Universal transparent artificial neural network-based fault section diagnosis models for power systems

Xuan Xie[1], Guojiang Xiong[1,*], Jun Chen[2], Jing Zhang[1]

[1] Guizhou Key Laboratory of Intelligent Technology in Power System, College of Electrical Engineering, Guizhou University, Guiyang, 550025, China

[2] Department of Electrical and Computer Engineering, Oakland University, Rochester, MI 48309, USA

Abstract

Fault section diagnosis (FSD) is significant for the power system dispatching. Artificial neural network (ANN)-based FSD method has strong fault tolerance but it looks like a black box and lacks the interpretability to the diagnosis outputs. In addition, when the topology of power systems changes, the ANN structure needs to be reconstructed and retrained, and thus has low adaptive capability. In order to tackle these challenges, in this paper, an ANN-based FSD method by constructing universal transparent diagnosis models is proposed. The diagnosis models are constructed for the transmission line, transformer, and bus types rather than for a specific power system section. They can express the logical relations among sections, protective relays (PRs) and circuit breakers (CBs) clearly and intuitively. In addition, fuzzy values are used to model the uncertainties of PRs and CBs, and to determine the inputs of diagnosis models. Furthermore, the differential evolution algorithm is employed to optimize the network parameters of diagnosis models. The proposed method is verified on the IEEE 30-bus test system and an actual local power system in Jilin Province of China.

***Keywords**:* Artificial neural network, differential evolution, fault section diagnosis, transparent model

[*] Corresponding author. E-mail address: gjxiongee@foxmail.com

1. Introduction

Fault section diagnosis (FSD) is one of the essential parts of power system dispatching [1]. When a fault event occurs in a power system, protective relays (PRs) will operate and trip the corresponding circuit breakers (CBs) off to remove the fault. Then, a large amount of alarm information will rush to the dispatching center within a short time. With the continuous expansion of power systems, the dispatchers are under increasing pressure to quickly locate the faulty section(s). Therefore, turning to an efficient FSD method is helpful for dispatchers to recover the normal operation of systems quickly.

Up to now, the FSD methods based on intelligent technologies have attracted more and more attention. Compared with the traditional methods, the intelligent technology-based FSD methods have higher diagnostic accuracy and faster diagnosis speed. A large number of intelligent technology-based FSD methods have been proposed, such as the expert systems (ES) [2, 3], analytical models [4-7], graphical models (Petri nets (PN) [8, 9], Bayesian networks (BN) [10, 11], Spiking Neural P systems (SNPS) [12], etc.) and artificial neural networks (ANN) [13]-[17].

The ES method uses rule matching to obtain the diagnosis results. Kaluđer *et al.* [2] proposed a fuzzy set expert system to improve the processing ability of uncertain fault information and applied to distributed power systems. Ma *et al.* [3] presented an expert system with learning ability by introducing a back propagation algorithm into the expert system. The ES method can fully use the experts' empirical knowledge and has a fast reasoning speed. However, if a faulty scenario is not included in the rule-based knowledge base, it may misdiagnose and it is not suitable for complex power systems.

For the analytical models, its essence is to transform the FSD into a 0-1 integer mathematical model and then employs optimization algorithms to solve the model. The merits of the analytical models include strict logic explanation, tight mathematical deduction and easy implementation [18]. But they have the deficiency that unknown complex faults may be identified inaccurately because constructing an accurate analytical model is a hard task. Oliveira *et al.* [4] proposed a staged FSD framework which can reduce the pressure brought by the excessive model input to the diagnosis process. Sobhy *et al.* [5] introduced artificial bee colony algorithm to solve the FSD problem. This algorithm has few parameters that need to determine and can solve

complex faults effectively. Xiong *et al.* [6] proposed a binary coded brain storm optimization algorithm to avoid the transcoding process in solving the FSD problem, which can improve the diagnosing efficiency. Alroayaee [7] introduced a teaching-learning-based optimization algorithm to establish a fast diagnosis model for large power systems.

The methods based on graphical models mainly build a causal model to describe the logical relations among sections, PRs, and CBs. They have strong interpretability and avoid the training samples. However, their fault tolerance is relatively low and they are inapplicable to complex fault scenarios. Kiaei *et al.* [8] checked the consistency of the local unity with external information coming from the adjacent areas to improve the fault tolerance rate. Xu *et al.* [9] replaced the event-by-event temporal reasoning with a layer-by-layer temporal reasoning. Cai *et al.* [10] proposed a dynamic Bayesian network -based fault diagnosis methodology. Mansour *et al.* [11] constructed a simplified FSD system based on Bayesian networks with noisy-OR/AND nodes to estimate the faulty sections of large power stations and transmission lines. Peng *et al.* [12] introduced an intuitionistic fuzzy set into the SNPS to enhance the interpretation and fault tolerance.

Compared with the aforementioned methods, the greatest strength of the ANN-based methods lies in the high fault tolerance in dealing with complex fault scenarios. They have strong learning ability and fast diagnosis speed, and can achieve high diagnosis accuracy as long as the training samples are sufficient and reliable. Therefore, they have gained significant attentions in the field of power system FSD. The existing ANN-based methods can be classified as the following three modeling approaches. The first one is the centralized diagnosis approach [13, 14], which takes the entire power system as a whole and diagnoses all sections simultaneously. Although its principle is simple, it suffers the curse of dimensionality when applied to large-scale power systems. The second modeling approach is the divisional diagnosis approach [15], for which the power system is firstly divided into several small-scale sub-systems. Then for each sub-system, the centralized diagnosis approach is employed to construct a corresponding ANN-based sub-diagnosis model to conquer the curse of dimensionality. However, the tie lines connecting regional sub-systems need to be diagnosed repeatedly. What's more, if the topology structure of power system changes, the related ANN-based sub-diagnosis model needs to be adjusted and trained again, which makes the adaptive capability that

adapts to the change of topology structure inadequate. The last one is the section-based diagnosis approach [16, 17]. Different from the abovementioned two diagnosis approaches, this approach constructs an ANN-based diagnosis model for each section. It can overcome both the curse of dimensionality and the inadequate adaptive capability simultaneously. However, the approaches in [16] and [17] also face up to some deficiencies, and one of them is low interpretability. The diagnosis models in [16] and [17] are typical three-layer fully connected networks, which makes the internal structure obsolete and thereby unable to express the logical relations among sections, PRs and CBs clearly and intuitively. Moreover, the number of neurons in the hidden layer is determined without a sufficient theoretical foundation. In this context, the present section-based diagnosis models could not interpret the diagnosis results clearly and consequently and are difficult to earn the trust of dispatchers.

In this paper, following the modeling idea of the section-based diagnosis approach, an improved FSD method by constructing universal transparent ANN-based diagnosis models is proposed. The main contributions are as follows:

(1) Three universal transparent ANN-based diagnosis models for different types of sections including transmission line, bus and transformer are constructed. They can express the logical relations among sections, PRs and CBs clearly and intuitively. They also have good adaptive capability that do not need to be reconstructed and retrained again even if the system topology changes.

(2) Fuzzy values are used to model the uncertainties of PRs and CBs, and further to determine the input neurons.

(3) The network parameter training of the diagnosis models is transformed into an optimization problem and the differential evolution (DE) algorithm is used to solve it.

(4) The proposed FSD method is firstly verified on the IEEE 30-bus power system, and then applied to an actual fault case occurred in Jilin power grid of China. Simulation results demonstrate that the proposed method is effective with high interpretability and adaptive capability.

2. Framework of the proposed FSD method

The framework of the proposed FSD method is presented in Fig. 1.

