A Differential Particle Swarm Optimization-based Support Vector Machine Classifier for Fault Diagnosis in Power Distribution Systems

Ming Yuan CHO, Thi Thom HOANG
Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan
thomht@ntu.edu.vn

***Abstract*—This paper proposes a new differential particle swarm optimization (DPSO) method for obtaining optimum support vector machine (SVM) parameters used for electrical fault diagnosis in radial distribution systems. Further, a multiple-stage DPSO-SVM classifier is developed to enhance classification accuracy in the fault diagnosis. Also, time-domain reflectometry (TDR) method with pseudo-random binary sequence (PRBS) excitation is utilized for generating the dataset required for validating this proposed approach. According to the characteristic of echo responses found in different types of faults, 12 features are extracted as input vectors for purposes of classification. The proposed fault diagnosis approach is tested on a typical radial distribution system to classify ten types of short-circuit faults accurately. Further, to demonstrate the superiority of the proposed DPSO algorithm, comparative studies of fault diagnosis are performed using SVM having parameters selected using cross-validation, GA and PSO. The overall classification accuracy obtained for fault diagnosis is 98.5%, which shows the effectiveness of the proposed approach.**

***Index Terms*—fault diagnosis, particle swarm optimization, power distribution lines, reflectometry, support vector machines.**

I. Introduction

In power systems, distribution networks deliver electrical energy from power-generating stations through transmission networks to consumers. Electrical faults are one of the most common undesirable phenomena which may interrupt the energy supply. Once an electrical fault occurs in any distribution systems, immediate fault classification plays an important role in post-fault analysis and power supply restoration. The accuracy of the fault type information not only assists the fault diagnosis system to locate the electrical faults promptly but also to ensure power quality as well as reliability of the system [1].

A variety of approaches have been developed to build an effective fault classifier in electrical distribution networks. These studies can be divided into three separate categories, as follows: (1) impedance based method [2-3], (2) travelling wave based method [4-5], (3) and artificial intelligence based method [6-7]. Among them, time-domain reflectometry (TDR) is one of the most popular methods for finding faults in distribution networks [8-9]. However, it is not a perfect fault classification method because of the complex characteristics of distribution systems, such as multi-branch topology, unbalanced operation and a widely varying range of loads [10-11]. Therefore, it requires other supporting pattern recognition algorithms to get reliable results. With the capacity of strong robust and nonlinear mapping, artificial neuron network (ANN) has been widely applied to solve the fault classification problem [12]. However, the shortcomings of over fitting and sinking into the local optimal are the major drawbacks of ANN. Compare to ANN, SVM has emerged as a powerful tool for fault classification because of the main advantage of high generalization and global optimization capability [13-15].

The performance of any SVM classifier is susceptible to the regularization parameter C and kernel function parameter such as gamma γ for the radial basis function (RBF) [16]. Noted that the error penalty parameter C controls the trade-off cost between the complexity of model and training error. Hence, set small or excessive values of C will reduce the generalization ability of SVM. A SVM classifier can achieve the best generalization capability with the best C value. Also, the RBF kernel parameter γ represents the distribution of training sample data, so it determines both the generalization capability and the accuracy of classification. Thus, the selection of SVM parameters plays an important role in improving classification accuracy as well as training speed. In [17], the grid search method (GSM) has been proposed to find the optimum parameters by attempting different values and selecting the values possessing the least testing error. However, this method is both time-consuming and unable to find the best parameters. To overcome these issues, various optimization approaches have been proposed in selecting the optimum parameters, including genetic algorithm (GA) [18-19] a combination of the cross-entropy (CE) method and a hill climbing type approach [20], an adaptive charged system search (ACSS) algorithm [21], a stochastic variable neighbourhood algorithm [22], artificial bee colony (ABC) algorithm [23]. Recently, particle swarm optimization (PSO)and differential evolution have also been performed as new methods with better performance in parameter optimization [24]. Although these optimization algorithm have resulted in better accuracy than non-optimizing methods, but they often get trapped in local optima.

To escape from the local optima, a novel differential particle swarm optimization (DPSO) algorithm is proposed in this paper to obtain a higher quality solution in optimization problems. Further, the DPSO-based SVM technique is capable of selecting the most optimum parameters in order to increase the fault classification

Digital Object Identifier 10.4316/AECE.2017.03007

1582-7445 © 2017 AECE

accuracy in power distribution systems. Besides, the result obtained by DPSO-SVM classifier is compared to that obtained by training SVM with cross-validation, GA and PSO which proves the superiority of the proposed DPSO algorithm. For generating simulated fault data, the TDR method with pseudo-random binary sequence (PRBS) excitation is utilized.

The rest of the paper is organized as follows. Section 2 shows basic theory of the proposed method, including TDR, SVM and the proposed differential particle swarm optimization algorithm. The DPSO-based SVM classifier and the fault diagnosis approach is developed in Section 3. Experimental results and discussion are given in Section 4. Finally, Section 5 presents the conclusion of the work.

II. Theory Of The Proposed Approach

The proposed method can be used to classify multiple fault types on multi-branch distribution networks. First, the TDR responses with different fault types are recorded by using Simulink software and MATLAB Toolbox. Next, the TDR curves along with the cross-correlation (CCR) function between the reflected wave and the incident wave can be used to train the SVM. Once being correctly trained, the SVM can classify faults from the measured TDR trace. Finally, a novel DPSO algorithm is developed to improve the ability of Support Vector Machines in classifying the fault types found in the distribution network.

A. Time-domain reflectometry

Time-domain reflectometry (TDR) is one of the most common methods used for fault classification and location. It uses a single pulse injection into a line or a cable and records echo responses which are caused by any impedance mismatches, including an electrical fault, tee joint or line terminal. Therefore, these obtained TDR curves are useful to identify the nature of any electrical fault.

Let us assume an enclosed coaxial distribution line can be modeled by an equivalent circuit, as shown in Figure 1.

The model characteristic impedance Z_0 and the propagation coefficient γ for the equivalent circuit in Figure 1 are given by:

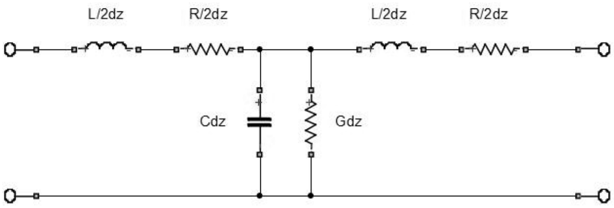

Figure 1. Equivalent model of a power distribution line

$$Z_C = \sqrt{\frac{(R + j\omega L)}{(G + j\omega C)}} \quad (1)$$

$$\gamma = \sqrt{(R + j\omega L)(G + j\omega C)} = \alpha + j\beta \quad (2)$$

$$\alpha = \sqrt{\frac{1}{2}\left[\sqrt{(R^2 + \mathrm{w}^2 L^2)(G^2 + w^2 C^2)} + (RG - w^2 LC)\right]} \quad (3)$$

$$\beta = \sqrt{\frac{1}{2}\left[\sqrt{(R^2 + \mathrm{w}^2 L^2)(G^2 + w^2 C^2)} - (RG - w^2 LC)\right]} \quad (4)$$

where α, β are the attenuation coefficient and the phase change coefficient, respectively.

It can be clearly seen from eqn. (2) that the TDR methods using a single pulse echo for fault diagnosis are inherently imprecise because of stimulus attenuation with fault distance and phase change distortion with frequency. In addition, the pulse width is one of the factors that affect the success rate of the reflectometry method. In [25], an improved TDR method, using incident pseudo-random binary sequence (PRBS) excitation can solve these problems by using cross-correlation (CCR) function between the reflected wave and incident wave given by eqn. (5) for fault diagnosis on transmission lines.

$$C_{xy}(k) = \frac{1}{L}\sum_{i=1}^{L} x(i)y(i+k) \quad (5)$$

where C_{xy} is the cross-correlation (CCR) function between the reflected wave y_i and the incident wave x_i.

