IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 16, 2019, accepted February 24, 2019, date of publication March 4, 2019, date of current version March 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902598

An Innovative Minimum Hitting Set Algorithm for Model-Based Fault Diagnosis in Power Distribution Network

QIUJIE WANG[1], TAO JIN[1,2], (Member, IEEE), AND MOHAMED A. MOHAMED[2,3], (Member, IEEE)

[1]Fujian Key Laboratory of New Energy Generation and Power Conversion, Fuzhou 350116, China
[2]Department of Electrical Engineering, Fuzhou University, Fuzhou 350116, China
[3]Electrical Engineering Department, Faculty of Engineering, Minia University, Minia 61519, Egypt

Corresponding authors: Tao Jin (jintly@fzu.edu.cn) and Mohamed A. Mohamed (dr.mohamed.abdelaziz@mu.edu.eg)

This work was supported by the Fuzhou Municipal Science and Technology Bureau Project of China under Grant 82318048.

ABSTRACT The distribution network plays a great role in the power system, and any fault in it may threaten the safe and stable operation of the power system. Hence, fault diagnosis has an important role in protecting the distribution network and maintaining power system stability. The model-based diagnosis (MBD) is one of the diagnostic methods that have some advantages compared with the common diagnosis methods. The latter, such as expert diagnosis systems depend on the professional experience and fault information. However, the inefficiency and incomplete calculation of the minimum hitting set (MHS) limits the performance of the MBD. In this paper, to overcome these limitations, an innovative MHS algorithm which considers the nature and characteristics of the distribution network is proposed. In the proposed algorithm, a new fitness function is constructed by a weighted combination method, which allows particles to move directly toward the MHS without the effect of non-hitting set particles. Finally, three cases studies and discussion on the distribution network are introduced to validate the advantages of the proposed algorithm, in terms of the computation efficiency and accuracy.

INDEX TERMS Distribution network, fault diagnosis, model based diagnosis, minimum hitting set algorithm, power system stability.

I. INTRODUCTION

Distribution network is an essential part of the power system which if disrupted, people's natural life and production will be seriously affected. Therefore it is important to locate and isolate faults timely and accurately to ensure the reliability and security of the distribution network [1]. Fault diagnosis incorporates detection, isolation and identification [2]. The main task of fault diagnosis is the fault detection, which check whether there is system malfunction or fault and decide the time when the fault happens. Moreover, fault isolation is to isolate the area of the faulty component, and fault identification is to decide the sort, form, and size of the fault.

Numerous impressive efforts have been exhibited toward developing of the fault diagnosis methods in distribution network. Among them, the explicit methods [3]–[8] such as the analytical models [3], [4] and the implicit methods [9]–[17] such as the expert systems [14], [15]. Although these methods offer incredible solutions to the fault diagnosis, it has some imperfections. Most of the implicit methods highly relied on the professional experiences of the fault which is quite troublesome. Accordingly, the faults outside experience are difficult to be diagnosed. Moreover, the time of response of these methods are normally not pertinent to a real-time environment because of the traditional knowledge representation and inference mechanism [14].

Artificial Neural Network (ANN) is a potential solution has been utilized for fault diagnosis [15], [16]. However, the practical application of ANN still remains issues such as moderate convergence, awful transparency and slow training process. This makes ANN not satisfactory approach in fault diagnosis.

To beat the previously mentioned imperfections, a model-based diagnosis (MBD) method is presented in this paper. This method has been applied in many fields [18]–[20] and some applications in power distribution network [21]–[23].

The associate editor coordinating the review of this manuscript and approving it for publication was Ruqiang Yan.

MBD have received impressive consideration in faults diagnosis, because of the expanding interest for safe, robustness and reliable operation of uncertain and complex dynamic systems [18]. In additional, the operation results demonstrated that MBD has better performance without fault experience; moreover, it can solve the systemic migration problem as well [19]. Furthermore, If the system is somewhat changed, just the model should be slightly changed which make the procedure faster [21].

In fault diagnosis field, estimation of the minimal hitting set (MHS) is the bottleneck step of MBD [23]. The MHS algorithm is used to find diagnoses which clarify all observed conflicts.

There are sex steps in MBD, the first three steps are carried out offline and the last three steps are carried out online. As the MHS calculation is the more complex and time-consuming in the last three steps, the performance of MBD mainly depends on the calculation efficiency and accuracy of the MHS.

There are many algorithms have been employed to calculate MHS. The hitting set tree (HS-TREE) algorithm [24], the binary hitting set tree (BHS-TREE) algorithm [25] and other algorithms have also been put forward [26]–[28]. The main shortcomings of these algorithms are: easily losing correct solution as a result of pruning, and low calculation efficiency [25]. Moreover, these algorithms are based on tree or graph theory, they need to generate numerous nodes. So that the calculations of MHS on these algorithms depend on the population size which in turn affects the run time and accuracy of calculations. Hence, the calculating time of these algorithms is unacceptable in the case of calculating MHS on a large scale of conflict sets. Therefore, the intelligent algorithms have been used to calculate the MHS, such as genetic algorithm (GA) [29], binary particle swarm optimization (BPSO) [30], improved differential evolution algorithm (IDEA) [31], immune genetic algorithm (IGA) [32] and series hybrid algorithm for BPSO and GA (BPSO-GA) [33]. These algorithms have the advantages of low difficulty, low complexity, fast convergence and global search capability. However, these algorithms are general, that is to say, the nature and characteristic of the problems to be solved are not taken into account. As a result, the efficiency and accuracy of MHSs calculations are greatly limited.

In order to improve the calculation efficiency and accuracy of MHSs, an innovative MHS algorithm is proposed in this paper. This algorithm takes into account the nature and characteristics of the distribution network. In the proposed algorithm a new fitness function is constructed by a weighted combination method. This function allows the particle iterate towards the MHS directly without the effect of non-HS particles. By analyzing the topology of the distribution network, the characteristic of no intersection between the minimum conflict sets (MCSs) of the distribution networks is derived. Based on this feature, a new MHS criterion is constructed, which could avoid the participation of MHS ensured strategy (MHSES). The new fitness function and MHS criterion could improve effectively the efficiency and accuracy of MHS calculation.

II. FUNDAMENTALS OF MBD

In this section, the relevant theories of the consistency-based MBD are presented. With these theories, the diagnosis steps of consistency-based MBD in distribution networks are organized. From the diagnosis steps, it can be clearly seen that the last three steps of the MBD are online in which the MHS calculations take up most of the run time. Therefore, there is a significant impact of the performance of MHS calculation on the application of MBD in fault diagnosis in distribution networks.

A. THE THEORY OF MBD

Some related definitions and theorems regarding MBD are introduced as the following:

Definition 1: A diagnosis system can be described by a three sets (SD, COMPS, OBS). The system description (SD) is a set of first order sentences. The system components (COMPS) are a set of constants. The system observation (OBS) is a finite set of first order sentences.

Definition 2: A set C is a conflict set (CS) of the system, if $C = [c_1, c_2, \ldots, c_n] \subseteq$ COMPS, and SD $\cup$ OBS $\cup$ $[\neg ab(c_1), \neg ab(c_2), \ldots, \neg \mathrm{ab}(c_n)]$ is inconsistent. Where, "ab" is "*abnormal*", "$\neg ab$" is "*normal*". If no subset of the CS is a CS for the system, the CS is a minimal conflict set (MCS).

Definition 3: A set H is called a hitting set (HS) for a MCS cluster if: $H \subseteq \cup C_i$, $\forall C_i \in MCSs$, $H \cap C_i \neq \emptyset$. If no proper subset of the HS is a HS, the HS is a minimal hitting set (MHS).

