# Forward and backward models for fault diagnosis based on parallel genetic algorithms[*]

Yi LIU[†], Ying LI, Yi-jia CAO, Chuang-xin GUO

(*School of Electrical Engineering, Zhejiang University, Hangzhou 310027, China*)

[†]E-mail: powersystemee@163.com

**Abstract:** In this paper, a mathematical model consisting of forward and backward models is built on parallel genetic algorithms (PGAs) for fault diagnosis in a transmission power system. A new method to reduce the scale of fault sections is developed in the forward model and the message passing interface (MPI) approach is chosen to parallel the genetic algorithms by global single-population master-slave method (GPGAs). The proposed approach is applied to a sample system consisting of 28 sections, 84 protective relays and 40 circuit breakers. Simulation results show that the new model based on GPGAs can achieve very fast computation in online applications of large-scale power systems.

**Key words:** Forward and backward models, Fault diagnosis, Global single-population master-slave genetic algorithms (GPGAs), Parallel computation

**doi:**10.1631/jzus.A0720087          **Document code:** A          **CLC number:** TM734

## INTRODUCTION

The fault diagnosis and identification (FDI) of an electric power system is a process of discriminating the fault system by tripping protective relays and circuit breakers. Fast and correct fault section diagnosis is very important to power system restoration. Up to now, a number of techniques have been developed, which include genetic algorithms (GAs) (Chantler *et al*., 2000), Boltzmann machine (Sun *et al*., 1998), expert systems (Cardozo and Talukdar, 1988; Park *et al*., 1997), Petri nets (Ren *et al*., 2005; Chuang *et al*., 2006), Bayesian networks (Chin, 2003; Zhu *et al.*, 2006), neural networks (Narendra *et al*., 1998), etc. The fault diagnosis of large-scale power systems, however, is still an unresolved problem because of the large amount of information received after fault, the need for high accuracy and high speed, the existence of hidden failures, and model's un-adaptability to topological changes.

Researchers have already utilized optimization techniques to perform diagnosis in power systems (Dysko *et al*., 1999; Huang S.J., 2002; Liu *et al*., 2007). But the computation speed for large-scale power systems is low. So, many methods have been developed to improve the optimization techniques. The 0-1 integer programming model based on GAs, Tabu-search and simulated annealing have been researched (Wen and Han, 1997). In this paper, the model based on GAs is adopted in the backward model, and the forward model reduces the searching scale of fault sections. In this way, the size of the problem can be greatly reduced and very fast computation can be achieved even for large-scale power systems.

## FORWARD AND BACKWARD MODELS FOR FAULT DIAGNOSIS

We first establish the mathematical model based on the relationships of sections, protective relays and

circuit breakers. This model is divided into the forward part and the backward part. The forward model finds the possible fault sections in a given power system and the backward model identifies the fault components (sections) from the possible fault sections. A simple example as shown in Fig.1 is used to illustrate how to determine the mathematical model.
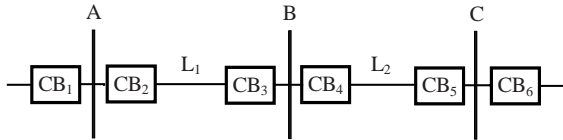


**Fig.1  A simple example consisting of 5 sections, 6 circuit breakers and 15 protective relays**

In this system there are 5 sections, 6 circuit breakers and 15 protective relays. The 5 sections are buses A, B, C and lines $L_1$, $L_2$ (their statuses are denoted by $s_1 \sim s_5$; $s_i=0$ corresponds to the normal state and $s_i=1$ the fault state). The 6 circuit breakers are $CB_1 \sim CB_6$ (their statuses are denoted by $c_1 \sim c_6$; $c_i=0$ corresponds to the non-operational state, and $c_i=1$ the operational state). The 15 protective relays are Am, Bm, Cm, $L_1$Am, $L_1$Bm, $L_2$Am, $L_2$Cm, $L_1$Ab, $L_1$Bb, $L_2$Bb, $L_2$Cb, $L_1$As, $L_1$Bs, $L_2$Bs, $L_2$Cs (their statuses are denoted by $r_1 \sim r_{15}$; $r_i=0$ and 1 correspond to the non-operational and operational states, respectively). The detailed description of the protection system and the operating logics of the main protective relays and backup protective relays can be found in (Wen and Chang, 1996).

