Distributed Power Grid Fault Diagnosis Based on Naive Bayesian Network and D-S Evidence Theory

Xiaoqin Liu1,a, Xing Yang1, Chengyu Li1, Jinsong Liu2, Fengwei Zhang3

1School of Information and Control Engineer Liaoning Shihua University, Fushun Liaoning 113001, China
2Shanxi Jingneng Lv Lin Electric Power Co.,Ltd ,Lv Lin Shanxi 033000,China
3State grid huludao electric power supply company,Huludao Liaoning 125000,China

aqinbuluoge@163.com

Abstract. On the basis of Bayesian network, evidence theory is introduced, a kind of distributed power network fault diagnosis method based on Naive Bayesian network and D-S evidence theory is proposed. Firstly, the fault area is determined by the real-time connection method, the fault region is segmented by the butterfly segmentation method. Secondly, according to the historical fault samples, the decision table is established, and knowledge reduction based on Rough Set Theory. The naive Bayesian network model is constructed according the best combination, and the probability of each node is trained. Finally, D-S evidence fusion is performed on the fault component diagnosis information in the overlap between the subnets. The simulation results show that the proposed method can reduce the complexity of modeling and improve the fault tolerance of the system in the condition of incomplete information, and has good diagnosis results.

1. Introduction

When the power grid fails, a large amount of information will flow into the dispatching center. The lack of information itself and the fault of staff may lead to the inaccurate transmission of information, which will affect the final diagnosis results. With the development of artificial intelligence technology, fault diagnosis technology based on artificial intelligence method has become a research hotspot of current scholars, such as expert system [1], fuzzy theory [2], artificial neural network [3], Petri net [4-6], rough set theory [7], information theory [8]. Applying Bayesian network based on probability reasoning to power grid fault diagnosis can solve the uncertainty problem caused by incomplete information in the reasoning process, and improve the fault tolerance of the system [9,10]. In reference [11], rough set theory and naive Bayesian network are combined to solve the problems of uncertainty and low fault tolerance of traditional fault diagnosis methods, and the diagnosis effect is not good when the power grid connecting line problem occurs. In reference [12], D-S evidence theory is applied to Bayesian network diagnosis model, which can solve the problem of overlapping area fault in power grid division, but the establishment of Bayesian network model is more complex.

This paper attempts to solve these problems by combining naive Bayesian network with D-S evidence theory. Firstly, the real-time wiring method is used to reduce the diagnosis area; secondly, the butterfly segmentation method is used to divide the fault area into several subnets with similar
calculation burden; then, the decision table is established according to the historical fault samples, and the optimal attribute reduction combination is obtained based on the attribute reduction algorithm of rough set, so as to build the naive Bayesian network model of each subnet; finally, for those in the electric field D-S evidence fusion is carried out for the fault in the overlapped part of the network, and the most accurate diagnosis conclusion is obtained by analyzing the fault feature diagnosis results in the overlapped area.

2. **Naive Bayesian classifier**

It is assumed that the conditions among the attribute variables of naive Bayesian classifier are independent, and each attribute node is only related to the class node C, as shown in Figure 1.

![Figure 1. Naive Bayesian classifier](image)

Because of the reduction of network layers, the complexity of building Bayesian network model is reduced exponentially, which can be expressed as follows:

$$P(C_k \mid x_1, x_2, \ldots, x_n) = \frac{P(C_k) \sum_{i=1}^{n} P(x_i \mid C_k)}{P(x_1, x_2, \ldots, x_n)}$$  \hspace{1cm} (1)

In formula (1), $P(x_1, x_2, \ldots, x_n)$ for all fault classes are constant, only the one with $P(C_k) \sum_{i=1}^{n} P(x_i \mid C_k)$ the largest probability value needs to be obtained and screened. For prior probability $P(C_k)$, the following formula can be used in calculation:

$$P(C_k) = \frac{N_{C_k}}{N}$$  \hspace{1cm} (2)

