Research Article

A Framework for Diagnosing the Out-of-Control Signals in Multivariate Process Using Optimized Support Vector Machines

Tai-fu Li,[1] Sheng Hu,[2] Zheng-yuan Wei,[2] and Zhi-qiang Liao[3]

[1] School of Electrical and Information Engineering, Chongqing University of Science and Technology, Chongqing 401331, China
[2] School of Mathematics and Statistics, Chongqing University of Technology, Chongqing 400054, China
[3] School of Electronic Engineering, Xi'an Shiyou University, Xi'an 710065, China

Correspondence should be addressed to Tai-fu Li; litaifuemail@126.com

Multivariate statistical process control is the continuation and development of unitary statistical process control. Most multivariate statistical quality control charts are usually used (in manufacturing and service industries) to determine whether a process is performing as intended or if there are some unnatural causes of variation upon an overall statistics. Once the control chart detects out-of-control signals, one difficulty encountered with multivariate control charts is the interpretation of an out-of-control signal. That is, we have to determine whether one or more or a combination of variables is responsible for the abnormal signal. A novel approach for diagnosing the out-of-control signals in the multivariate process is described in this paper. The proposed methodology uses the optimized support vector machines (support vector machine classification based on genetic algorithm) to recognize set of subclasses of multivariate abnormal patters, identify the responsible variable(s) on the occurrence of abnormal pattern. Multiple sets of experiments are used to verify this model. The performance of the proposed approach demonstrates that this model can accurately classify the source(s) of out-of-control signal and even outperforms the conventional multivariate control scheme.

1. Introduction

Statistical process control (SPC) is one of the most effective tools in total quality management (TQM), which is used to monitor manufacture process variation. Control charts are the most widely applied SPC tools used to reveal abnormal variations of monitored measurements, as well as to locate their assignable causes [1, 2]. According to the state of control chart, we can easily be informed whether the manufacture process is in in-control state or not; quality practitioners or engineers search for the assignable causes and take some necessary corrections and adjustments to bring the out-of-control process back to the in-control state [3].

In the multiply quality diagnose field, there are often two or more quality characteristics which should be monitored at the same time. A nature idea is that we can maintain a separate control chart for each characteristic. The traditional SPC technology is based on the process observation data meeting the independent and identically distributed. However, many manufacturing processes do not meet this assumption. It would result in high error discrimination phenomenon when the characteristics are highly correlated. Hotelling is the first to recognize the defects if we simply expand the univariate control chart to multivariate process; thus the concept of multivariate quality was proposed in 1947. Hotelling's T^2 statistic [4–7] might be the most common tool in multivariate analysis for identifying whether the whole process is in out-of-control state. The primary possible causes for the mean shifts result from the introduction of new workers, machines, material, or methods, a change in the measurement method or standard, and so forth. The T^2 test statistic of the mean vector is not only considered the volatility of the mean, but also the correlation among the variables. it is the optimal test statistic for detecting a general shift. Besides T^2 control chart, Mason et al. [8] proposed a cause-selecting procedure by using the decomposition of the T^2 statistic. The user can see the contribution of each variable by decomposing the T^2 statistic. Moreover, this decomposition allows the user to observe which variable(s) with a significant contribution

is (are) the cause of abnormal processes. The drawbacks of this method are additional computation and its sensitivity to the number of variables. Other multivariate control charts, such as multivariate exponentially weighted moving average (MEWMA) control chart, and multivariate cumulative sum (MCUSUM) control chart have been proposed to detect mean shifts. Lowry et al. [9] proposed a multivariate extension of the exponentially weighted moving average (EWMA) control chart. They compared their method to a multivariate cumulative sum (MCUSUM) control chart based on the average running length (ARL) performance and their results showed that the MEWMA chart was similar to the MCUSUM chart in detecting mean shifts of a multivariate normal distribution.

Besides the traditional multivariate statistical process control (MSPC) technology, many scholars have been trying to diagnose the abnormal process in multivariate process via data mining and artificial intelligence especially the neural networks [10, 11]. Neural networks (NNs) have excellent noise tolerance in real time, requiring no hypothesis on statistical distribution of monitored measurements. These important features make NNs promising and effective tools that can be used to improve data analysis in manufacturing quality control applications. In the past two decades, various NNs with different structures and learning algorithms have been adopted widely in quality control [12]. Multilayer perceptron (MLP) NN [13], learning vector quantization (LVQ) NN [14], probabilistic NN [15], adaptive resonance theory (ART) NN [16], modular NN [17], backpropagation NN (BPN) [18], and so forth, were used to detect abnormal signals and (or) identify some basic abnormal patters such as shift, trend, cycle, and mixture patterns. These NN-based models have shown their efficiency and effectiveness in SPC.

Summarizing the multiple variables of interest in one single statistic does not release source(s) of the out-of-control signals; that is, it cannot tell which one of the variables or a combination of variables has caused the out-of-control signals. You know, the knowledge of separation out-of-control variables can be helpful to narrow the set of possible assignable causes, which is more rapid identification of the particular causes and reducing of the adjustment cost.

