

Research Article

Hybrid Support Vector Regression and Autoregressive Integrated Moving Average Models Improved by Particle Swarm Optimization for Property Crime Rates Forecasting with Economic Indicators

Razana Alwee, Siti Mariyam Hj Shamsuddin, and Roselina Sallehuddin

Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Correspondence should be addressed to Razana Alwee; razana@utm.my

Received 2 April 2013; Accepted 8 May 2013

Academic Editors: P. Agarwal, S. Balochian, V. Bhatnagar, and J. Yan

Copyright © 2013 Razana Alwee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Crimes forecasting is an important area in the field of criminology. Linear models, such as regression and econometric models, are commonly applied in crime forecasting. However, in real crimes data, it is common that the data consists of both linear and nonlinear components. A single model may not be sufficient to identify all the characteristics of the data. The purpose of this study is to introduce a hybrid model that combines support vector regression (SVR) and autoregressive integrated moving average (ARIMA) to be applied in crime rates forecasting. SVR is very robust with small training data and high-dimensional problem. Meanwhile, ARIMA has the ability to model several types of time series. However, the accuracy of the SVR model depends on values of its parameters, while ARIMA is not robust to be applied to small data sets. Therefore, to overcome this problem, particle swarm optimization is used to estimate the parameters of the SVR and ARIMA models. The proposed hybrid model is used to forecast the property crime rates of the United State based on economic indicators. The experimental results show that the proposed hybrid model is able to produce more accurate forecasting results as compared to the individual models.

1. Introduction

Quantitative forecasting methods are classified into causal and time series models. The causal models are based on the relationship between the variable to be forecasted and independent variables. A linear relationship is typically used in the causal models. Regression, econometric models, and input-output models are examples of some of the causal models. The time series models are models that use historical data to estimate the future which can be categorized into linear and nonlinear models. Most of the linear models are statistical models such as exponential smoothing, moving average, and autoregressive integrated moving average (ARIMA). However, the nonlinear models consist of statistical models such as bilinear models, the threshold autoregressive (TAR) models, and autoregressive conditional heteroscedastic (ARCH) as well as nonstatistical models such as artificial neural networks (ANN) and support vector regression (SVR).

The causal models such as regression and econometric models are commonly used in crime rates forecasting. The causal models can describe the causal relationship between the crime variable and other explanatory variables. However, the development of the causal models is quite complex and requires theoretical assumptions about the relationship between the explanatory variables. Therefore, the time series model has been considered as a promising alternative tool for crime rates forecasting. The application of time series models for crime rates forecasting is still scarce. Standard time series models usually require a substantial number of observations. However, insufficient crime data makes the standard time series models less suitable for crime rates forecasting. Therefore, a new model that suits small data set is needed to improve the crime rates forecasting performance.

There are two types of time series models, namely, linear and nonlinear models. The linear models can only model the linear relationship, while the nonlinear models only

model the nonlinear relationship. In the literature, there is no single model that can predict well in all conditions. Therefore, many researchers have used a hybridization of linear model with nonlinear model as an approach to time series forecasting [1]. The hybrid linear and nonlinear models are not only capable of modeling the linear and nonlinear relationships, but are also more robust to changes in time series patterns [2]. Artificial neural networks (ANNs) and support vector regression (SVR) are two nonlinear models usually being employed, while ARIMA, seasonal autoregressive integrated moving average (SARIMA), autoregression (AR), exponential smoothing, moving average, and multiple linear regression are usually used to represent linear model in hybridization of linear and nonlinear model. Several examples of hybrid time series models that have been proposed in the literature are ARIMA and ANN [1–16], ARIMA and SVR [17–23], seasonal autoregressive integrated moving average (SARIMA) and SVR [24, 25], autoregression (AR) and ANN [26], exponential smoothing and ANN [27], ARIMA and genetic programming (GP) [28], exponential smoothing, ARIMA and ANN [29], and multiple linear regression (MLR) and ANN [30].