For the distribution systems, it is not easy to extract fault information on the branched network from many reflections in the reflectometry trace recorded. In this study, a multi-layer SVM classifier is proposed as a supporting technique for TDR method to identify the fault types in multi-branch distribution networks. The reflected responses and the CCRs between the reflected wave and the incident wave are used as input feature vectors for the training phase.

B. Support Vector Machine

A support vector machine (SVM) was first mentioned by Vapnik in 1995, and it has become one of the most optimal techniques for data classification. It has a solid theoretical foundation based on a combination between the structural risk minimization principle and statistical machine learning theory (SLR). The main advantages of SVM are the global optimization and high generalization ability it possesses with a limited number of samples. Further, it overcomes over-fitting problems and provides sparse solutions in comparison to existing methods such as artificial neuron network (ANN) and refined genetic algorithm (RGA) in fault classification.

For classification problem solutions, SVM includes training and testing data that are comprised of many samples. In the training phase, each sample will consist of two attributes, called the feature and the label. The goal of the SVM classifier is to create a model which can accurately predict the class label of unknown data.

Let us assume that we have a set of training data, (x_i, y_i) , i=1,2…m, where $x_i \in R_n$ are feature vectors and $y_i \in (-1,+1)$ are label vectors. A binary classification problem can be posed as an optimization problem in the following way;

$$\text{Minimize } \frac{1}{2}\|\mathrm{w}\|_2^2 + C\sum_{i=1}^{m}\xi_i \quad (6)$$

Subjected to

$$y_i\left(\mathrm{w} \times x_i\right) + b \ge 1 - \xi_i, \xi_i \ge 0, i = 1, ..., m \quad (7)$$

where C is the regularization parameter or the penalty parameter; ξ_i is the penalizing relaxation variables. Eqn. (7) can be elaborated as:

$$\mathrm{w} \times \phi(x_i) + b \ge +1 \text{ if } y_i = +1 \quad (8)$$

$$\mathrm{w} \times \phi(x_i) + b \ge -1 \text{ if } y_i = -1 \quad (9)$$

It is to be noted that the nonlinear classifier may be denoted in the input space as:

$$f(x) = sign(\sum_{i=1}^{m} \alpha_i^* \times y_i \times (x_i, y_i) + b^*) \quad (10)$$

where f(x) is the decision function; m is the number of support vector, α_i the Lagrangian multipliers; b^* is the bias, and $K(x_i,y_i)$ is the kernel function.

From eqn. (10), it can be concluded that SVM is decided by training patterns and kernel function. Therefore, the selection of an appropriate kernel function is important to SVM. In this paper, the following radial basis function (RBF) is used as the kernel function;

$$K\left(x, y\right) = \exp\left(-\gamma \left\|x - y\right\|^2\right) \quad (11)$$

where γ is the kernel function parameter.

From eqns. (7), (10) and (11), it is observed that the performance of SVM is dependent on regularization parameter C and kernel function parameter γ. In order to get optimal classification performance, these two SVM parameters must be selected with due diligence. In this work, DPSO-based technique is applied in order to optimize these two parameters.

C. *Proposed Differential Particle Swarm Optimization*

Differential particle swarm optimization (DPSO) is a modified version of classical particle swarm optimization (PSO). The classical PSO suffers from getting trapped into local minima. To overcome this issue a new modification in the classical PSO is proposed in order to obtain a higher quality solution for addressing fault diagnosis problems given as part of this work. The concepts of the classical PSO and the proposed DPSO are discussed below in this section.

1) Particle Swarm Optimization

Particle swarm optimization (PSO) is inspired by social and cooperative behavior displayed by various species to fill their needs in multi-dimensional search space. The algorithm is guided by personal experience (Pbest), overall experience (Gbest) and the present movement of the particles used to decide their next positions in the search space. Further, the experiences are accelerated by the two factors c_1 and c_2, and by two random numbers r_1 and r_2 generated between [0, 1], whereas the present movement is multiplied by an inertia factor w. Mathematically, updated positions of each particle in the search space can be expressed using the two equation discussed below.

The initial population (swarm) of size N and dimension D is denoted as $\mathbf{X} = [\mathbf{X_1}, \mathbf{X_2}, ..., \mathbf{X_N}]^T$, where 'T' denotes the transpose operator. Each individual (particle) $\mathbf{X_p}$ (p = 1, 2, ..., N) is given as $\mathbf{X_p} = [X_{p,1}, X_{p,2},...,X_{p,D}]$. Also, the initial velocity of the population is denoted as $\mathbf{V} = [\mathbf{V_1},\mathbf{V_2},...,\mathbf{V_N}]^T$. Thus, the velocity of each particle $\mathbf{X_p}$ (p = 1, 2, ..., N) is given as $\mathbf{V_p} = [V_{p,1},V_{p,2}, ..., V_{p,D}]$. The index p varies from 1 to N whereas the index q varies from 1 to D.

$$V_{p,q}^{k+1} = w \times V_{p,q}^{k} + c_1 r_1 (Pbest_{p,q}^{k} - X_{p,q}^{k}) + c_2 r_2 (Gbest_q^k - X_{p,q}^{k}) \quad (12)$$

$$X_{p,q}^{k+1} = X_{p,q}^{k} + V_{p,q}^{k+1} \quad (13)$$

In eqn. (12), $Pbest^k_{p,q}$ represents personal best q^{th} component of p^{th} individual, whereas $Gbest^k_q$ represents q^{th} component of the best individual of population up to iteration k. Figure 2 shows the search mechanism of PSO in the multi-dimensional search space.

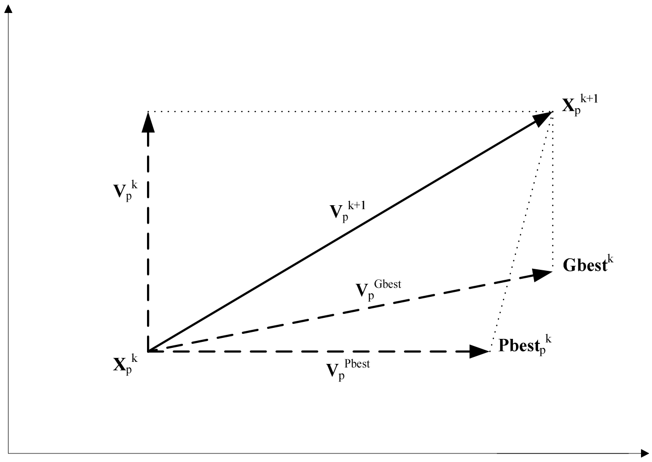

Figure 2. The classical PSO search mechanism of p^{th} particle at k^{th} iteration

The initial Pbest of each particle is their initial position, whereas the initial Gbest is the initial best particle position among randomly initialized population. The Pbest and Gbest of the swarm are updated as follows;

At iteration k

$$\text{If } f(\mathbf{X}_p^{k+1}) < f(\mathbf{Pbest}_p^k) \text{ then } \mathbf{Pbest}_p^{k+1} = \mathbf{X}_p^{k+1}$$
$$\text{else } \mathbf{Pbest}_p^{k+1} = \mathbf{Pbest}_p^k \quad (14)$$

$$\text{If } f(\mathbf{X}_p^{k+1}) < f(\mathbf{Gbest}^k) \text{ then } \mathbf{Gbest}^k = \mathbf{X}_p^{k+1}$$
$$\text{else } \mathbf{Gbest}^{k+1} = \mathbf{Gbest}^k \quad (15)$$

where f(·) is the objective function of minimization.

Repeat updating procedure until a stop condition is reached, such as a pre-specified number of iteration is met. Once terminated, the $\mathbf{Gbest}^k$ and $f(\mathbf{Gbest}^k)$ are to be reported as the solution of PSO technique. More details about the basic conceptualization of PSO can be found in [26-30].

