Article citation info:
GUO YM, WANG XT, LIU C, ZHENG YF, CAI XB. Electronic system fault diagnosis with optimized multi-kernel SVM by improved CPSO. Eksploatacja i Niezawodnosc – Maintenance and Reliability 2014; 16 (1): 85–91.

Yang-Ming GUO
Xiang-Tao WANG
Chong LIU
Ya-Fei ZHENG
Xiao-Bin CAI

ELECTRONIC SYSTEM FAULT DIAGNOSIS WITH OPTIMIZED MULTI-KERNEL SVM BY IMPROVED CPSO

DIAGNOZA USZKODZEŃ UKŁADU ELEKTRONICZNEGO Z WYKORZYSTANIEM WIELOJĄDROWEJ MASZYNY WEKTORÓW NOŚNYCH (SVM) ZOPTYMALIZOWANEJ PRZY POMOCY POPRAWIONEGO ALGORYTMU CPSO

Electronic systems' safety operation has become a key issue to complex and high reliability systems. Now more emphasis has been laid on the accuracy of electronic system fault diagnosis. Based on the characteristics of the electronic system fault diagnosis, we design a multi-classification SVMs model to attain better fault diagnosis accuracy, which utilizes multi-kernel function consisting of several basis kernel functions to enhance the interpretability of the classification model. In order to optimize the performance of multi-classification SVMs with multi-kernel, we improve the Chaos Particles swarm Optimization (CPSO) algorithm to achieve the optimum parameters of SVMs and the multi-kernel function. For the improved CPSO algorithm, a modified Tent Map chaotic sequence is used to strengthen the search diversity, and an effective method is embedded to the stander PSO algorithm which can ensure to avoid premature stagnation and obtain the global optimization values. The fault diagnosis simulation results of an electronic system show the proposed optimization scheme is a feasible and effective method and it can significantly improve the fault diagnosis accuracy of the multi-kernel SVM.

***Keywords**: electronic system; fault diagnosis; support vector machine; chaos particles swarm optimization; multi-kernel.*

Bezpieczeństwopracy układów elektronicznych stało się kluczowym zagadnieniem w odniesieniu do złożonych układów o wysokiej niezawodności. Obecnie coraz większy nacisk kładzie się na trafność diagnozy uszkodzeń układów elektronicznych. Na podstawie charakterystyki diagnozy uszkodzeń układów elektronicznych, opracowaliśmy model wielokryterialnej klasyfikacji SVM pozwalający osiągnąć lepszą trafność diagnozy uszkodzeń. Model wykorzystuje funkcję wielojądrową składającą się z kilku bazowych funkcji jądrowych pozwalającą na zwiększenie interpretowalności modelu klasyfikacyjnego. Aby zoptymalizować działanie modelu wielokryterialnej klasyfikacji SVM wykorzystującego funkcję wielojądrową, udoskonaliliśmy algorytm Optymalizacji Metodą Chaosu-Roju Cząstek (CPSO), co pozwoliło osiągnąć optymalne parametry SVM i funkcji wielojądrowej. W poprawionym algorytmie CPSO wzmocniono różnorodność wyszukiwania poprzez wykorzystanie chaotycznej sekwencji generowanej przez zmodyfikowaną mapę tent, a także włączono do standardowego algorytmu PSO efektywną metodę pozwalającą uniknąć przedwczesnej stagnacji oraz uzyskać globalne wartości optymalizacji. Wyniki symulacji diagnozy uszkodzeń systemu elektronicznego pokazują, że proponowany system optymalizacji może być wykorzystywany jako skuteczna metoda umożliwiająca znaczne zwiększenie trafności diagnozy uszkodzeń z wykorzystaniem wielojądrowej SVM.

***Słowa kluczowe**: układ elektroniczny, diagnoza uszkodzeń, maszyna wektorów nośnych, optymalizacja metodą chaosu-roju cząstek; funkcja wielojądrowa.*

1. Introduction

Electronic diagnosis is always an important research field of fault diagnosis. With the rapid development of the electronic technology in recent years, [21] more and more electronic systems become critical components of the whole system, and their safety also become the key issue to the system reliability. Thus, according to the information collected from the test ports, inferring the state condition, determining the fault location, forecasting the future failure and then giving the necessary maintenance tips have great significances to accomplishing the missions successfully.