Where, $k = 1, 2, \ldots, m, N$ represents the total number of all training samples, and $N_{C_k}$ represents the number of samples of fault types $C_k$ in all training samples. When the number of attributes is very large, when calculating the maximum posterior probability $P(X \mid C_k)$, it will increase the calculation cost, and in order to reduce the impact of this situation. Each attribute of naive Bayesian network $x_1, x_2, \ldots, x_n$ is assumed to be conditionally independent from each other, only related to fault type $C$, and the formula is as follows:

$$P(X \mid C_k) = P(x_1, x_2, \ldots, x_n \mid C_k) = \sum_{i=1}^{n} P(x_i \mid C_k)$$  \hspace{1cm} (3)

Among them, the probability represented $P(x_i \mid C_k)$ by symbols can be calculated according to the training samples, namely:
\[ P(C_k) = \frac{N^h_{C_k}}{N_{C_k}} \]  

(4)

In the formula, \( N^h_{C_k} \) represents the number of samples satisfying both fault characteristics \( C_k \) and condition attributes \( x_i \) in the training samples \( P(C_k) = \frac{N^h_{C_k}}{N_{C_k}} \). If it does not exist in the calculation \( N^h_{C_k} \), i.e. formula (4) cannot be met, the following formula is used:

\[ P(C_k) = \frac{1/N}{N_{C_k} + N_{x_i}/N} \]  

(5)

According to the above calculation principle, for the probability \( P(X|C_i)p(C_i) \) of each fault class of Bayesian network, the one with the largest probability value \( C_i \) corresponds to the fault class of diagnosis.

3. Rough set theory

The definition of discernible matrix is as follows: discernible matrix of decision table \( S \) \( M_D = \{m_{ij}\} \). Is a \( n \times n \) matrix of order, the elements of row \( i \), column \( j \) are obtained by comparing row \( i \) and row \( j \) of the decision table, and are defined as follows:

\[ m_{ij} = \begin{cases} 
\{ a \in A : f(a_i, a_j) \neq f(a_j, a_i) \} & f(a_i, d) \neq f(a_j, d) \\
0 & f(a_i, d) = f(a_j, d) \\
-1 & f(a_i, a_j) = f(a_j, a_i) \text{ and } f(a_i, d) \neq f(a_j, d) 
\end{cases} \]  

(6)

Equation (6) shows that three values constitute the elements in the discernible matrix: when the two attributes in the compared decision table have different conditional attribute values and different decision attribute values, different conditional attribute values are assigned to the element; when the two attributes in the compared decision table have the same decision attribute values, the element is assigned 0; when the two attributes in the compared decision table have the same decision attribute values, the element is assigned 0; when the two attributes in the compared decision table have the same decision attribute values, the element is assigned 0. The two attributes in the decision table have the same condition attributes, but the decision attributes are different, which indicates that the two attributes are in conflict. Assign -1 to this element.

In general, multiple attribute reduction combinations can be obtained through the above reduction process. In order to get the best attribute reduction combination from these combinations, this paper calculates the average mutual information of each reduction combination, and selects a group corresponding to the minimum average mutual information as the best attribute reduction combination. Mutual information based on information entropy is defined as [13].

If the information of equivalence relation \( G(U|I(G)) = \{x_1, x_2, ..., x_n\} \) is expressed by \( H(G) \), and the condition of equivalence relation \( Q(U|I(Q)) = \{y_1, y_2, ..., y_n\} \) relative to equivalence relation \( G \) is expressed by \( H(Q/G) \), then the mutual information between equivalence relation \( G \) and \( Q \) is defined as:

\[ I(Q,G) = H(Q) - H(Q/G) = \sum P(x,y) \log_2 \frac{P(x,y)}{P(x)P(y)} \]  

(7)
4. Fault diagnosis method of distributed power grid

4.1. Basic ideas
When the fault occurs, the real-time connection method is used to determine the fault area, and the butterfly grid segmentation method is used to segment the fault area. According to the historical fault samples of power grid, the decision table is established. Based on the rough set theory, the knowledge reduction of the decision table is carried out, and the optimal reduction combination of each subnet is obtained, and then the naive Bayesian network model is established. Analyze the subnets and fault areas. For the fault features that are not overlapped in subnets, analyze and reason them directly. For the fault features that are overlapped in subnets, do parallel diagnosis in subnets respectively, and then do D-S evidence fusion.

