2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)
July 6-11, 2014, Beijing, China

A Novel Algorithm to Solve the Minimal Hitting Sets in MBD

Jianfang Xu, the School of Electrical Engineering
Southwest Jiaotong University
Chengdu Sichuan 610031 ,China
e-mail: xujianfang324018@163.com

Zhigang Liu , the School of Electrical Engineering
Southwest Jiaotong University
Chengdu Sichuan 610031 ,China
e-mail: liuzg@swjtu.cn

Chenxi Dai , the School of Electrical Engineering
Southwest Jiaotong University
Chengdu Sichuan 610031 ,China
e-mail: 307298634@qq.com

***Abstract*—Although the advantages of model-based fault diagnosis are becoming more and more obvious in system fault diagnosis, the computation efficiency of minimal hitting sets is the main bottleneck. To improve the calculating speed of existed algorithms and solve the problem of incomplete minimal hitting sets, this paper proposes an improved differential evolution algorithm. Introduction of binary to differential evolution algorithm. The new method takes advantage of the rapid differential evolution algorithm, and adds the minimal assurance strategy to hitting sets in the evolutionary process. Accordingly, the computational efficiency and accuracy of minimal hitting sets are guaranteed. The algorithm is used in the classical digital circuits and traction substation fault diagnosis, and the simulation results show that the speed and accuracy of calculation minimal hitting set is improved. And the efficiency of model-based fault diagnosis is enhanced. In addition, the algorithm can be widely applied to all kinds of system or element fault diagnosis.**

***Keywords*—minimal hitting sets, model-based fault diagnosis, binary differential evolution algorithm**

I. INTRODUCTION

Model-based diagnosis technology, as an intelligent reasoning technology [1], was proposed to overcome the shortcomings of traditional methods which are greatly dependant on professional knowledge, are difficult to system migration and maintenance, cause combinatorial explosion and contradiction of rules when system is huge and lack diagnosis ability to small probability events, etc. Calculating the minimal conflict set and the minimal hitting set is the key step of model based diagnosis [2]. Since the hitting set is the core of system fault diagnosis, its calculation efficiency and accuracy directly affect the efficiency of system fault diagnosis. Many methods have been employed to calculate the minimal hitting set so far, such as HS-tree [3,4], BHS-tree[5], Boolean algebra[6], logic arrays[7], genetic algorithm[8,9], etc. These methods have the following shortcomings: easily losing correct solution because of pruning[3], low calculation efficiency[3,4,5] and low accuracy of random search algorithm[8,9].

Now, it is proved that the differential evolution (DE) algorithm[10] performs better than other evolution algorithms and has advantages of low difficulty, low complexity, high robustness, fast convergence and global search capability[11]. This paper improves the differential evolution algorithm and utilizes it to calculate the minimal hitting set. So far, DE algorithms have achieved great progress. Besides the standard differential evolution algorithm[10], many improved algorithms such as the adaptive operator DE[12], DE combining with simplex eugenic strategy[13], DE combining particle filter[14], are utilized in various fields[15]. Considering that binary numbers 0, 1 in the binary DE algorithm can respectively represent normal and fault status of the system, the improved binary DE algorithm is utilized to calculate the minimal hitting set in this paper. Liu[16] and Hu[17] used model based diagnosis method in diagnosis of traction power transformer substations and power distribution network lines with higher efficiency than other methods based on expert systems and solved problems of system migration and maintenance. However, the efficiency of the minimal hitting set calculation did not satisfy the actual demand well.

Considering the above facts, an improved differential evolution algorithm by introduced binary is proposed to calculate the minimal hitting sets. This algorithm utilizes the correspondence relation between binary numbers and the system components status and adds the minimal hitting sets assurance strategy in the process of evolution. The proposed method achieves higher efficiency of the minimal hitting sets calculation and guarantees that the calculation results are the minimal hitting sets. Then this algorithm is employed in diagnosis of traction power transformer substations to verify the efficiency and accuracy of the method.

II. FUNDAMENTALS OF MODEL BASED DIAGNOSIS

The model based diagnosis is realized by two steps: (1) establishing the system model using first-order logic statement based on the system components and their connection relation, (2) deducing the normal expected behavior of the system through logic reasoning. If there are differences between actual and expected behaviors, it means the system has faults[1,2,3]. This method plays an important role in promoting the research on artificial intelligence. Some related concepts and theorems about model-based diagnosis[3] are introduced as follows:

Definition 1: defining a system as a triple (SD, COMPS, OBS).