**Forward mathematical model**

Take line $L_1$ ($s_4$) as an example. The protective relays which have relationships with $L_1$ are $L_1$Am ($r_4$), $L_1$Bm ($r_5$), $L_1$Ab ($r_8$), $L_1$Bb ($r_9$), and $L_2$Cs ($r_{15}$). Based on the operating logics of the protective relays, the mentioned relays are triggered and $CB_2$, $CB_3$, $CB_5$ operate when there is a fault on $L_1$. The following function reflects the causality relationship between a fault on $L_1$ and the trips of $L_1$Am ($r_4$), $L_1$Bm ($r_5$), $L_1$Ab ($r_8$), $L_1$Bb ($r_9$), $L_2$Cs ($r_{15}$), $CB_2$, $CB_3$, and $CB_5$:

$$s_4^* = 1 - (1 - r_4 c_2)(1 - r_5 c_3)(1 - r_8 c_2)(1 - r_9 c_3)(1 - r_{15} c_5). \quad (1)$$

We call it the possible state function of $L_1$. $s_4^*=1$ and 0 correspond to the possible fault state and the normal state of $L_1$, respectively. The function can be

simplified as

$$s_4^* = 1 - (1 - r_4)(1 - r_5)(1 - r_8)(1 - r_9)(1 - r_{15}). \quad (2)$$

Possible state functions of other sections are

$$\begin{cases} s_1^* = 1 - (1 - r_1)(1 - r_{12}), \\ s_2^* = 1 - (1 - r_2)(1 - r_{11})(1 - r_{15}), \\ s_3^* = 1 - (1 - r_3)(1 - r_4), \\ s_5^* = 1 - (1 - r_6)(1 - r_7)(1 - r_{10})(1 - r_{11})(1 - r_{12}). \end{cases} \quad (3)$$

Inputting the state information of protective relays and circuit breakers into the forward mathematical model, we can obtain the possible fault sections. These sections will be identified in the next process.

**Backward mathematical model**

The model proposed by (Wen and Chang, 1996) is adopted as the backward mathematical model. The fault diagnosis process is to find the best interpretation of the alarm of protective relays and circuit breakers. It is denoted by the minimum of the following function:

$$E(S) = \sum_{k=1}^{n_r} | r_k - r_k^*(S) | + \sum_{j=1}^{n_c} | c_j - c_j^*(S,R) |, \quad (4)$$

where $n_r$ and $n_c$ are the number of protective relays and circuit breakers in the given power system, respectively; $r_k^*(S)$ is the expectation of the $k$th protective relay in the given power system. As $r_k^*(S)=1$, the $k$th protective relay operates, while $r_k^*(S)=0$ not. $c_j^*(S, R)$ is the expectation of the $j$th circuit breaker in the given power system. As $c_j^*(S, R)=1$, the $j$th circuit breaker operates, while $c_j^*(S, R)=0$ not.

Also, we use the example in Fig.1 to illustrate how to determine $r_k^*(S)$ and $c_j^*(S, R)$.

There are three kinds of protective relays: main protective relays, primary backup protective relays and secondary protective relays. It is easy to obtain the expectation functions of main protective relays based on their operating logics:

$$\begin{cases} r_1^*(S) = s_1, \ r_2^*(S) = s_2, \ r_3^*(S) = s_3, \\ r_4^*(S) = r_5^*(S) = s_4, \ r_6^*(S) = r_7^*(S) = s_5. \end{cases} \quad (5)$$

