An Algorithm Based on Probes to Locate Fault Nodes in Electric Power Data Network

Bo Fu[1, a], Peng Liu[1, b] and Hongwei Li[1, c *]

[1]Zhuhai Power Supply Bureau of Guangdong Power Grid Corporation, Zhuhai City, Guangdong Province, 519000

[a]446137166@QQ.com, [b]94205293@QQ.com, [c]94205293@QQ.com

*The corresponding author

Keywords: Electric power data network; Fault location; Probes; Bayes

Abstract: Fault location is a key part of electric power data network. The existing algorithms collect alarm information to locate fault, which gain low accuracy. IP network usually uses probe to monitor service performance which can locate fault rapidly. However this method does not be applied to electric power data network. Accordingly this paper apply probe method to electric power data network, and then proposes an algorithm to locate fault in electric power data network. The proposed method analyses the degree of business affected and calculates the impact factor of nodes firstly, then according the nodes detection to gain a sparse matrix, finally uses this sparse matrix gets a totally matrix. At last, the proposed algorithm uses the impact factor of nodes and totally matrix as the put in factors of Bayes network model to gain fault locations. The simulation results show that compared to the existing algorithms IBSD and IHU, our algorithm gains better accuracy and lower error rate.

Introduction

Fault location is the key part of electric power data network, the existing methods using network management equipment to collect and analyses alarm information for fault locating. In IP network, active probes are usually used to locate faults, which can gain service performance effectively, guarantee the service performance of every access network. When abnormal conditions is monitored, active probe method can locate fault immediately according to the information monitored [1]. Although probe method is common used in IP network, apply probe method to electric power data network has never been used yet.

Electric power data network using Virtual Private Network (VPN) to connect every dispatching center, using IP route in private channel to support public power information system such as SCADA/EMS scheduling automation system and power billing system and relaying protection management information system and dynamic pre-alarm monitoring system etc. Meanwhile it can also support media business such as voice video and network communication [2-4].

In electric power data network, the business flow is stable basically, which is different with IP network and electric integrate data network. In the analysis of each node involved in every business, gathering the impact degree of each node fault and then compute the influence factor of each node. This influence factor indicates the impact degree of fault node can influence the whole network [5-7]. When deploying the probes in electric power data network, considering node degree and impact factor, after detecting faults using impact factors to compute the fault probability of nodes, therefore this paper analysis the impact factor of node fault in advance to improve the accuracy of fault node location [8-9].

Self-healing Capability of Electric Communication Network

The business of electric power data network is stable and the important business following the N-1 principle, that is to say, every important business has one path at last and the importance of business can be expressed by the number of paths that can be chosen. Fig. 1 is using a small part of nodes in
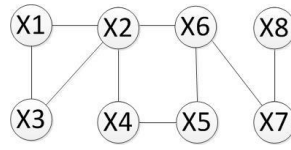
network to build model.

Figure 1.　Network nodes connection

There are several definitions in Fig. 1:

Because of the model scale supposed is small, the number of business is limited. Assuming 5 businesses in the model, S1-S5. The power communication business has its own characteristics and types, in order to discriminate the load mode and content of this business, besides specific the focus of attention, which can easily analysis business in the way of resources loaded by business, we indicate the business S1-S5 as resource intersection, such as the following formula.

S1：X1∩X2∩X5　　S2：X3∩X2∩X6　S3：X6∩X7∩X8

S4：X4∩X6∩X7　　S5：X3∩X6∩X8

Then, the impact factor of business influenced is analysis:

Because of the stability of business and the difference of impact degree of the fault, we should gain the impact factor of business firstly, and providing the influence of standard business impact degree, combining the possible circumstances in actually running, according to the supposed model, we can gain the impact factors of business as follows.

Table 1 Impact factor of business

Influence	Description
None (L1)	Business running normal
Potential influence (L2)	Business running normal, existing potential fault
Performance influenced (L3)	Business can be running, the service quality is poor
Severely influenced (L4)	Business cannot be running

According to analysis the influence of each node fault to business, the result is shown in Table.2.

Table 2 Influence of each node

	V1	V2	V3	V4	V5	V6	V7	V8
S1	L4	L3	L1	L2	L4	L2	L1	L1
S2	L2	L3	L4	L2	L2	L4	L1	L1
S3	L1	L1	L1	L1	L1	L4	L3	L4
S4	L1	L2	L1	L4	L2	L3	L4	L1
S5	L2	L2	L4	L2	L2	L3	L3	L4

The influence factors of nodes can be compute as follows:

When the network equipment is fault, we can measure the influence of electric communication network equipment to traditional business. The node fault influences not only business but also the relevant nodes. By using complicate network theory, considering the internal characteristics and external attributes, we can get the influence elements.