Where SD is the system description, which is represented by a first-order predicate formula; COMPS is the system component and a finite constant set; OBS is an observation set and a finite set represented by the first-order predicate formula. The unitary predicate AB indicates the abnormal status. When $c \in COMPS$ is abnormal, AB(c) is true, contrarily, AB(c) is false.

Definition 2: the component set $\{c_1, c_2, ..., c_n\} \subseteq COMPS$.

when $SD \cup OBS \cup \{\sim AB(c_1), \sim AB(c_2), ..., \sim AB(c_n)\}$ is inconsistent, this set is a conflict set (CS). If any subset of a conflict set is not a conflict set, the conflict set is called the minimal hitting set (MCS).

Definition 3: for a set cluster, when $H \subseteq \cup_{S \in C} S$ and $H \cap S \neq \varnothing$ ($S \in C$), the set H is called a hitting set of C. Similarly, if any subset of a hitting sat is not a hitting set (HS), the hitting set is a minimal hitting set (MHS).

Theorem 1: $D(\Delta, COMPS\text{-}\Delta)$ is a consistent diagnosis of the system only if Δ is the minimal hitting set of the minimal conflict set of the system.

III. BINARY DIFFERENTIAL EVOLUTION ALGORITHM FOR CALCULATING MINIMAL HITTING SET

The computation efficiency of minimal hitting sets is the main bottleneck in model-based fault diagnosis. DE algorithm has strong global search capability and higher efficiency, and binary differential evolution algorithm is introduced in this paper to calculate the minimum hitting set of minimal conflict sets. Firstly, the minimal conflict sets are mapped to D-dimensional binary sets represented by 0 and 1, where D represents the size of the problem, 0 indicates that the system components corresponding to that dimension is normal, and 1 indicates that the corresponding component is failure. The initial population is a NP × D dimension binary matrix represented by 0 and 1 randomly, where NP represents the population size. The fitness function constructed in this paper is adopted to judge whether the population individuals are hitting sets. Then, saving hitting sets, updating population states, obtaining all hitting sets, and finding out the minimal hitting sets according to the introduced assurance strategy of minimum hitting set until the iteration termination condition is satisfied.

A. Algorithm description

1) Population initialization

In order to create the initial point of optimized search, the population must be initialized. Normally, the initial population is generated by given boundary constraints randomly. Generally, a NP × D order 0-1 random matrix is generated as the initial population U in the binary differential evolution algorithm.

2) Mutation operation

The DE/rand/1 method is selected to mutation in the mutation operation:

$$V_i^{g+1} = U_{i1}^g + F(U_{i2}^g + U_{i3}^g) \quad (1)$$

where i、i1、i2、i3 are not equal to each other, V_i^{g+1} is mutation vector.

Three stochastic individuals *Ui1 (g)*, *Ui2 (g)*, *Ui3 (g)* in the g-th generation are randomly selected to be mutated, and Hamming distance is used to represent the difference vector between *Ui2(g)* and *Ui3(g)* [16].

Letting the Hamming distance between *Ui2(g)* and *Ui3(g)* be d, and F is the scaling factor, the distance after scaling is $d' = F \times d$. After mutation, the Hamming distance between the resulting mutation vector and *Ui1(g)* is m, i.e. *Ui1(g)* is the number of the needed mutation position, and the calculation formula is as follows:

$$m = \begin{cases} (\text{int})d' + 1 & \text{if} \quad rand < d' - (\text{int})d' \\ (\text{int})d' & otherwise \end{cases} \quad (2)$$

Where, *rand* is the stochastic number with Uniform Distribution in [0,1].

3) Crossover Operation

Crossover operation is adopted to improve the diversity of the population. The crossover is to cross the mutation vector *Vi(g+1)* and the target vector *Ui(g)*, generating the test vectors.

$$W_{ij}(g+1) = \begin{cases} V_{ij}(g+1) & rand_{ij} < CR \text{ or } j = l_i \\ U_{ij}(g) & else \end{cases} \quad (3)$$

Where, CR is the cross-factor, $rand_{ij}$ is the stochastic number with Uniform distribution in [0,1], l_i is an integer selected from the sequence [1,2, ..., D] randomly, which is used to ensure that the test individual at least have one component provided by *Vi(g+1)*.