Theorem 1: A system component $\Delta \subseteq$ COMPS is a diagnosis candidate of the system if Δ is a MHS of the system.

B. DIAGNOSIS PROCESS OF MBD FOR DISTRIBUTION NETWORK

According to relevant theoretical knowledge of MBD, the steps of MBD for distribution networks are as follows [21], [22]:

1) Establishing the system model. MBD model is established for distribution network by describing the analytic redundancy relationships (ARRs) in normal case, ground fault case and break fault case of the buses and feeder sections.
2) System division. The location of distribution circuit breaker, sectional switch and connection switch is chosen as the measurement point. Furthermore, the electrical quantity is collected by the existing distribution automation device feeder terminal units (FTUs). According to the position of the measurement points, the distribution network is divided into multiple subsystems.

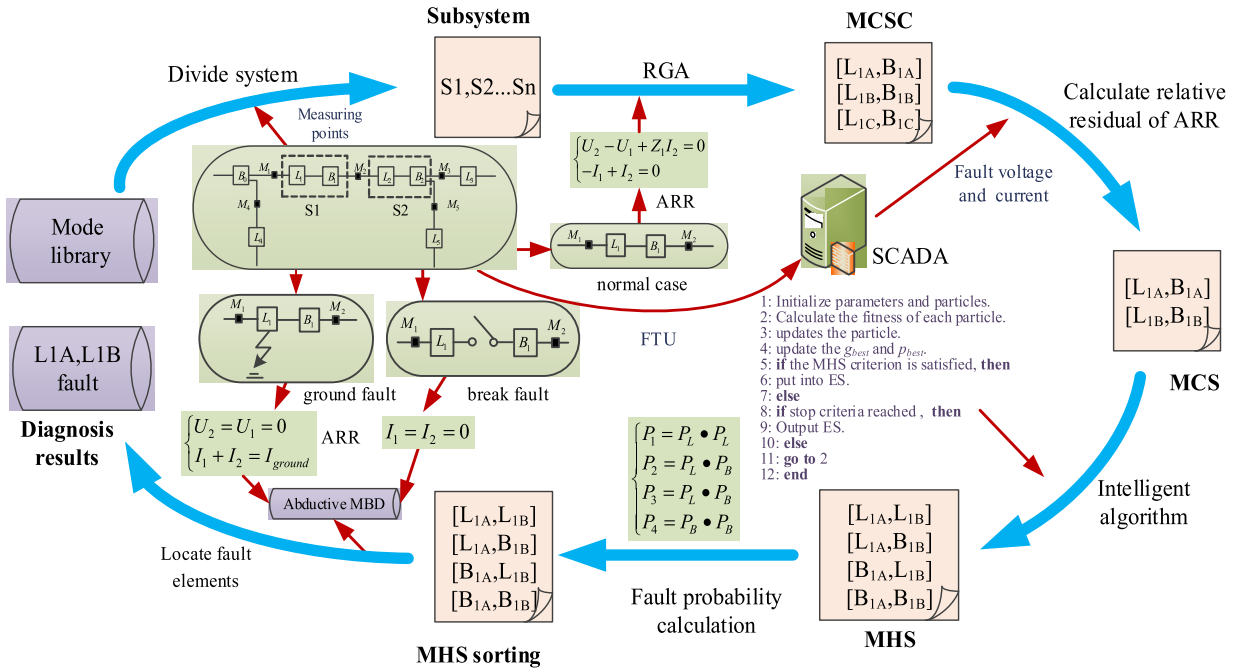


FIGURE 1. The steps of MBD for distribution network.

3) Searching for the MCSCs. According to the information implied in the ARRs in normal case, the MCSCs in each subsystem are searched off-line by the relation guiding algorithm (RGA).
4) Identify the MCSs. The FTUs at each measurement points upload the fault electrical quantity to the SCADA system. The fault electrical quantity and the parameters of the buses and feeder sections are counted into the corresponding ARRs, and the relative residual of each ARR is calculated. If the relative residuals of the ARR are larger than the prescribed threshold, the MCSC corresponding to the ARR is the MCS.
5) MHSs calculation. The intelligent algorithm uses the fitness function and the MHS criterion to find all the MHSs.
6) Locating fault elements. The failure probability of a MHS is the product of the failure probability of all components it contains. When the failure probability of each MHS is calculated and ranked, the MHS with maximum failure probability is the diagnostic result.

The steps of MBD for distribution networks can be described as shown in Fig. 1 [23]. During faults, the performances of the last three online steps determine the efficiency and accuracy of MBD. Among them, identifying MCSs, locating fault elements are simple formula calculation, and MHSs calculation is complex iteration calculation. Therefore, the latter is the more complex and time-consuming, which in turn determines the efficiency and accuracy of MBD to some extent [31].

III. PROPOSED MHS CALCULATION SCHEME

This section proposes a new fitness function and MHS criterion considering the nature and characteristic of the distribution network, which fundamentally simplifies MHS algorithms. In order to implement MHS calculation, an intelligent optimization algorithm is necessary. Therefore, the BPSOGA is introduced.

A. FITNESS FUNCTION ANALYSIS

The fitness function adopted in reference [35] is:

$$\max Fit1(x) = \frac{N_{cx}}{N_c} \tag{1}$$

where N_{cx} is the number of MCS in the MCS cluster C that intersect with the current particle x, N_c is the number of MCS in the MCS cluster C.

The fitness function adopted in reference [31] is:

$$\min Fit2(x) = 1 - \frac{N_{cx}}{N_c} \tag{2}$$

The above two fitness functions are essentially equal and can effectively avoid the influence of non-HS particles on the iterations. However, they only allow the particles iterate towards the HS rather than the MHS.

The fitness function adopted in reference [33] is:

$$\max Fit3(x) = \frac{N_{cx}}{L_x} \tag{3}$$

where L_x is the length of the particle x, which is the number of "1" in the particle. The fitness function just allows particles iterate toward the MHS, but it is easily affected by non-HS particles during iterations.

In order to implement that the particles iterate to the MHS directly without being affected by non-HS particles, a new fitness function is constructed as follows:

$$\min Fit4(x) = (1 - \frac{N_{cx}}{N_c}) + \omega L_x \tag{4}$$

where $1 - N_{cx}/N_c$ makes the particles iterate in the direction of the HS, and ωL_x allows the particle to iterate to the MHS on the basis of the HS.

To avoid the effect of the non-HS particles, the magnitude order of ωL_x should be smaller than $1 - N_{cx}/N_c$. ω is used to reduce the progressive relation from HS to MHS by one order of magnitude in fitness value. Thus, the two progressive relations are effectively fused, that is the non-HS, HS, and MHS particles can be distinguished in the fitness value. Therefore, the weight coefficient ω is crucial to the particle iteration performance. Assuming the most complex fault in the distribution network is triple three-phase fault. Then the maximum value of N_c is 9 and the maximum value of L_x is 18, which are illustrated in Table 3. Then, the minimum interval $1/N_c$ is 1/9. In order to distinguish HS from MHS particles in fitness value, the maximum value of ωL_x should be less than the minimum interval $1/N_c$. Therefore, the range of ω is obtained $\omega \leq (1/N_{c,\max})/L_{x,\max} = 0.0062$. In order to make it more reliable to distinguish HS from MHS in fitness value, the choice of ω should leave some margin. At the same time, in order to help readers better understand the proposed fitness function, we tentatively set ω at 0.001.