2) Differential Particle Swarm Optimization

The proposed differential particle swarm optimization (DPSO) considers an additional feature in the classical PSO. The additional feature is the opinion of one of the particles selected randomly from the swarm. The randomly-scaled difference of the particle and its opinion-giver particle is included in the velocity equation of the particle necessary to escape from local minima. Mathematically, the concepts of DPSO can be expressed as follows.

$$V_{p,q}^{k+1} = w \times V_{p,q}^{k} + c_1 r_1 (Pbest_{p,q}^{k} - X_{p,q}^{k}) + c_2 r_2 (Gbest_q^k - X_{p,q}^{k}) + c_3 r_3 (X_{l,q}^{k} - X_{p,q}^{k}) \quad (16)$$

$$X_{p,q}^{k+1} = X_{p,q}^{k} + V_{p,q}^{k+1} \quad (17)$$

In eqn. (16), c_3 is the scaling factor and r_3 is a randomly-generated random number, whereas l represents the expert particle corresponding to target particle p (l varies from 1 to N but $l \neq p$). Note that the introduction of three random numbers (r_1, r_2, and r_3) is to mimic the unpredictable behavior of nature swarms. Generally, three random numbers represent three separate calls, and most implementations use these random numbers uniformly distributed between 0 and 1. Thus, the pulling forces of pbest and gbest would vary between 0 and 1 with the uniform probability in the optimization procedure. Figure 3 shows the search mechanism of the proposed DPSO in a multidimensional search space.

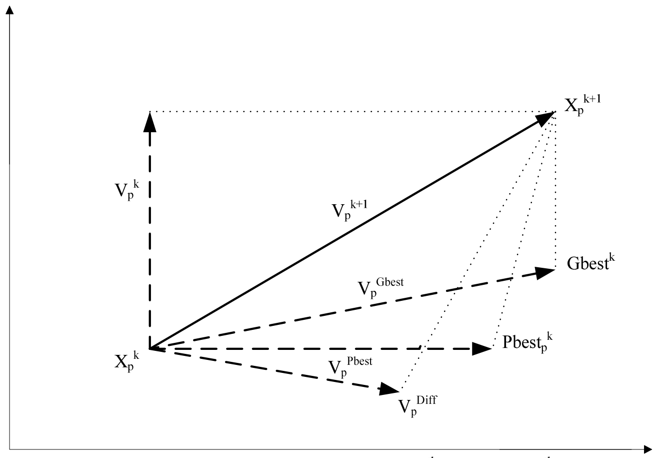

Figure 3. Proposed DPSO search mechanism of pth particle at kth iteration in a multi-dimensional search space

The proposed DPSO algorithm can be expressed using the following steps:

1) Set w, c_1, c_2 and c_3 parameters
2) Initialize positions X and velocities V of each particle of population
3) Evaluate fitness of each particle $F_p^k = f(X_p^k)$, $\forall p$ and find the best particle index b
4) Select $Pbest_p^k = X_p^k$, $\forall p$ and $Gbest^k = X_b^k$
5) Set iteration count $k = 1$
6) Update velocity and position of each particle using eqns. (16) and (17)
7) Evaluate updated fitness of each particle $F_p^{k+1} = f(\boldsymbol{X}_p^{k+1})$, $\forall p$ and find the best particle index $b1$
8) Update Pbest of each particle $\forall p$
   If $F_p^{k+1} < F_p^k$ then $Pbest_p^{k+1} = X_p^{k+1}$, else $Pbest_p^{k+1} = Pbest_p^k$,
9) Update Gbest of population
   If $F_{b1}^{k+1} < F_b^k$ then $Gbest^{k+1} < Pbest_{b1}^{k+1}$ and set $b = b1$ else $Gbest^{k+1} < Gbest^k$
10) If $k < Maxite$ then $k = k+1$ and go to step 6 else go to step 11
11) Optimum solution obtained and so print the results $Gbest^k$

A detailed flowchart of proposed DPSO considering the above steps is shown in Figure 4.

III. Developed DPSO Based SVM For Fault Diagnosis

Since TDR methods are inherently imprecise, they should require other supporting techniques to achieve reliable results. In this work, a DPSO encoding-based SVM classifier is developed to improve the performance of the reflectometry method used to identify the fault types in the radial distribution system. The overall structure of SVM classifier is shown in Figure 5 in which DPSO is performed to optimize these two parameters (C and γ) of the SVM classifier. For this, the data acquisition for data preprocessing is mentioned first.

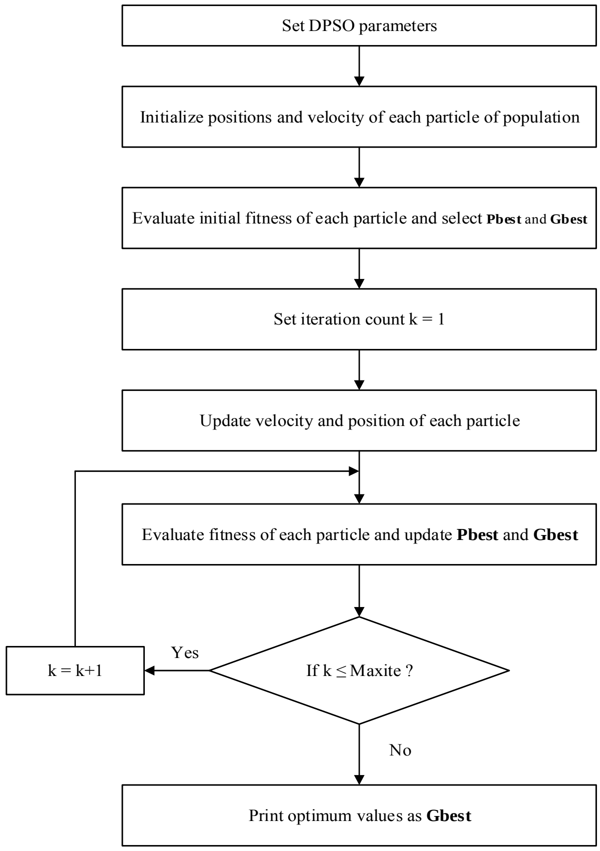

Figure 4. Flowchart of the proposed DPSO

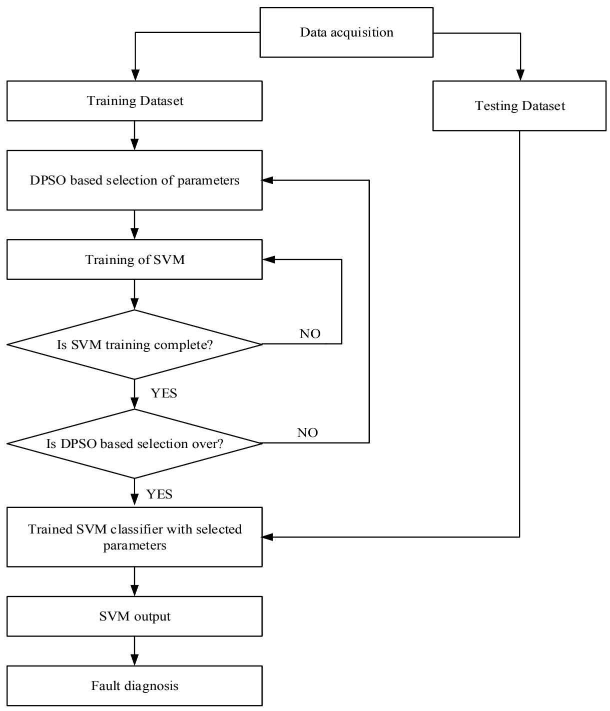

Figure 5. Flowchart of the proposed approach

A. Data acquisition

To obtain a suitable dataset for classification process, a PRBS disturbance is injected into the distribution feeder under test. Once a fault occurs in the distribution feeder, it causes to produce a reflected signal that travels between the fault location and the substation. These reflected responses are recorded and then they are cross-correlated with the incident impulse by eqn. (4) in order to reduce the impact of noise as well as to surmount amplitude attenuation.

It is worth noting that, for each of the fault type specified, the magnitudes of the feedback waves are different at the shortage time; as a result, the peaks of the CCR are not found to be the same.

Based on the above analysis, the reflected voltage and current magnitudes along with the peaks of CCR between reflected and incident waves are chosen as the input features of the SVM classifier, and the fault types are chosen as the output features. Therefore, the total number of derived features is 12 and comprises a feature vector $V=[v_1\ v_2 \ldots v_{12}]^T$. The corresponding meanings of the feature vector are expressed as follows:

a) v_1-v_6 are the reflected voltage values (v_a, v_b, v_c) and reflected current values (i_a, i_b, i_c).

b) v_7-v_{12} are the peaks of CCR functions between reflected and incident waves of voltage and current.