In electric power data network, internal characteristics of fault node can influence other nodes in the following aspects, such as the degree of node, the betweenness and tightness. The external attributes is the influence of fault node affect to each node. By quantifying these indicators, we can get the number of the influence factor of each node.

In order to corresponding to the internal characteristics, we should quantitative external characteristics, and specify a value for each level, as follows.

Table 3 Integrated influence degree of node fault to business

L1	L2	L3	L4
0	0.3	0.7	1

The influence degree of node fault is shown in Eq.1.

$$C_{d(V)} = \frac{\sum_{i=1}^{n} Li}{Sn} \times 100\%$$

(1)

In which v is a specific node in network, Sn is the business number which node can influence, Li is the quantitative number that business influenced by node. The integrated influence number of fault node is the average number that fault node influence each business, which can be a external influence of a node to the whole electric power data network.

Using $G = (V, E)$ indicates the electric power data network, the influence elements of node v0 is defined in Eq.2.

$$IF(v_0) = \alpha Fin(v_0) + \beta Fout(v_0)$$

(2)

Fin(v0) is the internal influence element of node, α is the external influence element of node, β is internal influence coefficient, is external influence coefficient, $\alpha + \beta = 1$, α and β can be set up according to the actual situation and the network structure. Because of the consideration of influence of external business, the proportion of external is bigger, then β is bigger, according to our experience, $\alpha = 0.2$, $\beta = 0.8$.

$$Fin(v_0) = \frac{K(v_0)}{\max(K(v))}, v \in V$$

(3)

K(v) is the internal influence element, which can be degree, betweenness, tightness and so on. Because we assume the number of α is small, which internal attribute is chosen makes no different to the result. max(K(v)) is the normalization factor, in order to eliminate the effect of network size to the result and make sure the number of internal influence elements are between interval [0,1]. Fout(v0) is the node fault influence number Cd(V0), by using which we can figure out the number of node influence element.

Finally we use Bayes network model and put node influence element into the probability formula to calculate the location of fault nodes:

In order to reduce network load, we probe a few probe into network firstly, and then use end-to-end probe predication technology to find the rest of probes, choose nodes communicate to probes selected from the rest of probes randomly. Additionally speaking, the probes selected are not

be put totally. Only a few probes are put into network. By using probes put into network, we can get a sparse matrix. The way to formulate the sparse matrix is: the rows of sparse matrix is beginning nodes of probes selected, the columns of sparse matrix is end nodes of probes selected. If the results returned from probes is positive, then the number of sparse matrix is -1. If the node is fault, the number of sparse matrix is 1. No probes return 0. The number of sparse matrix is 1, -1 or 0.

Algorithm MMMF is the maximum margin matrix algorithm. According to the sparse matrix, the complete matrix is obtained by predicting all the elements which are 0 using probes. MMMF is a process to synchronize learning feature vector v and coefficient vector u, then the matrix of the original matrix is decomposed into low norm. The low rank constraint is the dimension of the feature space and the prediction to the column of matrix Y is the prediction to low dimension space. Using the dependence of elements in matrix, we can get a feature vector v and a coefficient vector u, and then using the multiplication of these two vectors to form an complete matrix, which is an approximate evaluation of incomplete matrix to represent the detection of all probes.
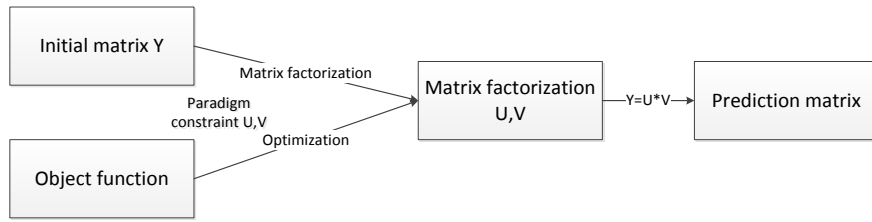

Figure 2.   MMMF algorithm

After predicting matrix, using Bayes network model to locate fault nodes. Information getting by probes is the states every elements after fault. Network fault location has the uncertain character. Whereas the Bayes model is good at finding reason in uncertain situation. Besides Bayes model using graphs to present variables joint probability distributions and conditional independence, which can extremely reduce fault location cost. Bayes network model join the nodes and probes, which form a two tier direct directed graph.
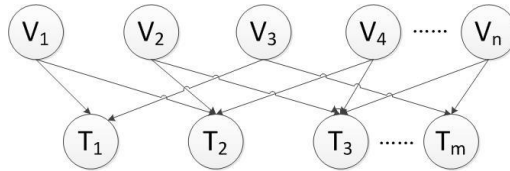

Figure 3.   The relationship between nodes and probes

The equipment node in network are regarded as father nodes, probes are regarded as son nodes. The probe results are depend on father nodes, when and only when all father nodes are right the probe result is right. By compute the probability of all nodes to find the most likely failure of the nodes.