There are a few multivariate studies found with neural networks in the literature, such as an effective detection out-of-control signals or identification source(s) of out-of-control signals by mean shift in multivariate processes. El-Midany et al. [19] developed artificial neural networks to recognition control chart pattern recognition in multivariate process. Yu et al. [1, 20] proposed a selective neural network (NN) ensemble approach (DPSOEN, discrete particle swarm optimization) to monitor and diagnose out-of-control signals in bivariate process. Extensive experiments demonstrate that this method is effective. Salehi et al. [21] presented a hybrid learning-based model for on-line analysis of out-of-control signals in multivariate manufacturing processes. The main contributions of this work are recognizing of the type of unnatural pattern and classification of major parameters for shift, trend and cycle and for each variable simultaneously by proposed hybrid model. Kim et al. [22] attempted to integrate state-of-the-art data mining algorithms including artificial neural networks, support vector regression, and multivariate adaptive regression splines with SPC techniques to achieve efficient monitoring in multivariate and autocorrelated processes. Simulation results from various scenarios indicated that data mining model-based control charts perform better than traditional time-series model-based control charts. Cheng et al. [23] developed using of SVR to predict the magnitudes of process shift. Feed-forward neural networks with different training algorithms and CUSUM-based estimator are used as benchmarks for comparison.

Research indicates that identifying source(s) of out-of-control signals using neural network has been proved to be an effective and useful tool in multivariate manufacturing processes, which has better estimation capabilities than CUSUM and MEWMA. However, ANN classifier just has the judgment and classification ability only after it is studied via machine learning. It is training based on the principle of empirical risk minimization, which leads to a long training time, a poor generalization ability, and easiness of fallng into local minimum. It is not suitable for the quality diagnosis when the process is dynamic and changeable. Although support vector machines has shown excellent generalization performance in a number of applications, one problem that faces the user of an SVM is how to choose a kernel and the specific parameters for that kernel. Applications of an SVM therefore require a search for the optimum settings for a particular application. The kernel functions map the original data into higher-dimension space and make the input data be set linearly separable in the transformed space. The choice of kernel functions is highly problem dependent and it is the most important factor in support vector machine applications. In this work, the RBF kernel is used as the kernel function of the SVM because it tends to give better performance.

In this study, considering the robust recognition power of support vector machine and the global search capability of the genetic algorithm, an optimized SVM approach named GA-SVM (support vector machine classification based on genetic algorithms) was developed; a framework of control chart pattern recognition of multivariate observation data is proposed for classifying source(s) of out-of-control signals in multivariate processes. The SVM classification is used to recognize selected subclasses of multivariate abnormal patterns, identify the variate(s) that is (are) responsible for the occurrence of the abnormal pattern, while the role of GA is that it optimizes SVM parameters such as kernel function parameters and the penalty factor.

The remainder of this paper is organized as follows: the patterns generation of multivariate process and support vector machine algorithms are presented in Section 2; Section 3 describes the diagnosis SVM classifier based on GA in the process mean for its content; in Section 4 we use numerical example and series tests to verify the proposed method; conclusions are presented in Section 5.

2. Methodology

This section provides a practical overview for the patterns generation of multivariate process and support vector

machine algorithms. It is preparing for the diagnosis model to be proposed in later chapter.

2.1. The Patterns Generation of Multivariate Process.

According to multivariate quality control theory [24–26], the T^2 statistic can reflect the state of the correlation structure and the mean vector of the multivariate process data. The overall quality of a multivariate process can be monitored by comparing the above statistic against a positive UCL. These charts are easy to construct as long as the process parameters (i.e., μ_0 is the mean vector and $\sum_0$ is the covariance matrix) are known. However, the covariance matrix $\sum_0$ is usually unknown in many quality process control fields; we could not assume the value arbitrarily. In this case, the covariance matrix $\sum_0$ should be estimated via limit sample. The mean vector μ_0 and the covariance matrix $\sum_0$ are sufficient to characterize any multivariate normal distribution. Thus, random multivariate vectors, such as $\mathbf{X}_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]$, represent a vector of observations from P quality characteristics of independent normal distributions at time i. In the multivariate statistical process control, the mean vector of T^2 statistic is thus obtained using the following form:

$$T^2 = n(\mathbf{X} - \overline{\mathbf{X}})^T \mathbf{S}^{-1} (\mathbf{X} - \overline{\mathbf{X}}). \tag{1}$$

Random samples of size are collected regularly from the process. Thus, represents the mean vector for each of the n multivariate vectors, such as $[x_{i1}, x_{i2}, \ldots, x_{ip}]$ for $i = 1$ to n. The term $(\mathbf{X} - \overline{\mathbf{X}})$ is a number of quality characteristics (P) represented by 1 column vector and $(\mathbf{X} - \overline{\mathbf{X}})^T$ is the transpose of $(\mathbf{X} - \overline{\mathbf{X}})$, that is, a 1 by P row vector. Finally, $\mathbf{S}^{-1}$, which is a P by P matrix, is the inverse of the S matrix. Therefore, the T^2 statistic is the result of a 1 by 1 matrix (i.e., it is a scalar value). Each normal multivariate mean vector must result in a value of T^2 that is less than or equal to a UCL given by $T^2_{\alpha,p}$, where α is the upper tail of the T^2 distribution that acts as the type I error, when the confidence limit equals $[100 \times (1 - \alpha)]$ percent and with P degrees of freedom. Any vector that results in T^2 plotting above $T^2_{\alpha,p}$ limit is discarded from the simulated data set. The sample size, n, is kept to 1, so that the vector value of each simulated multivariate point is used directly to represent one sample observation.

