A fault detection model for microgrid detection based on Bayesian network and association rule mining

Wen Hao Zhu¹,², Qiyi Guo¹*

¹Department of Electrical Engineering, Tongji University, Shanghai 200092, China
²Schneider Shanghai Low Voltage Terminal Apparatus Co., Ltd. Shanghai 201109, China

Received 1 June 2014, www.cmnlv

Abstract

In view of the problem that current fault detection methods exist large error in microgrid detection, this paper presents a model based on Bayesian network and association rule mining. It firstly adopts Hash technology to optimize Apriori algorithm and remove the undesired candidate item set, conducts data mining of original data set, introduces Bayesian network for sample training to reduce detection error, and finally obtains microgrid detection result. Simulation results show that the proposed fault detection model based on Bayesian network and association rule mining is efficient in microgrid fault detection with detection error far less than that of traditional algorithm.

Keywords: microgrid fault detection, association rules mining, frequent item set optimization, bayesian network, error optimization

1 Introduction

Under the development background of power electronics and modern control theory, microgrid was first proposed by professor R.H. Lasseter in 2001 [1]. It is a new type of web based energy supply and management technology which provides a convenient access to renewable energy system and realizes demand side management and the maximum use of available energy [2]. With the rapid computing ability, computer is used to study the microgrid fault detection which requires collaboration between schedulers and other sections, quickly and accurately give out the analysis results, and then make a reasonable operation decision for schedulers [3].

Because of the complex system faults, traditional mathematical method is unable to describe a model with variety of fault types [4]. The improvement of artificial intelligence technology, especially in simulating human beings to handle all kinds of problems, makes it widely accepted by researchers and successfully applied in fault detection of electric systems. C.A. Petri put forward a Petri network mathematical model, suitable for small scale power grids fault detection with stable structure [5]. It still has many limits in large scale power grids whose equipment numbers always change. Z. Pawlak applied rough set theory into grid fault detection in order to reduce recorded fault samples and insignificant condition attributes [6]. LA Zadeh proposed a grid fault detection model based on fuzzy theory to handle the uncertain problems [7]. However it is still difficult when there are so many uncertainties. Pitts applied artificial neural network into fault detection to make clear the logical relationship between fault type and fault signal [8]. The artificial neural network requires to be trained with slow convergence speed before using. Therefore, it is available only when the system structure does not change. Lee proposed a model based on Bayesian network [9]. Compared with traditional fault detection methods, this model has clear logic, fast solving speed, strong learning ability and good fault freedom, but it still exists shortcomings. For example, the conditional probability distribution of each node is not given by Bayesian network but usually determined by statistical analysis technique as a matter of experience.

On the basis of traditional fault detection methods, this paper puts forward a microgrid fault detection model based on Bayesian network and association rule mining, and optimizes the frequent item sets of the association rule mining algorithm.

2 The improvement of association rules mining algorithm

2.1 APRIORI ALGORITHM DESCRIPTION

Apriori algorithm is a basic algorithm searching for frequent item sets which uses an iterative method of layer-by-layer searching, namely searching \((k + 1)\) - item set with \(k\) - item set [10]. Apriori algorithm adopts the priori knowledge of frequent item set properties, firstly generates frequent 1 - item set \(I_1\), then frequent 2 - item set, until some \(r\) frequent item set \(I_r\) is empty.

In \(k\) -th circulation, the collection \(C_k\) of candidate \(k\) - item sets are first generated. Each candidate is generated based on the properties of frequent \((k - 1)\) - item sets. That is, let \(L_{k-1}\) be a frequent \((k - 1)\) - item set, then the candidate \(k\) - item set \(L_k\) is generated and the transaction \(T\) is used to check whether \(L_k\) is frequent. First, \(L_k\) is put into \(T\), then \(L_k\) is removed from \(T\) and the next \(L_k\) is put into \(T\), until all \(L_k\) is removed.

* Corresponding author’s e-mail: qiyi@263.net
by a $(k-2)$ - connection between two frequent sets that have only one different item belonging to $L_{k-1}$. The item sets in $C_k$ are the candidate sets of frequent item sets, and the last frequent set $L_k$ must be a subset of $C_k$. Each element should be verified in the transaction database in order to determine whether it should be added into $L_k$, which is a limit of algorithm performance due to its significant I/O load.

The basic idea of Apriori algorithm is: firstly, search out all the item sets whose frequency should be same to the predefined minimum support at least; secondly, the item set generates strong association rule which must satisfy the minimum support and credibility.

2.2 APRIORI ALGORITHM IMPROVEMENT

Before generating item sets $C_k$ each step, Apriori algorithm requires corresponding frequent $k-1$ - item sets, and searches the database to calculate the support of each candidate item set so as to generate frequent item sets, which requires much time and space. Therefore, the generation of even a small candidate item set plays an important role in improving the efficiency of searching the frequent item set. However, in Apriori algorithm, candidate set $C_k$ is generated jointly with $L_{k-1}$, where $C_k$ has a large potential. Hash technology can remove those unnecessary candidates to decrease the potential of $C_k$, thus lower the cost of time and space and increase the algorithm efficiency.