To better understand the difference between these fitness functions, the MCS cluster (C) obtained from distribution network diagnosis is set as follows:

$$C = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \tag{5}$$

The particle swarm (X_m) in the mth iteration is set as follows:

$$X_m = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{6}$$

where $x_1 - x_4$ are the non-HS particles, x_5 is the HS particle, x_6 is the MHS particle. The above mentioned fitness functions are used to calculate the fitness value of each particle, and the results are shown in Table 1.

TABLE 1. Fitness value comparison.

Particle	$Fit1(x)$	$Fit2(x)$	$Fit3(x)$	$Fit4(x)$
x_1	0.333	0.667	1	0.667
x_2	0.667	0.333	2	0.334
x_3	0.667	0.333	2	0.334
x_4	0.667	0.333	1	0.335
x_5	1	0	1	0.003
x_6	1	0	1.5	0.002

As shown in Table 1, both of the fitness functions $Fit1(x)$ and $Fit2(x)$ are not able to distinguish the HS particle x_5 and the MHS particle x_6 by the fitness value. $Fit3(x)$ is able to distinguish the HS particle x_5 and x_6 by the fitness values, but it is easily affected by non-HS particles x_2 and x_3. $Fit4(x)$ is not only able to distinguish the HS particle x_5 and the MHS particle x_6 from the fitness value, but also not affected by all the non-HS particles.

B. MHS CRITERION

The criterions corresponding to the fitness functions$Fit1$, $Fit2$, and $Fit3$ are as the following, respectively:

$$\begin{cases} Fit1 = 1 \Rightarrow HS \\ Fit1 < 1 \Rightarrow non\text{-}HS \end{cases} \tag{7}$$

$$\begin{cases} Fit2 = 0 \Rightarrow HS \\ Fit2 > 0 \Rightarrow non\text{-}HS \end{cases} \tag{8}$$

$$\begin{cases} Fit3 = 1 \Rightarrow MHS \\ Fit3 < 1 \Rightarrow non\text{-}MHS \end{cases} \tag{9}$$

The formulas (7) and (8) are able to identify the particles x_5 and x_6, but not able to distinguish the particles x_5 and x_6. Formula (9) is able to identify the particle x_6, but not able to distinguish the particles x_2 and x_4. To deduce the MHS criterion that can identify the MHS particles without being affected by non-HS particles, the topology of the distribution network needs to be analyzed. The multi-branch node (three or more) is an important part of the distribution network, which determines the complexity of the topology of the distribution network. Therefore, a distribution network consisting of two or three branch nodes is taken as an example to analyze the characteristics of the MCS for distribution network as shown in Fig. 2.

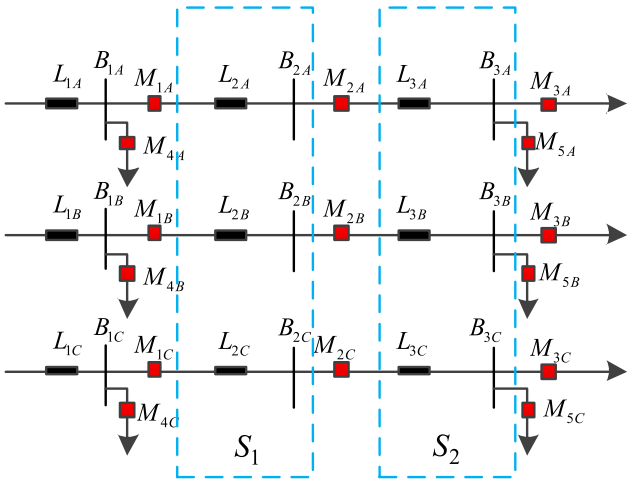


FIGURE 2. Typical topology structure of distribution network.

Where L_i is the feeder section, Z_i is the corresponding impedance, B_i is the bus, M_i is the measuring point, U_i is the corresponding voltage, and I_i is the corresponding current. By the basic principle of MBD, the area between the adjacent measuring points is an independent subsystem. Therefore, the system can be divided into two subsystems S_1 and S_2 as shown in Fig. 2.

Under normal condition, the A-phase ARRs of the subsystem S_1 and S_2 can be described according to Kirchhoff's voltage law as the following:

$$U_{2A} - U_{1A} + Z_{2A}I_{2A} = 0 \quad (10)$$

$$U_{3A} - U_{2A} + Z_{3A}(I_{3A} + I_{5A}) = 0 \quad (11)$$

The corresponding minimum support environments or MCSCs are $[L_{2A}, B_{2A}]$ and $[L_{3A}, B_{3A}]$. When faults occur in phase A of subsystem S_1 and S_2, the obtained MCSs are $[L_{2A}, B_{2A}]$ and $[L_{3A}, B_{3A}]$. The set of components of all MCSs is $[L_{2A}, B_{2A}, L_{3A}, B_{3A}]$, as a result, the corresponding binary codes of the two MCSs are [1100] and [0011], respectively. Obviously, there is no intersection between the two MCSs of phase A.

Correspondingly, the B-phase ARRs of the subsystem S_1 and S_2 are described as follows:

$$U_{2B} - U_{1B} + Z_{2B}I_{2B} = 0 \quad (12)$$

$$U_{3B} - U_{2B} + Z_{3B}(I_{3B} + I_{5B}) = 0 \quad (13)$$

The corresponding MCSCs are $[L_{2B}, B_{2B}]$ and $[L_{3B}, B_{3B}]$. When faults occur in phase B of subsystem S_1 and S_2, the obtained MCSs are $[L_{2B}, B_{2B}]$ and $[L_{3B}, B_{3B}]$. The set of components of all MCSs is $[L_{2B}, B_{2B}, L_{3B}, B_{3B}]$, as a result, the corresponding binary codes of two MCSs are [1100] and [0011], respectively. Also, there is no intersection between the two MCSs of phase B.

In the same way, the C-phase ARRs of the subsystem S_1 and S_2 are described as follows:

$$U_{2C} - U_{1C} + Z_{2C}I_{2C} = 0 \quad (14)$$

$$U_{3C} - U_{2C} + Z_{3C}(I_{3C} + I_{5C}) = 0 \quad (15)$$

The corresponding MCSCs are $[L_{2C}, B_{2C}]$ and $[L_{3C}, B_{3C}]$. When faults occur in phase C of subsystem S_1 and S_2, the obtained MCSs are $[L_{2C}, B_{2C}]$ and $[L_{3C}, B_{3C}]$. The set of components of all MCSs is $[L_{2C}, B_{2C}, L_{3C}, B_{3C}]$, as a result, the corresponding binary codes of the two MCSs are [1100] and [0011], respectively. Also, there is no intersection between the two MCSs of phase C.

When faults occur in phase A of subsystem S_1 and phase B of subsystem S_2, the corresponding AARs are presented in equations (10) and (13) and the corresponding MCS cluster are $[L_{2A}, B_{2A}]$ and $[L_{3B}, B_{3B}]$. Similarly when faults occur in phase A of subsystem S_1 and phase C of subsystem S_2, the corresponding AARs are introduced in equations (10) and (14) and the corresponding MCS cluster are $[L_{2A}, B_{2A}]$ and $[L_{3C}, B_{3C}]$.

When faults occur in phase B of subsystem S_1 and phase A of subsystem S_2, the corresponding AARs are as in equations (12) and (11) and the corresponding MCS cluster are $[L_{2B}, B_{2B}]$ and $[L_{3A}, B_{3A}]$. When faults occur in phase B of subsystem S_1 and phase C of subsystem S_2, the corresponding AARs are as in equations (12) and (15) and the corresponding MCS cluster are $[L_{2B}, B_{2B}]$ and $[L_{3C}, B_{3C}]$.