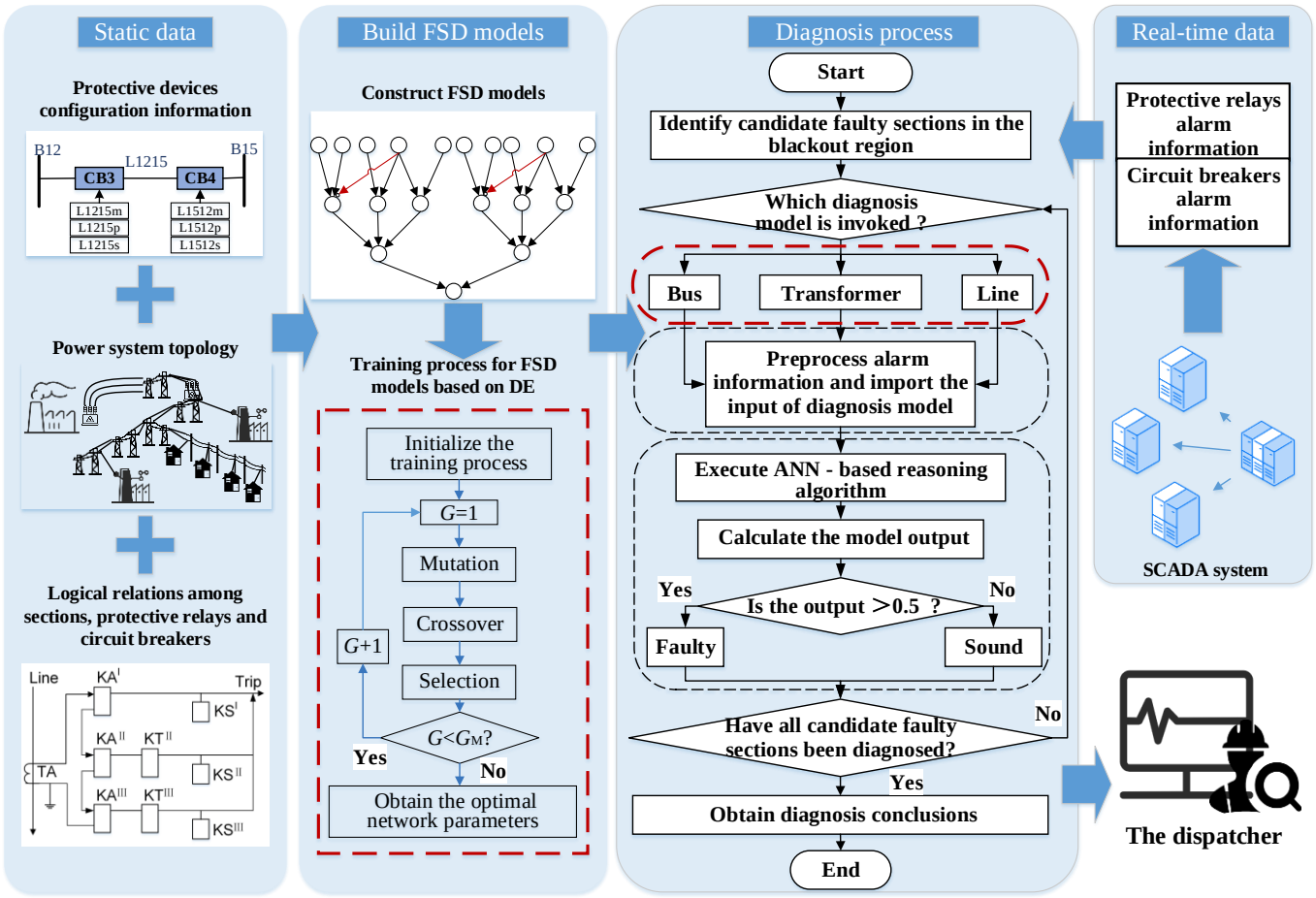

Fig. 1 Framework of the proposed ANN-based FSD method

Three universal transparent ANN-based FSD models are firstly constructed for transmission lines, transformers, and buses according to the relay protection system and logical relations among sections, PRs and CBs. Then training data sets are collected to train the network parameters of diagnosis models using the DE algorithm. When a fault event occurs, candidate faulty section(s) in the blackout region are firstly identified. For each candidate faulty section, the corresponding diagnosis model is invoked. The power system topology data coupling with the involved alarm information detected by the SCADA system are simultaneously used to determine the inputs of the diagnosis model. Then the ANN-based reasoning algorithm is executed to calculate the model outputs and thus to obtain the diagnosis conclusion.

3. Universal transparent ANN-based diagnosis models

3.1 Logical relations in the relay protection system

When a fault occurs on a section, it will be detected quickly by the PRs associated with the section. Then the PRs will operate and send signals to trip off the corresponding CBs and thus to remove the fault. This process is presented in Fig. 2.

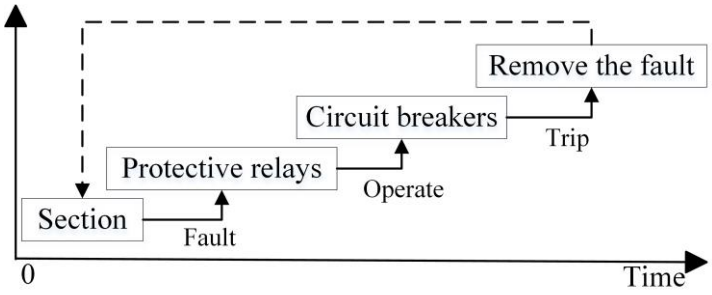

Fig. 2 Fault removing process

To describe the logical relations among sections, PRs, and CBs more specifically, a partial power system shown in Fig. 3 is taken for example. For the relay protection system, the main PR (MPR), primary backup PR (PBPR), secondary backup PR (SBPR) are denoted as m, p, and s, respectively. CBXY (CBYX) denotes the CB that between buses BX and BY, and connected to the BX (connected to BY), like CB2210 and CB1022 between buses B22 and B10.

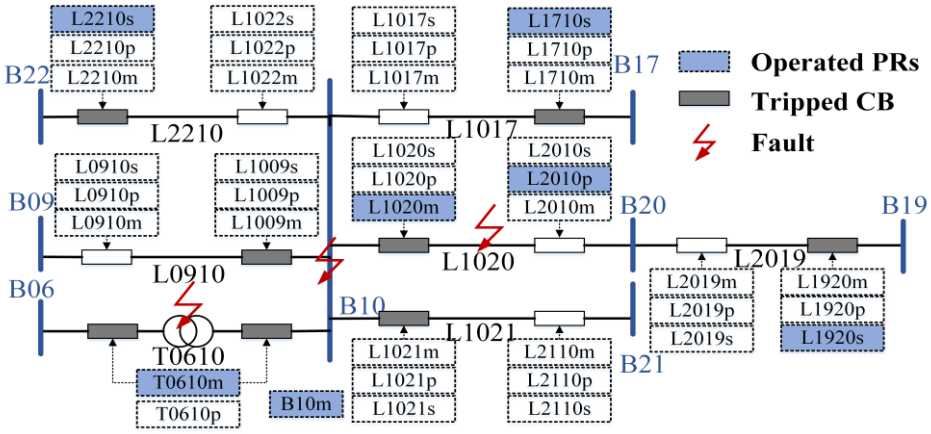

Fig. 3 A partial power system

As shown in Fig. 3, multiple faults occur in this system. The collected operating information includes B10m, T0610m, L1020m, L2010p, L1710s, L1920s, L2210s, CB1006, CB0610, CB1009, CB1021, CB1020, CB1710, CB2210, and CB1920. The candidate faulty sections are L1020, T0610, and B10.

According to the relay protection principle, the logical relations are as follows:

(1) In case of a fault, the MPRs will operate quickly and trip off the corresponding CBs. For example, when the fault occurs on L1020, the MPRs L1020m at the sending side operates first and trips off CB1020.

(2) If the MPRs fail to operate or the CBs fail to be tripped off, the PBPRs will operate to trip off the corresponding CBs. For example, L2010m fails to operate and then L2010p will operate to trip off CB2010.