In Apriori algorithm, when the transaction in $D$ and item in $I$ is very large, the item in $L_1$ may be large. If supposed to be $m$, then the number of candidate 2 - item set $C_2$ generated jointly is $\frac{m(m-1)}{2}$. Therefore, the calculation of support will have a very large calculated amount.

The main goal of using Hash technology is to solve this conflict. This paper adopts a two dimensional hash function to avoid it.

Each item $I_k (k=1,2,\ldots,m)$ in item set $I = \{I_1, I_2, \ldots, I_k, \ldots, I_m\}$ is given a sequential value, 1, 2, .., $m$. $\text{order}(x)$ and $\text{order}(y)$ represents the sequential value of item $x$ and $y$ of candidate 2 - item set. Then two dimensional Hash function is,

$$h_1(x, y) = \left(\left|L\right| \times \text{order}(x) + \text{order}(y) - \frac{x(x-1)}{2}\right) \mod p_1,$$  \hspace{1cm} (1)

$$h_2(x, y) = \left(\left|L\right| \times \text{order}(x) + \text{order}(y) - \frac{x(x-1)}{2}\right) \mod p_2,$$  \hspace{1cm} (2)

Here,

$$p_1 \cdot p_2 \geq (1 - \min \sup \) C_k^2,$$  \hspace{1cm} (3)

where $p_1 \neq p_2$ and they are relatively prime numbers.

If the value of $p_1$ and $p_2$ is large, the Hash table will cover much more space. If they are too small, it will cause conflicts. Therefore, the value of $p_1$ and $p_2$ should be adjusted with the item number and the defined minimum support.

$$H(x, y) = H(h_1(x, y), h_2(x, y)),$$  \hspace{1cm} (4)

$h_1(x, y)$ and $h_2(x, y)$ denote the subscripts of $H(x, y)$. When the hash number is projected to a unit, then count number is added 1.

This paper adopts the function

$$|L| \times \text{order}(x) + \text{order}(y) - \frac{x(x-1)}{2},$$

where $\left|L\right|$ is the item number is unique with two dimensional Hash table, which largely decreases the Hash conflict.

At the same time of scanning database and generating $L_1$, all the 2 - item set of each transaction is counted with two dimensional Hash function. After the database is scanned, not only $L_1$ but also a two dimensional Hash table is acquired where the value of each unit is the accumulation of a count. If the count number is equal to or larger than minsup, then 2 - item set belongs to $L_2$, otherwise, it is not the frequent 2 - item set.

3 Microgrid fault detection based on Bayesian method and association rule mining

3.1 SAMPLE TRAINING BASED ON BAYESIAN NETWORK

Simply speaking, if there is a training sample, then we can learn from ours ample data set and classify new data to a class. This Bayesian theory is agreeable with our task.

$$s_{\text{MAP}} = \arg \max_{s \in \mathcal{S}} P(s \mid x_1, x_2, \ldots, x_n).$$  \hspace{1cm} (5)

It can be expressed with Bayesian equation.

$$s_{\text{MAP}} = \arg \max_{s \in \mathcal{S}} P(x_1, x_2, \ldots, x_n \mid s) P(s) = \frac{\arg \max_{s \in \mathcal{S}} P(x_1, x_2, \ldots, x_n \mid s) P(s)}{P(x_1, x_2, \ldots, x_n)} .$$  \hspace{1cm} (6)

Then we should learn the value estimated from Equation (6) according to the training data. In fact, it is easy to predict each $P(s_i)$ by only calculating the frequency of each target value $s_i$ in training data.
However, if there is not a considerable training data collection, we cannot predict different \( P(x_1, x_2, ..., x_n | s) P(s) \) with this method. Because the number of this item is equal to the product of possible example number and possible target value. This method requires obviously large training item sets. In order to acquire a reasonable estimation, each sample should appear several times in the example space.

Bayesian algorithm has a simple suppose that if the target value of sample is given, then the joint probability of \( x_1, x_2, ..., x_n \) is equal to the product of the probability of each single attribute.

\[
P(x_1, x_2, ..., x_n | s) = \prod_i P(x_i | s).
\]

Adding it to Equation (6), then we obtain the method in Bayesian algorithm.

\[
s_{ab} = \arg \max_{s} P(s) \prod_i P(x_i | s),
\]

where \( s_{ab} \) is the target value calculated from Bayesian algorithm. In Bayesian algorithm, the number of different \( P(x_i | s) \) - item estimated from training data is only product of different target value number and attribute value number, which is far less than the amount from \( P(x_1, x_2, ..., x_n | s) \).