It is a little difficult to determine the expectation function of backup protective relays. Take $L_1Ab$ ($r_8$) as an example. $L_1Ab$ should be triggered when there is a fault on $L_1$ and $L_1Am$ ($r_4$) fails to operate. So the expectation function can be obtained as

$$r_8^*(S) = s_4(1 - r_4). \qquad (6)$$

Other functions can be obtained in the same way:

$$\begin{cases} r_9^*(S) = s_4(1-r_5), \quad r_{10}^*(S) = s_5(1-r_6), \\ r_{11}^*(S) = s_5(1-r_7), \\ r_{12}^*(S) = 1-[1-s_2(1-c_3)][1-s_5(1-c_3)(1-c_4)], \ (7) \\ r_{13}^*(S) = s_1(1-c_2), \quad r_{14}^*(S) = s_4(1-c_5), \\ r_{15}^*(S) = 1-[1-s_2(1-c_4)][1-s_4(1-c_3)(1-c_4)]. \end{cases}$$

The expectation functions of circuit breakers are more difficult to determine than those of protective relays, because the state of a circuit breaker has relationships with several sections and protective relays. For example, the expectation of $CB_2$ is dependent on the state of A ($s_1$), B ($s_2$), $L_1$ ($s_4$), $L_2$ ($s_5$), Am ($r_1$), $L_1Am$ ($r_4$), $L_1Ab$ ($r_8$) and $L_1As$ ($r_{12}$). Based on the operating logics of circuit breakers, we determine the expectation functions of circuit breakers as follows:

$$\begin{cases} c_1^*(S,R) = s_1r_1, \\ c_2^*(S,R) = \max\{s_1r_1, \ s_4r_4, \ s_4(1-r_4)r_8, \\ \quad \{1-[1-s_2(1-c_3)][1-s_5(1-c_3)(1-c_4)]\}r_{12}\}, \\ c_3^*(S,R) = \max\{s_2r_2, \ s_4r_5, \ s_4(1-r_5)r_9, \ s_1(1-c_2)r_{13}\}, \\ c_4^*(S,R) = \max\{s_2r_2, \ s_5r_6, \ s_5(1-r_6)r_{10}, \ s_3(1-c_5)r_{14}\}, \\ c_5^*(S,R) = \max\{s_3r_3, \ s_5r_7, \ s_5(1-r_7)r_{11}, \\ \quad \{1-[1-s_2(1-c_4)][1-s_4(1-c_3)(1-c_4)]\}r_{15}\}, \\ c_6^*(S,R) = s_3r_3. \end{cases}$$

$$(8)$$

The first key point in this paper is to reduce the scale of fault sections greatly by adding the forward mathematical model into the fault diagnosis process. For example, there are $2^{100}$ possible fault cases in a system consisting of 100 sections. By using the forward mathematical model, when a general fault occurs, only a few cases need to be considered. Compared to many methods developed by other

researchers (Wen and Chang, 1996; Huang Y.C., 2002; Hu *et al.*, 2005), the model presented in this paper is better for reducing the scale of fault section estimate.

Since fault diagnosis has become a mathematical problem, the best way to resolve it is to use GAs. Here, the second key point is to parallel the GAs for speeding up the fault diagnosis process.

## PARALLEL GENETIC ALGORITHMS

There are a few approaches to applying parallel computation to power system problems, such as optimal reactive power dispatch (ORPD), load forecasting (LF), optimal power flow calculation (OPFC), etc. Parallel genetic algorithms (PGAs) perform the GAs in parallel and have three different versions: global single-population master-slave genetic algorithms (GPGAs), massively parallel genetic algorithms (MPGAs), and distributed genetic algorithms (DGAs). Here, we adopt GPGAs in computation because it is easier to meet the requirement of programming in MPI (message passing interface) environment.

There is only one population in GPGAs, but evaluation of fitness is distributed among several processors. Here, selection and mating are global: each individual may compete and mate with any other. The evaluation of the individuals is usually parallelized, because the fitness of an individual is independent of the rest of the population and there is no need to communicate with each other during this phase. Communication occurs only as each slave receives its subset of individuals to evaluate fitness when the slaves return the fitness values. The GPGA is synchronous if the algorithm stops and waits for the fitness values for all the population before proceeding into the next generation. Synchronous GPGAs have the same properties as the sequential GAs, but it is faster if the algorithm spends most of the time on the evaluation process.