4.2. Fault diagnosis process and some relate technologies of distributed power grid
The fault diagnosis flow of distributed power grid based on Naive Bayesian network and D-S evidence theory is shown in Figure 2. For example, when the power grid has faults, the topology of the system before and after the faults will be identified based on the circuit breaker information collected in real time, combined with the real-time wiring method, and the two will be compared, so as to find the passive network generated after the faults. The fault components must be in these passive networks, and these passive networks will be divided into fault areas, which will greatly reduce the fault component's Number of diagnosis objects, improve diagnosis speed.

4.3. Knowledge extraction based on Rough Set Theory
The decision table is established according to the historical fault samples of power grid. Firstly, for a local power grid, the training sample set is selected according to the historical fault information. Then, taking the action information of protection and circuit breaker as the condition attribute and the fault component as the decision attribute, the decision table is established; according to the knowledge reduction method of rough set, the optimal attribute reduction combination is obtained.

4.3.1. Establishment of naive Bayesian network model. The decision attribute in the optimal reduction combination decision table is used as the parent node of the naive Bayesian network, and the condition attribute in the optimal reduction combination decision table is used as the child node to establish the
naive Bayesian network; the prior probability of each parent node is calculated by formula (2); the condition probability of each child node is calculated by formula (4) and formula (5) according to the situation.

4.3.2. D-S evidence theory. Construction of identification framework: in the distributed power grid fault model, since D-S evidence theory is only used to fuse the diagnosis results of overlapping components in each subnet, the identification framework is the overlapping components in the power grid. If butterfly segmentation is adopted, the corresponding identification framework components are the line components and transformer components;

The construction of basic reliability function: D-S evidence theory is applied to the fusion of diagnosis results of subnets. The evidence of the theory is composed of the diagnosis results of Bayesian networks of subnets. Each subnet is equivalent to each evidence element in the evidence space; The combination rule of two reliability functions: suppose that sum is the basic probability assignment function $m_1$ and $m_2$ of independent events in sample space $\Omega$, and its kernel is sum respectively $\{A_1, A_2, \ldots, A_n\}$ and $\{B_1, B_2, \ldots, B_m\}$, then the composite rule of evidence theory is used to express the composite $m = m_1 + m_2$ as follows:

$$m(A) = k^{-1} \sum_{A_1 \cap A_2 = A} m_1(A_1)m_2(A_2)$$

(8)

Where $m(\phi) = 0$, at that time of $A \neq \phi$, $K$ can be expressed as the formula (9).

$$k = 1 - \sum_{A_1 \cap A_2 = \phi} m_1(A_1)m_2(A_2)$$

(9)

5. Example

5.1. Distribution network structure diagram

Figure 3 shows a simple distribution network. The system has six areas, Sec1, sec2, Sec3, Sec4, sec5 and sec6, which are respectively equipped with overcurrent protection CO1, CO2, CO3, Co4, CO5 and CO6. Sec1 is equipped with distance protection rr1 to provide back-up protection for sec2 and Sec3. Sec5 is equipped with distance protection RR5 to provide back-up protection for sec2 and Sec4. Similarly, sec6 is equipped with distance protection RR6 to provide back-up protection for Sec3 and Sec4. CB1, CB2, CB3, CB4, CB5 and CB6 are circuit breakers. The grid is divided by the method of ground butterfly [14], as shown in Figure 3.

Figure 3. Simple distribution example and butterfly partition diagram
5.2. Establish decision table based on grid fault samples
The butterfly partition method is used to divide the fault area into several subnets with similar calculation burden. According to the corresponding historical fault samples, a decision table is established for each subnet with the same method. Table 1 is a decision table for the divided subnet M1, which is composed of seven grid fault samples. The decision table is established with the protection and breaker action information as the condition attribute and the fault area as the decision attribute S. The other two subnets m2 and M3 establish decision tables in the same way.