B. SVM parameter optimization and the proposed fault diagnosis approach

The performance of SVM is susceptible to the kernel function parameter γ and the regularization parameter C, so these parameters must be carefully selected to increase the classification accuracy.

In this paper, DPSO technique is used to optimize the parameters of the SVM classifier. Performance is measured according to the classification accuracy on unseen testing data. In the learning stage, the DPSO-based encoding SVM model is trained based on structural risk minimization to minimize the training error. While training error improvement occurs, penalty parameter C and kernel function parameter g are regulated by means of DPSO. The regulated parameters with minimal error are reported as the most suitable parameters. As a result, the optimal parameters (C and γ) are to be obtained. Once the optimized parameters of the SVM are obtained, then it is used for the retraining of the SVM model. After the training phase, the SVM classifier is ready to identify new samples in the testing phase. The testing set is also chosen by means of the above parameter selection from the original dataset obtained by the TDR responses.

As such, two different DPSO-based SVM fault diagnosis approaches are proposed. The two approaches are named as follows;

a) Single-stage DPSO-SVM

b) Multiple-stage DPSO-SVM

In single-stage DPSO-SVM, the fault dataset are given to the optimized SVM classifier which classifies the fault types directly into any of the ten types of faults discussed earlier. As the fault types can be directly identified using single SVM structure, it is termed as single-stage DPSO-SVM. The overall structure of single-stage DPSO-SVM is shown in Figure 6.

In multiple-stage DPSO-SVM, the fault dataset are given to the optimized SVM classifier which initially classifies the fault types into four possible fault groups (LG, LL, LLG, LLL) in the first stage and then each fault group is again classified into any fault type of that group in the second stage. Thus, in the first stage, fault groups are identified and then each group is classified into their corresponding fault types. So, in the first stage, four groups are identified by SVM1 whereas in the second stage the first three groups are classified into their corresponding fault types by use of SVM2, SVM3 and SVM4 respectively. The fourth group is to be left alone as it is already a fault type (ABC fault). As in this case, the fault type can only be identified using a structure in two stages, so it is termed as multiple-stage DPSO-SVM. The overall structure of multiple-stage DPSO-SVM classifier is shown in Figure 7.

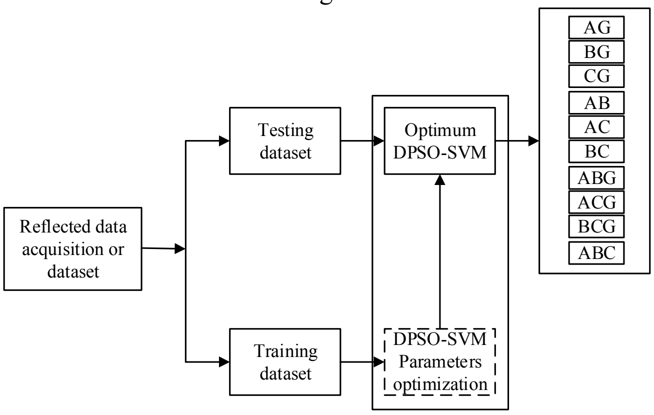

Figure 6. The overall structure of the proposed single-stage DPSO-SVM classifier for fault diagnosis

IV. TEST RESULTS AND DISCUSSION

The proposed approach has been tested by means of classifying faults on a typical two-branch radial distribution system, as shown in Figure 8.

The sample system consists of a radial primary feeder, two distribution transformers and several loads. Two distribution transformers in the sample system are used to reduce the voltage on the distribution line to the level of customers that are distributed along a feeder. Their ratings are 500 kVA, 0.22kV and j1.89%, and their phases are connected as delta and grounded wye connections, respectively. These distribution transformers are operated in a full-load condition with 0.8 lagging power factor. The main feeder and laterals are constructed by means of overhead lines whose impedances are shown in Table I.

TABLE I. PARAMETERS OF DISTRIBUTION LINE IN THE SAMPLE SYSTEM

Items	Positive- and zero-sequence resistances (Ohms/km)	Positive- and zero-sequence inductances (H/km)	Positive- and zero-sequence capacitances (F/km)
Feeder	[0.247 0.309]	[1.321e-3 2.473e-3]	[5.398e-9 4.656e-9]
Lateral	[0.471 0.561]	[1.538e-3 2.692e-3]	[9.882e-9 7.106e-9]

A. Training and testing samples

The proposed algorithm is implemented on a dataset obtained by TDR method using PRBS exciation (as mentioned in subsection II.A). In the work, a 127 bit PRBS disturbance with frequency f =1MHz is injected into the distribution feeder under test. The reflected responses are caused by any electrical fault on feeder and laterals and then they are cross-correlated with the incident impulse. Thus, each sample has 12 features extracted from the reflected signals and CCRs between reflected and incident waves. For this, ten types of faults created at distances of 10, 20, 30, . . . , 100% of the first lateral length. The fault resistance values are varied over the values 1, 5, 20, 30 and 60Ω during the simulation. As such, the samples are generated for 10 types of faults on the first lateral over 100 locations with varying 5 impedance values. For each type of fault, the number of samples generated is 100 × 5 = 500 patterns.

Training and testing sets are randomly selected from this dataset, where 4000 and 1000 are used for training and testing respectively. This dataset can be found at http://thom-project.webnode.vn/services.

Table II only gives a few of the dataset created by simulation of all ten types of short circuit faults on the first lateral, located at distances of 3km and 4km from the substation for desire of brevity.

The obtained dataset is inputted into a multi-class SVM for purpose of fault classification (as expressed in subsection II.B). Then, the proposed DPSO algorithm is applied to optimize the SVM parameters in order to increase the classification accuracy.

For this, parameter selection of DPSO algorithm plays an important role in achieving the best performance of the algorithm.

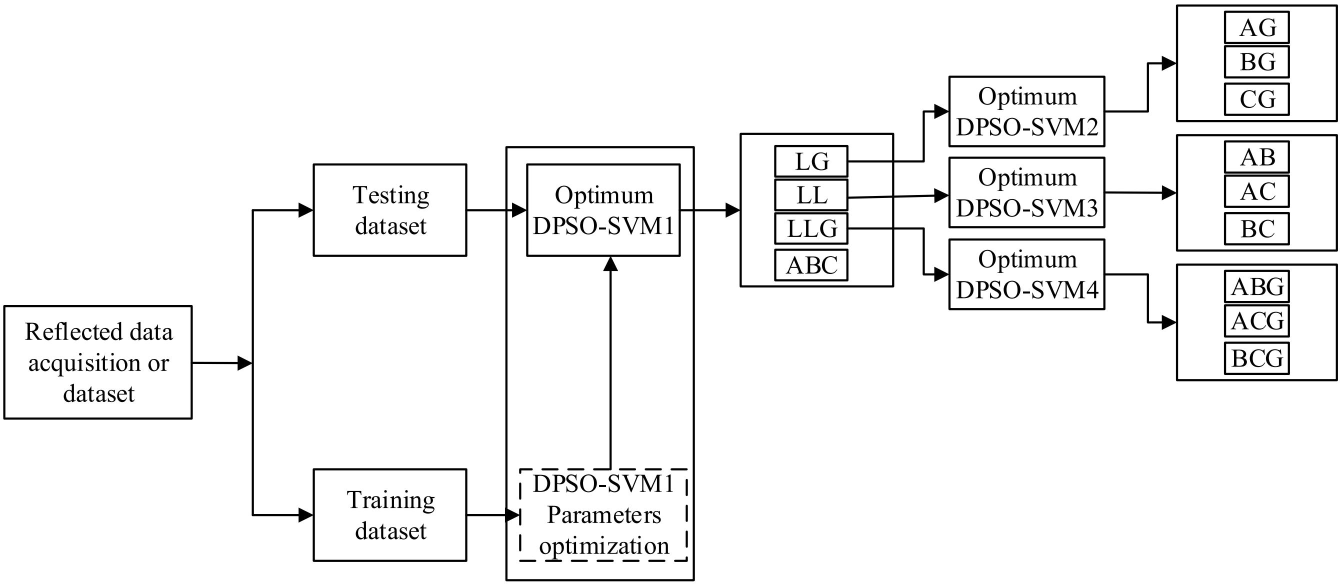

Figure 7. The overall structure of the proposed multiple-stage DPSO-SVM classifier for fault diagnosis