When faults occur in phase C of subsystem S_1 and phase A of subsystem S_2, the corresponding AARs are as in (14) and (11) and the corresponding MCS cluster are $[L_{2C}, B_{2C}]$ and $[L_{3A}, B_{3A}]$. When faults occur in phase C of subsystem S_1 and phase B of subsystem S_2, the corresponding AARs are shown in (14) and (13) and the corresponding MCS cluster are $[L_{2C}, B_{2C}]$ and $[L_{3B}, B_{3B}]$. Obviously, there isn't a mutual element between the MCSs in the respective MCS clusters. Moreover, it can conclude that there isn't intersection between MCSs among the different systems.

During faults occur in phase A and B of subsystem S_1, the corresponding AARs are as in (10) and (12) and the corresponding MCS cluster are $[L_{2A}, B_{2A}]$ and $[L_{2B}, B_{2B}]$. When faults occur in phase A and C of subsystem S_1, the corresponding AARs are as in (10) and (14) and the corresponding MCS cluster are $[L_{2A}, B_{2A}]$ and $[L_{2C}, B_{2C}]$. When faults occur in phase B and C of subsystem S_1, the corresponding AARs are as in (12) and (14) and the corresponding MCS cluster are $[L_{2B}, B_{2B}]$ and $[L_{2C}, B_{2C}]$. Also, there isn't a mutual element between the MCSs in the respective MCS clusters.

It can be concluded that there is no intersection between MCSs within a single subsystem. Therefore, the MCSs of the distribution network fault diagnosis possess the characteristic of no intersection between all the MCSs. According to this characteristic, the length of MHS L_{MHS} can be obtained as:

$$L_{MHS} = N_c \quad (16)$$

Therefore, the MHS criterion for the distribution network diagnosis is defined as:

$$\begin{cases} Fit4(x) > \omega N_c \Rightarrow non\text{-}MHS \\ Fit4(x) = \omega N_c \Rightarrow MHS \end{cases} \quad (17)$$

C. MHS CALCULATION PROCESS

In this paper, the new proposed fitness function and MHS criterion is used to calculate the MHS by BPSOGA. In BPSOGA, the integration between GA and BPSO is used to improve the performance of MHS calculation. In the calculation, if a particle is determined to be MHS, it will be put into an elite set (ES). The following steps details the specific calculation process of BPSOGA:

1. Dividing the population into two subpopulations of the same number.
2. The subpopulation#1 is initialized according to BPSO and subpopulation#2 is initialized according to GA.
3. The two subpopulations update the particles respectively.
4. The subpopulation#1 calculates the fitness of each particle and updates the global best (gbest1) and best particle (pbest1).
5. Similarly, the subpopulation#2 calculates the fitness of each particle and updates (gbest2) and (pbest2).
6. Then comparing gbest1 and gbest2 and updates the gbest.
7. Similarly, comparing pbest1 and pbest2 and updates the pbest.
8. If the MHS criterion is satisfied, then put into ES.

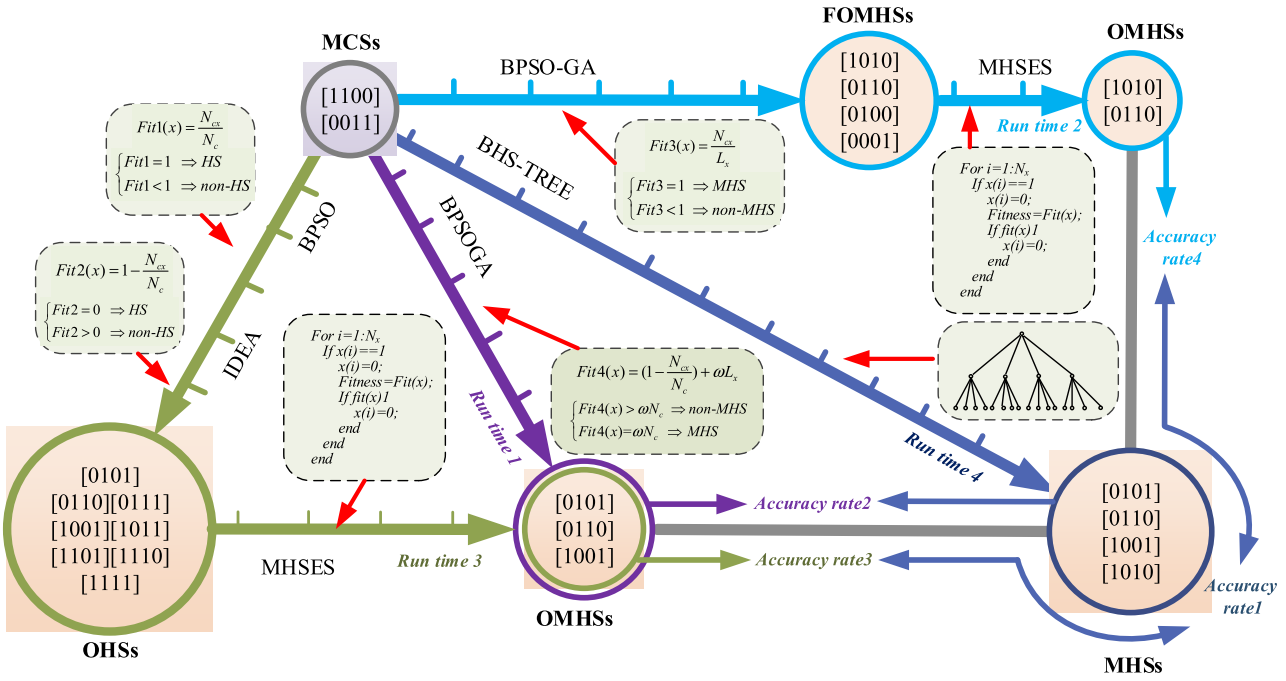


FIGURE 3. Algorithms analysis diagram.

9. If stop criteria is reached, then output ES. If not, then update the particles again.

IV. COMPARISON AND ANALYSIS OF THE ALGORITHMS

To validate the performance and advantages of proposed BPSOGA algorithm, the MHSs outcomes are compared with that from BHS-TREE, IDEA, PSO-GA, and PSO [35], [36]. The process of calculation of MHS for various algorithms is depicted in Fig. 3. The MCS cluster is set as [1100], [0011]. As shown in Fig. 3 the length of the axis represents the run time, and the size of the circle represents the number of sets. OMHSs are the obtained number of the MHS, and MHSs is the theoretical number of MHS. The accuracy rate can be defined as OMHS/MHSs.

As obvious in Fig. 3, the run time of BPSOGA adopting the proposed fitness function (4) and the new MHSs criterion (17) is the shortest of all. However, its accuracy rate is 3/4, which is not as good as that of BHS-TREE's. For PSO-GA, the fitness function (3) and MHSs criterion (9) allow the particles to iterate toward the MHS, but it is easily affected by non-HS particles during iteration. Therefore, the first obtained MHSs (FOMHSs) contain non-HSs. As a result, it consumes a lot of time on MHSES to remove the non-HSs and the accuracy rate is the lowest one of all, 2/4. For BPSO and IDEA, the fitness function (1) and (2) and MHSs criterion (7) and (8), respectively, can only allow the particles iterate towards the HS rather than the MHS. Therefore, they consumes a lot of time to remove the superset of MHS of the obtained HSs (OHSs) by MHSESs. As a result, the total consumed run time is longer than that of BPSOGA, and accuracy rate is 3/4. For the BHS-TREE, although its tree search algorithm can obtains all the MHSs (i.e. the accuracy rate is 4/4), its run time is the longest of all the algorithms. In general, it can be seen that the overall performance of BPSOGA is the best of all the algorithms.