4) Select operation

The one to one greedy selection method is adopted in the binary differential evolution algorithm. After the competition of the test individuals *Wi(g+1)* after mutating and crossing and initial population individual *Ui(g)*, the one which has optimum fitness is selected as the subsidiary. Expressed by the formula:

$$U_i(g+1) = \begin{cases} W_i(g+1) & \text{if } f(W_i(g+1)) \leq f(U_i(g)) \\ U_i(g) & else \end{cases} \quad (4)$$

B. Construct the fitness function

The calculation process of the minimal hitting sets can be divided into two steps: getting hitting sets and extracting the hitting sets of which any of subset is not a hitting set. If a fitness value is used to simultaneously measure both of the two sets, the resulting set may not be a hitting set. Hence we need to consider the two sets separately. The number of sets which have intersection of the conflict set cluster and population individual is taken as fitness value. And the assurance strategy is used to ensure the hitting sets are minimal hitting sets. In order to limit the fitness value in [0,1], the fitness function can be expressed as:

$$f(U_i) = \frac{h}{p} \tag{5}$$

Where h represents the number of set which has intersection between the conflict set cluster and the population individual, p is the number of conflict sets contained in conflict set cluster.

C. Minimal hitting sets assurance strategy

To further improve the efficiency of the binary search algorithm differential evolution and ensure that the final result is minimal hitting sets, the assurance strategy of minimal hitting sets is given. The basic method can be described as: extracting single hitting sets one by one from hitting set cluster H obtained by searching, and scanning the binary set, when the corresponding element is 1, set it to 0. Then evaluate their fitness by the fitness function structured above, if its fitness value still be 1, then it's hitting set, keeping the change. Otherwise reverting to the original value. The operation ends until all of the binary sets are scanned, and all of the minimal hitting sets are obtained. The basic process of the method is shown in Fig.1.

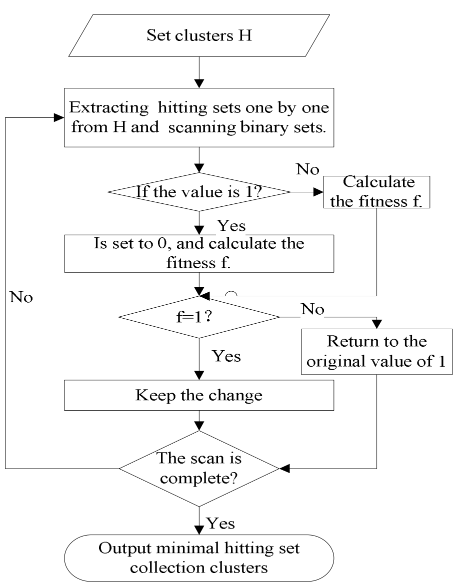

Fig.1. The basic process of minimum guaranteed policy

D. The process to calculate the minimum set using Binary differential evolution algorithm

Based on the above description, the basic steps of the algorithm to calculate the minimum hitting sets are as follows:

1) Determining the control parameters of differential evolution algorithm and mutation strategy, generating initial population randomly, evolution algebra g=1. DE control parameters include population size NP, mutation factor F, crossover factor CR, the problem dimension D, the maximum evolution algebra G, and termination conditions.

2) Evaluating each individual of the population. That is calculating the fitness value, extracting hitting sets and taking it into set H.

3) Judging whether the termination condition is satisfied. If so, turn to step *6)*, otherwise, continue *4)*.

4) Conducting mutation and crossover operation, obtaining temporary population and evaluating it, calculating the fitness value.

5) Conducting selection operation, getting a new population, evolution algebra g=g+1, and then return to step *2)*.

6) Using assurance strategy of minimal hitting sets to traverse hitting set cluster H, ensure that all sets in H are minimal hitting sets, output H finally.