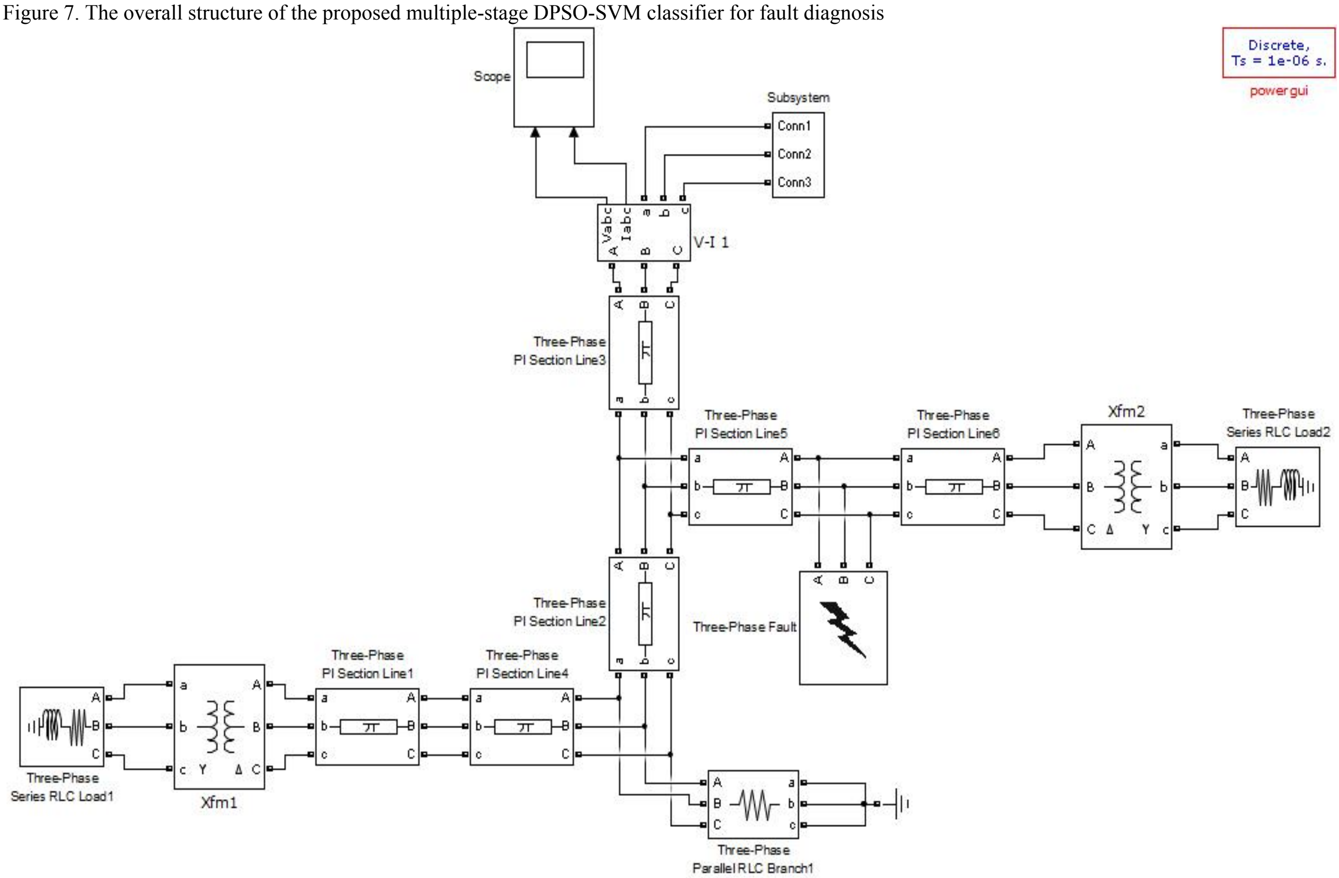

Figure 8. A two-branched distribution line diagram of the sample system

B. Parameter selection for the proposed DPSO-SVM classifier

In the following procedure is adopted to obtain the best parameters of the proposed DPSO-based algorithm for optimizing SVM parameters. For different population sizes (pop =10, 20 and 30).

TABLE II. DATASET OF TEN TYPES OF FAULT LOCATED AT DISTANCES OF 3KM AND 4KM FROM THE SUBSTATION

	v_a	v_b	v_c	i_a	i_b	i_c	cc-v_a	cc-v_b	cc-v_c	cc-i_a	cc-i_b	cc-i_c
AG	1.9197	-0.3071	0.1245	4.9815	-0.7968	0.3232	0.5941	-0.0950	0.0385	0.0502	-0.0080	0.0033
	0.6990	-0.1118	0.0453	1.5998	-0.2559	0.1038	3.5687	-0.5708	0.2315	3.0765	-0.4921	0.1996
BG	1.4521	0.7277	0.5122	3.7681	1.8884	1.3290	0.4494	0.2252	0.1585	0.0380	0.0190	0.0134
	0.5287	0.2650	0.1865	1.2101	0.6064	0.4268	2.6995	1.3528	0.9521	2.3271	1.1662	0.8208
CG	0.4648	4.5783	3.1718	0.0857	0.8445	0.5851	0.0275	0.2711	0.1878	0.0237	0.2331	0.1615
	0.0880	0.8668	0.6005	0.2284	2.2492	1.5582	0.0272	0.2683	0.1858	0.0023	0.0227	0.0157
BCG	-8.2016	9.6684	16.2648	-2.5267	2.9785	5.0107	-0.1137	0.1340	0.2254	-0.1137	0.1340	0.2254
	-4.1309	4.8697	8.1921	-0.7620	0.8983	1.5112	-0.2446	0.2884	0.4852	-0.2104	0.2480	0.4172
ACG	-1.2835	2.5576	4.8025	-0.9796	1.9519	3.6650	-1.6907	3.3688	6.3257	-1.4240	2.8375	5.3279
	-1.4241	1.7834	3.8278	-1.0868	1.3610	2.9212	-1.8757	2.3491	5.0419	-1.5799	1.9786	4.2466
ABG	-1.1327	0.0679	2.6912	-2.9393	0.1763	6.9832	-0.3506	0.0210	0.8329	-0.0296	0.0018	0.0704
	-2.0970	0.1258	4.9821	-1.6003	0.0960	3.8021	-2.7621	0.1657	6.5623	-2.3265	0.1395	5.5272
AB	-7.4589	-4.8688	17.7206	-1.3759	-0.8981	3.2688	-0.4417	-0.2883	1.0495	-0.3798	-0.2479	0.9024
	-1.4121	-0.9218	3.3549	-3.6643	-2.3918	8.7055	-0.4370	-0.2853	1.0383	-0.0369	-0.0241	0.0877
AC	-1.0143	-1.2113	7.8915	-0.7741	-0.9244	6.0225	-1.3360	-1.5955	10.3945	-1.1253	-1.3439	8.7550
	-1.5121	-7.9329	40.5259	-0.4658	-2.4439	12.4847	-0.0210	-0.1099	0.5616	-0.0210	-0.1099	0.5616
BC	2.0444	-4.2356	23.5915	0.3771	-0.7813	4.3518	0.1211	-0.2508	1.3972	0.1041	-0.2157	1.2013
	0.1409	-0.2920	1.6262	0.3225	-0.6682	3.7220	0.7195	-1.4907	8.3028	0.6203	-1.2851	7.1576

Legends:
AG, BG and CG are single phase to ground faults;
BCG, ACG and ABG are double line to ground faults;
AB, AC and BC are line to line faults; ABC is three phase fault.
v_a, v_b, v_c, i_a, i_b and i_c are magnitudes of the reflected voltage and current, respectively.
cc-v_a, cc-v_b, cc-v_c, cc-i_a, cc-i_b and cc-i_c are CCR between reflected signal and incident signal, respectively.