(3) When both the MPRs and PBPRs fail to operate or the corresponding CBs fail to be tripped off, the SBPRs will operate to trip off the corresponding CBs. For example, when both L2010m and L2010p fail to trip off CB2010, then the SBPR L1920s will operate to trip off CB1920.

(4) According to the operating sequence of PDs, if the CBs corresponding to the MPRs and PBPRs succeed in being tripped off, then the secondary backup PDs in adjacent sections will not operate. Therefore, the isolated regions are not extended.

(5) When the transmission network suffers from a fault, there should be at least one PD at each side operate to remove.

3.2 Universal transparent ANN-based diagnosis models

According to the logical relations described in the above subsection, the universal transparent ANN-based diagnosis models can be constructed as follows.

(1) Submodules in diagnosis models

Based on the basic principles of relay protection system, there are two types of diagnosis submodules. As shown in Fig. 4(a), for the MPR (PBPR), the input neurons are activated and the diagnosis results are transmitted to the neuron σ_{P3}. Weights w_1 and w_2 represent the contributions of MPR (PBPR) and CB, respectively. In Fig. 4(b), the red reverse arc is an inhibiting arc indicating that when the CB1 that corresponds to the MPRs (PBPRs) operates, the trust degrees of the SBPR and CB2 will be reduced. w_3, w_4, and w_5 are weights.

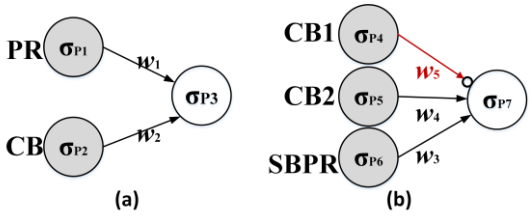

Fig. 4 Two types of diagnosis submodules. (a) Diagnosis submodule of MPRs (PBPRs); (b) Diagnosis submodule of SBPRs. CB1 and CB2 denote the CBs corresponding to the MPRs (PBPRs) and SBPRs.

(2) Diagnosis model for transmission lines

For a transmission line, the PRs consist of MPRs, PBPRs, and SBPRs. The diagnosis model contains two parts, the sending side and the receiving side as shown in Fig. 5. Each side has three diagnosis submodules according to the configured protection scheme. In addition, the CB corresponding to the MPR and PBPR plays an inhibiting effect on the operating of SBPR, like the reverse inhibiting arc w_{10}.

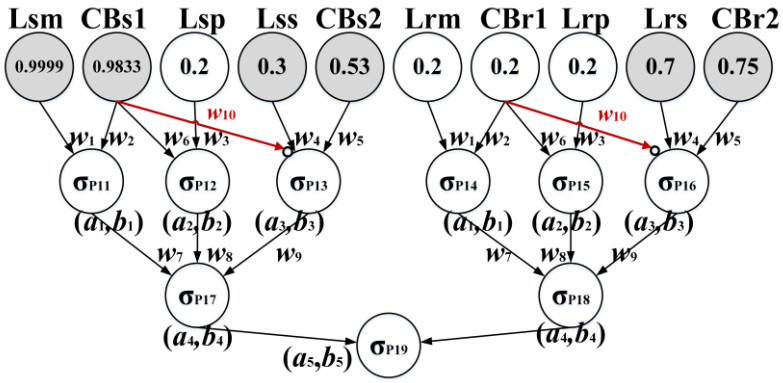

Fig. 5 Diagnosis model for transmission lines. Lsm (p, s)/Lrm (p, s) denotes the MPR (PBPR, SBPR) on the sending/receiving side; CBs1(2)/CBr1(2) denotes the CB corresponding to MPR/PBPR(SBPR) on the sending/receiving side.

Take the fault on L1020 for example. The operated PDs are highlighted in gray. On the sending side, the MPR Lsm (L1020m) operates to successfully trip off CBs1 (CB1020) and thus to remove the fault. CB1020 has an inhibiting effect on the SBPR (Lss). On the receiving side, the MPR Lrm (L2010m) refuses to operate and then the PBPR Lsp (L2010p) operates to trip off CB2010. However, CB2010 fails to be tripped off. Then the corresponding SBPR Lrs (L1920s) operates to trip off CBr2 (CB1920) successfully and thereby to remove the fault.

It is worth noting that, when the sending side or the receiving side of a line has a lot of adjacent sections, it is inefficient and impracticable for operation dispatching if we construct a diagnosis submodule for each SBPR and its corresponding CB installed at each adjacent section. For example, on the sending side of L1020, there are five adjacent lines /transformers as shown in Fig. 3. According to the traditional methods, we need to construct five diagnosis submodules for these adjacent sections. It is obvious

that the diagnosis model has a large structure. Moreover, when one adjacent line /transformer is out of operation or back into operation, we need to reconstruct the diagnosis model again. To avoid such problems, we take all the adjacent sections as a whole and construct only one submodule for all the SBPRs and their corresponding CBs installed at these sections as shown in Fig. 5. Obviously, the structure of this diagnosis model is very simple. Specific preprocessing method for calculating the initial value of input neurons to reduce the work of reconstructing is introduced in section 4.

(3) Diagnosis model for transformers

The model of relay protection system of transformers is similar to that of transmission lines. The main difference is that a line has two MPRs (PBPRs) corresponding to the sending side and the receiving side while a transformer has only one MPR (PBPR). This MPR (PBPR) can operate to trip off all the CBs installed at different sides of a transformer. The universal transparent ANN-based diagnosis model for transformers is provided in Fig. 6. For the SBPRs and their corresponding CBs installed at the adjacent sections, the handling approach similar to that of lines is also adopted.

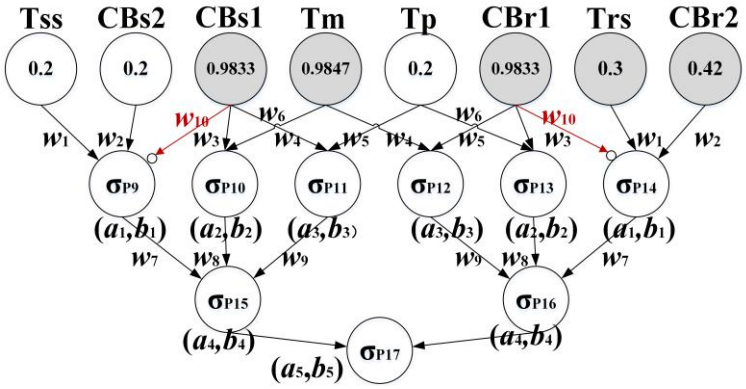

Fig. 6 Diagnosis model for transformers

(4) Diagnosis model for buses

Different from transmission lines and transformers, the relay protection system of buses does not contain the PBPR. Besides, buses have another representative characteristic, i.e., they have many direct connected lines /transformers. Similar to the handling approach for the adjacent sections of lines, we also take all the connected sections of a bus as a whole and construct only one submodule for the SBPRs and their corresponding CBs as shown in Fig. 7. In addition, for the CBs corresponding to the

MPR, they are also taken as a whole and using one input neuron to represent them. It is clear that buses have a concise diagnosis model.

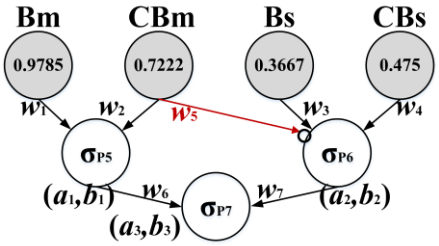

Fig. 7 Diagnosis model for buses

It can be seen from Fig. 5 to Fig. 7 that the structures of the constructed universal transparent ANN-based diagnosis models are determined by the logical relations among sections and PDs. They are no longer three-layer fully connected networks. The diagnosis models for lines and transformers have two hidden layers while the model for buses has one. In addition, adjacent layers are not fully connected but partially connected.