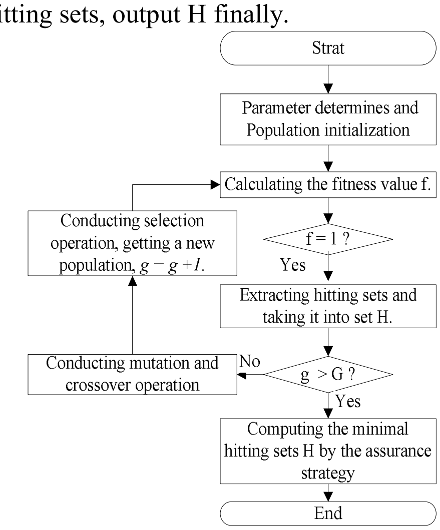

Fig.2. The process to calculate the minimum set using Binary differential evolution algorithm

IV. Comparative Analysis of Algorithms

Because it is the most classical method that the calculation method to the minimal hitting set based on the HS-tree algorithm. Therefore, in order to verify the efficiency of the proposed method, *HS*-tree method proposed in literature[3] was adopted for comparison.. Let the number of the minimum conflict sets be $n = 9$, the value is $<1,2, ..., m>$, $<2,3, ..., m+1>$, ..., $<n, n+1, ..., n+m-1>$. Test machine configuration: CPU T5850 2.16GHz, memory 2.00G, Windows XP, MATLAB R2009a. Run time of the two methods is compared in Table 1.

TABLE 1. THE COMPARISON OF RUN TIME BETWEEN IMPROVED BINARY DE ALGORITHM AND HS-TREE ALGORITHM

The number of minimal conflict set	The number of elements of conflict set	Run time of Improved Binary DE /s	Run time *HS*-tree /s
9	3	0.33	0.61
9	4	0.47	0.87
9	5	0.62	1.25
9	6	0.74	1.94
9	7	0.77	2.67
9	8	0.82	3.51
9	9	0.86	4.43

Minimal hitting sets in binary form	Minimal hitting sets restored
0 0 0 1 0	{A1}
0 1 0 0 1	{M2, A2}
0 1 1 0 0	{M2, M3}
1 0 0 0 0	{M1}

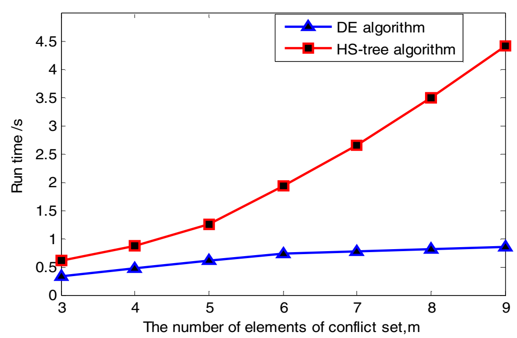

Fig.3. The comparison of run time between Improved Binary DE algorithm and HS-tree algorithm

According to Tab.1 and Fig.3, it is obvious that the improved binary differential evolution algorithm proposed in this paper has significant advantage in computational efficiency.

V. Examples of Applications

A. Application In model-based fault diagnosis of digital circuit

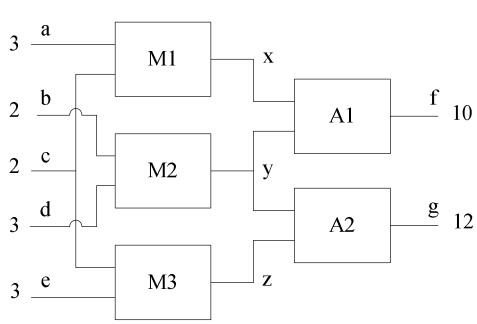

Fig.4. Classical digital circuit

The classical digital circuit is shown in Fig.4. The circuit comprises three multipliers M1, M2, M3, and two adders A1, A2. These components are all with two inputs and one output, where a, b, c, d, e, f, g are the observing variable, and x, y, z are unobservable variables. System COMPS={M1, M2, M3, A1, A2}, the inputs of the system are shown in Fig.3. If the components of the system are all normal, then the output of the digital circuit f=12, g=12. Now the actual output of the circuit is f=10, g=12, illustrating that several components of the circuit system are failure. According to model-based fault diagnosis we can infer that the candidate minimum conflict sets are {M1, M3, A1, A2} and {M1, M2, A1}.

Now the introduced binary differential evolution algorithm was used to calculate minimal hitting sets of minimal conflict sets. The configuration of test machine is the same as that of chapter IV. The control parameters used in the algorithm are as follows: D =5, NP=30, F=0.6, CR=0.5, G=100. The minimal conflict sets are transmitted into binary form: {10111} and {11010}, and the minimal hitting sets obtained by the improved binary DE algorithm are shown in Table 2.