Table 1. Fault diagnosis decision table

| Sample | Condition properties (fault information) | Fault area |
|--------|------------------------------------------|------------|
|        | CB1 | CB2 | CB3 | CO1 | CO2 | CO3 | RR1 |          |
| 1      | 1   | 0   | 0   | 1   | 0   | 0   | 0   | Sec1     |
| 2      | 0   | 1   | 0   | 0   | 1   | 0   | 0   | Sec2     |
| 3      | 0   | 0   | 1   | 0   | 0   | 1   | 0   | Sec3     |
| 4      | 1   | 0   | 0   | 0   | 1   | 0   | 1   | Sec2     |
| 5      | 1   | 0   | 0   | 0   | 0   | 1   | 1   | Sec3     |
| 6      | 1   | 0   | 0   | 0   | 0   | 0   | 1   | Sec2/3   |
| 7      | 0   | 0   | 0   | 0   | 0   | 0   | 0   | No       |

5.3. Attribute reduction based on discernible matrix and information entropy
It can be seen from table 1 that there are 7 samples in the decision table established for the subnet M1 according to the historical fault samples, so a 7×7 order discernible matrix can be constructed accordingly. According to the discernible matrix reduction rules and formula (6), the discernible matrix elements are assigned according to the conditions in the subnet M1 decision table s, and the discernible matrix $M'_D$ is obtained as follows.

Table 2. The discernible matrix $M'_D$

|          | CB1,CB2,C | CB1,CB3,C | CO1,RR1,CO2 | CO1,RR1,CO3 | CO1,RR1,CO3 | CB1,CB2,C |
|----------|------------|------------|--------------|--------------|--------------|------------|
| 0        | CO1,CO2    | CO1,CO3    | CB1,CB2,CO2  | CB1,CB3,CO3 | CB1,CO1     | CB2,CO2    |
|          | CB2,CB3,C  | CO2,CO3    | CB1,CB3,CO2  | CB1,CO1     | CB3,CO3     | CB1,RR1    |
|          | O1,CO3     | CB1,CO1    | CB1,CB2,CO2  | CB1,CO3     | CB1,RR1,C   |
|          | CB2,CO3    | CO3        | CB1,CO3     | CB1,RR1,C   |
|          | CB1,CO1    | 0          | CO2,CO3     | CB1,RR1,C   |
|          | CB1,CO3    | 0          | CO3         | CB1,RR1,C   |
|          | 0          | 0          | 0           | CB1,RR1,C   |
|          | 0          | 0          | 0           | CB1,RR1,CO3 |

Find the core attributes CO2 and CO3, then find the combination without core attributes according to the method of knowledge extraction, and get the reduced discernible matrix $M'_D$ as follows:

Find the core attributes CO2 and CO3, then find the combination without core attributes according to the method of knowledge extraction, and get the reduced discernible matrix $M'_D$ as follows:
Table 3. The reduced discernible matrix $M_D'$

|   |   |   |   |   | CO1, RR1, CB1, CO1 |
|---|---|---|---|---|-------------------|
|   |   |   |   |   | 0                 |
| 0 | 0 | 0 | 0 | 0 | 0                 |
| 0 | 0 | 0 | 0 | 0 | 0                 |
| 0 | 0 | 0 | 0 | 0 | 0                 |
| 0 | 0 | 0 | 0 | 0 | 0                 |

Three sets of attribute combinations $M_D'$ without kernel attributes can be clearly obtained:

1. CO1 RR1
2. CB1 CO1
3. CB1 RR1

After the transformation of conjunctive normal form and disjunctive normal form, the final disjunctive form and core attribute form attribute reduction combination:

1. (CO1, CB1, CO2, CO3)
2. (RR1, CB1, CO2, CO3)
3. (CO1, RR1, CO2, CO3)

According to the combination of the three attribute reduction, the corresponding three reduction decision tables can be obtained. The corresponding reduction decisions table (CO1, CB1, CO2, CO3) is listed, as shown in Table 4.

Table 4. Attribute reduction

| sample | CO1 | CB1 | CO2 | CO3 | Fault area |
|--------|-----|-----|-----|-----|------------|
| 1      | 1   | 1   | 0   | 0   | Sec1       |
| 2      | 0   | 0   | 1   | 0   | Sec2       |
| 3      | 0   | 0   | 0   | 1   | Sec3       |
| 4      | 0   | 1   | 1   | 0   | Sec2       |
| 5      | 0   | 1   | 0   | 1   | Sec3       |
| 6      | 0   | 1   | 0   | 0   | Sec2/3     |
| 7      | 0   | 0   | 0   | 0   | No         |

The basic steps of selecting the optimal reduction combination from three reduction combinations are as follows: first, calculate the mutual information between two attributes in each group’s attribute reduction combination, then add them to get the product mean value, and get the average mutual information of each group. Finally, compare the average mutual information of each group, and the smallest is the most simplified form.