2.2. A Brief Introduction to Support Vector Machine Algorithms.

The support vector machine method is a new and promising classification and regression technology [27–31]. The basic idea of an SVM can be stated briefly as follows. An SVM initially maps the input vectors into a high-dimensional feature space, either linearly or nonlinearly, which is relevant to the selection of the kernel function. The input or feature vectors in the feature space are then classified linearly by a numerically optimized hyperplane, separating the two classes (this can be extended to multiclass). The SVM training always seeks a global optimized solution and the hyperplane depends only on a subset of training examples [32–35].

Let $\{\widetilde{x}_i, y_i\}_{i=1}^n, \widetilde{x} \in R^m, y_i \in [-1, 1]$, be the training set with input vectors $\widetilde{x}$ and output y_i. Here, n is the number of sample observations, m is the dimension of each observation, and y_i is the known target. The algorithm is to seek the hyperplane $\widetilde{w}\widetilde{x}_i + b = 0$, where $\widetilde{w}$ is the vector of hyperplane and b is a bias term, to separate the data from two classes with maximal margin width $2/\|\widetilde{w}\|^2$, and all the points under the boundary are named the support vector. In order to optimize the hyperplane, the SVM solves the following optimization problem:

$$\begin{aligned} \text{Min} \quad & \Phi(\widetilde{x}) = \frac{1}{2}\|\widetilde{w}\|^2, \\ \text{S.t.} \quad & y_i(\widetilde{w}^T \widetilde{x}_i + b) \geq 1, \quad 1 = 1, 2, \ldots, n. \end{aligned} \tag{2}$$

It is difficult to solve (2). Thus, SVM transforms the optimization problem to be dual problem by the Lagrange method. The value of α in the Lagrange method must be nonnegative real coefficients. Equation (2) is transformed into the following constrained form:

$$\text{Max} \quad \Phi(\widetilde{w}, b, \xi, \alpha, \beta) = \sum_{i=1}^N \alpha_i - \frac{1}{2}\sum_{i=1,j=1}^N \alpha_i \alpha_j y_i y_j \widetilde{x}_i^T \widetilde{x}_j, \tag{3}$$

$$\text{s.t.} \quad \sum_{j=1}^N \alpha_j y_j = 0 \quad 0 \leq \alpha_i \leq C, \ i = 1, 2, \ldots, N.$$

In (3), C is the penalty factor and determines the degree of penalty assigned to an error. It can be viewed as a tuning parameter, which can be used to control the trade-off between maximizing the margin and the classification error. In order to separate two classes exactly, add as lack variable ξ in the Lagrange equation to make $y_i(\widetilde{w}^T \widetilde{x}_i + b) \geq 1 - \xi_i, i = 1, 2, \ldots, N, \xi_i \geq 0$. An objective of slack variable is to increase the flexible buffer of boundary.

In general, it could not find the linear separate hyperplane for all application data. For problems that cannot be linearly separated in the input space, the SVM uses the kernel method to transform the original input space into a high-dimensional feature space, where an optimal linear separating hyperplane can be found. The common kernel functions are linear, polynomial, radial basis function (RBF), and sigmoid. Although several choices for the kernel function are available, the most widely used kernel function is the RBF, which is defined as [36]

$$\mathbf{K}(x_i, x) = \exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right), \tag{4}$$

where σ is the parameter of the kernel function. Consequently, the RBF is employed in this study. Also, in this study, we used multiclass SVM method to build the diagnosis model.

3. Diagnosis Model Based on Optimized Support Vector Machine in the Process

The parameters that must be determined are the kernel parameter γ, the regularization parameter C, and ε. Kernel
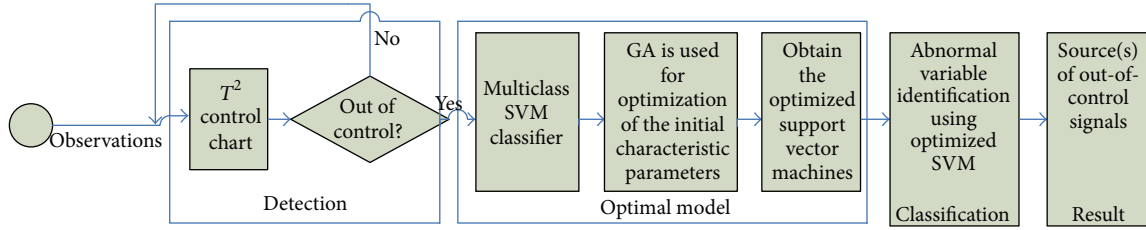
FIGURE 1: Detection and classification by optimized SVM model.

parameter defines the structure of the high-dimensional feature space. The kernel parameter is selected through genetic algorithm (GA). The regularization parameter C should be chosen with caution to avoid overfitting.