TABLE 2: MINIMAL HITTING SET TABLE

The run time of searching process is 0.165630 s. Finally, according to the model-based fault diagnosis knowledge, the fault components and fault type are inferred.

B. Applications in Traction Substation Fault Diagnosis

The main electrical wiring diagram of an AT traction substation is shown in Fig.5. The constituent elements of the system can be expressed as COMPS={L1_A, L1_B, L1_C, T1_T1F1, T1_T2F2, L3_T1, L3_T2, L3_F1, L3_F2, L3_N, B_T1, B_T2, B_F1 , B_F2, B_N, GND}. The meanings of the symbols are shown in Table 3.

TABLE 3: THE MEANINGS OF THE ELEMENT COMPOSITION REPRESENTED BY THE SYMBOL

Component symbols	The meaning of symbols
T1{ T1_T1F1, T1_T2F2}	1T traction transformer and its two independent single-phase transformer.T1 represents the independent component object of the second-layer structure an abstract model;T1_T1F1 and T1_T2F2 represent the first layer structure with an abstract model of independent component object.
L1{L1_A,L1_B, L1_C }	The three-phase cable of the power grid to the power supply side of the transformer, where the breakers and disconnectors as part of a three-phase cable.
B{ B_T1,B_T2,B_F1,B_F2,B_N }	Traction bus and its various single-phase busbar and neutral bus between the in-phase busbar disconnectors, as part of the phase busbars.
L3{L3_F1, L3_F2, L3_N, L3_T1, L3_T2}	Part of the 1T traction transformer substation is connected to the bus line side, where the isolation switch as cable
GND	Track and ground, as the reference node of the entire main wiring system of traction substation.

It is assumed that in the traction substation bus bar B_F1 and connecting line L1_C is shorted to ground, which are shown in Fig.5. Then according to the theory of model-based fault diagnosis the minimal conflict sets clusters can be obtained: MIC = {{L1_A, L1_B, L1_C, T1_T1F1, T1_T2F2}, {B_F1}, {L1_A, L1_C, L3_T1, L3_F1, T1_T1F1}, {B_N, B_T1, L1_A, L3_N, L3_T1, T1_T1F1}}. The binary form of the minimal conflict sets cluster is shown in Table 4.

TABLE 4: MINIMAL CONFLICT SETS AND THE CORRESPONDING BINARY FORM

Minimal conflict sets	The corresponding binary form
{L1_A,L1_B,L1_C,T1_T1F1,T1_T2F2}	1111100000000000
{B_F1}	0000000000001000
{L1_A,L1_C,L3_T1,L3_F1,T1_T1F1}	1011010001100010
{B_N,B_T1,L1_A,L3_N,L3_T1,T1_T1F1}	1011010100000000

Using the introduced binary differential evolution algorithm is used to search and get the minimum hitting set of the minimal conflict set. In this case the control parameters are: D=16, NP =80, F=0.6, CR=0.5, G=100. The run time of the search process is 0.710775 s, and 11 minimum hitting sets are obtained, as shown in Table 5.

Converting the minimal hitting sets in the form of binary in Table 5 into the form of element symbols according to the conversion strategy. The minimal hitting sets cluster is HS={{B_F1, L1_A}, {B_F1, L1_C}, {B_F1, T1_T1F1}, {B_F1 , L1_B, L3_T1}, {B_F1, L3_T1, T1_T2F2}, {B_N, B_F1, L1_B, L3_F1}, {B_N, B_F1, L3_F1, T1_T2F2}, {B_F1, B_T1, L1_B, L3_F1}}. Finally, according to the minimal hitting sets clusters, using the theory and technology of models-based fault diagnosis the fault elements and type can be inferred efficiently and accurately.