However, electronic system fault diagnosis is complexity and difficulty in many cases. For example, one fault phenomenon always shows with several fault modes. We need to solve multi-classification fault diagnosis problem. In recent years, some intelligent classification methods, such as artificial neural network (ANN) [10], Support Vector Machine (SVM) [5], etc. have been applied to electronic system fault diagnosis. As a state-of-the-art learning method based on the statistical learning theory, SVM not only characterizes a simple model structure, but also has excellent classification in solving learning problem with small training sample set. [9, 30] It can solve the problems of "over fitting", local optimal solution and low-convergence rate existing in ANN. Moreover, SVM has better generalization performance than ANN due to its risk minimization principle. [15] Thus, SVM has received more extensive attention and achieved superior performance in electronic system fault diagnosis. However, SVM also has two main drawbacks in practical application of multi-classification fault diagnosis.[15, 34]

Firstly, because SVM is originally designed for binary-class classification, we must combine several binary-SVMs for the multi-classification with a suitable structure, such as "one-against-rest", "one-against-one", "decision directed acyclic graph (DDAG)", etc. [4, 16, 34] But the performance does not seem satisfactory as much as the binary classification. This may be due to that, whatever the scheme is, each binary-SVM must train at least two class fault mode data, i.e., we must look at least two different class data as same class. Generally, an appropriate kernel for one class fault mode data does not always conduct to the others. So, it is difficult to attain good results by the reported schemes.

Secondly, it is difficult to select the appropriate parameter values of the SVM model, which have a great impact on the generalization capability and model accuracy. Researchers proposed many optimization methods to solve the problem, such as genetic algorithm, particles swarm optimization (PSO) algorithm, etc. [17]. But these methods involve too many human factors or requirements, for example, the kernel function should be continuously differentiable, and the results of SVM classifier are prone to failing into the local minimum.

In order to overcome above shortcomings, we propose a new scheme to improve the classification performance of electronic system fault diagnosis. This scheme includes two parts: One is to design a appropriate multi-classification SVMs model using several multi-kernel SVMs, which depends on the characteristics of electronic system fault diagnosis and can mine the information in the data more effectively. The other one is to improve the chaotic particles swarm optimization (CPSO) algorithm to optimize the parameters of the classifier, which can avoid the premature stagnation and ensure to obtain the best parameters values combination. The feasibility and efficiency of the proposed scheme for electronic system fault diagnosis are verified via application experiments.

The remainder of the paper will be structured as follows: Section 2 gives a brief introduction of SVM and stander PSO algorithm; Section 3 proposes the scheme of electronic fault diagnosis with multi-kernel SVM and the improved CPSO algorithm; Section 4 shows application experiments of electronic system fault diagnosis; and Section 5 gives the conclusions.

2. Related work review

2.1. Support vector machines

As a machine-learning algorithm, SVM integrates the optimal separating hyper-plane with the kernel method. Its resolution has good generalization capability, and the generalization capability is independence of the particular sample distribution. The performance of SVM is mainly referred to its generalization capability, namely the capability of recognizing the new data, and availability to the situation of small samples.[9, 15, 30] SVM is more suitable for electronic system fault diagnosis, because electronic system always shows nonlinear, complexity and diversity features.

Consider n training data x_i and the corresponding labels y_i ($i = 1, 2, ..., n$). In the simplest form, SVM will yield a hyper-plane that separates the training data by a maximal margin. The data lying on one side of the hyper-plane are labeled as +1, and the other data lying on the other side are labeled as −1. The training data that lie closest to the hyper-plane are called support vectors. In the case of linearly data, it is possible to determine the hyper-plane $w^T x + b$=0 that separates the given data, where w and b are used to define the position of the separating hyper-plane. It is easy to find that the parameter pair (w, b) corresponding to the optimal hyper-plane is the solution to the following optimization problem:

$$\begin{aligned} &minimize: L(w) = \frac{1}{2}\|w\|^2 \\ &subject\ to: y_i(w^T x_i + b) \geq 1 \qquad i = 1, 2, ..., n \end{aligned} \tag{1}$$

For linearly non-separable cases, there is no such a hyper-plane that is able to classify every training sample correctly. So the optimization idea is generalized via the concept of soft margin. The new optimization problem thus becomes:

$$\begin{aligned} &minimize: L(w, \xi_i) = \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}\xi_i \\ &subject\ to: y_i(w^T x_i + b) \geq 1 - \xi_i \qquad i = 1, 2, ..., n \end{aligned} \tag{2}$$

where ξ_i is called slack variables related to the soft margin, and c is the tuning parameter used to balance the margin and training error. Both optimization problems (see Eq. (1) and Eq.(2)) can be solved by the Lagrange multipliers α_i that transform them to quadratic programming problems.