In this paper, in order to achieve the goal of diagnosis on the mean vector, as well as identification and location of the abnormal variables in multivariate process, firstly, we are assuming that the abnormal factors in the process only make the changeable of the mean vector, instead of the changeable of the covariance matrix. Under this premise, a diagnosis of the out-of-control signals in the multivariate process based on hotelling T^2 chart and optimized support vector machines is proposed in this paper. The structure of the diagnosis model is briefly illustrated in Figure 1.

The scheme can be briefly introduced as follows: the detection part and the diagnostic process. When an out-of-control signal is detected by T^2 statistic, more observations from the process are collected and they are regarded as a chromosome which was then performed with binary code. A multiclass support vector machine classifier will be used to recognize a random chromosome and recognition accuracy is considered as the fitness function to evaluate the fitness of individual feature. By the operations of selection, crossover, and mutation, with GA self-adaptive optimizing for penalty parameter and kernel parameter, we obtain the optimal model, which is finally used for identifying the source(s) of the out-of-control signal, as well as identification and location of the abnormal variables.

In Figure 1, we identify and locate the abnormal variables via translating them into a pattern recognition problem. For a p-dimensional manufacturing process, there are cases: the mean of each variable is either normal or abnormal. Under these circumstances, the mean vector of p-dimensional totally has 2^p normal (abnormal) states. In other words, only one normal state and $(2^p - 1)$ abnormal state exist. Once the T^2 control chart is out-of-control, it must be one of the $(2^p - 1)$ abnormal states. In order to distinguish between these 2^p states, we define this abnormal state as $(2^p - 1)$ patterns which should be identified; we use optimized SVM model to identify the abnormal signal, which is obtained via genetic algorithm.

Taken the bivariate process as an example, the reference mean vector is $\mu = (\mu_1, \mu_2)^T$ and covariance matrix $\sum$; the abnormal state of the mean vector only has three circumstances that should be identified: (1) the first variable is out of control and the second variable is in control $(\mu_1 + \delta_1, \mu_2)$; (2) the first variable is in control, the second variable is out of control $(\mu_1, \mu_2 + \delta_2)$; (3) both the two variables are out

of control $(\mu_1 + \delta_1, \mu_2 + \delta_2)$. If denoting the out-of-control variable as 1 and in-control variable as 0, then the target vectors for these cases were values (0,0,0,0), (0,1,0,0), (0,0,1,0), and (0,0,0,1), which are used for identifying that the normal, the first and the second, or the both quality characteristics are in out of control state, where the mark (0,0,0,0) represents the normal patterns and (0,1,0,0), (0,0,1,0), and (0,0,0,1) represent the three abnormal patterns. When an out-of-control signal is detected by T^2 statistic, and if the optimal SVM model detects the signal (0,1,0,0) pattern, it indicates that in this process, the first variable is in control while the second variable is out-of-control; once the optimal SVM model detect the signal (0,0,0,1) pattern, it indicates that in this process, both the two variables are out of control, and so on. In the idea mentioned above, the optimal SVM model is not only a substitute for hotelling T^2 control chart, but also identification and location of the abnormal variables when an out-of-control signal is detected by the T^2 control chart.

4. Simulation and Analysis

In this section, an example derived from a bivariate instance of the chemistry process of Montgomery is given to illustrate the use of the proposed approach. The parameters of this process represent process target values of X_1 and X_2; the reference mean vector $\overline{X}$ and the reference variance matrix S are obtained by the first fifteenth observation data. Thus, the X is

$$X = \begin{bmatrix} 10 \\ 10 \end{bmatrix}, \tag{5}$$

and the S is

$$S = \begin{bmatrix} 0.7986 & 0.6973 \\ 0.6973 & 0.7343 \end{bmatrix}. \tag{6}$$

In the following part, we would compare the estimation capabilities of different statistical approaches, such as BPNN, SVM, and optimized SVM.