V. CASE STUDY AND DISCUSSION

In order to test the performance and sensitivity of the proposed algorithm using the new fitness function and MHS criterion, three cases have been conducted. Furthermore, the proposed algorithm is applied to the MBD to discuss its adaptability in fault diagnosis of distribution network [37].

The proposed fitness function and MHS criterion has been applied on the original algorithms GA, BPSO, IDEA and IGA to become improved algorithms IGA, IBPSO, IIDEA, and IIGA, respectively.

The parameters of GA algorithm are set as follows: Maximum iterations G_{max}, population size N_{pop}, population dimension D, mutation rate $P_m = 0.2$, crossover rate $P_c = 0.9$.

The parameters of BPSO algorithm are set as follows: G_{max}, N_{pop}, D, inertia weight $\omega = 0.9$, social learning factor $c_1 = 2.5$, self-learning factors $c_2 = 1.5$.

The parameters of IDEA algorithm are set as follows: G_{max}, N_{pop}, D, $P_m = 0.5$, $P_c = 0.6$, and termination constraint condition $T = 20$ [38].

The parameters of IGA algorithm are set as follows: G_{max}, N_{pop}, D, $P_m = 0.1$, $P_c = 0.8$, inversion rate $P_{in} = 0.1$, and immune probability $P_{im} = 0.2$.

The four algorithms are compared under the same values of G_{max} and N_{pop}. The values of G_{max} and N_{pop} depend on the number of elements in MCS cluster, i.e. the population

dimension D, which ranges from 20 to 500. For example, when n = 2 and m = 1, MCS cluster is {[1], [2]}, apparently D = 2. Therefore, values of G_{max} and N_{pop} in four algorithms are set to $G_{max} = N_{pop} = 20$.

The MHS computation outcomes of the improved algorithms have been compared with the original algorithms in terms of run time and accuracy rate. According to the characteristic that there is no intersection between the MCSs in the distribution network, the tested MCS cluster is set as $[1, 2, \ldots, m]$, $[m+1, m+2, \ldots, 2m], \ldots, [nm+1, nm+2, \ldots, (n+1)m]$. Where n is the number of MCSs in the MCS cluster and m is the number of elements in a single MCS. The three cases are:

Case A: using $n = 2, m = 1, 2, \ldots, 5$. The run time and accuracy rate of the algorithms are shown in Fig. 4, Fig. 5, respectively.

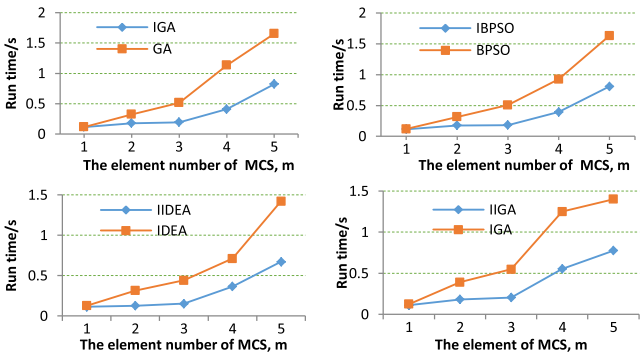


FIGURE 4. The run time of the algorithms when n = 2.

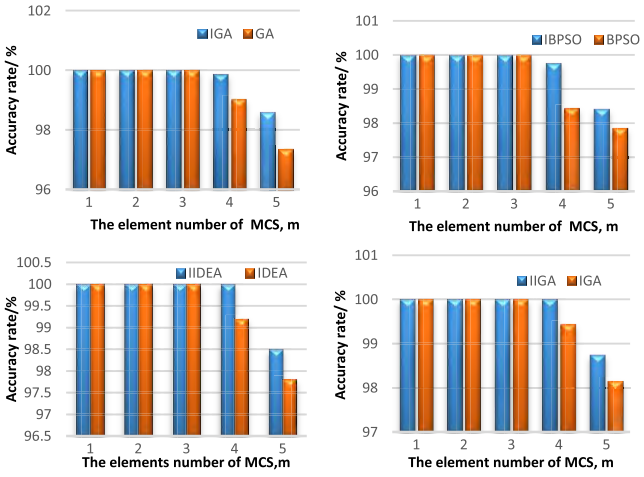


FIGURE 5. The accuracy rate of the algorithms when n = 2.

Case B: using $n = 3, m = 1, 2, \ldots, 5$. The run time and accuracy rate of the algorithms are shown in Fig. 6, Fig. 7, respectively.

Case C: using $n = 4, m = 1, 2, \ldots, 5$. The run time and accuracy rate of the algorithms are shown in Fig. 8, Fig. 9, respectively.

As shown in Fig. 4, Fig. 6 and Fig. 8, the run time of the improved algorithms is less than that of the original algorithms. In additional, the accuracy rate of the improved algorithms is higher than that of the original algorithms as shown in Fig. 5, Fig.7 and Fig. 9. Moreover, with the increase of the n and m, the advantages of improved algorithms become greater.

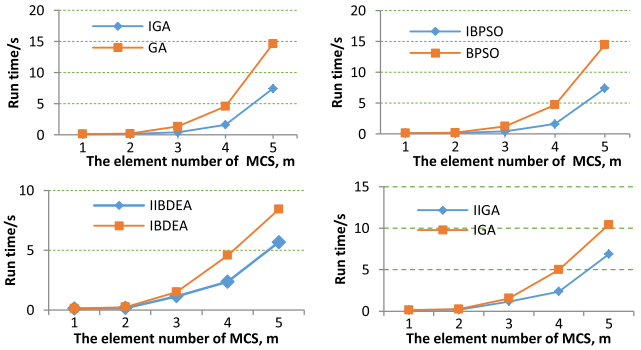


FIGURE 6. The run time of the algorithms when n = 3.

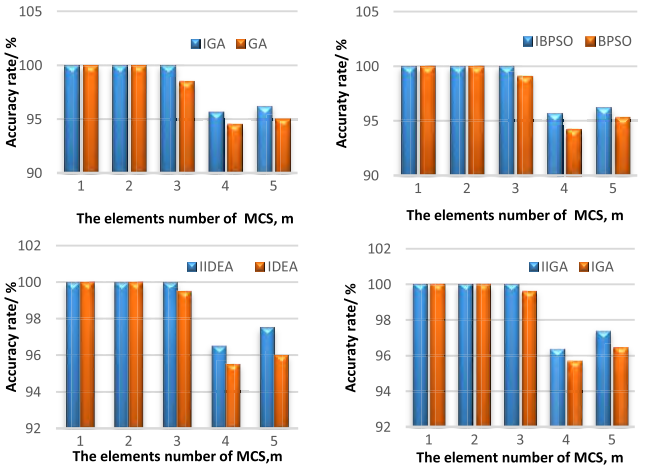


FIGURE 7. The accuracy rate of the algorithms when n = 3.