3.3 Activation function of hidden layer neurons

In the ANN-based diagnosis models, the transmission of alarm information is synchronized with the transmission of nerve impulses. An impulse is used to transmit alarm information from the current-layer neurons to the next-layer neurons by activating the current-layer neurons. In an ANN, an activation function is used to simulate the process of nerve impulse. The sigmoid function is a commonly used activation function. Traditionally, the activation functions of different hidden layer neurons of ANN are consistent. In this work, we hope to obtain a group of activation functions with similar characteristics to simulate the logical relations among sections, PRs and CBs. To this end, a modified two-parameter sigmoid function is defined as follows:

$$m.sigmoid(x) = \frac{a}{1+e^{-b(x-0.5)}} \tag{1}$$

where a is used to control the maximum value of the activation function, and b is used to control the changing rate. This function can achieve different mapping effects by tuning the parameters a and b, and thus to satisfy the transmission requires of different hidden layer neurons.

4. Fuzzy representation of the diagnosis models and input data preprocessing

4.1 Fuzzy representation of the FSD models

In the actual power system, the alarm information received by the SCADA system sometimes is not entirely correct. Some information may be distorted or lost in the transmission process. In addition, the PDs have the possibilities of refused operating or mal-operating. To simulate the uncertainties and reduce the influence of sample errors on FSD results, fuzzy numbers are used to represent the initial values of PDs.

The operation statistics of PDs in 220kV and above AC system of State Grid Corporation of China in ten years (2006 – 2015) [19] are taken as the confidence levels of main PDs. In general, the confidence levels of primary backup PDs are lower than those of main PDs but higher than those of secondary backup PDs. For the undetected information, the confidence levels are set to be 0.2. The detailed fuzzy initial values of different PRs and corresponding CBs are tabulated in Table 1.

Table 1 Confidence levels of PDs

	PDs	Line	Bus	Transformer
Main PDs	PRs	0.9999 (0.2)[a]	0.9785 (0.2)	0.9847 (0.2)
	CBs	0.9833 (0.2)	0.9833 (0.2)	0.9833 (0.2)
Primary backup PDs	PRs	0.8 (0.2)	-	0.8 (0.2)
	CBs	0.85 (0.2)	-	0.85 (0.2)
Secondary backup PDs	PRs	0.7 (0.2)	0.7 (0.2)	0.7 (0.2)
	CBs	0.75 (0.2)	0.75 (0.2)	0.75 (0.2)

a Numbers in brackets denote the confidence levels of undetected information.

4.2 SBPRs and CBs preprocessing

As described in section 2, to reduce the complexity of diagnosis models and improve the model universality, for the adjacent sections of transmission lines, transformers, and buses, we take them as a whole and construct only one input neuron for the SBPRs and one for the corresponding CBs. For the CBs corresponding to the MPR of buses, we also use only one input neuron to represent them. For these involved input neurons, a weighted average method is presented to calculate their values. For the transmission lines and transformers, the initial values of the input neurons corresponding to the SBPRs and corresponding CBs are calculated as follows:

$$\gamma_{\mathrm{Xss}}\left(\gamma_{\mathrm{Xrs}}\right)=\frac{n_{\mathrm{Xs,o}}\times a_{\mathrm{Xs,o}}+n_{\mathrm{Xs,c}}\times a_{\mathrm{Xs,c}}}{n_{\mathrm{Xs,o}}+n_{\mathrm{Xs,c}}} \tag{2}$$

$$\gamma_{\mathrm{CBss}}(\gamma_{\mathrm{CBss}}) = \frac{n_{\mathrm{CBs,o}} \times a_{\mathrm{CBs,o}} + n_{\mathrm{CBs,c}} \times a_{\mathrm{CBs,c}}}{n_{\mathrm{CBs,o}} + n_{\mathrm{CBs,c}}} \tag{3}$$

where the subscript X denotes the sections such as lines and transformers; the subscript s(r) denotes the sending (receiving) side; $n_{\mathrm{Xs,o}}$ ($n_{\mathrm{Xs,c}}$) and $n_{\mathrm{CBs,o}}$ ($n_{\mathrm{CBs,c}}$) represent the operated (undetected) number of SBPRs and CBs; $a_{\mathrm{Xs,o}}$ ($a_{\mathrm{Xs,c}}$) and $a_{\mathrm{CBs,o}}$ ($a_{\mathrm{CBs,c}}$) denote the initial fuzzy values of operated (undetected) SBPRs and CBs as shown in Table 1.

For the SBPRs and their corresponding CBs of buses, the initial values of the input neurons are calculated as follows:

$$\gamma_{\mathrm{Bs}} = \frac{n_{\mathrm{Bs,o}} \times a_{\mathrm{Bs,o}} + n_{\mathrm{Bs,c}} \times a_{\mathrm{Bs,c}}}{n_{\mathrm{Bs,o}} + n_{\mathrm{Bs,c}}} \tag{4}$$

$$\gamma_{\mathrm{CBs}} = \frac{n_{\mathrm{CBs,o}} \times a_{\mathrm{CBs,o}} + n_{\mathrm{CBs,c}} \times a_{\mathrm{CBs,c}}}{n_{\mathrm{CBs,o}} + n_{\mathrm{CBs,c}}} \tag{5}$$

where $n_{\mathrm{Bs,o}}$ ($n_{\mathrm{Bs,c}}$) and $n_{\mathrm{CBs,o}}$ ($n_{\mathrm{CBs,c}}$) represent the operated (undetected) number of SBPRs and CBs; $a_{\mathrm{Bs,o}}$ ($a_{\mathrm{Bs,c}}$) and $a_{\mathrm{CBs,o}}$ ($a_{\mathrm{CBs,c}}$) denote the initial fuzzy values of operated (undetected) SBPRs and CBs as shown in Table 1.

For the CBs corresponding to the MPR, the initial value of the input neuron is calculated as follows:

$$\gamma_{\mathrm{CBm}} = \frac{n_{\mathrm{CBm,o}} \times a_{\mathrm{CBm,o}} + n_{\mathrm{CBm,c}} \times a_{\mathrm{CBm,c}}}{n_{\mathrm{CBm,o}} + n_{\mathrm{CBm,c}}} \tag{6}$$

where $n_{\mathrm{CBm,o}}$ ($n_{\mathrm{CBm,c}}$) represents the operated (undetected) number of CBs; $a_{\mathrm{CBm,o}}$ ($a_{\mathrm{CBm,c}}$) denotes the initial fuzzy values of operated (undetected) CBs as shown in Table 1.

5. Training method for diagnosis models with differential evolution

One of the prominent advantages of ANN is its strong learning ability. It is also the key factor that makes the FSD method possess high fault tolerance in dealing with complex fault scenarios. However, the traditional training method based on back propagation can be stuck in the local optimum, leading to the failure of training. To find accurate values for the models' involved network parameters, the training task is transformed into an optimization problem and the DE algorithm is employed as solution method in this work to improve the training results.

5.1 DE algorithm

DE is a powerful population-based metaheuristic algorithm [20]. It has a simple structure and can deal with non-convex optimization problems well with strong global search ability. The main operations of DE are as follows:

(1) Initialization

A population with *NP* individuals in a *D*-dimensional space can be encoded as a *NP*×*D* matrix. The *n*-th individual can be expressed as follows:

$$\boldsymbol{x}_{n,G} = [x_{1,n,G}, x_{2,n,G}, ..., x_{D,n,G}] \tag{7}$$

where n=1, 2, …, *NP*; *G* is the number of iterations. $x_{m,n,G}$ is generally initialized as follows:

$$x_{m,n,G} = rand(0,1) \cdot (x_{m,\max} - x_{m,\min}) + x_{m,\min} \tag{8}$$

where rand(0,1) is a random number in (0,1); m=1, 2, …, *D*; $x_{m,\max}$ and $x_{m,\min}$ are the upper and lower bounds, respectively.