For the applications, linear SVM does not meet satisfactory performance, non-linear SVM is more often applied. The basic idea of designing a non-linear SVM model is to map the input vector $x \in R^n$ into a high-dimensional feature space to solve a non-linear classification problem. Here the mapping function $\varphi(x)$, called kernel function, is selected in advance. The kernel function can perform a non-linear mapping to a high-dimensional feature space by replacing the inner product for non-linear pattern problem, which performs the non-linear mapping. The kernel functions are Mercer functions which meet Mercer condition, and the approximating feature map for the Mercer kernel is $k(x, y) = \varphi(x)^T \varphi(y)$. The main basis kernel functions are listed as follows:

(1) Linear kernel function: $k(x, y) = (x \cdot y)$

(2) Polynomial kernel function: $k(x, y) = (s(x \cdot y) + \gamma)^d$

(3) RBF kernel function: $k(x, y) = \exp(-\frac{\|x - y\|^2}{2\sigma^2})$

(4) Sigmoid kernel function: $k(x, y) = \tanh(s(x \cdot y) + \gamma)$

The learning algorithm for a non-linear classifier SVM follows the design of an optimal separating hyper-plane in a feature space. The procedure is the same as associated with hard and soft margin classifier SVMs in the x-space. Using the chosen kernel function, the Lagrangian is maximized in the corresponding high-dimensional feature space as follows:

$$\begin{aligned} &maximize: L(\alpha) = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j k(x_i, x_j) \\ &subject\ to: \sum_{i=1}^{n}\alpha_i y_i = 0, \qquad \alpha_i \geq 0, \qquad i = 1, 2, ..., n \end{aligned} \tag{3}$$

where the constraints must be revised by a non-linear soft margin classifier SVM. The unique difference between these constraints

and the separable non-linear classifiers is in the upper bound c and the Lagrangian multipliers α_i. The constraints of the optimization problem become as follows:

$$subject\ to: \quad \sum_{i=1}^{n} \alpha_i y_i = 0, \qquad c \geq \alpha_i \geq 0, \qquad i = 1, 2, ..., n \tag{4}$$

This way, the influences of the training data will be limited and remained on the wrong side of a separating non-linear hyper-plane. The non-linear SVM classifier is described below:

$$f(x) = \text{sgn}(\sum_{i=1}^{l} \alpha_i y_i K(x, x_i) + b) \tag{5}$$

where l presents the number of support vectors.

2.2. Stander particles swarm optimization algorithm

Compared with the other optimization algorithms, PSO algorithm is a kind of global search algorithm with many merits, such as simple concept, fast convergence rate, etc. It has been successfully applied in many fields. [6, 28] The stander PSO algorithm was proposed by James Kennedy and Eberhart, which is derived from the simulations of the birds in finding foods [12, 25, 32]. The basic idea of the stander PSO algorithm is to optimize the solution to every problem as a particle which searches the optimal value by sharing the historical information and the social information amount the particle individuals. Each particle flights with a certain speed in the D-dimensional search space, and uses the fitness function to judge the merits of particles, particle flying experience of its own and other particles flying experiences. Then the speeds and best location of the particle group are adjusted dynamically, finally the optimal solution of optimization problems is given.

Assume m particles and form themselves into a particles swarm in a D-dimensional search space, $x_i = (x_{i1}, x_{i2}, ..., x_{iD})$ denotes the position of the i-th particle, and $v_i = (v_{i1}, v_{i2}, ..., v_{iD})$ denotes the velocity of the i-th particle. The best position of a particle is $p_i = (p_{i1}, p_{i2}, ..., p_{iD})$, and the best position of the whole swarm is $p_g = (p_{g1}, p_{g2}, ..., p_{gD})$. Therefore the position and velocity of the particles in the particles swarm can be expressed as follows:

$$\begin{cases} v_{id}(k+1) = \mathrm{w} \cdot v_{id}(k) + rand(0, c1) \cdot [p_{id}(k) - x_{id}(k)] + rand(0, c2) \cdot [p_{gd}(k) - x_{id}(k)] \\ x_{id}(k+1) = x_{id}(k) + v_{id}(k+1) \qquad d = 1, 2, ..., D \end{cases} \tag{6}$$

where $c1$ and $c2$ are acceleration constants which, respectively, stand for the weights of the accelerations, and by which, a particle flies towards the individual local best position or the best global position; $rand(0, c1)$ and $rand(0, c2)$ are the random numbers evenly distributed in $[0, c1]$ and $[0, c2]$ respectively. If $c1 = 0$, the particle only has self-experience which means that its convergence rate may be fast, and it is easy to fall into the local optimum. If $c2 = 0$, the particle only has social experience which means that all particles in a swarm become moving by themselves without interaction, and the probability of finding a solution is very low. The velocity v_{id} is generally condition by $v_{id} \in [-v_{\max}, v_{\max}]$ to prevent the particles from flying out of the solution area. ω is an inertial weight which denotes the influence of the previous velocity of a particle upon its current velocity. The bigger ω means the bigger velocity v_{id} and search space for the particles, which helps to find the new solution space. The smaller ω means the smaller velocity v_{id}, which helps to find a better solution in the current solution space. ω is always defined as follows [35]:

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{Iter_{\max}} \times Iter \tag{7}$$

where ω_{max} is the initial inertia weight factor, ω_{min} is the final inertia weight factor, $Iter$ is the current iteration number, and $Iter_{\max}$ is the maximum iteration number.