The MPI approach is chosen to implement GPGAs in this paper. Fig.2 shows an example of MPI based parallel computation using *n* processors. First, processor 0 reads and preprocesses all the data before scattering them to other processors' memory. After processor 0's initialization work, all the processors calculate the fitness of individuals with the data sent

by processor 0 through the message passing interface (Obviously, processor 0 already has the required data in its memory, so there is no communication inside it). When the calculations are done, the results are sent back to processor 0's receiving buffer to be dealt with. Then processor 0 decides whether to continue computing or to output the final results and terminate the program. In this section, we will first briefly outline the transmitter optimization, and then consider the Lloyd algorithm for the codebook construction.

Assume there are $n$ processors running this program simultaneously. The GPGA flow can be described as follows:

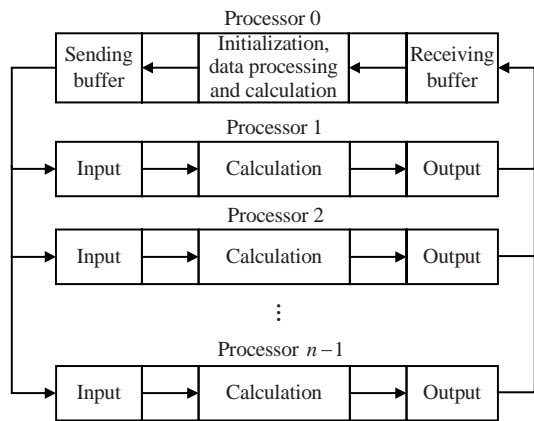Step 1: Processors $0\sim(n-1)$ initialize GPGA parameters and MPI environment.



**Fig.2 An example of parallel computation based on MPI**

Step 2: Processor 0 reads power system data and generates a population of chromosomes, each of which stands for a candidate solution to the FDI problem. Let the size of chromosome population be $p$.

Step 3: Let $m=p/n$, $m$ is an integer. If $p$ cannot be divided exactly, let $m=m+1$, thus $mn>p$.

Step 4: Processor 0 partitions the population into $n$ parts, each of which has $m$ or $m-1$ chromosomes. Then processor 0 sends these parts to other processors. Each processor receives $m$ or $m-1$ chromosomes from processor 0.

Step 5: Each processor calculates the fitness values of its own chromosomes based on the method in the previous section.

Step 6: Processor 0 gathers the fitness values of all chromosomes from processors $1\sim(n-1)$ and produces the next generation based on GA parameters.

Step 7: If the generation number meets the requirement, go to Step 8; otherwise, go to Step 4.

Step 8: Processor 0 outputs the solution with the best fitness value in the last iteration. All processors exit in MPI environment.

## TEST RESULTS

A power system as shown in Fig.3, which consists of 28 sections, 84 protective relays and 40 circuit breakers, is used to test the new mathematical model for the FDI problem.
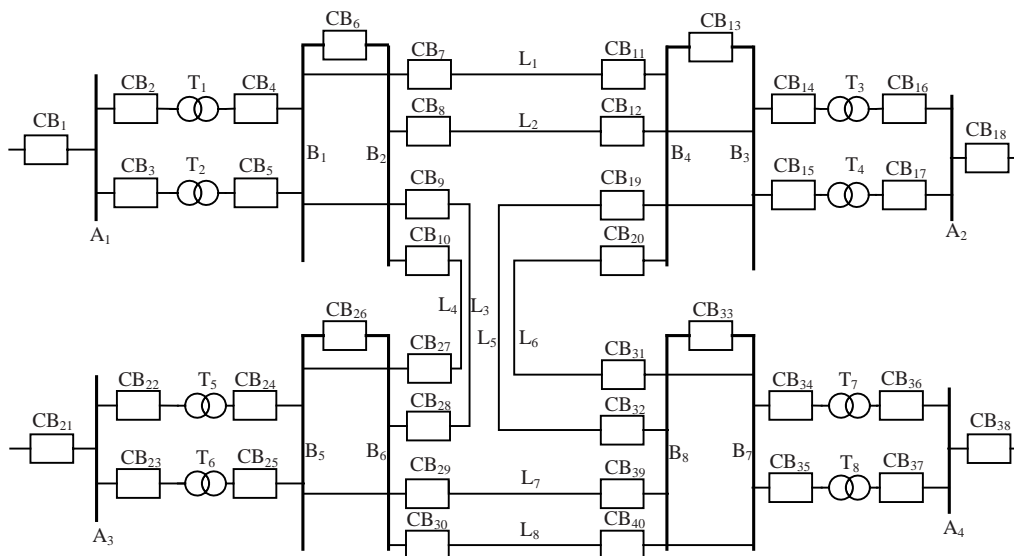