A hybridization of ARIMA and ANN models as linear and nonlinear model is extensively studied by researchers since it produces promising results. However, this hybridization requires sufficient data to produce a good model. Furthermore, ANN models suffer from several problems such as the need for controlling numerous parameters, uncertainty in solution (network weights), and the danger of over fitting. Support vector regression (SVR) was proposed by Vapnik [31] in order to overcome the drawback of ANN. SVR is a nonlinear model to solve regression problems and has been used by researchers as an alternative model to ANN [17–23]. A hybridization of ARIMA and SVR has been successfully applied in time series forecasting such as stock market [17, 20], electricity price [22], and power load [18, 19]. There are four factors that contributed to the success of SVR which are good generalization, global optimal solution, the ability to handle nonlinear problems, and the sparseness of the solution. This has made SVR a robust model to work with small training data, nonlinear, and high-dimensional problems [32]. Despite the advantages, SVR also has some limitations. For example, SVR model parameters must be set correctly as it can affect the regression accuracy. Inappropriate parameters may lead to overfitting or underfitting [33]. Genetic algorithm (GA) and particle swarm optimization (PSO) are among the approaches that have been used by researchers to estimate the SVR parameters. However, PSO is easier to implement as compared to GA, because it does not require evolution operators such as crossover and mutation [34]. As for linear model, ARIMA is the preferred choice by researchers for hybridization with nonlinear model due to its ability to model various smoothing models such as simple autoregressive (AR), a simple moving average (MA), and a combination of AR and MA (ARMA model) [23]. ARIMA has high forecasting performance when the data set is large and linear. However, it is not robust for small data sets and nonlinear data. Therefore, several improvements have been proposed to improve the performance of ARIMA [35].

Asadi et al. [36] suggested the use of PSO to estimate the parameters of ARIMA model for small data sets.

Most hybrid models use a sequence of linear and nonlinear models, but there are also hybrid models that use a sequence of nonlinear and linear models. It depends on which component is more dominant, either linear or nonlinear. The dominant component needs to be modeled first. However, the linear and nonlinear components in the data can have interaction. It is quite difficult to determine which component is more dominant. Modeling linear patterns using linear model will change the nonlinear patterns, and vice versa. However, in comparison with ANN model, the SVR is able to keep the linear components undamaged [22]. A hybridization of SVR and ARIMA model with a sequence of nonlinear and linear models has been found to outperform the existing neural network approaches, traditional ARIMA models, and other hybrid models, such as ARIMA and ANN (with a sequence of linear and nonlinear models and a sequence of nonlinear and linear models), as well as ARIMA and SVR (with a sequence of linear and nonlinear models) [22]. Therefore, this study uses a sequence of nonlinear and linear models by combining SVR with ARIMA model.

The hybrid linear and nonlinear model has never been used for crime rates forecasting. Most of the models used for crime rates forecasting are linear. As for real-world data, crime rates may have linear and nonlinear components, where the use of linear model may not be adequate to forecast the crime rates. Therefore, this study attempts to propose a hybrid time series model that can work well with limited data that consist of both linear and nonlinear components. SVR is used as nonlinear model and ARIMA is employed as linear model. First, the SVR is used to model the nonlinear component. After that, the remaining from the SVR model, which represents the linear component, is modeled by using ARIMA. In order to overcome the drawbacks of the SVR and ARIMA model, particle swarm optimization (PSO) is used to estimate the parameters of SVR and ARIMA models. PSO has the ability to escape from local optima, easy to implement, and has fewer parameters to be adjusted [34]. There are several factors that influence the property crime rates, among these are economic indicators. Therefore, this study uses three economic indicators, namely, unemployment rate, gross domestic product, and consumer price index as inputs to the proposed hybrid model.

The remainder of this study is organized as follows. Related work on crimes and economic indicators is first discussed in Section 2. In Section 3, brief explanations on the support vector regression, ARIMA, particle swarm optimization, and the proposed hybrid model are described. Section 4 describes the data set and model evaluation employed in this study. The determination parameters of model and the analysis of the results are presented in Sections 5 and 6, respectively. Finally, a brief conclusion is drawn in Section 7.

2. Related Work on Crimes and Economic Conditions

Economic conditions are often considered to be related to crimes, especially property crimes. In the literature, many

studies have been done by researchers in order to relate the economic conditions with property crimes. The unemployment rate is often selected by the researchers in their studies to represent the economic conditions [37]. A study using a country level panel data set from Europe found that unemployment has a positive influence on property crimes [38]. Meanwhile, another study based on UK annual regional data has discovered that unemployment is an important explanatory variable for crimes motivated by economic gain [39]. Results produced by some other studies also found significant relationship between the unemployment and property crimes. Among the findings is that motor vehicle theft is significantly associated with the unemployment rate [40] and is also cointegrated with male youth unemployment [41]. Another finding shows that unemployment has a positive effect on burglary, car theft, and bike theft [42].