3. Optimization multi-kernel SVM by improved CPSO

For SVM, the penalization parameter c, kernel function and its parameters are the main factors which influence the classification performance [3, 29]. Therefore, in order to obtain good generalization capability of SVM, one of the main issues is to select the appropriate c, kernel function and its parameters. In this section, aiming at the characters of electronic fault diagnosis, we apply multi-kernel SVM to enhance the generalization capability and exploit more discriminative information in sample data. In addition, we propose an improved CPSO, called ICPSO, to jointly optimize the parameters of multi-kernel SVM.

3.1. Improved CPSO algorithm

According to Ref. [7], the performance of standard PSO algorithm mainly depends on the number of particles and initial parameters. We can see from Eq.(6) that current individual local position is attracted not only by its own current local best position but also by the best global position. If both of the local best position and the best global position are local optima, the particles will repeat the same search path, called premature stagnation. The Eq. (6) doesn't offer a method of jumping out of the premature stagnation. In this paper, an improved CPSO (ICPSO) algorithm is proposed to enhance the search diversity and to overcome the premature stagnation.

3.1.1. Initialization of particles

The initialization of the particles of stander PSO algorithm always adopts a random distribution strategy, but it is difficult to ensure better ergodicity of the initial particles swarm. Chaos is an universal phenomenon of non-linear systems [18]. In general, chaotic motion is not haphazard, and the chaotic variable has three main properties, such as ergodicity, randomness and sensitivity to initial conditions. Using these characteristics, the chaos optimization algorithm is utilized to solve the problem of the particles initialization.

Different chaos map functions have great different influences on chaotic optimization search. Logistic map is more often used in research and application currently. Ref. [2, 19, 20, 24] point out that Tent map has better ergodic homogeneity and higher iterative evolution velocity than Logistic map via mathematical analysis and simulation verification. In this paper, we modify stander Tent map function and use it for the chaotic search. The stander equations of Tent map function are shown as follows:

$$x_{n+1} = \begin{cases} 2(x_n + 0.1 \times rand(0,1)) & 0 \le x_n \le 0.5 \\ 2(1 - x_n + 0.1 \times rand(0,1)) & 0.5 < x_n \le 1 \end{cases} \quad (8)$$

Here we modify Eq.(8) with the following constrained condition. If $x_n = \{0, 0.25, 0.5, 0.75\}$ or $x_n = x_{n-m}$, $m = \{0,1,2,3,4\}$, then $x_n \to x_n + 0.1 \times rand(0,1)$. In practice, because the optimization ranges of every parameters are different, we will carry interval [0, 1] of the chaotic variables to define the rang of the particles swarm solution vectors.

3.1.2. Judgment of premature stagnation

In the iteration process of each particle, if the premature stagnation happens, i.e., the best global position of the whole swarm p_g is unchanged within N times iterations, which means that the particles swarm has already or will be fall into local optima. Here, the value of N can be set in advance according to the scale of the problem or the experimental results. Generally, the bigger N means the standard of the judgment of premature stagnation is more lax. Base on this idea, some researchers[8] applied to embed a premature stagnation counter , called K, into the PSO algorithm, K is used to count the number of stagnation. If p_g is same as the previous value, then K adds 1, otherwise K is cleared to zero. When K is up to the limit time N, a mechanism will run to help the particles to jump out of the local optimum. Obviously, this process will consume more computation time in practice, thus we propose to leave this judgment process out, this means that p_g will be replaced by a new value at each iteration process step.

According to the Eq.(6), the next location of particles depends on the current position and speed. The current speed is the key of the particles swarm optimization since it makes the particles to have better movement and enhance the diversity of the particles swarm. The current speed depends on three main factors, called previous speed, the best position of each particle p_i and the global best position p_g. It is obvious that p_g plays a vital role in information exchange among the particles and the convergence rate increase. With the pulling of p_g, the stander PSO algorithm may lose their diversity of the particles swarm and the premature convergence is more likely to happen during the evolutionary process. In order to enhance the diversity of the particles swarm and avoid premature convergence. We develop p_g as follows at each evolution step.

$$p_g \to \overline{p}_g = \frac{1}{m}\sum_{i=1}^{m} p_i \quad (9)$$

Eq.(9) indicates that the global best place p_g will be replaced by the center of all individual best places $\overline{p}_g$. Then the new equations of the position and velocity of the particles in the swarm are shown as follows:

$$\begin{cases} v_{id}(k+1) = \omega \cdot v_{id}(k) + rand(0,c1) \cdot [p_{id}(k) - x_{id}(k)] + rand(0,c2) \cdot [\overline{p}_{gd}(k) - x_{id}(k)] \\ x_{id}(k+1) = x_{id}(k) + v_{id}(k+1) \qquad d = 1,2,...,D \end{cases} \quad (10)$$