(2) Mutation

The mutation operation is adopted to generate a donor vector $\boldsymbol{v}_{n,G}$ =[$v_{1,n,G}$, $v_{2,n,G}$, …, $v_{D,n,G}$] using the following method:

$$\boldsymbol{v}_{n,G} = \boldsymbol{x}_{t_1,G} + F \cdot (\boldsymbol{x}_{t_2,G} - \boldsymbol{x}_{t_3,G}) \tag{9}$$

where t_1, t_2 and t_3 are mutually exclusive integers randomly selected from [1, *NP*], which are also different from the base vector index n; *F* is a scaling factor used to control the mutation proportion.

(3) Crossover

A crossover rate *CR* is used to determine the sources of a trial vector $\boldsymbol{u}_{n,G}$ =[$u_{1,n,G}$, $u_{2,n,G}$, …, $u_{D,n,G}$]. The detailed crossover method is as follows:

$$u_{m,n,G} = \begin{cases} v_{m,n,G}, \text{if } (rand(0,1) < CR \text{ or } m == m_{rand} \\ x_{m,n,G}, \text{otherwise} \end{cases} \tag{10}$$

where m_{rand} is a random integer in {1, 2, …, *D*}.

(4) Selection

The selection operation is used to compare the trial vector $\boldsymbol{u}_{n,G}$ with the target individual $\boldsymbol{x}_{n,G}$ and choose the better one to go to the next iteration:

$$\boldsymbol{x}_{n,G+1} = \begin{cases} \boldsymbol{u}_{n,G}, \text{if } f(\boldsymbol{u}_{n,G}) \le f(\boldsymbol{x}_{n,G}) \\ \boldsymbol{x}_{n,G}, \text{if } f(\boldsymbol{u}_{n,G}) > f(\boldsymbol{x}_{n,G}) \end{cases} \tag{11}$$

where $f(\cdot)$ is the objective function.

5.2 Application of DE to diagnosis model training

The trained network parameters include the weights and sigmoid function parameters. We transform the diagnosis model's training into a minimization optimization problem whose objective function is the deviation between the actual outputs and the ideal outputs of training samples. The mean square error (MSE) is used as the deviation evaluation function:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(AO_i - IO_i)^2$$
$$st.\begin{cases} w_{l,\min} \le w_l \le w_{l,\max} \\ a_{j,\min} \le a_j \le a_{j,\max} \\ b_{k,\min} \le b_k \le b_{k,\max} \end{cases} \tag{12}$$

where N is the number of training samples; AO and IO are the actual output and ideal output, respectively.

5.3 Diagnosis models' training and validating

Due to the space limitation, we only provide the training and validating results for the line type section to show the effectiveness of the proposed method. A total number of 60 samples are employed as the training samples. The training convergence curve presented in Fig. 8 shows that training MSE is only 7.8679E-04. Fig. 9 shows that the actual outputs of the training samples agree well with the ideal outputs over the whole samples, which indicates that the training method based on DE is effective.

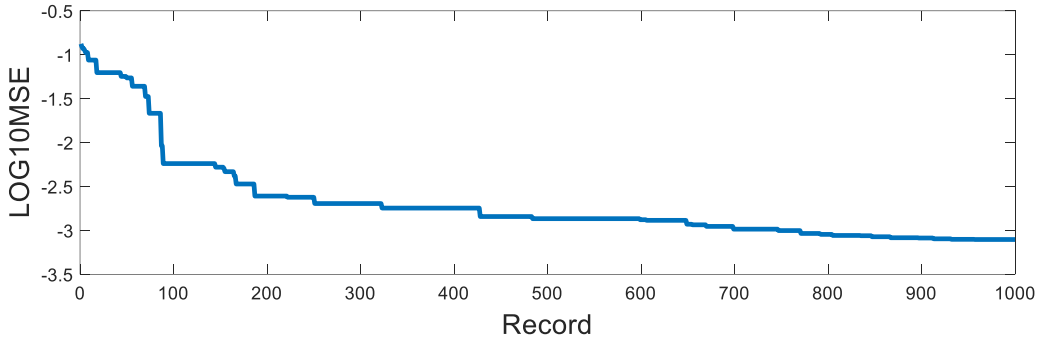

Fig. 8 Training convergence curve

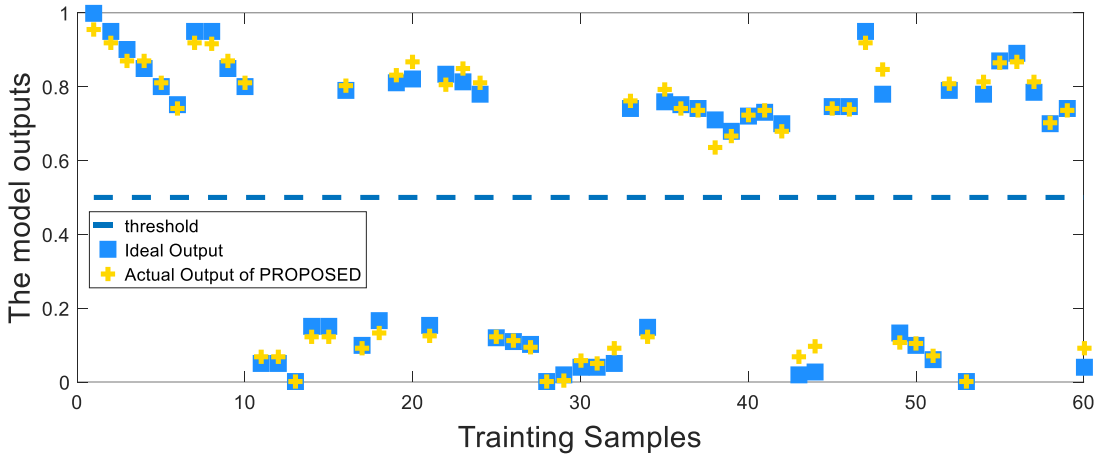

Fig. 9 Distribution of the training samples and their actual outputs

Besides, we use another 20 samples to validate the diagnosis model. As shown in Fig. 10, the actual outputs of the validating samples are also highly consistent with the ideal outputs over the whole samples. In fact, the validating MSE is only 8.1371E-04, demonstrating that the trained diagnosis model has robust generalization.

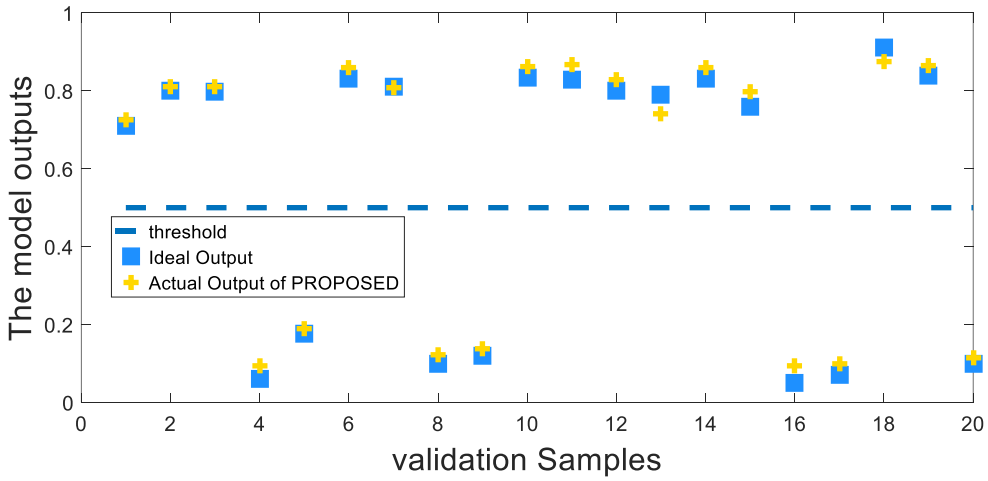

Fig. 10 Distribution of the validating samples and their actual outputs

After the training, the obtained values for the involved network parameters of the line diagnosis model are presented in Table 2. For the transformers and buses, the optimized values for the network parameters of their diagnosis models are presented in Table 3 and Table 4, respectively.