Unemployment, especially among youth and young adults, is also found to influence crimes. According to a study on the United States arrest data, unemployment has a positive relationship with theft crimes among youth and young adults (16–24 years) [43]. Another study investigated the relationship between crime with male adult (26–64 years) and youth (16–25 years) unemployment in Britain [44]. The results indicate that youth unemployment and adult unemployment are both significantly and positively related to burglary, theft, fraud, and forgery as well as total crime rates.

In addition to unemployment, other economic indicators such as consumer price index, gross domestic product, and consumer sentiment index were also studied by the researchers to examine the relationship between economic conditions with crimes. Several researchers used the consumer price index to measure the inflation [45, 46]. Inflation reduces the purchasing power and increases the cost of living. A study on the impact of inflation rate on crime in the United States using the modified Wald causality test found that the crime rate is co integrated with inflation and unemployment rates [45]. Further, another study which examined the linkages between inflation, unemployment, and crime rates in Malaysia revealed that inflation and unemployment are positively related to the crime rate [46]. Meanwhile, for gross domestic product, a study to explain the relationship between national crime rates with social and economic variables has found that robbery and homicide have significant negative relationship with gross domestic product [47].

As a conclusion, the economic conditions do have an influence on the property crime rates. Therefore, this study attempts to employ the economic conditions to forecast the property crime rates. The economic conditions will be represented by three economic indicators, namely, unemployment rate, consumer price index, and gross domestic product. These economic indicators are used as input to the proposed hybrid model for forecasting property crime rates.

3. Methodology

In this section, explanations on support vector regression, ARIMA, and particle swarm optimization are summarized as a basis to describe the proposed hybrid model.

3.1. Support Vector Regression (SVR). Support vector regression (SVR) is a nonlinear model to solve regression problems. SVR training process is similar to solving the linearly constrained quadratic programming problems that provides a unique optimal value and there is no local minimum problem. The solution is sparse, as only essential data are used to solve the regression function. Lagrangian multipliers are introduced to solve the problem. The SVR model is given by formula [48]

$$f(\mathbf{x}) = (\mathbf{z} \cdot \phi(\mathbf{x})) + b, \tag{1}$$

where \mathbf{z} is weight vector, b is a bias value, and $\phi(\mathbf{x})$ is a kernel function. SVR used ε -insensitivity loss function which can be expressed as formula

$$L_\varepsilon(f(\mathbf{x}) - y) = \begin{cases} |f(\mathbf{x}) - y| - \varepsilon, & \text{if } |f(\mathbf{x}) - y| \geq \varepsilon, \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where ε is the region for ε -insensitivity. Loss is accounted only if the predicted value falls out of the band area. The SVR model can be constructed to minimize the following quadratic programming problem:

$$\begin{aligned} \min: & \quad \frac{1}{2} \mathbf{z}^T \mathbf{z} + C \sum_i (\xi_i + \xi_i^*), \\ \text{subjected to} & \quad y_i - \mathbf{z}^T \mathbf{x}_i - b \leq \varepsilon + \xi_i, \\ & \quad \mathbf{z}^T \mathbf{x}_i + b - y_i \leq \varepsilon + \xi_i^*, \\ & \quad \xi_i, \xi_i^* \geq 0, \end{aligned} \tag{3}$$

where $i = 1, 2, \dots, n$ is the number of training data, $(\xi_i + \xi_i^*)$ is the empirical risk, $(1/2)\mathbf{z}^T \mathbf{z}$ is the structure risk preventing overlearning and lack of applied universality, and C is the regularization parameter. After selecting proper regularization parameter (C), width of band area (ε) and kernel function (K), the optimum of each parameter can be resolved through Lagrange function. The commonly used kernels are linear kernel, polynomial kernel, radial basis function (RBF), or Gaussian kernel and sigmoid kernel. Formulas (4), (5), (6), and (7) are the equation for linear kernel, polynomial kernel, RBF kernel [49], and sigmoid kernel [50], respectively. Consider

linear kernel,

$$K(x_i, x_j) = x_i^T x_j, \tag{4}$$

polynomial kernel,

$$K(x_i, x_j) = (1 + x_i \cdot x_j)^d, \tag{5}$$

RBF kernel,

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \tag{6}$$

sigmoid kernel,

$$K(x_i, x_j) = \tanh[v(x_i, x_j) + \alpha]. \quad (7)$$

The type of kernel function influences the parameters of SVR kernel. The kernel function and parameters of SVR kernel function should be set properly because it can affect the regression accuracy. Inappropriate parameters may lead to over-fitting or under-fitting [33]. This study uses the RBF kernel function because it suits most forecasting problems [51]. The RBF kernel is also effective and has fast training process [52]. For the RBF kernel function, there are three important parameters to be determined [53].