From Eq.(10), we can see that the new improved algorithm not only remains the simplicity of the standard PSO algorithm, but also makes the particles to have the capability of jumping out of local optimal position. $\overline{p}_g$ improves the diversity of the particles swarm, and it plays the same effect as the p_g. In fact, this scheme is also accordance with the social and psychological habits, that is, particles not only want to move closer to the best but also hope to be the "majority". Here the "majority" is $\overline{p}_g$, the center of all individual best places. This way, the new ICPSO algorithm can reduce the invalid iteration effectively and improve the convergence rate greatly as well as the optimization accuracy.

3.2. Optimization of multi-kernel SVM

3.2.1. Multi-kernel SVM

The kernel function and corresponding kernel parameters are the key issue affecting the model prediction accuracy. An efficient kernel function should represent sample data adaptively. General kernel methods use a single kernel function and choose consistent parameters for the whole sample data sets. In reality, the distributions of the sample data in the different mapping space are different. So MKL was proposed by Lanckriet et al. [13]

MKL is an active research topic in the field of machine learning and it provides a more flexible framework than a single kernel. MKL mines information in data more adaptively and effectively, especially in enhancing the interpretability of the classification function and improving its performance.

In the MKL framework, a combined kernel function is defined as the weighted sum of the individual basis kernels. MKL aims to optimize combining weights while training the SVM-based methods [11, 31]. Though researchers proposed a variety of methods of integrate multiple kernels, linear convex combination of basis kernels is still one of the most frequently used approaches. With this method, each basis kernel can exploit the full set of features, or use a subset of features. So, using the equations described by Sonnenburg et al. [26], we consider the combined kernel as follows:

$$k(x_p, x_q) = \sum_{j=1}^{m} \mu_j k_j(x_p, x_q)$$

$$subject\ to: \quad \begin{cases} \sum_{j=1}^{m} \mu_j = 1 \\ \mu_j \ge 0 \qquad j = 1,2,...,m \end{cases} \quad (11)$$

where m is the number of basis kernels, μ_j is the combining weight of the j-th basis kernel. According to the properties of kernel function[22], k is Symmetric Positive Semidefinite Matrix, i.e., $k \ge 0$. Afterward, all kernel matrices k_j are normalized by replacing $k_j(x_p, x_q)$ by the following equation to get unit diagonal matrices:

$$k_j(x_p, x_q) \to k_j(x_p, x_q) \Big/ \sqrt{k_j(x_p, x_p) k_j(x_q, x_q)} \quad (12)$$

Based on MKL and Schur complement lemma [11, 27], the solution problem of μ_j can be cast to the form of quadratically constrained

quadratic program (QCQP). The objective of QCQP is convex in μ_j and Lagrange multipliers α_i. Such a QCQP problem can be solved efficiently by the interior point methods. The obtained dual variables can be used to solve the optimal kernel coefficients.

Although some reported optimization software packages, such as MOSEK [1], can solve the primal and dual problems simultaneously, it is also complex and not unfit for application. In this paper, because the combination of several basis kernels is a linear combination, we propose to use joint optimization method to determine the combining weights of the new multi-kernel. This way can not only reduce the computation complexity comparing with MOSEK, but also obtain the best parameters' values combination.

3.2.2. Optimization steps of multi-kernel SVM

In the ICPSO algorithm, the chaotic sequence is used to initiate individual position, which strengthens the diversity of search, and the stander PSO algorithm with new global best place is mainly used to perform a global search. The process of ICPSO is given below.

Step 1: Initialize each parameter, such as population size m, the stopping criterion (maximum number of iterations) $Iter_{\max}$, displacement genes $c1$ and $c2$, weight ω_{max} and ω_{min}, maximum velocity $v_{\max}$. Let the maximum iteration counter $J=0$;

Step 2: Assign the initial location and velocity with the Eq. (8) randomly;

Step 3: Evaluate the fitting degree of each particle, and let initial individual best position to be p_i and let initial global best position to be p_g;

Step 4: Update the position and speed of each particle based on Eq.(6) and Eq.(7);

Step 5: Calculate the new fitting degree of each particle and replace p_g by $\overline{p}_g$ according to Eq. (9);

Step 6: The maximum iteration counter J plus one, and if $J < Iter_{\max}$, return to step 4. Otherwise end the iteration computation and out put the current p_g.