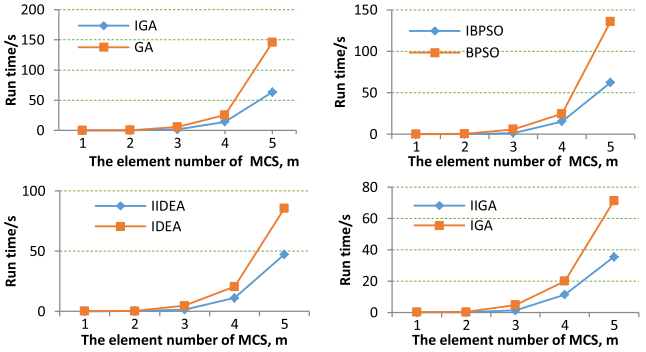


FIGURE 8. The run time of the algorithms when n = 4.

In order to test the adaptability of the proposed MHS algorithm in fault diagnosis of distribution network, a 14-nodes distribution network model is set up in PSCAD as shown in Fig. 10.

The model contains buses $B_1 - B_{14}$ and feeder sections $L_{01} - L_{1314}$, and their impedances are $Z_{01} - Z_{1314}$. Furthermore, there are 18 measuring points, and the current and

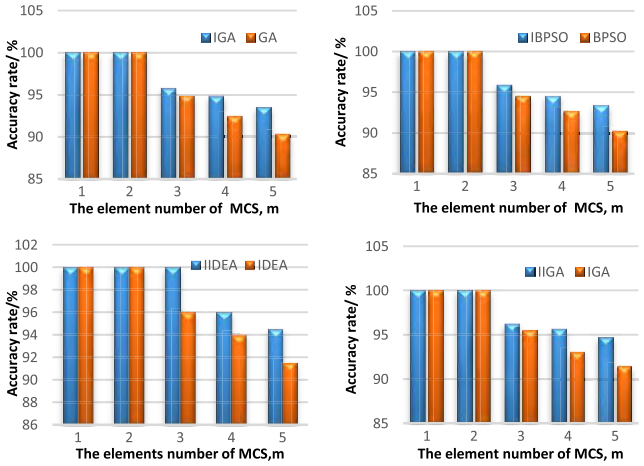


FIGURE 9. The accuracy rate of the algorithms when n = 4.

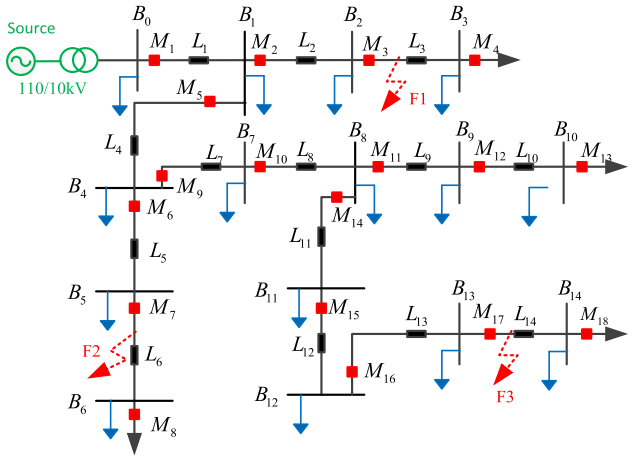


FIGURE 10. 14-nodes distribution network.

voltage information of each switch are collected through FTU. The measuring point 1 is the circuit breaker, the measuring points 8 and 18 are the contact switches, and the rest of the measuring points are the section switches.

By the steps of the MBD, the distribution network is firstly divided into 14 subsystems by 18 measuring points before the fault occurred. Then, the ARRs of each subsystem are obtained, and the MCSCs for each subsystem are searched. The results of ARRs and MCSCs are shown in Table 2.

In order to obtain MCS cluster, single one-phase, single two-phase and single three-phase short circuit faults are set up in L_3. Double one-phase, double two-phase and double three-phase are set up in L_3 and L_6 respectively. Triple one-phase, triple two-phase and triple three-phase are set up in L_3, L_6 an L_{14} respectively. The MCS cluster and corresponding binary form are shown in Table 3.

The MHSs are calculated under the above mentioned faults using BHS-TREE algorithm, BPSO algorithm and the proposed BPSOGA algorithm, respectively. The parameters of BPSOGA set as follows: G_{max}, N_{pop}, D, $P_m = 0.2$, $P_c = 0.7$, reinsertion factor $K = 0.9$. $\omega = 0.9$, $c_1 = 1.5$, and $c_2 = 1.5$. The value of G_{max}, N_{pop} are set as the same as the above algorithms. Then, the accuracy rate and run time of these algorithms are obtained, as shown in Fig. 11 and Fig. 12, respectively.

TABLE 2. ARRs and MCSCs of subsystems.

Subsystem	ARR	MCSC
…	…	…
S_3	U_{4A}-U_{3A}+$Z_{3A}I_{4A}$=0	[L_{3A},B_{3A}]
	U_{4B}-U_{3B}+$Z_{3B}I_{4B}$=0	[L_{3B},B_{3B}]
	U_{4C}-U_{3C}+$Z_{3C}I_{4C}$=0	[L_{3C},B_{3C}]
…	…	…
S_6	U_{8A}-U_{7A}+$Z_{6A}I_{8A}$=0	[L_{6A},B_{6A}]
	U_{8B}-U_{7B}+$Z_{6B}I_{8B}$=0	[L_{6B},B_{6B}]
	U_{8C}-U_{7C}+$Z_{6C}I_{8C}$=0	[L_{6C},B_{6C}]
…	…	…
S_{14}	U_{18A}-U_{17A}+$Z_{14A}I_{18A}$=0	[L_{14A},B_{14A}]
	U_{18B}-U_{17B}+$Z_{14B}I_{18B}$=0	[L_{14B},B_{14B}]
	U_{18C}-U_{17C}+$Z_{14C}I_{18C}$=0	[L_{14C},B_{14C}]

TABLE 3. MCS cluster and corresponding binary form.

Fault type	MCS cluster	Corresponding binary form
Single one-phase	[L_{3A},B_{3A}]	[11]
single two-phase	[L_{3A},B_{3A}] [L_{3B},B_{3B}]	[1100][0011]
single three-phase	[L_{3A},B_{3A}] [L_{3B},B_{3B}] [L_{3C},B_{3C}]	[110000][001100] [000011]
Double one-phase	[L_{3A},B_{3A}] [L_{6A},B_{6A}]	[1100][0011]
Double two-phase	[L_{3A},B_{3A}] [L_{3B},B_{3B}] [L_{6A},B_{6A}] [L_{6B},B_{6B}]	[11000000][00110000] [00001100][00000011]
Double three-phase	[L_{3A},B_{3A}] [L_{3B},B_{3B}] [L_{3C},B_{3C}] [L_{6A},B_{6A}] [L_{6B},B_{6B}] [L_{6C},B_{6C}]	[110000000000][001100000000] [000011000000][000000110000] [000000001100][00000000011]
Triple one-phase	[L_{3A},B_{3A}] [L_{6A},B_{6A}] [L_{14A},B_{14A}]	[110000][001100] [000011]
Triple two-phase	[L_{3A},B_{3A}] [L_{3B},B_{3B}] [L_{6A},B_{6A}] [L_{6B},B_{6B}] [L_{14A},B_{14A}] [L_{14B},B_{14B}]	[110000000000][001100000000] [000011000000][000000110000] [000000001100][00000000011]
Triple three-phase	[L_{3A},B_{3A}]	[1100000000000000]
	[L_{3B},B_{3B}]	[0011000000000000]
	[L_{3C},B_{3C}]	[0000110000000000]
	[L_{6A},B_{6A}]	[0000001100000000]
	[L_{6B},B_{6B}]	[0000000011000000]
	[L_{6C},B_{6C}]	[0000000000110000]
	[L_{14A},B_{14A}]	[0000000000001100]
	[L_{14B},B_{14B}]	[0000000000000011]
	[L_{14C},B_{14C}]	

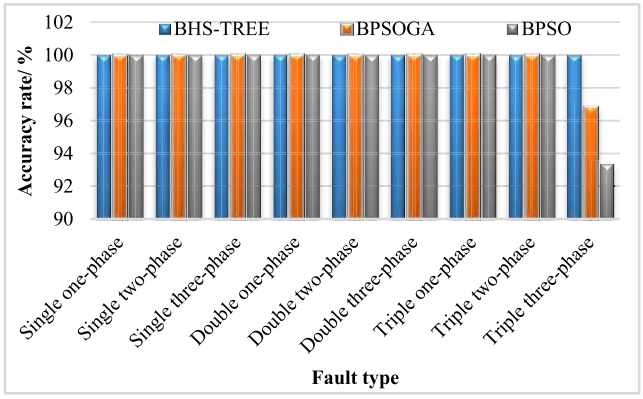


FIGURE 11. The accuracy rate of the algorithms under faults.