- (i) Regularization parameter C : C is parameter for determining the tradeoff cost between minimizing training error and minimizing model complexity.
- (ii) Kernel parameter (γ): γ represents the parameter of the RBF kernel function.
- (iii) The tube size of ϵ -insensitive loss function (ϵ): ϵ is the approximation accuracy placed on the training data points.

These parameters must be set correctly, in order to produce accurate estimation model. In this study, these parameters are determined through particle swarm optimization (PSO). The explanation on how PSO is used to estimate the parameters of SVR is given in Section 3.3.1.

3.2. Autoregressive Integrated Moving Average (ARIMA). Autoregressive integrated moving average (ARIMA) model was introduced by Box and Jenkins and has become one of the most popular models in forecasting [17]. The ARIMA model is a stochastic model for time series forecasting where the future value of a variable is a linear function of past observations and random errors, expressed as

$$y_t = \theta_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q}, \quad (8)$$

where y_t is the actual value and ε_t is the random error at time t , and ϕ_i ($i = 1, 2, \dots, p$) and θ_j ($j = 0, 1, 2, \dots, q$) are model parameters. Integers, p and q are referred to as order of the model and random errors, ε_t , are assumed to be independently and identically distributed with a mean of zero and a constant variance of σ^2 [2].

ARIMA model is developed using Box-Jenkins methodology that involves the following three iterative steps [2].

(i) *Model Identification.* At this step, data transformation should be done if necessary, to produce a stationary time series. Stationary time series is needed because the ARIMA model is developed with the assumption that the time series is stationary. Mean and autocorrelation structure are constant over time for stationary time series. Therefore, for a time series that exhibit trends and heteroscedasticity, differentiation and power transformation are necessary to change the time series to be stationary. Next, autocorrelation (ACF)

and partial autocorrelation (PACF) are calculated from the data and compared to theoretical autocorrelation and partial autocorrelation for the various ARIMA models to identify the appropriate model form. The selected model is considered as a tentative model. Steps (ii) and (iii) in turn will determine whether the model is adequate [54].

(ii) *Parameter Estimation.* Once the tentative model is identified, parameters in ARIMA model can be estimated using the nonlinear least square procedure.

(iii) *Diagnostic Checking.* The last step in model development is to check whether the model is adequate. Model assumptions about the errors must be met. Several diagnostic statistics and plots of the residual can be done to check the goodness of fit of the tentative model to the historical data. Among plots that can be very useful are histogram, normal probability plot, and time sequence plot. Residual autocorrelations should be small where chi-square test can be used to test the overall model adequacy. However, if the model is considered inadequate, a new tentative model should be identified and steps (ii) and (iii) will be repeated again.

Once a satisfactory model is produced, the three-step development process is no longer repeated and selected model will be used for forecasting purposes. In this study particle swarm optimization (PSO) as suggested by Asadi et al. [36] is used to estimate the parameters of ARIMA model. The explanation on how PSO is used to estimate the parameters of ARIMA model is given in Section 3.3.2.

3.3. Particle Swarm Optimization (PSO). Particle swarm optimization (PSO) is one of stochastic optimization methods introduced by Kennedy and Eberhart [55]. This method is based on the natural evolution process which uses swarming strategies in bird flocking and fish schooling. PSO is a population-based which consists of particles. Initially, the particles are randomly generated. Each particle has a position and velocity, which represents a potential solution to a problem in D -dimensional space. The position and velocity of i th particle are denoted by $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ and $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$, respectively. While solving the search problem, each particle explores the search space by moving in previous direction, its previous best particle ($pbest$), and the best solution for the entire population ($gbest$). The velocity and position of each particle are updated by using the following [56]:

$$v_{ij}(t+1) = w \cdot v_{ij}(t) + c_1 \cdot \text{rand1}_{ij} \cdot (pbest_{ij}(t) - x_{ij}(t)) + c_2 \cdot \text{rand2}_{ij} \cdot (gbest_j(t) - x_{ij}(t)), \quad (9)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1), \quad (10)$$

where $v_{ij}(t)$ is the velocity of i th particle at iteration t , $x_{ij}(t)$ is the position of i th particle at iteration t , $j = 1, 2, \dots, D$, is the dimension of the search space, w is the inertia weight

to balance the global and local search abilities of particles, rand1_{ij} and rand2_{ij} are two uniform random numbers generated independently within the range of $[0, 1]$, c_1 and c_2 are two learning factors which control the influence of the social and cognitive components, $pbest_{ij}(t)$ is the best previous position yielding the best fitness value for i th particle at iteration t , and $gbest_j$ is the global best particle by all particles at iteration t . After changing the position of the particle, the particle's fitness value is evaluated. The $pbest$ and $gbest$ are updated based on the current position of the particles. As this process is repeated, the whole population evolves towards the optimum solution.

The following are the steps in PSO implementation [57].

Step 1. Initialize the positions and velocities of all the particles randomly in the D -dimensional search space by uniform probability distribution function.

Step 2. Evaluate the fitness values of the particles.

Step 3. Update $pbest$ for each particle; if the current fitness value of the particle is better than its $pbest$ value, set the $pbest$ equal to the current position of the particle.

Step 4. Update $gbest$; if the current fitness value of the particle is better than the $gbest$ value, then set $gbest$ equal to the current position of the particle.

Step 5. Update the velocity and position of each particle using (9) and (10), respectively.

Step 6. Repeat Steps 2 to 5, until stopping criteria are met, such as a sufficient good fitness value or a maximum number of iterations.

The explanations on how PSO is used to estimate the parameters of SVR and ARIMA models are given in Sections 3.3.1 and 3.3.2, respectively.

3.3.1. PSO for SVR Parameters Estimation (PSOSVR). Since there are three parameters to be estimated, the i th particle is represented by the three-dimensional vectors, $X_i = (x_{i1}, x_{i2}, x_{i3})$ and $V_i = (v_{i1}, v_{i2}, v_{i3})$, where the first, second, and third dimensions of the vectors refer to C , γ , and ϵ , respectively. In this study, the fitness is defined by k -fold cross-validation, where $k = 5$. In k -fold cross-validation, the training data set is divided into k subsets of equal size. One subset is used for validation. The regression function is built with a given set of parameters (C, γ, ϵ) using the remaining $k-1$ subsets. The performance of the parameter set is measured by the root mean square error (RMSE) on the validation set. Each subset is used once for validation and the process is repeated k times. The average of RMSE on the validation set from 5 trials is used as a measure of fitness. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (11)$$

where n is the number of validation data; y_t is the actual value and \hat{y}_t is the predicted value.

3.3.2. PSO for ARIMA Parameters Estimation (PSOARIMA). A hybrid model of PSO and ARIMA was proposed by Asadi et al. [36] to estimate the parameters of ARIMA model. This method is efficient for cases where inadequate historical data is available. The implementation of this method involves two main steps. First, an ARIMA model is generated by applying the Box and Jenkins method. Second, the PSO model is used to estimate the ARIMA parameters. The data set is divided into training and testing data set. The training data set is used to estimate the ARIMA model. However, testing data set is used to evaluate the estimation results. In this study, the fitness is defined by sum square error (SSE) as follows:

$$\text{SSE} = \sum_{t=1}^n (y_t - \hat{y}_t)^2, \quad (12)$$

where n is the number of training data; y_t is the actual value and \hat{y}_t is the predicted value.

3.4. The Proposed Model. The proposed hybrid model consists of a nonlinear model, SVR, and a linear model, ARIMA. According to Zhang [2], it is reasonable to consider a time series as the composition of a linear autocorrelation structure and a nonlinear component, as

$$y_t = N_t + L_t, \quad (13)$$

where N_t denotes the nonlinear component and L_t denotes the linear component. These two components are estimated from data using the following two steps. First, SVR model is used to model nonlinear components in the data. Second, the residual from the nonlinear model is modeled using linear model, ARIMA. Let r_t denote the residual which is represented by

$$r_t = y_t - \hat{N}_t. \quad (14)$$

The residual represents linear components that cannot be modeled by SVR model. The SVR and ARIMA parameters are estimated by applying PSO, as described previously. Forecasting results from SVR and ARIMA models will be combined to represent the forecasting results of the proposed hybrid model. The combined forecast is shown by the formula (15)

$$\hat{y}_t = \hat{N}_t + \hat{L}_t. \quad (15)$$

Figure 1 shows the flowchart for the proposed hybrid model, PSOSVR_PSOARIMA.