TABLE 5: MINIMAL HITTING SETS AND THE CORRESPONDING SET OF COMPONENTS IN BINARY FORM

Minimal hitting sets of binary form	Corresponding minimum hitting sets
0000100100001010	{T1_T2F2,L3_F1,B_F1,B_N}
0000100100101000	{T1_T2F2,L3_F1,B_T1,B_F}
0000100101001000	{T1_T2F2,L3_F1,L3_N,B_F}
0000110000001000	{ T1_T2F2,L3_T1,B_F1}
0001000000001000	{T1_T1F1,B_F1}
0010000000001000	{L1_C, B_F1}
0100000100001010	{L1_B, L3_F1, B_F1,B_N }
0100000100101000	{ L1_B, L3_F1, B_T1,B_F1}
0100000101001000	{ L1_B, L3_F1, L3_N, B_F1}
0100010000001000	{ L1_B,L3_T1,B_F1}
1000000000001000	{L1_A, B_F1}

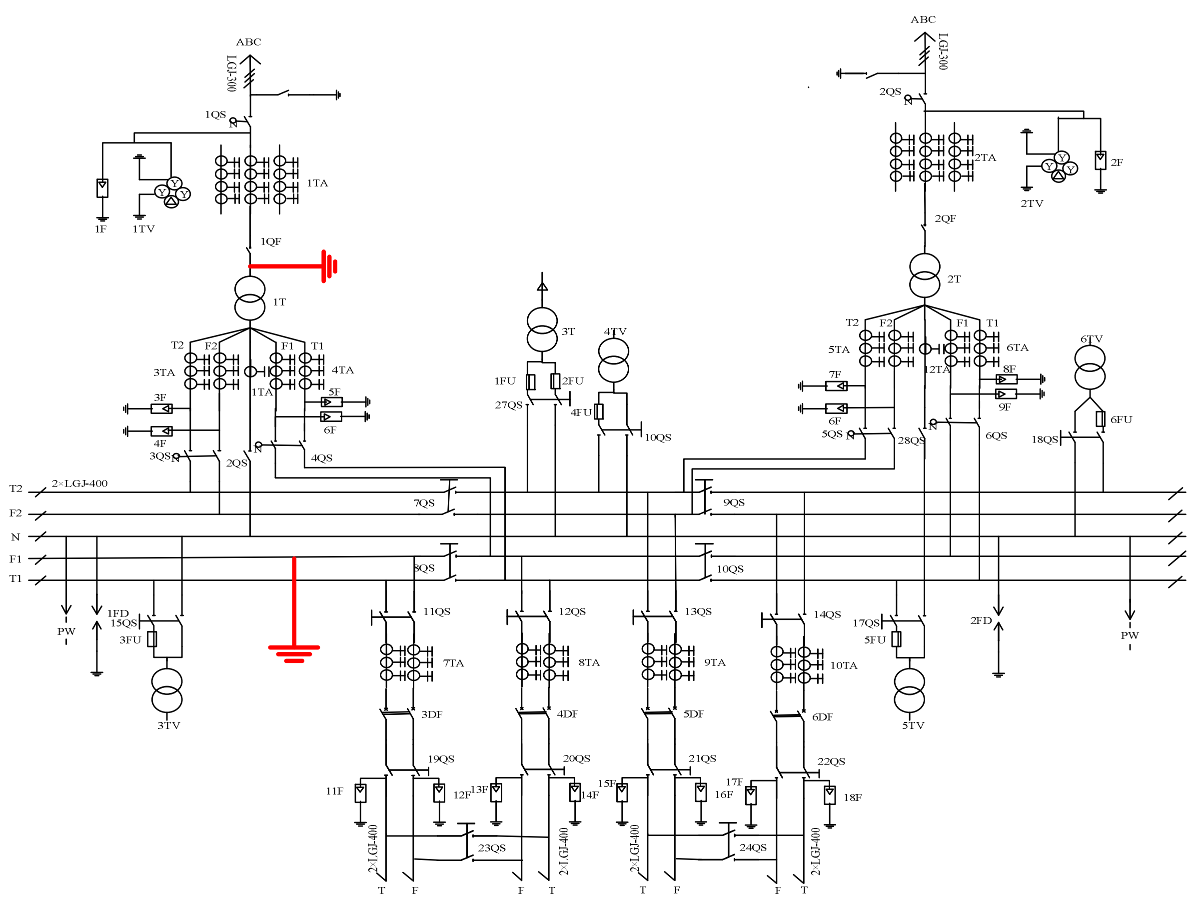

Fig.5. The main electrical wiring diagram of an AT traction substation

VI. CONCLUSION

The improved binary differential evolution algorithm is proposed in the paper. Simulation results showing that the minimal hitting sets of the minimal conflict sets could be calculated accurately and efficiently by the algorithm. In addition, the assurance strategy of minimal hitting sets proposed in this paper ensures the final output results are all minimal hitting sets. The efficiency and accuracy are improved by the algorithm in the fault diagnosis of traction substation. Therefore, the algorithm is accurate and efficient to calculate the minimal hitting sets of model-based diagnosis. In addition, the other advantages of the algorithm need to be further researched.