4. Application experiments of electronic system fault diagnosis

4.1. Case representation

Electronic system faults always exist multiple fault modes simultaneously. A local circuit of a certain electronic system is shown in Fig.1. In order to obtain the sample data for information fusion, we measure 7 voltage values with many times at the measuring points called A, B, C, D, E, F and G (with pentagrams representation). In this case, F1 indicates normal mode, F2, F3, and F4 indicate three fault modes respectively: circuit board damaged mode, chip burned mode and pins broken mode.

4.2. Design and optimization of multi-classification SVMs

Because the extension of SVM to multi-classification problems is not straightforward, multi-classification SVMs should be designed in this case. Ref. [37] addressed the existing representative multi-class classification methods with SVM and compared their merits and defects systematically. The popular methods applying SVMs to multi-class classification problems decompose the multi-class classification problems into many binary-class classification problems and incorporate many binary-class SVM [14, 23, 33]. For example, an N-class classification problem needs $N(N-1)/2$ binary-class classification SVMs with the "one-against-one" approach, while N SVMs

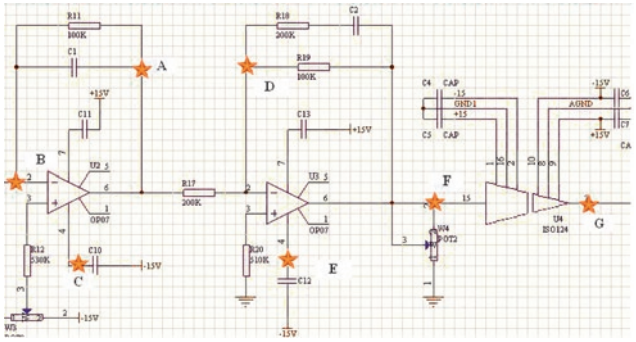

Fig. 1. Local circuit with faults of certain electronic system

for the "one-against-Rest" approach. Although the "one-against-one" approach demonstrates superior performance, it may require prohibitively-expensive computing resources in many real-world problems. The "one-against-rest" approach shows less accuracy, and demands heavy computing resources, especially for the real-time applications.

Based the conclusions of Ref. [37], we establish a multilayered classification structure in this paper. Firstly, we classify the normal mode and fault modes using SVM1 to achieve the purpose of fault detection, and then we use the "DDAG" approach to execute fault pattern recognition, such as recognizing circuit board damaged mode and chip burned mode via SVM2, recognizing chip burned mode and pins broken mode via SVM3, recognizing circuit board damaged mode and pins broken mode via SVM4. The multi-classification SVMs model is shown in Fig. 2.

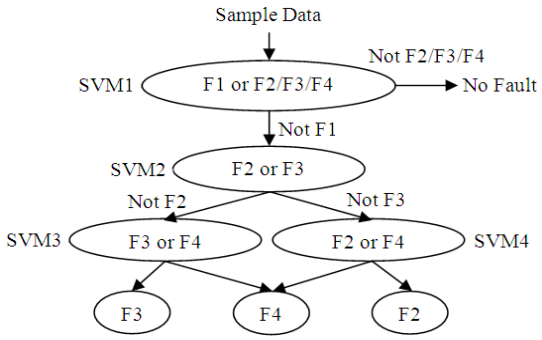

Fig. 2. Fault class flow chart with multi-classification SVMs

Based on Fig.2, we choose appropriate basis kernel functions firstly, and then use the proposed ICPSO algorithm to achieve the optimal parameters' values combination of the multi-kernel SVM. Here we choose the average test set accuracy (TSA) [11] for the sample data as evaluation criteria (fitting function). The TSA is defined as follows:

$$\sigma = \frac{y_t}{y_t + y_f} \tag{13}$$

where y_t and y_f represent the number of true and false classification respectively. That SVM has higher TSA means the model is a better model. The main steps of fault diagnosis are shown as follows

Step 1: Choose sample data via unitary processing and divide them into two parts: training set and testing set;

Step 2: Jointly optimize the parameters of each multi-kernel SVM using the ICPSO algorithm and establish the multi-classification multi-kernel SVMs;

Step 3: Train each multi-kernel SVM of multi-classification SVMs;

Step 4: Verify fault diagnosis accuracy of the multi-classification multi-kernel SVMs using the testing set.

4.3. Experiments and discussion

We collect 1000 voltage values as the sample data from the seven test points (see Fig.1). After unitary processing, we divide them into two parts: 500 data as training set and 500 data as testing set. The number of each fault mode is shown in Tab.1.

Table 1. Number of the fault modes in the sample data

Type	Normal	Fault Mode		
		Circuit Board Damaged	Chip Burned	Pins Broken
Training Set	112	138	150	100
Testing Set	146	121	113	120

In this study, we perform two experiments. Experiment I compares the multi-classification multi-kernel SVMs on fault detection rate (FDR) and fault detection accuracy (FDA) using ICPSO, CPSO and GA respectively. Here we choose one linear kernel function which represents global information and one RBF function which represent local information [36] as basis kernel functions. Experiment II compares single RBF SVM with multi-kernel SVM on FDR and FDA. The multi-kernel SVM has same basis kernel functions as experiment I. Their parameters are optimized by ICPSO in experiment II.