TABLE III. TOP FIVE LEAST VALUES OF MSE FOR DIFFERENT VALUES OF VARIOUS PARAMETERS OF THE PROPOSED DPSO-SVM

Algorithm	**Top five parameters**	**Pop = 10**				**Pop = 20**				**Pop = 30**			
		c1	c2	c3	MSE	c1	c2	c3	MSE	c1	c2	c3	MSE
DPSO	1	1.5	1.5	0.04	0.035	1.5	1	0.03	0.035	0.5	1.5	0.01	0.035
	2	1.5	2	0.01	0.035	1.5	1.5	0.01	0.035	0.5	1.5	0.05	0.035
	3	1.5	2.5	0.02	0.035	1.5	1.5	0.04	0.035	0.5	2	0.04	0.035
	4	2	1.5	0.01	0.035	1.5	2	0.02	0.035	1	1.5	0.05	0.035
	5	2	2	0.04	0.035	1.5	2.5	0.03	0.035	1	2	0.03	0.035

a. Inertia weight is taken in between 0.1 to 0.5 (randomly at each iteration);
b. Acceleration factors (c1 and c2) are varied from 0.5 to 2.5 (in step of 0.5)
c. Scaling factor c3 is varied from 0.01 to 0.05 (in step of 0.1)
d. Maximum iteration is set to 1000.

A summary of the parameter selection results for DPSO is given in Table III which five sets of important parameters and the corresponding values of mean square error (MSE) are included in the table for desire of brevity.

From Table III, it is clear that the best set of parameters for the DPSO algorithm are c1 = 1.5, c2 = 1.5, c3 = 0.04 and population size pop = 10. Although, the values of MSE for the pop = 20 and pop = 30 are the same with different parametric values than that of pop = 10. However, the parameters corresponding to pop = 10 are selected since they take less execution time because of the smaller population size. It is to be noted that similar testing has been performed in selecting the acceleration coefficients c1 and c2 for PSO and crossover rate (CR) and the mutation fraction (MF) for GA. The best values of c1 and c2 for PSO have been obtained as 1.5 and 2.5, whereas the best values of CR and MF for GA have been obtained as 0.8 and 0.01, respectively.

Furthermore, the 5-fold cross-validation approach has been considered to select optimum SVM parameters for SVM classifier.

C. Experimental results

In this subsection, results of the two types of DPSO-SVM fault diagnosis approaches discussed in Section 3 are given. Further, to show the effectiveness of the proposed DPSO-trained SVM, the results obtained by cross-validation, PSO and GA trained SVM have been compared.

For the single-stage model, the results of the fault classification made for the sample system using the SVM classifier whose parameters are optimized by cross-validation, PSO, GA and the proposed DPSO techniques are given in Table IV.

TABLE IV. RESULTS OF SINGLE-STAGE SVM CLASSIFIER USING VARIOUS OPTIMIZATION TECHNIQUES

Classifier	**C**	**γ**	**Classification accuracy (%)**
Cross-validation-SVM	181.0193	1.1212	93.00
PSO-SVM	97.2221	8.2346	95.08
GA-SVM	3.4218	3.1067	96.42
DPSO-SVM	2.0817	4.2174	96.50

As can be seen from Table IV, the optimum values of C and γ of the SVM classifier are 2.0817 and 4.2174 with testing data accuracy of 96.50% by means of the proposed DPSO. Further, DPSO results in the highest accuracy claim over any other SVM parameter optimizers. The accuracy rates of Cross validation-SVM, PSO-SVM and GA-SVM are 93%, 95.08% and 96.42%, respectively.

The convergence characteristic of the proposed DPSO is shown in Figure 9. From this figure, it can be observed that

MSE beyond 35 iterations is non-decreasing and thus the optimized SVM parameters can be obtained much sooner than the total training time taken (159.31 sec).

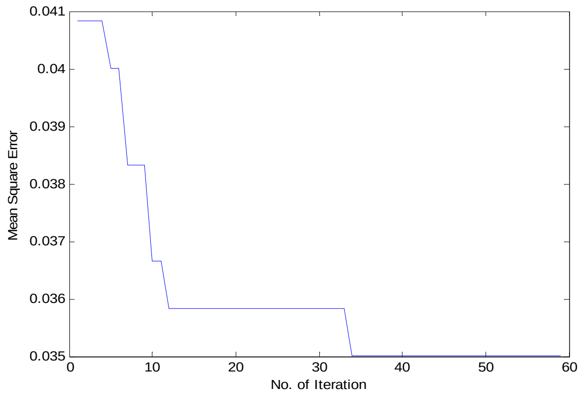

Figure 9. Convergence characteristic of the proposed DPSO in the multiple-stage SVM classifier

It is observed from Table IV that the proposed DPSO gives the highest classification accuracy for fault diagnosis. Hence, for multiple-stage model (Figure 6), only DPSO has been used to optimize parameters of SVMs. The optimized values of the SVM parameters of the proposed multiple-stage DPSO-SVM are given in Table V.

TABLE V. RESULTS OF MULTIPLE-STAGE SVM CLASSIFIER USING THE PROPOSED DPSO

Classifier	C	γ	Classification accuracy (%)
SVM1	1.0	5.6926	99.08
SVM2	42.41	0.0004	99.17
SVM3	77.97	0.0071	97.25
SVM4	42.61	0.0022	98.06
Overall accuracy			98.50

From Table V, it is observed that the classification accuracy of the proposed multiple-stage DPSO-SVM classifier in the first stage is 99.08%; whereas, in the second stage which possesses three SVMs, the accuracy rates are 99.17%, 97.25% and 98.06%, respectively. In multiple-stage DPSO-SVM classifiers, the overall classification accuracy is 98.50% which is significantly higher than what is obtainable with a single-stage DPSO-SVM classifier.

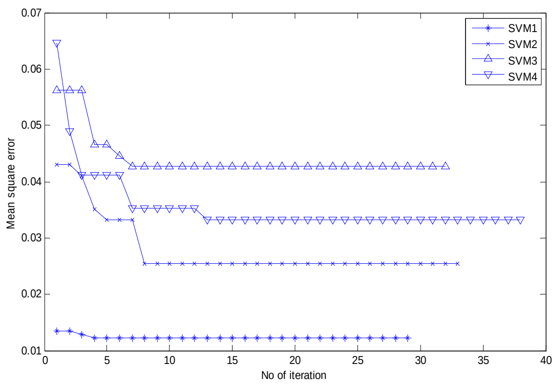

Figure 10. Convergence characteristic of the proposed DPSO in the multiple-stage SVM classifier

The convergence characteristic of the proposed multiple-stage DPSO-SVM classifier is shown in Figure 10. As can be seen from Figure 10, there are four different convergence characteristics of four SVM classifiers which have been trained independently. MSE beyond 15 iterations are non-decreasing in all these four curves whereas it takes about 35 iterations in single-stage DPSO-SVM classifier. Hence, the training time of the multiple-stage DPSO-SVM classifier is much faster than that of the single-stage DPSO-SVM classifier.

From these studies, it is clear that the proposed DPSO algorithm offers better SVM parameter optimization than with cross-validation, PSO and GA. Further, the proposed multiple-stage SVM classifier is better than any single-stage SVM classifier.

D. Effects of the results by varying training and testing datasets

In this subsection, the effects of different ration of diving training and testing dataset on the classification accuracy have been studied. Under this study, both the single-stage DPSO-SVM and the multiple-stage DPSO-SVM classifier have been utilized and four different training and testing dataset divisions have been considered for 5000 samples of dataset used in this paper. Under this, entire dataset have been divided into four different ratios of training and testing data samples. These four dataset division ratios from training to testing (training:testing) are as follows:

a) 60:40
b) 70:30
c) 80:20
d) 90:10.

The proposed DPSO-SVM classifiers have been applied to test the classification performance of these four dataset division.

Table VI and Table VII show the classification accuracy as well as the optimum SVM parameters obtained using the proposed single-stage DPSO-SVM and multiple-stage DPSO-SVM classifiers, respectively.