According to the parameters to be positioned including prior probabilities of all network nodes and results of probes, the conditional probabilities of all nodes can be computed. That is the probability of P(V1,V2,...Vi,...Vn,T1,T2,...Tj,...Tm), Vi is the ith node, Tj is the jth probe, n is the number of nodes in network, m is the probe number. The formula of P(V1,V2,...Vi,...Vn,T1,T2,...Tj,...Tm) is:

$P(V_1, V_2, \ldots V_i, \ldots V_n, T_1, T_2, \ldots T_j, \ldots T_m) =$

$P(T_1 \mid Pa(T_1)) P(T_2 \mid Pa(T_2)) \ldots P(T_m \mid Pa(T_m)) P(V_1) P(V_2) \ldots P(V_n)$

Among:

(1) $P(V_1, V_2, \ldots V_i, \ldots V_n, T_1, T_2, \ldots T_j, \ldots T_m)$ is the conditional probability of all nodes, which is influence of fault nodes to network.

(2) $Pa(T_j)$ is all the link probe T_j goes through, when the link has fault node,

p(Ti=1|Pa(Ti))=1,p(Ti=0|Pa(Ti))=0, otherwise, *p(Ti=1|Pa(Ti))=0,p(Ti=0|Pa(Ti))=1.*

(3) The prior probability is based on the past experience and the analysis of the probability of fault network nodes. Prior probability is an estimate number, which is different according to the network. In our model prior probability is 0.1, when the network is good prior probability is 0.9, if there is fault node *p (Xi)=0.1*IF (Xi)* that is to say, considering failure probability of fault node and influence element, besides, considering the character of Bayes network, we use prior probability multiply influence element to gain posterior probability, which can locate fault node more accurately.

(4) $P(V_1,V_2,\ldots,\ V_n,T_1,T_2,\ldots,T_m)=P(V_x)\ldots P(V_y)*0.9^{n-k}$, in which k is the number of fault nodes, *Vx...Vy* is fault nodes, *P(Vx)...P(Vy)* is the failure probability correspondingly, which is the number of *V=0*, then compute the probability of *V=0,V=1*. Then computer *the P(V1,V2,...Vi,...Vn,T1,T2,...Tj,...Tm)* in all circumstance, and the condition when *P(V1,V2,...Vi,...Vn,T1,T2,...Tj,...Tm)* is max, that is to say when *V=0* or *V=1,* all the nodes when *V=0* is the set of all the fault nodes.
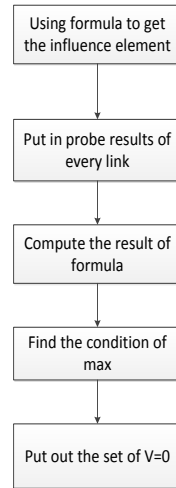

Figure 4.   Algorithm flow chart of fault location algorithm based on probes

Fig. 4 is the algorithm flow chart of fault node location algorithm based on influence element.

Step1: using formula in section 3 to compute the influence element of every node.

Step2: put in the probe result of every link.

Step3: according to probe result of every probe and the links of probes go through, using failure probability of nodes to get the result of formula P(T1|Pa(T1))P(T2|Pa(T2))...P(Tm|Pa(Tm))P(V1)P(V2)...P(Vn).

Step4: put out node set when V=0, which is the set of fault nodes.

Experiment Results

In order to estimate the proposed algorithm, we use a topology of electric power data network. There are 140 nodes in this topology, 760 packets and 56 faults. There is 90 business in network. First, we gain the relationship between nodes and business, then compute the influence element of every node. There are 320 probes in network, which have different probe links.

Compared algorithms are IBSD and IHU, our algorithm used simulation instances to illustrate its performance. In experiments the compared parameters are location accuracy and location error rate under different noise circumstance.

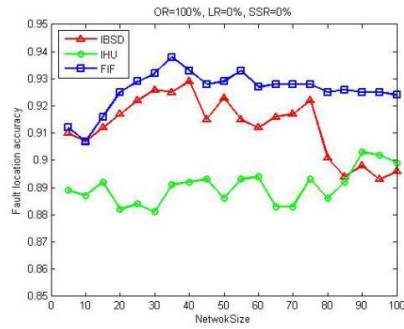(1) OR=100%, LR=0%, SSR=0%, no noise.