**Fig.3 A sample power system consisting of 28 sections, 84 protective relays and 40 circuit breakers**

The following is a very brief description of the system:

$s_1\sim s_{28}$ represent 28 sections: $A_1\sim A_4$, $T_1\sim T_8$, $B_1\sim B_8$, $L_1\sim L_8$; $c_1\sim c_{40}$ represent 40 circuit breakers: $CB_1\sim CB_{40}$; $r_1\sim r_{84}$ represent 84 protective relays. Among the 84 protective relays, $r_1\sim r_{36}$ represent 36 main protection relays: $A_1m\sim A_4m$, $T_1m\sim T_8m$, $B_1m\sim B_8m$, $L_1Sm\sim L_8Sm$, $L_1Rm\sim L_8Rm$; $r_{37}\sim r_{60}$ represent 24 backup protection relays: $T_1b\sim T_8b$, $L_1Sb\sim L_8Sb$, $L_1Rb\sim L_8Rb$; $r_{61}\sim r_{84}$ represent 24 second backup protection relays: $T_1s\sim T_8s$, $L_1Ss\sim L_8Ss$, $L_1Rs\sim L_8Rs$.

The operating logics of protective relays are described in the Appendix of (Wen and Chang, 1996) and not presented here. The forward and backward models of this system are built on the rules described in Section 2. We will study several fault cases as listed in Table 1.

Case 1: This is a very simple case. Input the alarm signal information into the forward model and obtain the result that the fault section is $B_1$ from the function

$$s_{13}^{*}=1-(1-r_{13})(1-r_{61}c_4)(1-r_{62}c_5)(1-r_{77}c_{11}) \\ \cdot (1-r_{78}c_{12})(1-r_{79}c_{28})(1-r_{80}c_{27}). \quad (9)$$

Here, $s_{13}^{*}$ represents the possible fault state of $B_1$. There is only one possible fault section and no need to identify it through the backward model.

Case 2: Possible fault sections are $L_5$, $L_7$, $B_7$, $B_8$, $T_7$ and $T_8$ from the forward model's computation. There are $2^6=64$ possible fault cases. It is easy to obtain the result by sending all 64 cases to the parallel computation processors. Hence, in the backward model we use GAs when the number of possible fault sections is more than 8 and the traversal method when the number is not more than 8. The final result is that the fault sections are $L_5$, $L_7$, $B_7$, $B_8$ and $T_8$.

Case 3: A more complex case is given here. Possible fault sections are $L_1$, $L_2$, $L_5$, $L_7$, $L_8$, $A_2$, $A_4$, $B_3$, $B_6$, $B_7$, $B_8$, $T_7$ and $T_8$ from the forward model's computation. The real fault ones are found from the 13 sections through the backward model based on GPGA. The final result is that the fault sections are $L_1$, $L_7$, $L_8$, $B_3$, $B_6$, $B_7$, $T_7$ and $T_8$. From the results, it can be seen that $B_3m$, $B_6m$, $B_7m$ fail to operate and $L_2Rb$ makes a wrong operation when the fault occurs.

The results of these cases are compared with traditional algorithms in Table 2 and the comparison shows that the new model can also deduce the correct results.

