependency relationship between decision making attribute and condition attribute to simplify the decision table. In addition, RS could be compatible with other artificial intelligence theory. Therefore, RS theory has a good prospect in the field of power system.

Reference