The same set of parameters is assigned for the two experiments: m=20, D=4 (represent the two combination coefficients of the Multi-kernel, the penalization parameter c and the width of RBF σ respectively), ω_{max}=1.2, ω_{min}=0.2, $c1 = c2$=2, $v_{\max}$=15 and $Iter_{\max}$=300. The results are reported in Tab. 2 and Tab. 3.

Table 2. Results of fault diagnosis experiment I (compare with ICPSO, CPSO and GA)

Algorithm	Fault Detection Rate (FDR)	Fault Detection accuracy (FDA)		
		Circuit Board Damaged	Chip Burned	Pins Broken
ICPSO	93.2%	94.2%	95.6%	92.5%
PSO	90.4%	89.3%	92.9%	90.8%
GA	90.0%	89.3%	91.2%	91.7%

Table 3. Results of fault diagnosis experiment II (compare with RBF and multi-kernel)

Algorithm	Fault Detection Rate (FDR)	Fault Detection Accuracy (FDA)		
		Circuit Board Damaged	Chip Burned	Pins Broken
Method with Multi-kernel	93.2%	94.2%	95.6%	92.5%
Method with RBF	87.2%	86.8%	91.2%	83.3%

Tab. 2 and Tab. 3 show that the multi-classification multi-kernel SVMs with optimization parameters by ICPSO have better FDR and FDA. This indicates that the proposed ICPSO algorithm is a better optimization method. Moreover, multi-kernel learning method can enhance the interpretability of the classification function and it is an effective approach to improve performance of SVM-based classifier considerably.

5. Conclusions

The reliability of electronic system has attracted more attention, and SVM is widely applied in the electronic system fault diagnosis. In many reality cases, the parameters of SVM and its kernel function have great impact with the classification accuracy and generalization capabilities. In order to attain satisfactory fault diagnosis results of electronic system, we design a multi-classification multi-kernel SVMs model to perform classification. We propose an improved CPSO algorithm to achieve the optimal parameters' values combination of the electronic fault diagnosis model. In the ICPSO algorithm, a modified Tent map chaotic sequence is used to initiate individual position, and a new scheme is embedded into the stander PSO algorithm to avoid premature stagnation. The feasibility and efficiency of the proposed scheme have been verified via the application experiments.

***Acknowledgments**: The authors would like to thank Prof. Ming J. Zuo, the Director of Reliability Research Laboratory at University of Alberta, for his helpful comments. This work is supported by the National Basic Research Program of China (973 Program), the National Natural Science Foundation of China (No. 61001023 and No. 61371024), Chinese Astronautics Support Technology Foundation, Aviation Science Foundation of China (No.2013ZD53051) and Xi'an Technology Plan Projects 2013.*

References

1. Andersen ED, Andersen AD. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. High performance optimization. Norewll, USA: Kluwer Academic Publishers, 2000.
2. Brian R, Sonja S. Structure of inverse limit spaces of Tent maps with nonrecurrent critical points. Glasnik matematickl 2007; 42(62):43–56.
3. Chapelle O, Vapnik V, Bousquet O, et al. Choosing kernel parameters for support verctor machines. Machine Learning 2002; 46(1):131–159.
4. Diao ZH, Zhao CJ, Guo XY, Lu SL. A new SVM multi-class classification algorithm based on balance decision tree. Control and Decision 2011; 26(1):149-152,156.
5. Drucker H, Wu D, Vapnik V. Support verctor machines for spam categorization. IEEE Transactions on Neural Networks 1999; 10(5):1048–1054.
6. Gaing ZL. Discrete particles swarm optimization algorithm for unit commitment. Power Engineering Society General Meeting, July, 2003.
7. Guo JQ, Li Y, Wang HS. Application of particles swarm optimization algorithms to determination of water quality parameters of river streams. Advances in science and technology of water resources 2007; 27(6):5.
8. Guo YM, Ma JZ, Xiao F, Tian T. SVM with optimized parameters and its application to electronic system fault diagnosis. 2012 IEEE International Conference on Prognostics and Health Management, June, 2012.
9. Hao Y, Sun JG, Yang GQ, Bai J. The application of support vector machines to gas turbine performance diagnosis. Chinese Journal of Aeronautics 2005; 18(1):15-19.
10. He Y, Ding Y, Sun Y. Fault diagnosis of analog circuitswith tolerances using artificial neural networks. IEEE Asia-Pacific Conference on Circuits and Systems, October, 2002.
11. Jian L, Xia ZH, Liang XJ, Gao CH. Design of a multiple kernel learning algorithm for LS-SVM by convex programming. Neural Networks 2011; 24: 476–483.