If one of the conditional probabilities calculated from this method is zero, no matter what other attribute probability is, the value is always zero. These attributes are independent so that the target value will be influenced by the continued product of them. In order to avoid the zero probability problem from limited samples, we take the method of m-estimate to optimize it so as to make the classification more accurately. M-estimate is defined as,

\[
P(m_i | n_i) = \frac{v_i + mp}{v + m}.
\]

Here, the number of samples in class \( n_i \) is denoted as \( v \); the sample number of \( m_i \) is denoted as \( v_i \) in the samples similar to class \( n_i \); the attribute corresponding to sample is \( m \); the parameter assigned by tester is \( p \). If training set is zero, then \( P(m_i | n_i) = p \). Therefore, the prior probability of attribute \( m_i \) in class \( n_i \) is indicated as \( p \). The equal value of sample can determine the relationship between probability \( p \) and \( v_i / v \).

### 3.2 Specific Process of Microgrid Fault Detection

This paper adopts a grid fault detection method based on association rule mining and Bayesian network, considering the redundancy of earlier fault data and tolerance of grid fault detection of uncertainties. The specific flow is shown as follows (Figure 1).

According to the flow chart, the specific algorithm is, 1) Taking the action information of protector and breaker as conditional attributes, the fault component as decision attribute, and select training sample set to build original decision table \( S \) according to past fault samples; input it to database to generate data resource as the object of association rule mining; 2) Interactively mining the original decision table with association rules, reducing the redundant objects and attributes. It is noted that if the support and confidence threshold is small, then large number of weak association rules will be generated and mining time is prolonged, otherwise some valuable association rules will be ignored. In practice, smaller sample is suggested to conduct iterative mining by continuing threshold adjustment to determine the appropriate threshold. Then the method is applied in mass data aimed by support and confidence; 3) Initializing Bayesian network; 4) Using improved Apriori algorithm in newly generated attribute set to mine the frequent item set satisfying support threshold; building association rule for each frequent item set and selecting the strong ones of which the minimum confidence and relevancy is larger than 1. Those whose support is less than minimum value are not considered so as to reduce the search space; 5) Building Bayesian model with the association rule from step (2):

\[
X = \{ x | \text{x is attribute out of classification in } D \},
\]

\[
\text{NewRules} = \{ \text{association rules satisfying the min sup in L} \}
\]

For each rule, \( X \Rightarrow C \) and extracting from \( \text{NewRules} \) in order; extracting the confidence corresponding to \( X \) and adding \( X \Rightarrow C \) to Bayesian network; 6) Taking the fault region \( C(c_1, c_2, ..., c_n) \), namely the decision attributes in decision table, as the father node in Bayesian network, and taking the conditional attributes of reduced association rule group \( \text{Rule group} \) as nodes to build Bayesian network; 7) Training each node and calculating the prior probability of each father node and the conditional probability of each child node; 8) Realtime fault detection.
The real time fault information $X(x_1, x_2, ..., x_n)$ is given, including complete information, incomplete information, and even wrong information, and they are diagnosed with trained Bayesian network to calculate the probability of possible fault region.

Then the obtained probability of possible fault region is sorted, and the fault region $C_k$ with largest probability is input as the final diagnosis result.

4 Simulation experiment

To verify the effectiveness of the improved algorithm, this paper conducted the simulation experiment. The experiment has seven microgrid fault samples with conditional attributes breaker B1, B2, B3, overcurrent protection C1, C2, C3, and distance protection R1. Each attribute is valued 0 or 1 representing the on or off of the breaker. The decision attributes are the fault area, usually circuit S1, S2, S3, as listed in Table 1.

A group of fault information is given and detected by the detection model based on Bayesian network and association rules mining. The results are listed in Table 2.

The error between detection result from current model and actual result is recorded with the curve as follows.

Seen from Table 2 and Figure 2, the detection model based on Bayesian network and association rule mining can be well applied in microgrid fault detection, with detection error far below that of Apriori algorithm and Bayesian algorithm.

5 Conclusions

Microgrid has the characteristics of bidirectional current and small capacity of short circuit so that traditional power distribution network is no longer appropriate for microgrid. In view of that, this paper put forward an improved detection model based on Bayesian model and association rule mining. Simulation experiment results show that this model is effective and greatly reduces the error of original algorithms in microgrid detection.

References

[5] Zhang M 2013 A kind scheme of design and implementation of remote fault detection Application of Electronic Technique 39(12) 126-8
[10]Du M 2013 Realtime SVDD algorithm based on sample reducing and application in electric fault detection Microelectronics and Computer 30(7) 96-90
Wenhao Zhu, 08.10.1977, China.

Current position, grades: doctoral student at Tongji University, China.
University studies: Master degree in Electronics and Communications Engineering from Fudan University, China in 2007.
Scientific interest: intelligent system in final distribution of smart grid, and communication security and stability.
Publications: 20 patents in intelligent low voltage apparatus area, 10 articles.

Qiyi Guo, 06.1961, China.

Current position, grades: professor and PhD supervisor in electronic information engineering institute at Tongji University, China.
Scientific interest: electrical equipment failure diagnosis in automation and information engineering.