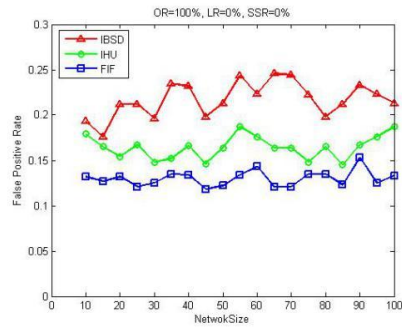Figure 5.   (1)   Figure 5.   (2)

Fig. 5(1) is the location accuracy of three algorithms, when the network scale is growing, the other two algorithms is stable except IBSD. In previous research we found that IBSD didn't consider the accuracy of large scale network, when the number of nodes in network is more than 75, the performance of IBSD is decreased sharply. Under no noise circumstance, the average locate accuracy of IBSD is 91.25%, IHU is 89.36%, our algorithm is 92.62%.

Fig. 5(2) is the error rate of fault location, IBSD is 12.72%, IHU is 10.0%, FIF is 8.24%. When the number of nodes is increase, the performance of these three algorithms changed little.

(2) OR=80%,LR=10%.SSR=1%, there is noise in this circumstance, the alarm rate dropped to 80%
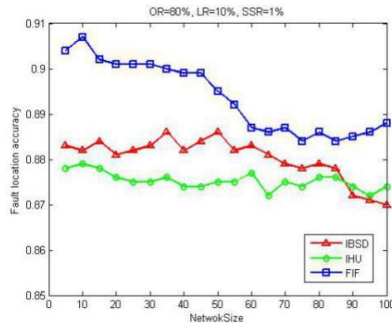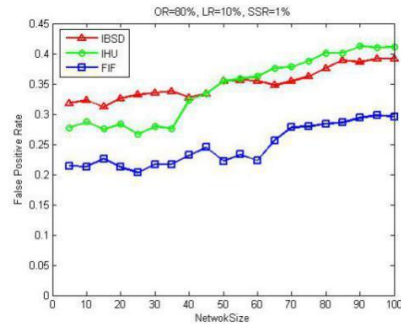

Figure 6.   (1)   Figure 6.   (2)

Fig. 6(1) compares the fault location accuracy of these three algorithms under noise circumstance, when the scale of network is more than 50, all these three algorithms show decrease trend. The average accuracy of IBSD is 88.03%, IHU is 87.53%, FIF is 89.35%. That is to say our algorithm is better than the other two algorithms.

Fig. 6(2) compares the fault location error rate of these three algorithms. Different from the other two algorithms, the error rate of our algorithm is decreased sharply when the scale of network is increasing. The average error rate of our algorithm is 35.05%, IBSD is 34.26%, FIF is 30.57%. When the scale of network is large, our algorithm is better than the other two algorithms.

Conclusion

Fault location is the key point of electric power data network. In this paper, we apply probe technology of IP network into electric power data network. A fault location algorithm based on probe is proposed, which uses Bayes network model and influence elements to calculate node failure probability. Experiment results show that our algorithm has better accuracy and error rate compared to existing algorithms especially in large scale electric power data network.

Acknowledgements

References

[1] Du Bao, The network toplogy and safe analysis of state electric power data network [J], China Electric Power Education. 2013(30).

[2] Li Xichun, Yao Weigang, He Zhiming, The apply and safety of electric power data network [J], Jiangsu Electrical Engineering. 2008(03).

[3] Zou Shifeng, Analysis of the status monitoring system based on the distributed processing of the main electrical equipment on the.Net platform [J], EDN China. 2013(10).

[4] Fan Yaoxue, Yu Zong, Research and design of trusted execution environment based on Windows system[J], Computer Engineering and Design, 2010(13).

[5] Bossart. J, Bean. E, "Metrics and benefits analysis and challenges for Smart Grid field projects," Energy tech, 2011IEEE, May2011.1-5.

[6] Xiang Yuan, Zhenxing Qian, Yang Zhou, "Discussion on the development trend of smart grid and its key technology," Electricity Distribution (CICED), 2012 China International Conference on Shanghai, Sept 2012.

[7] Liu Xiaosheng, Zhao Zhenfeng, Zhang Pengyu, Ren Huifen, "Study on reliability of anovel Electric Power Data Network for smart grid," Power Electronics and Motion Control Conference (IPEMC), vol. 3, 2012, pp.2305-2310.

[8] Natu. M, Sethi. A.S, "Active Probing Approach for Fault Localization in Computer Networks," End-to-End Monitoring Techniques and Services, 20064thIEEE/IFIP Work shop on, April 2006, 5-33.

[9] Di Giorgio. A, Liberati. F.A, "Bayesian Network-Based Approach to the Critical Infrastructure Interdependencies Analysis," Systems Journal, IEEE.vol.6, 2012, pp.510–519.