4.1. Examples Development and Parameter Setting. Training data is very critical in applications of SVM, which determines the recognition efficiency of the classifier work. In this study, the Monte Carlo simulation method was applied to generate the required data sets of normal and abnormal examples for training and testing. Multivariate simulation can provide only a simplified picture of the reality. The best alternative is, where possible, to collect various data from real-world manufacturing systems.
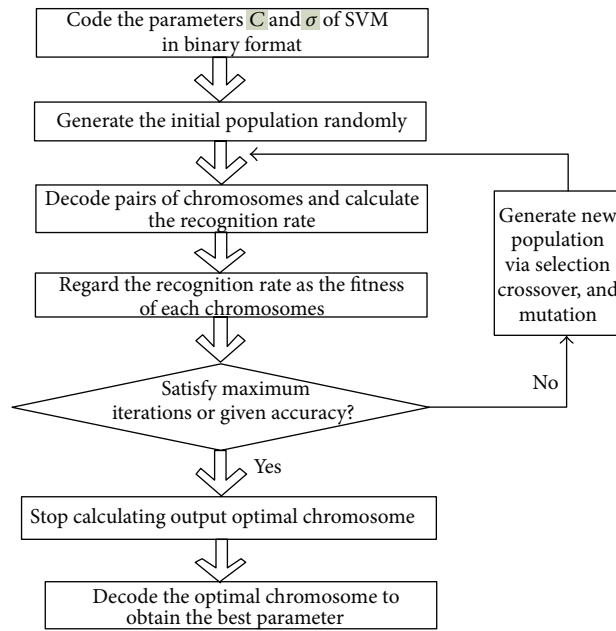
FIGURE 2: The architecture of parameter optimization using genetic algorithm.

In a bivariate SPC application, an input to the neural system should consist of a time-series window of bivariate vectors. In this study, five distinct types of shift (i.e., $\delta = 1.0\sigma, 1.5\sigma, 2.0\sigma, 2.5\sigma$, and 3.0σ) associated with the lth quality characteristic are considered, which cover evenly the range of whole process shifts. For generating training data sets, we shift the mean vector as $(\mu_1 + 1.0\sigma_{11}, \mu_2)$, $(\mu_1, \mu_2 + 1.0\sigma_{22})$, and $(\mu_1 + 1.0\sigma_{11}, \mu_2 + 1.0\sigma_{22})$ in this bivariate process, in order to verify the good ability of SVM in small sample; five hundred input vectors shifted 1.0 for each abnormal model were generated by using statistic, which were used as the training data set of the optimized SVM.

The number of SVM in the output layer is determined by the dimension of quality characteristics. The output vector consists of four elements and uses the largest value to identify the source(s) of the whole signals; in this study, the number of the output nodes is 4.

In order to evaluate the performance of optimized SVM, different sample vectors that shifted 1.0σ, 1.5σ, 2.0σ, 2.5σ, and 3.0σ were generated. Similar to generated training data sets, we applied T^2 statistic to detect the overall out-of-control signal for each sample vector with the type I error being equal to 0.5. We did so until we obtained 500 examples for each abnormal case, which were used as the testing data sets of optimized SVM.

4.2. Optimization Using Genetic Algorithm (GA). SVM is a powerful machine learning and data mining tool that has already been shown to possess classification power. However, the performance of SVM is greatly affected by the penalty (cost) parameter C and the variance of the RBF kernel function parameter σ.

Global optimization attempts to locate the absolutely best set of optimum conditions that results in the highest objective value. It is usually a very difficult problem. Traditional optimization methods such as gradient ascent/descent search in the direction of the local gradient vector and thus easily get stuck in a problem with a multimodal objective function. The optimization method used in our study is the genetic algorithm (GA), which is one of the most popular and widely used techniques for global optimization. Figure 2 presents the structure of the parameter optimization using genetic algorithm (GA).

4.3. Performance Analysis for the Bivariate Process. In this section, the influences of the key factors of optimized SVM model upon their generalization performance are analyzed by implementing the following experiments. The analysis can help us to obtain the suitable parameter setting to improve the generalization performance of the proposed model.

4.3.1. Test 1: The Selection Process of the Optimal Parameters. In this study the free parameters of SVM were selected following a genetic algorithm optimization experiment. The whole process was implemented in Matlab 2009b. In order to seek the optimal parameters, we chose $\delta = 1.5\sigma$ as types of shift for the three different volatilities; each type generates 100 samples. The relevant parameters of the genetic operations are set as follows: population size is 20, the maximum number of iterations is 200, the genetic generation gap and mutation probability are taken as 0.9 and 0.01, respectively, the single ε is 0.01, the chosen kernel was based on RBF with σ being 1.0834 and the regularization parameter C was set to 1.5136. Figures 3(a) and 3(b) show an example of the grid search result, where the x-axis and the y-axis are $\log 2C$ and $\log 2\sigma$, respectively. The z-axis is the accuracy performance. The findings of this experiment were that SVM is quite robust against parameter selections.

4.3.2. Test 2: Sensitivity Analysis to Different Model. Table 1 presents comparison of the performance among the raw

TABLE 1: Correct classification percentages (%) of the optimization SVM, the SVM, and the BPNN.