4. Data Set and Model Evaluation

This section describes the data set used and the model evaluation carried out in this study.

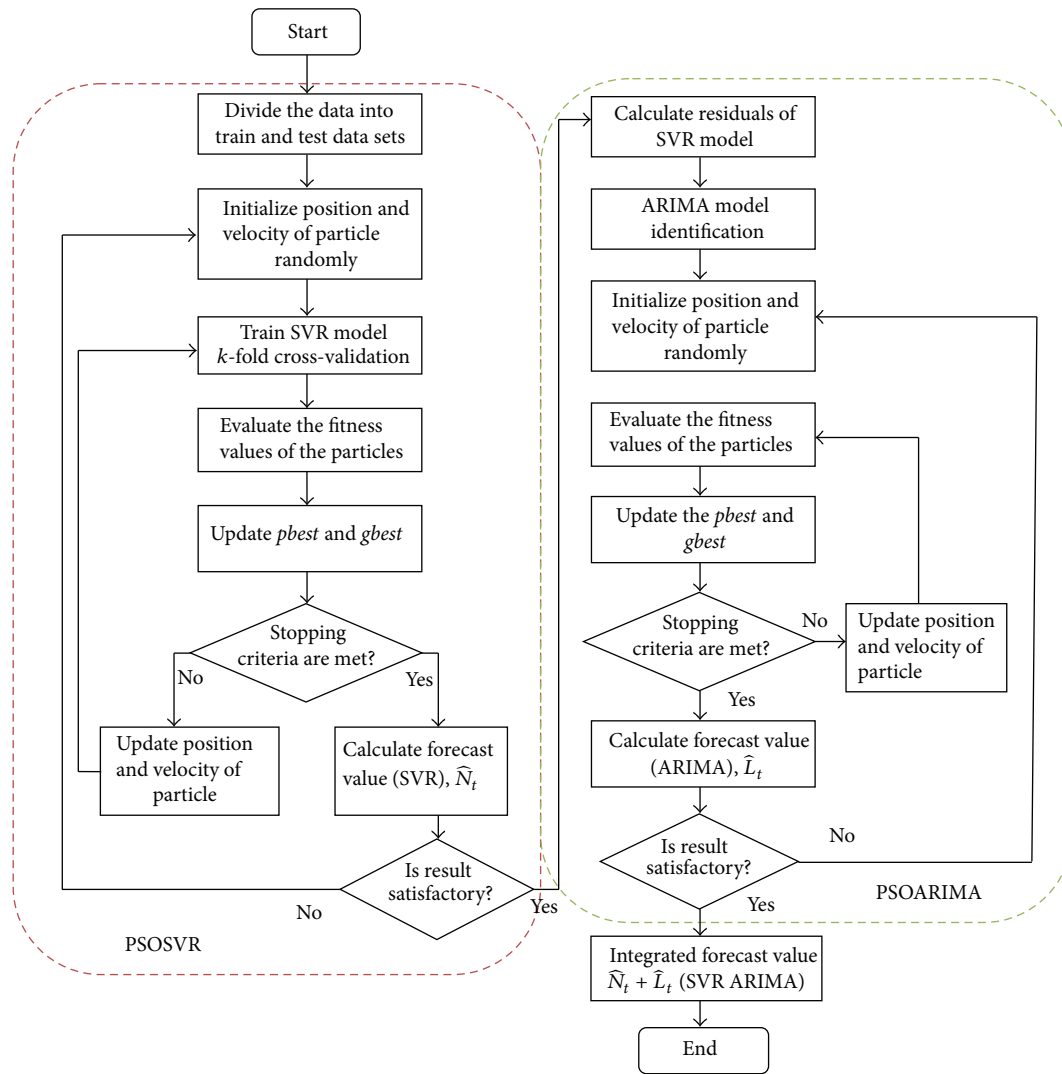


FIGURE 1: The Flow chart for the proposed hybrid PSOSVR and PSOARIMA model.