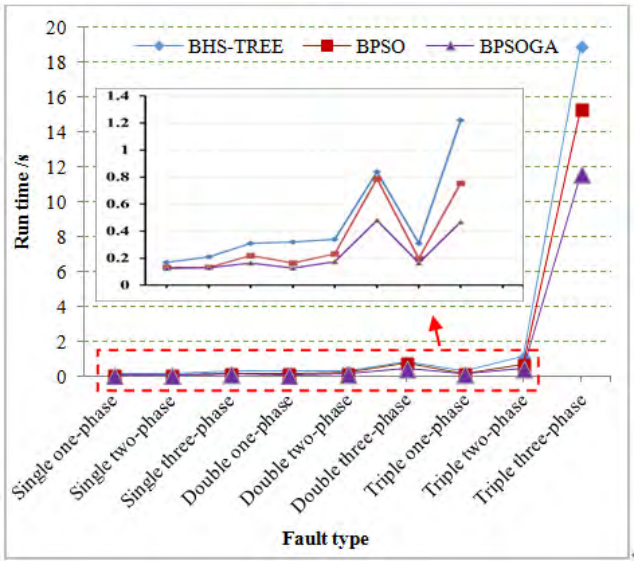


FIGURE 12. The run time of the algorithms under faults.

In Fig. 11, although the accuracy rate of BPSOGA can't keep 100% consistently like BHS-TREE, it is higher than BPSO. In Fig. 12, the run time of BPSOGA is shorter than that of BPSO and BHS-TREE under the same fault type. Therefore, the comprehensive performance of BPSOGA is the best of all.

VI. CONCLUSION

In this paper, an innovative MHS for MBD's algorithm is proposed for faults diagnosis in distribution network. This algorithm considers the nature and characteristic of the distribution network, which fundamentally simplifies the MHS. Three cases have been conducted in order to test the performance and sensitivity of the proposed algorithm. Furthermore, the proposed algorithm has been connected to the MBD to discuss its adaptability in fault diagnosis of distribution network. With analysis and comparison of this algorithm with other selected algorithm for validation, the overall performance in terms of run time and accuracy rate is the best. Moreover, with the proposed algorithm, MHS can be calculated more effectively and efficiently by utilizing the new fitness function and the MHS criterion.

REFERENCES

[1] F. Mahmood, L. Vanfretti, M. Pignati, H. Hooshyar, F. Sossan, and M. Paol, "Experimental validation of a steady state model synthesis method for a three-phase unbalanced active distribution network feeder," *IEEE Access*, vol. 6, pp. 4042–4053, 2018.

[2] S. Deng, C. Yuan, J. Yang, and A. Zhou, "Distributed mining for content filtering function based on simulated annealing and gene expression programming in active distribution network," *IEEE Access*, vol. 5, pp. 2319–2328, 2017.

[3] X. Zhang, J. Wei, S. Yue, and X. Zha, "An analytical method for mapping alarm information to model of power grid fault diagnosis," *IEEJ Trans. Elect. Electron. Eng.*, vol. 13, no. 6, pp. 823–830, 2018.

[4] J. Zhao *et al.*, "Power Grid Fault Diagnosis Based on Fault Information Coding and Fusion Method," in *Proc. 2nd IEEE Conf. Energy Internet Energy Syst. Integr.*, Oct. 2018, pp. 1–6.

[5] K. Chen, J. Hu, and J. He, "A framework for automatically extracting overvoltage features based on sparse autoencoder," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 594–604, Mar. 2018.

[6] Z. Wang, Z. Huang, C. Song, and H. Zhang, "Multiscale adaptive fault diagnosis based on signal symmetry reconstitution preprocessing for microgrid inverter under changing load condition," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 797–806, Mar. 2018.

[7] M. Kezunovic, T. Djokic, P. C. Chen, and V. Malbasa, "Improved transmission line fault location using automated correlation of big data from lightning strikes and fault-induced traveling waves," in *Proc. 48th Hawaii Int. Conf. Syst. Sci.*, Jan. 2015, pp. 2719–2728.

[8] A. Ghaderi, H. A. Mohammadpour, H. L. Ginn, and Y. J. Shin "High-impedance fault detection in the distribution network using the time-frequency-based algorithm," *IEEE Trans. Power Delivery*, vol. 30, no. 3, pp. 1260–1268, Jun. 2015.

[9] T. Lv and Q. Ai, "Interactive energy management of networked microgrids-based active distribution system considering large-scale integration of renewable energy resources," *Appl. Energy*, vol. 163, pp. 408–422, Feb. 2016.

[10] G. Liang, P. Liyuan, L. Ruihuan, Z. Fen, and W. Xin, "Fault location in distribution network with distributed generation based on neural network," in *Proc. China Int. Conf. Electr. Distrib.*, Sep. 2014, pp. 209–212.

[11] F. Ren, M. Zhang, and D. Sutanto, "A multi-agent solution to distribution system management by considering distributed generators," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1442–1451, May 2013.

[12] C. Gu, W. Yang, Y. Song, and F. Li, "Distribution network pricing for uncertain load growth using fuzzy set theory," *IEEE Trans. Smart Grid*, vol. 7, no. 4, pp. 1932–1940, Jul. 2016.

[13] R. H. Salim, R. A. Ramos, and N. G. Bretas, "Analysis of the small signal dynamic performance of synchronous generators under unbalanced operating conditions," in *Proc. IEEE PES General Meeting*, Jul. 2010, pp. 1–6.

[14] H. Song *et al.*, "Stochastic programming-based fault diagnosis in power systems under imperfect and incomplete information," *Energies*, vol. 11, no. 10, p. 2565, 2018.

[15] H. Su and F. A. Zhao "A novel substation fault diagnosis approach based on RS ANN ES," in *Proc. Int. Conf. Commun. Circuits Syst.*, vol. 3, Jun. 2006, pp. 2124–2127.

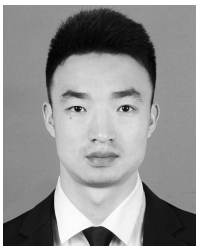
[16] N. K. Chanda and Y. Fu "ANN-based fault classification and location in MVDC shipboard power systems," in *Proc. North Amer. Power Symp.*, Aug. 2011, pp. 1–7.

[17] W. Guo, F. Wen, G. Ledwich, Z. Liao, X. He, and J. Liang, "An analytic model for fault diagnosis in power systems considering malfunctions of protective relays and circuit breakers," *IEEE Trans. Power Delivery*, vol. 25, no. 3, pp. 1393–1401, Jul. 2010.