The mutual information between the two attributes in the reduction combination mutual information is calculated from formula (9), and then the average value is obtained by adding and summing to the average mutual information of the reduction combination. The average mutual information results of the three reduction combinations are as follows:

The average mutual information of reduction combination (CO1, CB1, CO2, CO3) is 0.0772;
The average mutual information of reduction combination (rr1, CB1, CO2, CO3) is 0.0869;
The average mutual information of reduction combination (CO1, rr1, CO2, CO3) is 0.1192.

If the average mutual information of reduction combination (CO1, CB1, CO2, CO3) is the least, it is the best reduction combination.
5.4. naive Bayesian network modeling and node probability calculation

After the optimal reduction combination is obtained by the above reduction methods, five naive Bayesian network models can be established with the fault areas Sec1, sec2, Sec3, sec2 / 3 and no in the attribute reduction table as the parent nodes and the conditional attributes in the optimal attribute reduction combination as the child nodes, as shown in Figure 4. Set Q1 = Sec1, Q2 = Sec2, Q3 = Sec3, Q4 = Sec2 / 3, Q5 = No.

![Figure 4. Naive Bayesian network model](image)

According to formula (2), the prior probability of each parent node is calculated as shown in Table 5.

| Parent node | P(Q1) | P(Q2) | P(Q3) | P(Q4) | P(Q5) |
|-------------|-------|-------|-------|-------|-------|
| Prior probability | 1/7 | 2/7 | 2/7 | 1/7 | 1/7 |

The conditional probabilities of each sub node calculated by formulas (4) and (5) are shown in Table 6 below.

| Fault Q1 node | P(CO1=0|Q1) | P(CO1=1|Q1) | P(CB1=0|Q1) | P(CB1=1|Q1) | P(CO2=0|Q1) | P(CO2=1|Q1) | P(CO3=0|Q1) | P(CO3=1|Q1) |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| conditional probability | 1/13 | 1/10 | 1 | 1 | 1/9 | 1 | 1/9 | 1 |

| Fault Q2 node | P(CO1=0|Q2) | P(CO1=1|Q2) | P(CB1=0|Q2) | P(CB1=1|Q2) | P(CO2=0|Q2) | P(CO2=1|Q2) | P(CO3=0|Q2) | P(CO3=1|Q2) |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|
| conditional probability | 1 | 1/15 | 1/2 | 1/2 | 1/19 | 1 | 1 | 1/16 |

| Fault Q3 node | P(CO1=0|Q3) | P(CO1=1|Q3) | P(CB1=0|Q3) | P(CB1=1|Q3) | P(CO2=0|Q3) | P(CO2=1|Q3) | P(CO3=0|Q3) | P(CO3=1|Q3) |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|
| conditional probability | 1 | 1/15 | 1/2 | 1/2 | 1 | 1/16 | 1/19 | 1 |

| Failed Q4 node | P(CO1=0|Q4) | P(CO1=1|Q4) | P(CB1=0|Q4) | P(CB1=1|Q4) | P(CO2=0|Q4) | P(CO2=1|Q4) | P(CO3=0|Q4) | P(CO3=1|Q4) |
|----------------|-------|-------|-------|-------|-------|-------|-------|-------|
| conditional probability | 1 | 1/8 | 1/10 | 1 | 1 | 1/9 | 1 | 1/9 |

| Fault Q5 node | P(CO1=0|Q5) | P(CO1=1|Q5) | P(CB1=0|Q5) | P(CB1=1|Q5) | P(CO2=0|Q5) | P(CO2=1|Q5) | P(CO3=0|Q5) | P(CO3=1|Q5) |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|
| conditional probability | 1 | 1/8 | 1/11 | 1 | 1/9 | 1 | 1/9 |
For a given distribution network example, if the pre-set fault area is Q2, when all the real-time fault information (CO1 = 0, CB1 = 0, CO2 = 1, CO3 = 0) is given, fault reasoning is carried out. According to the Bayes formula 