12. Kennedy J, Eberhart R. Particle swarm optimization. IEEE International Conference on Neural Networks, December, 1995.
13. Shen LC, Huo XH, Niu YF. Survey of discrete particles swarm optimization algorithm. Systems Engineering and Electronics 2008; 30(10): 1986–1990.
14. Lee Yoonkyung. Multicategory support vector machines, theory, and application to the classification of microarray data and satellite radiance data. Technical Report No.1063, 15 Sept. 2002.
15. Li H, Zhang YX. An algorithm of soft fault diagnosis for analog circuit based on the optimized SVM by GA. The Ninth International Conference on Electronic Measurement & Instruments (ICEMI'2009), August, 2009.
16. Li Q, Dong CL, Chen ZZ, He XL. Improved algorithm for kernel-based SVM. Computer Engineering and Application 2010; 46(10): 150–152.
17. Lian GY, Wang WG, Huang KL, Guo R. Research of optimization method for test selection based on particles swarm optimization algorithm. Computer Measurement & Control 2008; 16(10):1387–1389.
18. Liu B, Wang L, Jin YH, Tang F, Huang DX. Improved particles swarm optimization combined with chaos. Chaos, Solitons and Fractals 2005; 25: 1261–1271.
19. Liu JD. A spatiotemporal chaotic sequence based on coupled chaotic Tent map lattices system with uniform distribution. Journal of computers 2011; 6(2): 190–199.
20. Meng LQ, Guo JQ. Application of chaos particles swarm optimization algorithm to determination of water quality parameter of river steam. Journal of earth sciences and environment 2009; 31(2): 169–172.
21. Michael P, Rubyca J. A prognostics and health management roadmap for information and electronics-rich systems. IEICE Fundamentals Review 2010; 3(4): 25–32.
22. Scholkopf B, Smola A,Learning with Kernels. Cambridgeshire MIT Press, 2002.
23. Schwenker F. Hierarchical support vector machines for Multi-class pattern recognition. The fourth international conference on knowledge-based intelligent engineering systems & allied technologies, Augest, 2000.
24. Shan L, Qiang H, Li J, Wang ZQ. Chaotic optimization algorithm based on Tent map. Control and Decision 2005; 20(2): 179–182.
25. Shen LC, Huo XH, Niu YF. Survey of discrete particles swarm optimization algorithm. Systems Engineering and Electronics 2008; 30(10): 1986–1990.
26. Sonnenburg S, Räosch G, Schäer C, Schökopf B. Large scale multiple kernel learning. Journal of Machine Learning Research 2006; 7: 1531–1565.
27. Vandenberghe L, Boyd S, Semidefinite programming. SIAM Review 1996; 38: 49–95.
28. Van B, Engelbrecht A. A new locally convergent particles swarm optimizer. IEEE International Conference on Systems, Man and Cybernetics, October, 2002.
29. Vapnik V. An overview of statistical learning theory. IEEE Transactions on Neural Networks 1999; 10(5): 988–999.
30. Vapnik V. Statistical learning theory. New York: John Wily and Sons, 1998.
31. Wang HQ, Sun FC, Cai YN, Chen N. Ding LG, On multiple kernel learning methods. Acta Automatica Sinica 2010; 36(8): 1037–1050.
32. Wang M, Zhu YL. He XX. A survey of swarm intelligence. Computer Engineering 2005; 31(22):194–196.
33. Weston J, Watkins C. Multi-class support vector machines. Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, Egham, UK, 1998.
34. Xu QH, Shi J. Fault diagnosis for aero-engine applying a new multi-class support vector algorithm. Chinese Journal of Aeronautics 2006; 19(1): 175–182.
35. Yang CS, Chuang LY , Ke CH , Yang CH. Comparative particles swarm optimization (CPSO) for solving optimization problems. IEEE International Conference on Research, Innovation and Vision for the Future, July, 2008.
36. Zhang JF, Hu SS. Chaotic time series prediction based on multi-kernel learning support vector regression. Acta Physica Sinica (Chinese Physics) 2008; 57(5): 2708–2713.
37. Zhang M Zhang DX. Research on text categorization based on M-SVMs. Computer Technology and Development 2008; 28(3): 139–141, 156.

Yang-Ming GUO
Xiang-Tao WANG
Chong LIU
Ya-Fei ZHENG
Xiao-Bin CAI
Reliability and Maintenance Research Laboratory
School of Computer Science and Technology
Northwestern Polytechnical University
Youyi West Road 127, Xi'an Shaanxi, 710072
P. R. China
E-mail: yangming_g@nwpu.edu.cn (Y. M. Guo)