Table 2 Network parameters of diagnosis model for lines

Parameter	Value	Parameter	Value
w_1	0.3542	a_1	2.6416
w_2	0.2953	a_2	2.2062

w_3	0.8107	a_3	2.8350
w_4	0.2270	a_4	0.3266
w_5	0.4671	a_5	1.0714
w_6	0.7313	b_1	25.0000
w_7	0.2776	b_2	24.0762
w_8	0.4514	b_3	24.5498
w_9	0.3002	b_4	5.0000
w_{10}	0.2329	b_5	13.9411

Table 3 Network parameters of diagnosis model for transformers

Parameter	Value	Parameter	Value
w_1	0.3922	a_1	0.8890
w_2	0.5343	a_2	0.4790
w_3	0.4320	a_3	1.9727
w_4	0.2018	a_4	0.3386
w_5	0.2023	a_5	1.0742
w_6	0.6318	b_1	22.4791
w_7	0.8435	b_2	29.4561
w_8	0.9485	b_3	29.9863
w_9	0.4322	b_4	6.5478
w_{10}	0.2037	b_5	11.0528

Table 4 Network parameters of diagnosis model for buses

Parameter	Value	Parameter	Value
w_1	0.2024	a_1	2.3983
w_2	0.6302	a_2	0.9892
w_3	0.3691	a_3	0.9470
w_4	0.5708	b_1	11.5021
w_5	0.1281	b_2	11.4733
w_6	0.4839	b_3	7.3726
w_7	0.8034	-	-

6. Case studies

To verify the fault tolerance and feasibility of the proposed method, two different power systems including the IEEE 30-bus test system [17] and an actual Siping power system in Jilin province of China [6] are studied. The proposed method is compared with Petri nets [9] (don't consider the temporal information), SNPS [12], and RBFNN [17]. All the simulations are executed on a 2.1-GHz Intel(R) Core(TM) computer with 8.0-GB RAM under MATLAB 2016a.

6.1 Case 1 IEEE 30-bus test system

The IEEE 30-bus system is shown in Fig. 10. Due to space limitations, the device numbers of CBs are not given in Fig. 11.

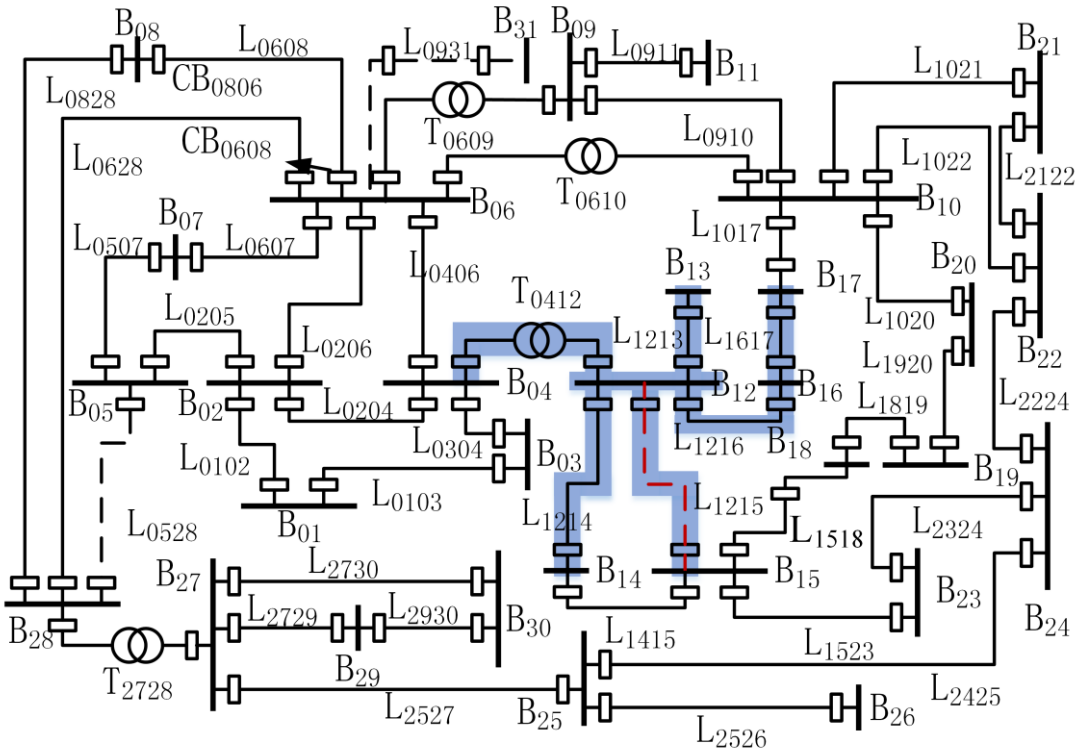

Fig. 11 IEEE 30-bus test system

For this test system, a complex fault event with multiple faulty sections is considered. The alarm informations detected by the SCADA system are presented in Table 5.

Table 5 Alarm information of case 1

Item	Alarm information	Item	Alarm information
1	MPR B12m	8	CB1214
2	MPR T0412m	9	CB1215

3	MPR L1216m	10	CB1216
4	SBPR L1312s	11	CB1312
5	SBPR L1716s	12	CB1512
6	CB0412	13	CB1716
7	CB1204	-	-

From the alarm information, we can see that the fault event involves multiple candidate faulty section. We take line L1216 for example to illustrate the diagnosis process. After analyzing the system topology, we know that there are four adjacent sections on the sending side and one on the receiving side. During the fault removing process, the MPR L1216m on the sending side operates and trips off CB1216 successfully. While on the receiving side, both the MPR L1612m and PBPR L1612s fail to operate. In this case, the SBPR L1617s operates to trip off CB1617 to remove the fault. When using the proposed diagnosis model to diagnose this candidate faulty section, we first identify the section type, i.e., the line type and invoke the corresponding diagnosis model. Then the initial values for the input neurons are preprocessed using the proposed weighted average method. Although the MPR L1216m has removed this candidate faulty section on the sending side, the complex fault event also triggers other second backup PDs like the SBPR L1312s and CB0412. Therefore, the initial values for the input neurons of second backup PDs are calculated as follows:

$$\gamma_{\mathrm{Lss}} = (1\times 0.7 + 3\times 0.2)/(1+3) = 0.325 \tag{13}$$

$$\gamma_{\mathrm{CBss}} = (3\times 0.75 + 1\times 0.2)/(1+3) = 0.6125 \tag{14}$$

$$\gamma_{\mathrm{Lrs}} = (1\times 0.7 + 0\times 0.2)/1 = 0.7 \tag{15}$$

$$\gamma_{\mathrm{CBrs}} = (1\times 0.75 + 0\times 0.2)/1 = 0.75 \tag{16}$$

Therefore, the values for the input neurons of the diagnosis model are [0.9999, 0.9833, 0.2, 0.325, 0.6125, 0.2, 0.2, 0.2, 0.7, 0.75]. Combining the input values with the parameters in Table 2, we can calculate the model output σ_{P19} quickly whose value is 0.87. Therefore, it can be concluded that this line is faulty because the output is larger than the threshold 0.5 [17].

For the other involved candidate faulty sections, their diagnosis outputs can be obtained quickly using the same method. The FSD results and computational time are

presented in Table 6. The diagnosis results are consistent with the results in Ref. [17], indicating that the proposed method is effective. Compared with other methods, the diagnosis outputs of the proposed method of the proposed method are the overall highest. In addition, the diagnosis speed of the proposed method is significantly faster than other models. This is mainly because the diagnosis model size of the proposed method is smaller than that of other methods. Taking L1216 for example, the proposed method's neurons: activation function number equals 19:9; the Petri nets' propositions: rule number equals 19:14; the IFSNPS's proposition neurons: rule neurons equals 35:19. Obviously, the proposed diagnosis model is more compact.