\[ P(Q_i | X) = \frac{P(X | Q_i)P(Q_i)}{P(X)} \]

since \( P(X) \) is constant for all fault classes. Then \( P(X) = a \), according to the real-time fault information, find the prior probability of the corresponding parent node and the conditional probability of the child node in Table 5 and table 6, then the fault probability corresponding to each fault feature can be obtained:

\[
P(Q_i | X) = a \times P(X | Q_i)P(Q_i) = a \times P(CO_i = 0 | Q_i) \times P(CB_i = 0 | Q_i) \times P(CO_2 = 1 | Q_i) 
\]

(10)

\[
P(Q_1 | X) = a \times \frac{1}{13} \times \frac{1}{10} \times \frac{1}{9} \times \frac{1}{7} = \frac{a}{8190} \]

(11)

(12)

(13)

The same goes for:

\[
P(Q_2 | X) = \frac{a}{7}, P(Q_3 | X) = \frac{3a}{4256} \]

(14)

\[
P(Q_4 | X) = \frac{a}{630}, P(Q_5 | X) = \frac{a}{63} \]

(15)

From \( \sum P(Q_i | X) = 1 \), \( a = 6.2 \). Bring back the above results:

\[
P(Q1 | X) = 0.0008, P(Q2 | X) = 0.8863, P(Q3 | X) = 0.0818, P(Q4 | X) = 0.0091, P(Q5 | X) = 0.00985 \]

Since the fault probability of area Q2 is the highest, it is determined that the fault occurs in area Q2, which is consistent with the preset fault area.

If the pre-set fault area is Q2, when the given part of real-time fault information (CO1 = 0, CB1 = 0, CO2 = 1, CO3 information is lost), the same method is used for fault reasoning:

\[
P(Q_i | X) = a \times P(X | Q_i) \times P(Q_i) \]

(16)

\[
P(Q_1 | X) = a \times P(CO_i = 0 | Q_i) \times P(CB_i = 0 | Q_i) \times P(CO_2 = 1 | Q_i) 
\]

(17)

(18)

(19)

The same goes for:

\[
P(Q_2 | X) = \frac{a}{7}, P(Q_3 | X) = \frac{a}{70} \]

(20)

\[
P(Q_4 | X) = \frac{a}{630}, P(Q_5 | X) = \frac{a}{63} \]

(21)
From \( \sum P(Q_i | X) = 1 \), we get \( a = 5.7223 \). Bring back:

\[
P(Q1|X) = 0.0007, \quad P(Q2|X) = 0.8176, \quad P(Q3|X) = 0.0818, \quad P(Q4|X) = 0.0091, \quad P(Q5|X) = 0.0908.
\]

Since the fault probability of area Q2 is the highest, it is determined that the fault occurs in area Q2, which is consistent with the preset fault area.

5.5. Distributed diagnosis and result fusion

When a fault occurs, the distributed diagnosis method is used to diagnose the fault probability of each area in each subnet, and the result is obtained: The probability of diagnosis of area sec2 in subnet M1 model is, and the probability of diagnosis of area sec2 in subnet M2 model is. The probability of Sec4 diagnosis in subnet M2 model is, and the probability of Sec4 diagnosis in subnet m3 model is. Then, according to the D-S evidence theory fusion rules, the evidence body M1 and M2 fusion results are shown in Table 7 below, and the results after fusion with the evidence body m3 are shown in Table 8 below.

| Table 7. Fusion results of D-S evidence theory 1 |
|------------------------------------------------|
| \( m_1 / m_2 \) | \( a \) | \( b \) | \( U \) |
|-----------------|------|------|------|
| \( a \) | 0.525 | 0 | 0.475 |
| \( b \) | 0.43365 | 0 | 0.39235 |
| \( U \) | 0.0882 | 0 | 0.0798 |
| 0.006 | 0.00315 | 0 | 0.00285 |

From equation 9 \( K=0.0882 \).