TABLE VI. RESULTS OF SINGLE-STAGE DPSO-SVM CLASSIFIER USING VARIOUS DATASET DIVISION PATTERNS

Dataset division (Training : testing)	C	γ	Classification accuracy (%)
60:40	100	1.3938	95.66
70:30	52.5271	1.1277	96.08
80:20	100	1.3468	96.05
90:10	47.4570	1.0864	96.42

From Table VI, it is observed that the performance of the single-stage DPSO-SVM classifier gives the best result in the training and testing dataset division of 90:10 ratio whereas from Table VII, it is observed that for the multiple-stage DPSO-SVM classifier, the performance is the best in the training and testing dataset division of 80:20 ratio. In other words, a higher training dataset gives better classification accuracy versus that with a lower dataset. Further, from Table VI and Table VII, it can be concluded that the multiple-stage DPSO-SVM classifier gives better classification accuracy than that obtained by the single-stage DPSO-SVM classifier.

The optimum convergence characteristics of the proposed single-stage DPSO-SVM and multiple-stage DPSO-SVM classifiers have been shown in Figure 11 and Figure 12, respectively. Subfigures of these two figures show the convergence characteristics of the proposed algorithms corresponding to different training and testing dataset divisions. From Figure 11 and Figure 12, it is observed that MSE is the lowest in the case of dataset division 90:10 whereas, the average MSE is the lowest in the case of dataset division of 80:20. Thus, the observations made from tabulated results are supported by the characteristics.

TABLE VII. RESULTS OF MULTIPLE-STAGE DPSO-SVM CLASSIFIER USING VARIOUS DATASET DIVISION PATTERNS

Dataset division (Training : testing)	SVM1		SVM2		SVM3		SVM4		Classification accuracy (%)
	C	γ	C	γ	C	γ	C	γ	
60:40	7.9849	5	47.2832	0.0057	4.6708	0.0763	100	0.0053	97.46
70:30	100	1.7848	62.0977	0.0015	92.7704	0.0091	1.0311	0.0032	97.89
80:20	16.4415	1.7493	21.4799	0.0034	53.1976	0.0116	52.9386	0.0583	98.42
90:10	56.4793	0.7384	4.8591	0	69.0436	0.0109	40.8192	0.0685	98.07

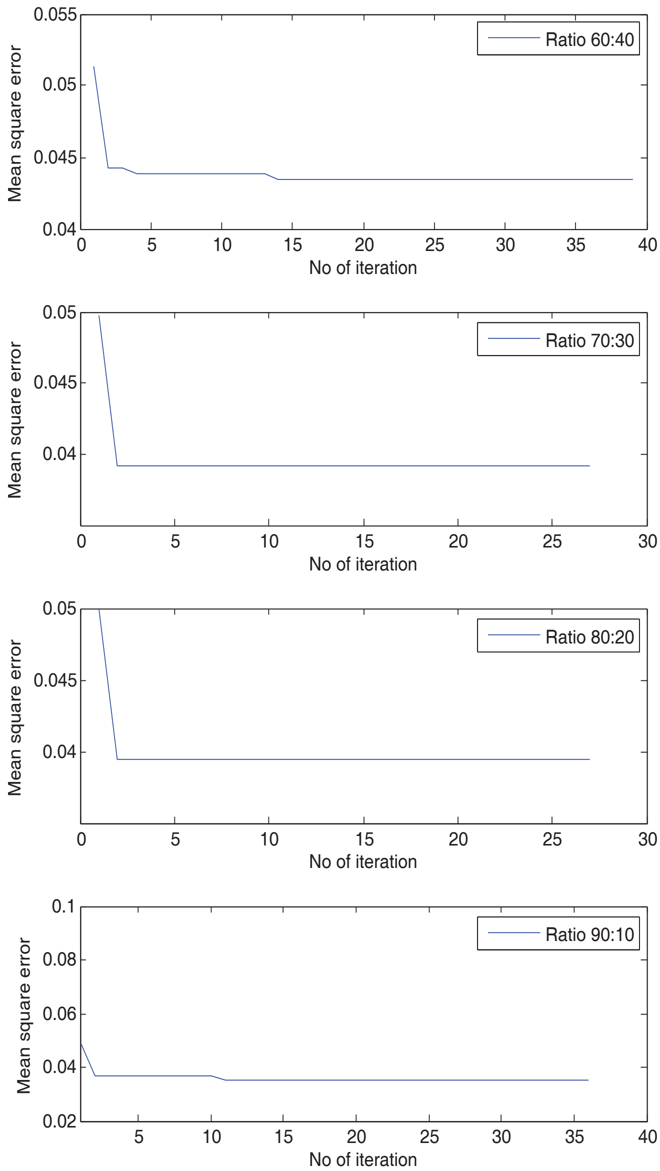

Figure 11. Convergence characteristic of the proposed single-stage DPSO-SVM classifier for different dataset division patterns

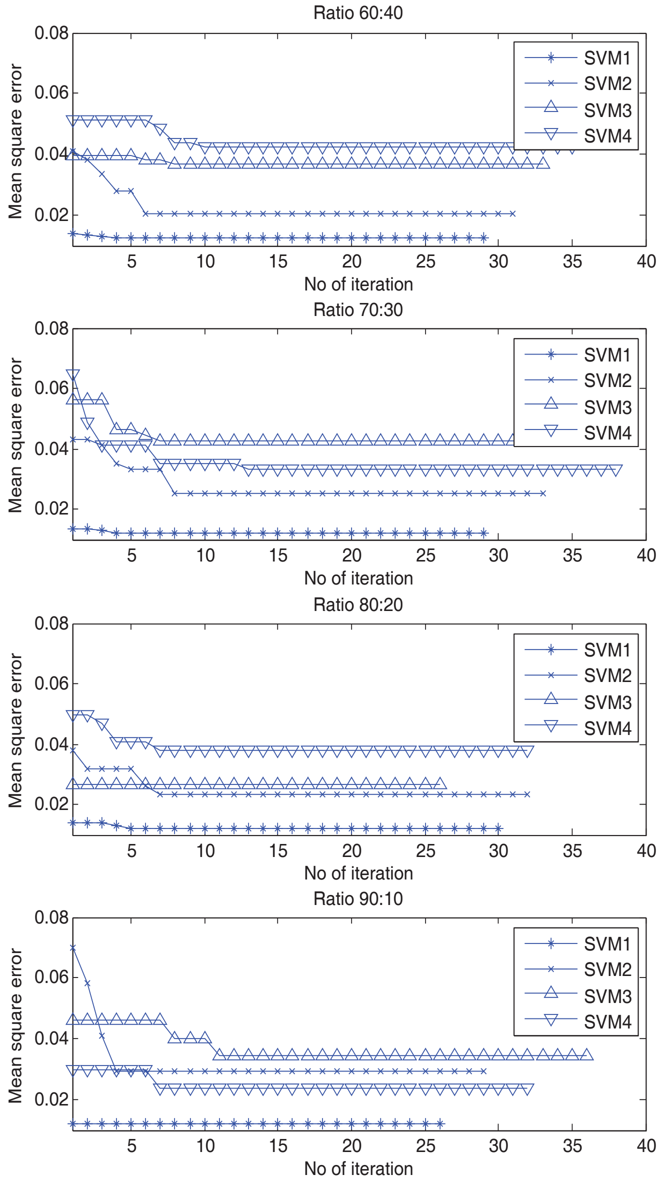

Figure 12. Convergence characteristic of the proposed multiple-stage DPSO-SVM classifier for different dataset division patterns

V. CONCLUSION

In this paper, a differential particle swarm optimization (DPSO) algorithm is proposed to improve the performance of SVM for the purpose of fault classification in the radial distribution network. The DPSO-based technique can optimize the parameters of SVM classifier in order to increase the classification accuracy. Further, a multiple-stage SVM classifier is introduced for purposes of better fault classification. Also, time-domain reflectometry (TDR) with pseudo-random binary sequence (PRBS) stimulus has been utilized for generating a reliable fault dataset. The proposed approach is tested successfully to identify ten types of electrical short-circuit fault in a typical radial distribution network. Then, the results have been compared with diagnosis results obtained from different methods, such as cross-validation, PSO and GA. The overall accuracy obtained in classifying fault types is 98.5%, which demonstrates the effectiveness of the proposed fault diagnosis approach.