Table 6 FSD results and computational time (ms) of case 1

Section		B12	T0412	L1216
Method		[0.9785, 0.8266, 0.3, 0.58]	[0.2, 0.2, 0.9833, 0.9847, 0.2, 0.9833, 0.3250, 0.4750]	[0.9999, 0.9833, 0.2, 0.325, 0.6125, 0.2, 0.2, 0.2, 0.7, 0.75]
Proposed	Result	0.9458	0.9371	0.8700
	Time	4.20E-02	1.65E-01	2.87E-01
Petri nets	Result	0.9050	0.7375	0.8750
	Time	2.39E+01	4.34E+01	3.86E-00
SNPS	Result	0.81	0.81	0.81
	Time	5.94E+01	6.08E+01	5.44E+01
RBFNN	Result	0.9293	0.9336	0.8738
	Time	-	-	-

6.2 Case 2 Adaptive capability to topology change

In the actual power system, the system topology changes frequently, and therefore the diagnosis methods need to have the capability to adapt to the change. Another case based on Case 1 where L1215 is out of service due to maintenance is used to verify the adaptive capability of the proposed method. The alarm information is tabulated in Table 7.

Table 7 Alarm information of case 2

Item	Alarm information	Item	Alarm information
1	MPR B12m	7	CB1204
2	MPR T0412m	8	CB1214
3	MPR L1216m	9	CB1216
4	SBPR L1312s	10	CB1312

5	SBPR Ll716s	11	CB1716
6	CB0412	-	-

We still take line L1216 for example. The number of adjacent sections on the sending side is changed from four to three due to the maintenance of L1215. The initial inputs related to the second backup PDs now are as follows:

$$\gamma_{\mathrm{Lss}} = (1\times 0.7 + 2\times 0.2)/(1+2) = 0.3667 \tag{17}$$

$$\gamma_{\mathrm{CBss}} = (2\times 0.75 + 1\times 0.2)/(1+2) = 0.5667 \tag{18}$$

$$\gamma_{\mathrm{Lrs}} = (1\times 0.7 + 0\times 0.2)/1 = 0.7 \tag{19}$$

$$\gamma_{\mathrm{CBrs}} = (1\times 0.75 + 0\times 0.2)/1 = 0.75 \tag{20}$$

Therefore, the values for all the input neurons are [0.9999, 0.9833, 0.2, 0.3667, 0.5667, 0.2, 0.2, 0.2, 0.7, 0.75]. After calculating the model output σ_{P7}, we get the value 0.8699 which also gives the same diagnosis conclusion.

For other sections, the diagnosis outputs and computational time are presented in Table 8. Obviously, the overall diagnosis accuracy of the proposed method is also the highest among the involved methods and it faster than other methods considerably. After the line L1215 quits, the Petri nets needs to reconstruct 10 parameter matrices and the IFSNPS needs to refresh five parameters, while the proposed method only needs to modify the model inputs, indicating that the proposed diagnosis method has good adaptive capability. At this point, the diagnosis model sizes of these three methods change to 19:9, 17:13, and 32:18, respectively.

Table 8 FSD results and computational time (ms) of system topology change

Section		B12	T0412	L1216
Method		[0.9785, 0.7875, 0.3, 0.53]	[0.2, 0.2, 0.9833, 0.9847, 0.2, 0.9833, 0.3667, 0.3833]	[0.9999, 0.9833, 0.2, 0.3667, 0.5667, 0.2, 0.2, 0.2, 0.7, 0.75]
Proposed	Result	0.9447	0.9368	0.8699
	Time	5.18E-02	8.65E-02	2.63E-01
Petri nets	Result	0.9000	0.7375	0.8750
	Time	3.58E+01	2.24E+01	2.24E+01
SNPS	Result	0.81	0.81	0.81
	Time	5.99E+01	8.00E+01	5.37E+01

RBFNN	Result	0.8840	0.9215	0.9152
	Time	-	-	-

6.3 Case 3 FSD of actual power system

To further verify the feasibility of the proposed method, an actual fault event occurred in the Siping power system in Jilin province of China [6] is tested. The related system and alarm information are given in Fig. 12 and Table 9, respectively.

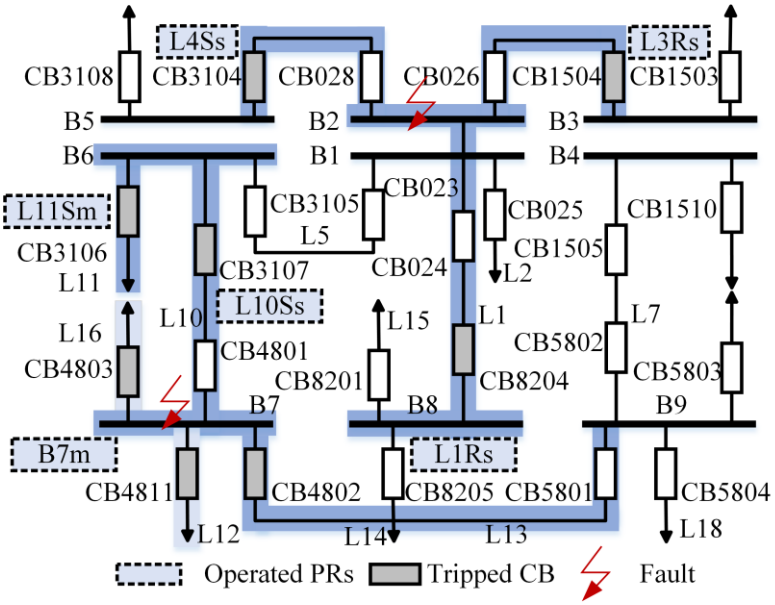

Fig. 12 Siping power system in Jilin province

Table 9 Alarm information of case 3

Item	Alarm information	Item	Alarm information
1	MPR B7m	8	CB3104
2	MPR L11Sm	9	CB8204
3	SBPR L1Rs	10	CB4802
4	SBPR L3Rs	11	CB4803
5	SBPR L4Rs	12	CB4811
6	SBPR L10Rs	13	CB3106
7	CB1504	14	CB3107

According to the alarm information, the candidate faulty sections can be obtained. When the fault occurs, the MPR B2m refuses to operate and thereby the SBPRs operate to trip off corresponding CBs. Then we can easily get the values of input neurons for the diagnosis model of B2, as shown in Table 10. It is clear that the proposed method is

significantly superior to other methods in both terms of diagnosis results and computational time. Besides, it is worth noting that the IFSNPS misdiagnoses the faulty section B2.

Table 10 FSD results and computational time (ms) of case 3

Section		B2	B7
Method		[0.2, 0.2,0.7,0.75]	[0.9785,0.7875,0.325,0.3375]
Proposed	Result	0.8481	0.9397
	Time	7.04E-05	7.67E-05
Petri nets	Result	0.825	0.9
	Time	3.76E-02	3.78E-02
IFSNPS	Result	0.441	0.81
	Time	5.48E-02	6.59E-02
Method in [6]	Result	Faulty	Faulty
	Time	-	-

At the same time, for bus B7, the MPR B7m operates quickly and trips off CB4811, CB4802 and CB4803 except for CB4801. Consequently, the SBPR L10Ss operates to trip off CB3107.While for the L11Ss and CB3106, they operate wrong. Therefore, the input values for B7 are as follows:

$$\gamma_{CBm} = (3\times 0.9833+1\times 0.2)/4=0.7875 \tag{21}$$

$$\gamma_{Bs} = \left(0.7+3\times 0.2\right)/4 = 0.325 \tag{22}$$

$$\gamma_{CBs} = (1\times 0.75+3\times 0.2)/4 = 0.3375 \tag{23}$$

After yielding the input values, we can achieve the model outputs quickly as shown in Table 10. Both buses B2 and B7 are faulty. Compared with the results in Ref. [6], it indicates that even if multiple serious faults occur in the actual power system, the proposed method can diagnose them accurately.