Model	Shift magnitude	Correct classification percentage (%)				
		(0 0 0 0)	(0 1 0 0)	(0 0 1 0)	(0 0 0 1)	Percentage
BPNN	(0.0 0.0)	470	14	13	3	94.00
	(1.0 0.0)	0	450	12	38	90.00
	(0.0 1.0)	0	3	452	45	90.38
	(1.0 1.0)	0	30	25	445	89.00
	(1.5 0.0)	0	460	1	39	92.00
	(0.0 1.5)	0	2	458	40	91.60
	(1.5 1.5)	0	27	23	450	90.00
	(2.0 0.0)	0	468	0	32	93.6
	(0.0 2.0)	0	5	465	30	93.0
	(2.0 2.0)	0	22	28	450	90.0
	(2.5 0.0)	0	470	0	25	94.0
	(0.0 2.5)	0	3	472	25	94.4
	(2.5 2.5)	0	20	22	458	91.6
	(3.0 0.0)	0	475	2	10	94.0
	(0.0 3.0)	0	3	471	12	94.2
	(3.0 3.0)	0	25	15	460	92.0
Aggregate						92.11
SVM	(0.0,0.0)	480	8	6	6	96.0
	(1.0 0.0)	0	468	12	20	93.6
	(0.0 1.0)	0	9	468	23	93.6
	(1.0 1.0)	0	16	22	462	92.4
	(1.5 0.0)	0	461	6	33	92.2
	(0.0 1.5)	0	3	472	25	94.2
	(1.5 1.5)	0	12	13	475	95.0
	(2.0 0.0)	0	474	8	18	94.8
	(0.0 2.0)	0	3	470	27	94.0
	(2.0 2.0)	0	10	22	468	93.6
	(2.5 0.0)	0	480	2	18	96.0
	(0.0 2.5)	0	9	472	19	94.4
	(2.5 2.5)	0	22	18	460	92.0
	(3.0 0.0)	0	480	2	18	96.0
	(0.0 3.0)	0	2	478	20	95.6
	(3.0 3.0)	0	15	13	472	94.4
Aggregate						94.23
optimized	(0.0,0.0)	500	0	0	0	100
SVM	(1.0 0.0)	0	480	1	19	96.0
	(0.0 1.0)	0	3	484	13	96.8
	(1.0 1.0)	0	8	10	482	96.4
	(1.5 0.0)	0	491	0	9	98.2
	(0.0 1.5)	0	0	488	12	97.6
	(1.5 1.5)	0	5	5	490	98.0
	(2.0 0.0)	0	490	1	9	98.0
	(0.0 2.0)	0	1	488	11	97.6
	(2.0 2.0)	0	7	8	485	97.0
	(2.5 0.0)	0	490	0	10	98.0
	(0.0 2.5)	0	0	492	8	99.4
	(2.5 2.5)	0	5	4	491	99.2
	(3.0 0.0)	0	494	0	6	98.8

TABLE 1: Continued.

Model	Shift magnitude	Correct classification percentage (%)				
		(0 0 0 0)	(0 1 0 0)	(0 0 1 0)	(0 0 0 1)	Percentage
	(0.0 3.0)	0	0	495	5	99.0
	(3.0 3.0)	0	7	6	487	97.4
Aggregate						97.96
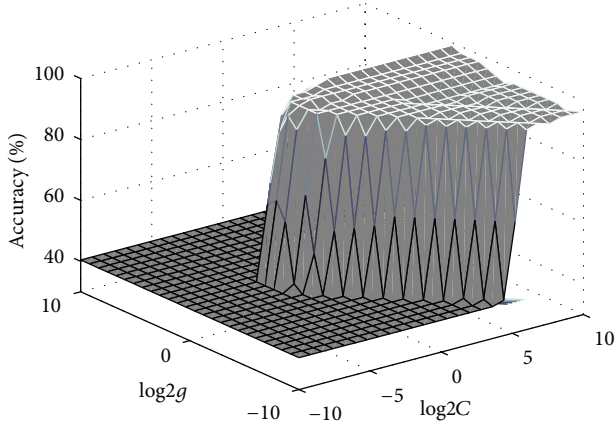
(a) A surface plot of grid search
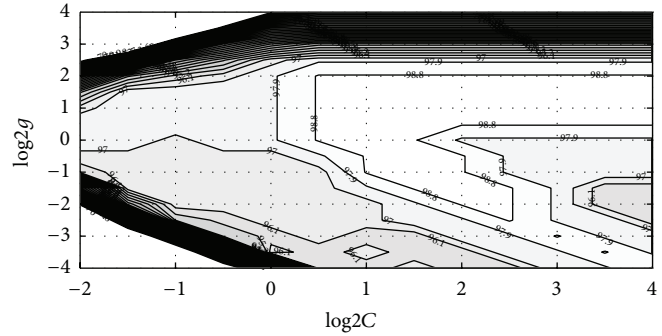
The selection process of optimal parameters using genetic algorithm
(best $C = 1.5136$, $g = 1.0834$)

(b) A contour plots of grid search

FIGURE 3: The selection result of the optimal parameter using genetic algorithm.

data-based BPNN model, the SVM model, and the optimized SVM model trained and tested using raw data. The overall total percentages of correct recognition of the three models are 92.11, 94.23, and 97.96, respectively. The optimized SVM shows the better identification performance in major cases in comparison with those of the BP model and SVM. This indicates that the type of classifier and the choice of nuclear parameter and penalty (cost) parameter can help to affect the recognition performance of classification. Thus, the genetic algorithm can strengthen the pattern feature of the out-of-control signals.