The parameters of GPGAs are set as follows: Population size=50, maximum permitted generations number=10, initial crossover probability=0.9, initial mutation probability=0.001, MPI processors number =10, and the stop criterion is that the maximum permitted generations number has been reached.

All 28 sections' state should be identified in traditional algorithms, while GPGAs only search the fault sections from the possible fault ones which are given in the forward model results here. Obviously, we can conclude that the workload of the new model is greatly reduced through the comparison in Table 3. In addition, we compared the calculation time of Case 3 with that of traditional algorithms. The new model based on PGAs consumes less than 1 s while other methods for FDI need more than 5 s. The simulation results show that the new model based on GPGAs can be applied to online applications in large-scale power systems.

**Table 1  Alarm signal list of fault cases**

| Case | Alarm signals | |
|------|---------------|---|
|      | Operated relays | Tripped circuit breakers |
| 1 | $B_1m$, $L_2Rs$, $L_4Rs$ | $CB_4$, $CB_5$, $CB_7$, $CB_9$, $CB_{12}$, $CB_{27}$ |
| 2 | $T_7m$, $T_8b$, $B_7m$, $B_8m$, $L_5Sm$, $L_5Rb$, $L_6Ss$, $L_7Rm$, $L_8Ss$ | $CB_{19}$, $CB_{20}$, $CB_{30}$, $CB_{32}\sim CB_{35}$, $CB_{39}$ |
| 3 | $T_3s$, $T_4s$, $T_7m$, $T_8m$, $T_7s$, $T_8s$, $L_1sm$, $L_1Rb$, $L_2Ss$, $L_2Rb$, $L_5Ss$, $L_5Rm$, $L_5Rs$, $L_6Ss$, $L_7Sb$, $L_7Ss$, $L_7Rm$, $L_7Rs$, $L_8Sm$, $L_8Ss$, $L_8Rb$, $L_8Rs$ | $CB_7$, $CB_8$, $CB_{11}$, $CB_{12}$, $CB_{14}$, $CB_{15}$, $CB_{19}$, $CB_{20}$, $CB_{29}$, $CB_{30}$, $CB_{34}\sim CB_{40}$ |

**Table 2  Results compared with traditional algorithms**

| Case | Forward model | Backward model | Traditional algorithms |
|------|---------------|----------------|------------------------|
| 1 | $B_1$ | $B_1$ | $B_1$ |
| 2 | $L_5$, $L_7$, $B_7$, $B_8$, $T_7$, $T_8$ | $L_5$, $L_7$, $B_7$, $B_8$, $T_8$ | $L_5$, $L_7$, $B_7$, $B_8$, $T_8$ |
| 3 | $L_1$, $L_2$, $L_5$, $L_7$, $L_8$, $A_2$, $A_4$, $B_3$, $B_6$, $B_7$, $B_8$, $T_7$, $T_8$ | $L_1$, $L_7$, $L_8$, $B_3$, $B_6$, $B_7$, $T_7$, $T_8$ | $L_1$, $L_7$, $L_8$, $B_3$, $B_6$, $B_7$, $T_7$, $T_8$ |

**Table 3  Comparison of searching scales between different models**

| Case | Forward and backward models | Traditional model | Improved model |
|------|------------------------------|-------------------|----------------|
| 1 | 1 | $2^{28}$=268 435 456 | $2^2$=4 |
| 2 | $2^6$=64 | $2^{28}$=268 435 456 | $2^{10}$=1024 |
| 3 | $2^{13}$=8192 | $2^{28}$=268 435 456 | $2^{18}$=262 144 |

CONCLUSION

In this paper, a mathematical model for FDI based on GPGAs has been developed. The model consists of forward and backward models which are built on the operating logics of protective relays and circuit breakers. The scale of possible fault sections is reduced greatly by the forward model and the genetic algorithm is paralleled in GPGAs. These two methods accelerate the FDI process. Simulation results show that the forward and backward models based on GPGAs is time and computation efficient for large-scale power systems.