ACKNOWLEDGMENTS

This study is partly supported by National Nature Science Foundation of China (No. U1134205, 51377136) in China.

REFERENCES

[1] De Kleer J, Mackworth A.K, Reiter R. Characterizing diagnoses and systems[J]. Artificial Intelligence, vol. 56, no.2 , pp. 197-222, 1992.
[2] De Kleer J, Williams B.C. Diagnosing multiple faults[J]. Artificial Intelligence, vol. 32, no. 1, pp. 97-130, 1987.
[3] Reiter R. A theory of diagnosis from first principles[J]. Artificial Intelligence, vol. 32, no. 1, pp. 57~96, 1987.
[4] Greiner R, Smith B A, Wilkerson R W. A Correction to the algorithm in reiter’s theory of diagnosis[J]. Artificial Intelligence, vol. 41, no. 1, pp. 79-88, 1989.
[5] Yunfei Jiang, Li Lin. Computing minimal hitting sets with Bisection Hs-tree algorithm[J]. Journal of Software, vol. 13, no. 22, pp. 2267-2274, 2002.
[6] Yunfei Jiang, Li Lin. Computing minimal hitting sets with Boolean algebra algorithm[J]. Chinese Journal of Computers, vol. 26, no. 8, pp. 919-924, 2003.
[7] Li Lin. Calculate the minimum hitting sets of model-based diagnostic with logic array[J]. Journal of Jinan University (Natural Science & Medicine Edition), vol. 23, no. 1, pp. 24-27, 2002.
[8] Li Lin, Yunfei Jiang. Computing minimal hitting sets with genetic algorithm[C]. Proc 13th Int’l Workshop on Principles of Diagnosis, 2002.
[9] Huang Jie, Chen Lin, Zou Peng. A genetic simulated annealing algorithm to solving minimal diagnostic[J]. Journal of Software , vol. 15, no. 9, pp. 1345-1350, 2004.
[10] Storn R, Price K. Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous spaces[J]. Technical Report, International Computer Science Institute. vol. 8, pp. 22-25, 1995.
[11] Vesterom J, Thomsen R. A comparative study of differential evolution, particle swarm optimization and evolutionary algorithms on numerical benchmark problems [C]. Congress on Evolutionary Computation, pp. 1980-1987, 2004.
[12] Xuefeng Yan, Yu Juan, Qian Feng, Junwei Ding. Estimated supercritical water oxidation kinetics parameters based on improved differential evolution algorithm[J]. Journal of East China University of Science and Technology (Natural Science Edition), vol. 3, no. 1, pp. 94-97, 2006.
[13] Fang Qiang. The differential evolutionary algorithm based on eugenic policy and its application in chemical[D]. Hangzhou: Zhejiang University. pp. 28-32, 2004.
[14] Zhang Hao, Zhanhua Huang, Daoyin Yu. Contour tracking with particle filter and differential evolution[J]. Opto-Electronic Engineering, vol. 33, no. 2, pp. 29-32, 2006.
[15] Jungang Yuan, Zhiguo Sun, Guangji Qu. Application of differential evolution algorithm[J]. Computer Engineering and Applications, vol. 43, no. 7, pp. 75-77, 2007.
[16] Qiu Wei, Jianhua Zhang, Liu Nian, Wenxia Liu. The dual population of binary differential evolutionary algorithms of transmission network expansion planning[J]. Proceedings of the Chinese Society of Universities for Electric Power System and its Automation. vol. 21, no. 6, pp. 56-61, 2009.
[17] Zhigang Liu, Zhiwei Han. Fault diagnosis of electric railway traction substation with model-based relation guiding algorithm[J]. Expert Systems with Applications, vol. 41, no. 4, pp. 1730-1741, 2014.
[18] Hu Fei, Zhigang Liu, Fuqiang Fan, Guan Long. Distribution network line fault diagnosis method based on model[J]. Automation of Electric Power Systems, vol. 36, no. 10, pp. 56-60,78, 2012.