From Table 1, it can be observed that the SVM shows similar performance to that of the BPNN model. Thus, the optimized SVM can be applied in some multivariate processes that have many quality characteristics (e.g., more than three characteristics), which can decrease the complexity of the structure of NNs and which helps to construct SVM with more simple structure. Moreover, the optimized SVM shows good performance in identifying the source(s) of out-of-control signals.

4.3.3. Test 2: Sensitivity Analysis to the Number of the Training Examples. Training sets of different sizes were generated to train optimized. Table 2 presents the test results of optimized SVM trained by different examples. From these results, some conclusions are drawn as follows.

(1) Increasing the training examples can improve the performance of GASVM up to a certain level of accuracy at any given value of shift magnitudes. This could be explained by the fact that with enough large training sets there is a better chance of true representation

of a problem space. However, any further increase of the training size after reaching such limits will not improve the performance of GASVM. Moreover, the larger training set results in higher time cost of training.

(2) Abnormal patterns with small shift magnitudes require more representation of the density distribution by larger training sets.

(3) Abnormal patterns with large shift magnitudes require smaller training data sets as they are easier to identify. This is due to the fact that abnormal patterns with large shift magnitudes have stronger features that separate easier themselves from other abnormal patterns.

5. Conclusion and Further Work

In this paper, we have proposed optimized SVM as an approach to estimate process shift size. Genetic algorithm is used to self-adaptively optimize penalty parameter and kernel parameter; then this optimized SVM modeling was introduced to identify the responsible variable(s) on the occurrence of abnormal pattern. The performances of the proposed approach were evaluated by estimating the various patterns of recognition accuracy using simulation. The proposed optimized SVM approach has been compared to ANN and SVM. Extensive comparisons show that the proposed approach presented in this paper offers a competitive alternative to existing control procedures. Our study reveals that SVM achieves the best performance. The results indicate

Table 2: Accuracy (%) of optimized for different number of the training examples.

Training examples	Shift magnitude					Aggregate
	1.0	1.5	2.0	2.5	3.0	
300	95.78	96.02	96.65	97.13	97.18	96.55
600	95.80	96.67	97.33	97.45	97.86	97.02
900	96.15	97.24	97.84	98.25	98.64	97.62
1200	96.64	97.82	98.06	98.50	98.87	97.98

that SVM is a promising tool in estimating the magnitude of mean shift. Future research can be performed in a number of areas. A further improvement would be to combine several other optimization algorithms in SVM, which could further improve the generalization performance of SVM. A second area would be the construction of input vector based on features extracted from original input observations.

Acknowledgments

References

[1] J. Yu and L. Xi, "A neural network ensemble-based model for on-line monitoring and diagnosis of out-of-control signals in multivariate manufacturing processes," *Expert Systems with Applications*, vol. 36, no. 1, pp. 909–921, 2009.

[2] S. Bersimis, S. Psarakis, and J. Panaretos, "Multivariate statistical process control charts: an overview," *Quality and Reliability Engineering International*, vol. 23, no. 5, pp. 517–543, 2007.

[3] J. Yu, L. Xi, and X. Zhou, "Identifying source(s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 1, pp. 141–152, 2009.

[4] R. Guh, "On-line identification and quantification of mean shifts in bivariate processes using a neural network-based approach," *Quality and Reliability Engineering International*, vol. 23, no. 3, pp. 367–385, 2007.

[5] F. B. Alt, "Multivariate quality control," in *The Encyclopedia of Statistical Sciences*, S. Kotz, N. L. Johnson, and C. R. Read, Eds., pp. 110–122, Wiley, New York, NY, USA, 1985.

[6] R. B. Crosier, "Multivariate generalizations of cumulative sun quality control schemes," *Technometrics*, vol. 30, no. 3, pp. 291–303, 1988.

[7] A. Hassan, M. S. Baksh, A. M. Shaharoun, and H. Jamaluddin, "Improved SPC chart pattern recognition using statistical features," *International Journal of Production Research*, vol. 41, no. 7, pp. 1587–1603, 2003.

[8] R. L. Mason, N. D. Tracy, and J. C. Young, "Decomposition of T^2 for multivariate control chart interpretation," *Journal of Quality Technology*, vol. 27, no. 2, pp. 99–108, 1995.

[9] C. A. Lowry, W. H. Woodall, C. W. Champ, and S. E. Rigdon, "A multivariate exponentially weighted moving average control chart," *Technometrics*, vol. 34, pp. 46–53, 1992.

[10] D. T. Pham and A. B. Chan, "Control chart pattern recognition using a new type of self-organizing neural network," *Proceedings of the Institution of Mechanical Engineers*, vol. 212, no. 2, pp. 115–127, 1998.

[11] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 79, no. 8, pp. 2554–2558, 1982.

[12] C. Wang, W. Kuo, and H. Qi, "An integrated approach for process monitoring using wavelet analysis and competitive neural network," *International Journal of Production Research*, vol. 45, no. 1, pp. 227–244, 2007.