REFERENCES

[1] J. G. M. S. Decanini, M. S. Tonelli-Neto, C. R. Minussi, "Robust fault diagnosis in power distribution systems based on fuzzy ARTMAP neural network-aided evidence theory," IET Gener. Transm. Distrib, vol. 6, pp. 1112-1120, Jul. 2012. doi:10.1049/iet-gtd.2012.0028

[2] E. C. Senger, G. Manassero, C. Goldemberg, E. L. Pellini, "Automated fault location system for primary distribution networks," IEEE Trans. Power Delivery, vol. 20, pp. 1332-1340, 2005. doi:10.1109/TPWRD.2004.834871

[3] Y. Liao, "Algorithms for power system fault location and line parameter estimation," 39th Southeastern Symposium on System Theory, Mercer University, Macon, GA, 2007. doi:10.1109/SSST.2007.352346

[4] F. H. Magnago, A. Abur, "Fault location using wavelets," IEEE Trans. Power Delivery, vol. 13, pp. 1475-1480, 1998. doi:10.1109/61.714808.

[5] A. Borgheti, S. Corsi, C. A. Nucci, M. Paolone, L. Pereto, R. Tinarelli, "On the use of continuous-wavelet transform for fault location in distribution power systems," Electrical Power and Energy Systems, vol. 28, pp. 608-617, 2006. doi:10.1016/j.ijepes.2006.03.001

[6] M. N. Pourahmadi, A. A. Safavi. "Path characteristic frequency-based fault locating in radial distribution systems using wavelets and neuron networks," IEEE Trans. Power Delivery, vol. 60, pp. 1654-1663, 2011. doi:10.1109/TPWRD.2010.2050218.

[7] M. Al. Shaher, M. M. Sabra, A. S. Saleh, "Fault location in multi-ring distribution networks using artificial neural network," Electric Power Systems Research, vol. 64, pp. 87-92, 2003. doi:10.1016/S0378-7796(02)00174-8

[8] J. P. Steiner, W. L. Weeks, H. W. Ng, "An automated fault locating system," IEEE Trans. on Power Delivery, vol. 7, pp. 967–978, 1992. doi:10.1109/TDC.1991.169511.

[9] S. Navaneethan, J. J. Soraghan, W. H. Siew, F. McPherson, P. F. Gale, "Automatic Fault Location for Underground Low Voltage Distribution Networks," IEEE Transactions on power delivery, Vol. 16, pp. 346-351, 2001. doi:10.1109/PESW.2002.985114

[10] J. Mora-Flórez, J. Cormane-Angarita, G. Carrillo-Caicedo, "K-means algorithm and mixture distributions for locating faults in power systems," Electric Power System Research, vol. 79, pp. 714-721, 2009. doi.org/10.1016/j.epsr.2008.10.011.

[11] H. Mokhlis, H. Mohamad, H. Li, Ab H. A. Bakar, "Voltage Sags Matching to Locate Faults for Underground Distribution Networks," Advances Electrical and Computer Engineering, vol. 11, No.2, pp. 43-48 2011. doi:10.4316/aece.2011.02007.

[12] J. J. Mathew, A. Francis, "HVDC Transmission Line Fault Location Using Wavelet Feeded Neural Network Bank," Science Technology & Engineering, vol. 2, pp. 1-6, 2013. doi:10.1109/ICSSSM.2015.7170211.

[13] X. Deng, R. Yuan, Z. Xiao, T. Li, K. L. L. Wanga, "Fault location in loop distribution network using SVM technology," Electrical Power and Energy Systems, vol. 65, pp. 254-261, 2015. doi.org/10.1016/j.ijepes.2014.10.010.

[14] X. Zhang, M. Zhang and D. Liu, "Reconstruction of faulty cable network using time domain reflectometry," Progress In Electromagnetics Research, vol. 136, pp. 457-478, 2013. doi:10.2528/PIER12121402.

[15] L. Ye, D. You, X. Yin, K. Wang, J. Wu, "An improved fault-location method for distribution system using wavelets and support vector regression," Electrical Power and Energy Systems, vol. 55, pp. 467-472, 2014. doi.org/10.1016/j.ijepes.2013.09.027.

[16] V. N. Vapnik, "The Nature of Statistical Learning Theory," Springer. Verlag, New York, pp. 1-14, 1995.

[17] C. Hsu, C. Chang, C. Lin, "A practical guide to support vector classification," Department of Computer Science, National Taiwan University, Tech. Report, pp. 1- 16, 2003. doi=10.1.1.224.4115.

[18] L. B. Jack and A. K. Nandi, "Fault detection using support vector machines and artificial neural networks, augmented by genetic algorithms," Mechanical Systems and Signal Processing, vol. 16, pp. 373-390, 2002. doi.org/10.1006/mssp.2001.1454.

[19] B. Samanta, K. R. Al-Balushi, and S. A. Al-Araimi, "Artificial neural networks and support vector machines with genetic algorithm for bearing fault detection," Engineering Applications of Artificial Intelligence, vol. 16, pp. 657-665, 2003. doi.org/10.1016/j.engappai.2003.09.006.

[20] Z. C. Johanyák, and O. Papp, "A hybrid algorithm for parameter tuning in fuzzy model identification," Acta Polytechnica Hungarica, vol. 9, pp. 153-165, 2012.

[21] R. E. Precup, R. C. David, E. M. Petriu, S. Preitl, and M. B. Raˇdac, "Novel adaptive charged system search algorithm for optimal tuning of fuzzy controllers," Expert Systems with Applications, vol. 41, pp. 1168-1175, 2014. http://dx.doi.org/10.1016/j.eswa.2013.07.110

[22] I. P. Solos, I. X. Tassopoulos, and G. N. Beligiannis, "Optimizing shift scheduling for tank trucks using an effective stochastic variable neighbourhood approach," International Journal of Artificial Intelligence, vol. 14, pp. 1-26, 2016.

[23] M. S. Kırana, and O. Fındık,"A directed artificial bee colony algorithm," Applied Soft Computing, vol. 26, pp. 454-462, 2015. http://dx.doi.org/10.1016/j.asoc.2014.10.020

[24] A. Basgumus, M. Namdar, G. Yilmaz, A. Altuncu, "Performance Comparison of the Differential Evolution and Particle Swarm Optimization Algorithms in Free-Space Optical Communications Systems," Advances Electrical and Computer Engineering, vol. 15, pp. 17-22, 2015. doi:10.4316/AECE.2015.02003.

[25] M. H. David, A. G. Richard, "A novel pulse echo correlation tool for transmission path testing and fault diagnosis," Journal of computers, vol. 11, pp. 31-39, 2006. doi:10.1109/DFTVS.2005.12.

[26] R. C. Eberhart, Y. Shi, J. Kennedy, "Swarm Intelligence," Morgan Kaufmann Publishers Inc, San Francisco, CA, 2001.

[27] R. Eberhart, Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in Proceedings Congress on in Evolutionary Computation, vol. 1, pp. 84-88, 2000. doi:10.1109@cec.2000.870279.

[28] M. Clerc, J. Kennedy, "The particle swarm - explosion, stability, and convergence in a multidimensional complex space," IEEE Transactions on Evolutionary Computation, vol. 6, pp. 58-73, 2002. doi:10.1109/4235.985692.

[29] M. N. Alam, B. Das, and V. Pant, "A comparative study of metaheuristic optimization approaches for directional overcurrent relays coordination," Electric Power Systems Research, vol. 128, pp. 39-52, 2015. doi.org/10.1016/j.epsr.2015.06.018.

[30] Y. del Valle, G. Venayagamoorthy, S. Mohagheghi, J.-C. Hernandez, and R. Harley, "Particle swarm optimization: Basic concepts, variants and applications in power systems," IEEE Transactions on Evolutionary Computation, vol. 12, pp. 171-195, 2008. doi:10.1109/TEVC.2007.896686