**References**

Cardozo, E., Talukdar, S.N., 1988. A distributed expert system of fault diagnosis. *IEEE Trans. on Power Syst.*, **3**(2):641-646. [doi:10.1109/59.192917]

Chantler, M., Pogliano, P., Aldea, A., Tornielli, G., Wyatt, T., Jolley, A., 2000. Use of fault-recorder data for diagnosing timing and other related faults in electricity transmission networks. *IEEE Trans. on Power Syst.*, **15**(4):1388-1393. [doi:10.1109/59.898117]

Chin, H.C., 2003. Fault section diagnosis of power system using fuzzy logic. *IEEE Trans. on Power Syst.*, **18**(1):245-250. [doi:10.1109/TPWRS.2002.807095]

Chuang, C.Y., Ke, Y., Chen, Y.L., 2006. Rule-expert knowledge-based Petri net approach for distribution system temperature adaptive feeder reconfiguration. *IEEE Trans. on Power Syst.*, **21**(3):1362-1370. [doi:10.1109/TPWRS.2006.876681]

Dysko, A., McDonald, J.R., Burt, G.M., Goody, J., Gwyn, B., 1999. Integrated Modelling Environment: A Platform for Dynamic Protection Modelling and Advanced Functionality. Proc. IEEE Transmission Distribution Conf., **1**:406-411. [doi:10.1109/TDC.1999.755386]

Hu, Z.H., Cai, Y.Z., Li, Y.G., 2005. Data fusion for fault diagnosis using multi-class support vector machines. *J. Zhejiang Univ. Sci.*, **6A**(10):1030-1039. [doi:10.1631/jzus.2005.A1030]

Huang, S.J., 2002. Application of immune-based optimization method for fault-section estimation in a distribution system. *IEEE Trans. on Power Del.*, **17**(3):779-784. [doi:10.1109/TPWRD.2002.1022803]

Huang, Y.C., 2002. Fault section estimation in power systems using a novel decision support system. *IEEE Trans. on Power Syst.*, **17**(2):439-444. [doi:10.1109/TPWRS.2002.1007915]

Liu, L., Logan, K.P., Cartes, D.A., Srivastava, S.K., 2007. Fault detection, diagnostics, and prognostics: software agent solutions. *IEEE Trans. on Vehic. Technol.*, **56**(4):1613-1622. [doi:10.1109/TVT.2007.897219]

Narendra, K.G., Sood, V.K., Khorasani, K., 1998. Application of a Radial Basis Function (RBF) neural network for fault diagnosis in a HVDC system. *IEEE Trans. on Power Syst.*, **13**(1):177-183. [doi:10.1109/59.651633]

Park, Y.M., Kim, G.W., Sohn, J.M., 1997. A Logic Based Expert System, LBES, for Fault Diagnosis of Power System. *IEEE Trans. on Power Syst.*, **12**(1):363-369. [doi:10.1109/59.574960]

Ren, H., Mi, Z.Q., Zhao, H.S., 2005. Power system fault diagnosis by use of encoded Petri net models. *Proc. CSEE*, **25**(20):44-49 (in Chinese).

Sun, Y.M., Jiang, H., Wang, D., 1998. Fault synthetic recognition for an EHV transmission line using a group of neural networks with a time-space property. *IEE Proc.-Gener.*, *Transm. Distrib.*, **145**(3):265-270. [doi:10.1049/ip-gtd:19981919]

Wen, F.S., Chang, C.S., 1996. A New Approach to Fault Section Estimation in Power Systems Based on the Set Covering Theory and a Refined Genetic Algorithm. Proc. 12th Power Systems Computation Conf., Dresden, Germany, **1**:358-365.

Wen, F.S., Han, Z.X., 1997. Probabilistic approach for fault section estimation in power systems based upon a refined genetic algorithm. *IEE Proc.-Gener.*, *Transm. Distrib.*, **144**(2):160-168. [doi:10.1049/ip-gtd:19970802]

Zhu, Y.L., Huo, L.M., Lu, J.L., 2006. Bayesian networks-based approach for power systems fault diagnosis. *IEEE Trans. on Power Delivery*, **21**(2):634-639. [doi:10.1109/TPWRD.2005.858774]