[13] D. T. Pham and E. Oztemel, "Control chart pattern recognition using neural networks," *Journal of Systems Engineering*, vol. 2, no. 4, pp. 256–262, 1992.

[14] D. T. Pham and E. Oztemel, "Control chart pattern recognition using learning vector quantization networks," *International Journal of Production Research*, vol. 32, no. 3, pp. 721–729, 1994.

[15] J. Plummer, "Tighter process control with neural networks," *AI Expert*, vol. 8, pp. 49–55, 1993.

[16] H. B. Hwarng and C. W. Chong, "Detecting process non-randomness through a fast and cumulative learning ART-based pattern recognizer," *International Journal of Production Research*, vol. 33, no. 7, pp. 1817–1833, 1995.

[17] C. Cheng, "A multi-layer neural network model for detecting changes in the process mean," *Computers and Industrial Engineering*, vol. 28, no. 1, pp. 51–61, 1995.

[18] R. Guh, "A hybrid learning-based model for on-line detection and analysis of control chart patterns," *Computers and Industrial Engineering*, vol. 49, no. 1, pp. 35–62, 2005.

[19] T. T. El-Midany, M. A. El-Baz, and M. S. Abd-Elwahed, "A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1035–1042, 2010.

[20] J. Yu, L. Xi, and X. Zhou, "Identifying source(s) of out-of-control signals in multivariate manufacturing processes using selective neural network ensemble," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 1, pp. 141–152, 2009.

[21] M. Salehi, A. Bahreininejad, and I. Nakhai, "On-line analysis of out-of-control signals in multivariate manufacturing processes using a hybrid learning-based model," *Neurocomputing*, vol. 74, no. 12-13, pp. 2083–2095, 2011.

[22] S. B. Kim, W. Jitpitaklert, S. Park, and S. Hwang, "Data mining model-based control charts for multivariate and autocorrelated processes," *Expert Systems with Applications*, vol. 39, no. 2, pp. 2073–2081, 2012.

[23] C. Cheng, P. Chen, and K. Huang, "Estimating the shift size in the process mean with support vector regression and neural networks," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10624–10630, 2011.

[24] D. S. Holmes and A. E. Mergen, "Improving the performance of the T^2 control chart," *Quality Engineering*, vol. 5, pp. 619–625, 1993.

[25] S. Bersimis, S. Psarakis, and J. Panaretos, "Multivariate statistical process control charts: an overview," *Quality and Reliability Engineering International*, vol. 23, no. 5, pp. 517–543, 2007.

[26] J. E. Jackson, "Multivariate quality control," *Communications in Statistics*, vol. 14, pp. 2657–2688, 1985.

[27] C. Wu, Y. Ken, and T. Huang, "Patent classification system using a new hybrid genetic algorithm support vector machine," *Applied Soft Computing Journal*, vol. 10, no. 4, pp. 1164–1177, 2010.

[28] V. Sugumaran, G. R. Sabareesh, and K. I. Ramachandran, "Fault diagnostics of roller bearing using kernel based neighborhood score multi-class support vector machine," *Expert Systems with Applications*, vol. 34, no. 4, pp. 3090–3098, 2008.

[29] E. Pourbasheer, S. Riahi, M. R. Ganjali, and P. Norouzi, "Application of genetic algorithm-support vector machine (GA-SVM) for prediction of BK-channels activity," *European Journal of Medicinal Chemistry*, vol. 44, no. 12, pp. 5023–5028, 2009.

[30] S. Pang, T. Ban, Y. Kadobayashi, and N. Kasabov, "Personalized mode transductive spanning SVM classification tree," *Information Sciences*, vol. 181, no. 11, pp. 2071–2085, 2011.

[31] S. Ch, N. Anand, B. K. Panigrahi, and S. Mathur, "Streamflow forecasting by SVM with quantum behaved particle swarm optimization," *Neurocomputing*, vol. 101, pp. 18–23, 2013.

[32] X. Zhou, W. Jiang, Y. Tian, and Y. Shi, "Kernel subclass convex hull sample selection method for SVM on face recognition," *Neurocomputing*, vol. 73, pp. 2234–2246, 2010.

[33] I. Hmeidi, B. Hawashin, and E. El-Qawasmeh, "Performance of KNN and SVM classifiers on full word Arabic articles," *Advanced Engineering Informatics*, vol. 22, no. 1, pp. 106–111, 2008.

[34] Z. Luo, P. Wang, Y. Li, W. Zhang, W. Tang, and M. Xiang, "Quantum-inspired evolutionary tuning of SVM parameters," *Progress in Natural Science*, vol. 18, no. 4, pp. 475–480, 2008.

[35] S. Ch, N. Anand, B. K. Panigrahi, and S. Mathur, "Streamflow forecasting by SVM with quantum behaved particle swarm optimization," *Neurocomputing*, vol. 101, pp. 18–23, 2013.

[36] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York, NY, USA, 2nd edition, 2001.