6.4 Performance comparison

From the results of the above cases， it is clear that the proposed method is suitable for kinds of FSD cases. The method can make intuitive logical reasoning when diagnosing faulty sections. The proposed method is qualitatively compared with some advanced methods, as shown in Table 11. Through comparative analysis, the proposed method is effective and has strong competitiveness.

Table 11 Comparisons between the proposed method and some other methods

FSD methods	Model scale	Model transparency	Fault tolerance	Adaptive capability to large-scale system	Adaptive capability to system topology change
Petri nets [9]	Good: model is constructed for each specific section	Excellent: model can express the logical relations among section, PRs and CBs	General: uses fuzzy numbers to simulate the uncertainties of PRs and CBs	Excellent: model is independent of the system scale	Good: model needs to be reconstructed according to the new topology
Bayesian network [11]	Good: similar to [9]	Excellent: similar to [9]	Good: model is trained but doesn't simulate the uncertainties of PRs and CBs	Excellent: similar to [9]	General: needs to collect new training samples to train models of relevant sections
SNP systems [12]	Good: similar to [9]	Excellent: similar to [9]	General: uses fuzzy numbers to simulate malfunction and rejection of PRs and CBs	Excellent: similar to [9]	Good: similar to [9]
Improved RBFNN [14]	Poor: model is constructed for entire system	Poor: model is like a black box	Good: similar to [11]	Poor: model scale is very large	Poor: needs to collect new training samples to train models of entire system
RBFNN [17]	General: model is constructed for a sub-system	Poor: similar to [14]	Good: similar to [11]	Good: large-scale system can be divided into small-scale sub-system	Poor: needs to collect new training samples to train models of relevant sub-system
Proposed method	Excellent: model is constructed for each type of section	Excellent: similar to [9]	Excellent: model is trained and simulates the uncertainties of PRs and CBs	Excellent: similar to [9]	Excellent: doesn't need to reconstruct model

7. Conclusion

In this paper, a FSD method by constructing universal transparent ANN-based diagnosis models for power systems is proposed. The proposed method has high diagnosis accuracy in short computational time, provides interpretable results for

dispatchers to make decision, and has strong adaptivity to system topology changes. Two power systems including the IEEE 30-bus test system and the Siping power system of China are employed to verify its feasibility and effectiveness, which clearly demonstrate the benefits of the proposed methods. In future, we will apply the proposed method to more complex fault events in actual power systems.

Acknowledgments: The authors would like to thank the editor and the reviewers for their constructive comments. This work was supported by the National Natural Science Foundation of China (Grant 51907035 and 52167007).

References

[1] M.M. Eissa, "Protection techniques with renewable resources and smart grids—A survey," Renewable and Sustainable Energy Reviews, vol. 52, pp. 1645-1667, Aug. 2015.
[2] S. Kaluđer, K. Fekete, L. Jozsa, and Z. Klaić, "Fault diagnosis and identification in the distribution network using the fuzzy expert system," Eksploatacja i Niezawodnosc - Maintenance and Reliability, vol. 20, no. 4, pp. 621-627, 2018.
[3] D. Ma, Y. Liang, X. Zhao, R. Guan, and X. Shi, "Multi-BP expert system for fault diagnosis of power system," Engineering Applications of Artificial Intelligence, vol. 26, no. 3, pp. 937-944, Apr. 2013.
[4] A. Oliveira, O. Araújo, G. Jr, A. Morais, and L. Mariotto, "A mixed integer programming model for optimal fault section estimation in power systems," International Journal of Electrical Power & Energy Systems, vol. 77, pp. 372-384, May. 2016.
[5] M. Sobhy, A. Abdelaziz, M. Ezzat, W. Elkhattam, A. Yadav, and B. Kumar. "Artificial Bee Colony Optimization Algorithm for Fault Section Estimation," Artificial Intelligence and Evolutionary Computations in Engineering Systems, pp.127-139, 2017.
[6] G. Xiong, D. Shi, J. Zhang, and Y. Zhang, "A binary coded brain storm optimization for fault section diagnosis of power systems," Electric Power Systems Research, vol. 163, no. 1, pp. 441-451, Jul. 2018.
[7] R. Alroayaee, " Fault section estimation in electric power systems using Teaching-Learning-Based-Optimization algorithm," International Research Journal of Engineering and Technology, vol. 5, no. 5, pp. 2689-2696, 2019.
[8] I. Kiaei, and S. Lotfifard, "A Two-Stage Fault Location Identification Method in Multiarea Power Grids Using Heterogeneous Types of Data," IEEE Transactions on Industrial Informatics, vol. 15, no. 7, pp. 4010-4020, 2019.
[9] B. Xu, X. Yin, X. Yin, Y. Wang, and S. Pang, "Fault diagnosis of power systems based on Temporal Constrained Fuzzy Petri Nets," IEEE Access, vol. 7, no. 1, pp. 276 - 285. Aug. 2019.
[10] B. Cai; Y. Liu; and M. Xie, "A Dynamic-Bayesian-Network-Based Fault Diagnosis Methodology Considering Transient and Intermittent Faults," IEEE Transactions on Automation Science and Engineering, vol. 14, no. 1, pp. 42-46, Jan. 2017.
[11] M. Mansour, M. Wahab and W. Soliman, "Bayesian networks for fault diagnosis of a large power station and its transmission lines," Electric Power Components and Systems, vol. 40, no. 8, pp. 845-863, Apr.2012.
[12] H. Peng, J. Wang, J. Ming, P. Shi, M. Pérez-Jiménez, W. Yu, and C. Tao, "Fault diagnosis of power systems using Intuitionistic Fuzzy Spiking Neural P Systems," IEEE Transactions on Smart Grid, vol. 9, no. 5, pp. 4777-4784, Sep.2018.
[13] P.P. Bedekar, S.R. Bhide, V.S. Kale, "Fault section estimation in power system using Hebb's rule and continuous genetic algorithm," Electrical Power and Energy Systems. vol. 33, no.3, pp. 457–465, 2011.
[14] Y. Luo, S. Ling, and Y. Cao, "Fault diagnosis of Electric Power Grid Based on Improved RBF Neural Network," TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 12, no. 9, pp. 4777-4784. 2014.

[15] G. Xiong, D. Shi, J. Chen, L. Zhu, and X. Duan, “Divisional fault diagnosis of large-scale power systems based on radial basis function neural network and fuzzy integral,” Electric Power Systems Research, vol.105, no.2, pp. 9-19, Jul. 2013.
[16] A. Flores , E. Quiles, E. García and F. Morant, “Fault Diagnosis of Electric Transmission Lines using Modular Neural Networks,” IEEE Latin America Transactions, vol. 14, no. 8, pp. 3663-3668, Aug. 2016.
[17] G. Xiong, D. Shi, L. Zhu, and X. Chen, “Fuzzy cellular fault diagnosis of power grid based on radial basis function neural network,” Automation of electric Power Systems, vol.38, no.5, pp. 59-65, 2014.
[18] G. Xiong, X. Yuan, A. W. Mohamed, and J. Zhang, “Fault section diagnosis of power systems with logical operation binary gaining-sharing knowledge-based algorithm,” International Journal of Intelligent Systems, vol. 37, no. 2, pp. 1057-1080, Feb. 2022.
[19] L. Zhang, D. Wang, Y. Liu, Z. Zhou, P. Lü, Z. Wang, Y. Li, Y. Liu, H. Shen, and G.Yang, “Analysis on Protective Relaying and Its Operation Conditions in 220 kV and Above AC System of SGCC in Past Ten Years,” Power System Technology, vol.41, no.5, pp. 1654-1659, May. 2017.
[20] R. Storn and K. Price, “Differential Evolution-A simple and efficient heuristic for global optimization over continuous Spaces,” Journal of Global Optimization, vol. 11, no.1, pp. 341–359, Mar. 1997.