From equation 8:

\[
m(a) = \frac{0.43365 + 0.39235 + 0.00315}{1 - K} = 0.9094
\]

\[
m(b) = \frac{0.0798}{1 - k} = 0.0875
\]

| Table 8. Fusion results of D-S evidence theory 2 |
|------------------------------------------------|
| \( m_3 / m \) | \( a \) | \( b \) | \( U \) |
|-----------------|------|------|------|
| \( a \) | 0.9094 | 0.0876 | 0.003 |
| \( b \) | 0.223 | 0.0215 | 0.000735 |
| \( U \) | 0.0866 | 0.06614 | 0.002265 |

From equation 9: \( K=0.223 \). From equation 8:
The fault diagnosis result of Sec2 area is 0.8837, and that of Sec4 area is 0.1137, so Sec2 is the fault area. Compared with the diagnosis results of single subnet M2 model, the diagnosis results obtained by D-S evidence theory fusion are more obvious, compared with the fault probability of non fault components, and the judgment result is more accurate.

6. Conclusion
This paper presents a distributed fault diagnosis method based on Naive Bayesian network and D-S evidence theory. The two methods complement each other's advantages, use rough set to reduce attributes, extract the best reduction attribute set, and then build naive Bayesian network, which reduces the modeling difficulty of Bayesian network. According to D-S evidence theory, the fault diagnosis results of network overlap are fused with evidence, which effectively improves the accuracy of fault diagnosis. The example analysis shows that the method can improve the fault tolerance of the system to a certain extent, and the model is simple and the diagnosis accuracy is high.

Acknowledgments
Fund project: Liaoning provincial education department project (L2019007)

Reference
[1] Yang Fan, Zhang Yanjun. Microgrid fault diagnosis method based on improved causal time series network [J]. Foreign electronic measurement technology, 2019, 38 (04): 28-33.
[2] Xu Biao, Yin Xianggan, Zhang Zhe, Chen Guoyan, Yang Wen. Power network fault diagnosis model based on topological primitive information fusion [J]. Journal of Electrical Technology, 2018, 33 (03): 512-522.
[3] Ling Zijun, Hu Chao, Tang Junsheng. Power grid fault diagnosis method based on Bayesian network and rough set [J]. Science and technology perspective, 2014 (31): 79-109
[4] Hao Li Gang, Chen Hao, Tan Fenglei. Power network fault diagnosis model based on weighted fuzzy time series Petri network [J]. Power engineering technology, 2019, 38 (03): 93-99.
[5] Xiong Biao. Application of Swarm Intelligence Algorithms in Power System Fault Diagnosis [D]. Anhui University of Technology, 2018.
[6] Zhou Jiehua. Fault diagnosis optimization model based on physical model of power grid information [J]. Electronic test, 2019 (12): 51-52.
[7] Feng Feng. Application of Fengfeng Bayesian Network and D-S Evidence Theory in Power Network Fault Diagnosis [D]. Lanzhou University of Technology, 2017.
[8] Yang Haibin. Fault Diagnosis of Power Grid Equipment Based on Bayesian Network [D]. Hebei University of Technology, 2015.
[9] He Wenbo, Guan Huanxin. Overview of power grid fault diagnosis technology [J]. Northeast Power Technology, 2018,39(05): 52-55.
[10] Du Ruijuan, Gui Jianting, Yang Jing, Xu Tao and Hao Mengfei. Application of Reduction Algorithms Based on Importance in Power Network Fault Diagnosis [J]. Journal of Nanyang Institute of Technology, 2017,9(02): 6-9.
[11] Chen Guoen, Li Wei, Zhang Zhongxiao. Research on fault diagnosis method and system architecture of power grid [J]. Electrical technology, 2018, 19 (01): 64-67.
[12] Wang Dazhi, Jiang Xuechen, Ningyi, Zhao Linlin. An improved analytic model for power grid fault diagnosis [J]. Journal of Northeast University (Natural Science Edition), 2016, 37 (08): 1065-1069.
[13] Luo Xiaohui, Tong Xiaoyang. Fault Diagnosis of Variable Structure Bayesian Network with Credibility [J]. Power Grid Technology, 2015, 39 (09): 2658-2664.
[14] Han Yingchun, Tong Xiaoyang. Power network fault diagnosis method based on dynamic
reasoning chain [J]. Power grid technology, 2017, 41 (04): 1315-1324.