[18] A. Sidhu, A. Izadian, and S. Anwer, "Adaptive nonlinear model-based fault diagnosis of Li-ion batteries," *IEEE Trans. Ind. Electron.*, vol. 62, no. 2, pp. 1002–1011, Feb. 2015.

[19] J. Schumann, K. Y. Rozier, T. Reinbacher, O. J. Mengshoel, T. Mbaya, and C. Ippolito, "Towards real-time, on-board, hardware-supported sensor and software health management for unmanned aerial systems," in *Proc. Annu. Conf. Prognostics Health Manage. Soc. (PHM)*, New Orleans, LA, USA, Oct. 2013.

[20] A. Scacchioli, G. Rizzoni, M. A. Salman, W. Li, S. Onori, and X. Zhang, "Model-based diagnosis of an automotive electric power generation and storage system," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 44, no. 1, pp. 72–85, Jan. 2014.
[21] Z. Gao, C. Cecati, and S. X. Ding, "A survey of fault diagnosis and fault-tolerant techniques-Part I: Fault diagnosis with model-based and signal-based approaches," *IEEE Trans. Ind. Electron.*, vol. 62, no. 6, pp. 3757–3767, Jun. 2015.
[22] L. Guan, Z. G. Liu, J. F. Xu, and Y. Wang, "Key issues with model-based diagnosis in distribution network," *Power Syst. Protection Control*, vol. 40, no. 20, pp. 145–150, 2012.
[23] Z. Liu and K. Hu, "A model-based diagnosis system for a traction power supply system," *IEEE Trans. Ind. Inform.*, vol. 13, no. 6, pp. 2834–2843, Dec. 2017.
[24] H. J. Eysenck and S. Rachman, *The Causes and Cures of Neurosis (Psychology Revivals): An introduction to modern behaviour therapy based on learning theory and the principles of conditioning*. London, U.K.: Routledge, 2013.
[25] R. Abreu and A. J. van Gemund, "A low-cost approximate minimal hitting set algorithm and its application to model-based diagnosis," in *Proc. 8th Symp. Abstraction, Reformulation, Approximation*, Oct. 2009, pp. 456–486.
[26] A. Feldman, G. M. Provan, and A. J. van Gemund, "Computing Minimal Diagnoses by Greedy Stochastic Search," in *Proc. AAAI* , vol. 8, Jul. 2008, pp. 911–918.
[27] X. Zhao and D. Ouyang, "Deriving all minimal hitting sets based on join relation," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 45, no. 7, pp. 1063–1076, Aug. 2015.
[28] A. Nyberg, "A generalized minimal hitting-set algorithm to handle diagnosis with behavioral modes," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 41, no. 1, pp. 137–148, Jan. 2011.
[29] N. Cardoso and A. Rui, "A kernel density estimate-based approach to component goodness modeling," in *Proc. 27th AAAI Conf. Artif. Intell.*, Jun. 2013, pp. 152–158.
[30] L. Guan, Z. Liu, S. He, and H. Yang, "Application of BPSO algorithm in model-based fault diagnosis of distribution network," *Electric Power Automat. Equip.*, vol. 33, no. 9, pp. 89–93, 2013.
[31] K. Hu, Z. Liu, K. Huang, C. Dai, and S. Gao, "Improved differential evolution algorithm of model-based diagnosis in traction substation fault diagnosis of high-speed railway," *IET Elect. Syst. Transp.*, vol. 6, no. 3, pp. 163–169, Sep. 2016.
[32] G. Zhou, W. Feng, B. Jiang, and C. Li, "Computing minimal hitting set based on immune genetic algorithm," *Int. J. Modelling, Identificat. Control*, vol. 21, no. 1, pp. 93–100, 2014.
[33] S. Gao, C. Dai, Z. Liu, and X. Geng, "Application of BPSO with GA in model-based fault diagnosis of traction substation," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2014, pp. 2063–2069.
[34] T. Jin and H. Li, "Fault location method for distribution lines with distributed generators based on a novel hybrid BPSOGA," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 10, pp. 2454–2463, Jul. 2016.
[35] J. Liu, D. T. Ouyang, and Y. Y. Wang, "Computing minimal hitting sets with particle swarm optimization combined characteristics learning," *Acta Electr. Sinica*, vol. 43, no. 5, pp. 841–845, 2015.
[36] M. A. Mohamed, A. M. Eltamaly, and A. I. Alolah, "Swarm intelligence-based optimization of grid-dependent hybrid renewable energy systems," *Renew. Sustain. Energy Rev.*, vol. 77, pp. 515–524, Aug. 2017.
[37] K. Hou *et al.*, "Research on practical power system stability analysis algorithm based on modified SVM," *Protection Control Modern Power Syst.*, vol. 3, no. 1, p. 11, Dec. 2018.
[38] J. K. Pattanaik, M. Basu, and D. P. Dash, "Opposition-based differential evolution for hydrothermal power system," *Protection Control Modern Power Syst.*, vol. 2, no. 1, p. 2, Dec. 2017.



QIUJIE WANG was born in Hubei, China, in 1988. He received the B.S. degree from the Wuhan University of Technology, in 2010, and the M.S. degree from China Three Gorges University, in 2016. He is currently pursuing the Ph.D. degree with the School of Electrical Engineering and Automation, Fuzhou University. His current research interests include the fault diagnosis of distribution networks and artificial intelligence in power systems.

TAO JIN (M'08) was born in Hubei, China, in 1976. He received the B.S. and M.S. degrees from Yanshan University, in 1998 and 2001, respectively, and the Ph.D. degree in electrical engineering from Shanghai Jiao Tong University, in 2005.

From 2005 to 2007, he held a postdoctoral position at Shanghai Jiaotong University. During this time, he was the in charge of a research group in the biggest dry-type transformer company in Asia, Sunten Electrical Co., Ltd., to develop a new transformer technology with distribution grid. From 2008 to 2009, he was a Research Scientist with Virginia Tech, Blacksburg, USA, where he was involved in the design and test of PMU technology and GPS/the internet-based power system frequency monitoring networks, and developing new solutions for power system dynamics wide monitoring and control. In 2010, he joined Imperial College London, U.K., as the European Union Marie Curie Research Fellow, where he was focused on electrical technologies related to smart grid. He is currently a Professor with the College of Electrical Engineering and Automation, Fuzhou University, China. He has published over 110 papers. He is a member of the IEEE Power and Energy Society and the IEEE Industrial Electronics Society. He is a special Committee Member of the Chinese Society of Electrical Engineering and the China Electrotechnical Society. He currently serves as an Associate Editor of the *China Measurement and Testing Technology*.

MOHAMED A. MOHAMED was born in Minia, Egypt, in 1985. He received the B.Sc. and M.Sc. degrees from Minia University, Minia, in 2006 and 2010, respectively, and the Ph.D. degree from King Saud University, Riyadh, Saudi Arabia, in 2016. He joined the College of Electrical Engineering and Automation, Fuzhou University, China, as a Postdoctoral Research Fellow, in 2018. He has been a Faculty Member of the Department of Electrical Engineering, College of Engineering, Minya University, since 2008. He has supervised a number of M.Sc. and Ph.D. thesis, and he was involved in a number of technical projects. He has published more papers and books@perio. His current research interests include the areas of renewable energy, energy management, power electronics, power quality, and optimization and smart grids. He joined the editorial board of some scientific journals and joined the steering committees of many international conferences.