4.1. *Data Set.* This study uses annual data of property crime rates, consumer price index for all urban consumers (Apparel), gross domestic product in United States natural log of billions of chained 2005 US Dollars, and unemployment rate (20 to 24 years) from 1960 to 2009 in United States. The crime rates are obtained from the Uniform Crime Reporting Statistics website (<http://www.ucrdatatool.gov/>), while economic indicators data are available on Economic Research Federal Reserve Bank of St. Louis website (<http://research.stlouisfed.org/>). Property crime data comprised property crime rate, vehicle rate, larceny-theft rate, and burglary rate. In addition to economic indicators, two crime indicators are also used in the model, which are one-year lagged property crime rate (PCR) and one-year lagged robbery rate. Data is divided into training and test data sets. The training data set is used to develop the models while the test data set is used to evaluate the forecasting performance of the models. In this study, 90 percent of the data will be used as

training (1961 to 2004) and 10 percent is as test data set (2005 to 2009).

4.2. *Model Evaluation.* The performance of the proposed hybrid model is evaluated using the test data set. The forecasting performance of the proposed hybrid model is evaluated using four types of evaluations.

(i) *Descriptive Statistics.* Graph of actual values and forecasting of testing data set is plotted in order to see the pattern of model predictions compared with the actual data patterns. A box plot diagram is used to check the error values. Box plot is used to see the dispersion of error values such as the position of median whether it is close to zero, and to ensure that there are no extreme values in error.

(ii) *Quantitative Error Measurements.* Four types of quantitative error measurements are conducted, namely, root mean

square error (RMSE), mean square error (MSE), mean absolute percentage error (MAPE), and mean absolute deviation (MAD). Formulas (16), (17), (18), and (19) are the equation for RMSE, MSE, MAPE, and MAD, respectively. Consider

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}, \quad (16)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2, \quad (17)$$

$$MAPE = \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times \frac{100}{n}, \quad (18)$$

$$MAD = \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{n}, \quad (19)$$

where n is the number of test data; y_t is the actual value and \hat{y}_t is the predicted value.

(iii) *Model Comparative Performance.* The performance of the proposed model is compared with individual model, PSOSVR, ARIMA, and PSOARIMA. A comparison with the hybrid model PSOSVR_ARIMA is also done to see the effects of PSO on the ARIMA model in the hybridization of SVR and ARIMA model.

(iv) *Hypothesis Test.* The hypothesis test is performed to prove that there is no significant means difference between the forecasting values and the actual data. Paired sample t -test is used in this study. There are two hypotheses, the null hypothesis (H_0) and the alternative hypothesis (H_1). Let μ_1 be the mean of actual data, μ_2 the mean of forecasting values of the forecasting model, and $(\mu_1 - \mu_2 = \mu_D)$ the difference of means. The hypothesis is that

$$\begin{aligned} H_0: \mu_D &= 0, \\ H_1: \mu_D &\neq 0. \end{aligned} \quad (20)$$

The test statistic is shown by formula (21), which is t distributed with $n_D - 1$ degrees of freedom. Consider

$$t = \frac{\bar{y}_D - \mu_D}{s_D / \sqrt{n_D}}, \quad (21)$$

where \bar{y}_D is a sample mean difference, s_D is a sample standard deviation of the difference, and n_D is a sample size. The mean of forecasting values is equal to the mean of actual values, if the hypothesis test fails to reject the null hypothesis. It indicates that the model is appropriately used as forecasting model since it represents the real situation.

5. Determination Parameters of Model

In PSO algorithm, population size is set to 5, maximum number of iterations is set to 50 and the value of c_1, c_2 is set

TABLE 1: The range values for the parameters (C, γ, ϵ).

Parameters	Range values
C	2^{-1} to 2^7
γ	2^{-4} to 2^2
ϵ	0.01 to 0.05

to 1. The inertia weight, w , initially is set to 1.4, and its value is decreased along with the iterations according to [58]

$$Weight = \frac{(Weight - 0.4) * (Maxiter - iter)}{Maxiter} + 0.4, \quad (22)$$

where $Maxiter$ is the maximum iteration and $iter$ is the current iteration. The range values of SVR parameters (C, γ, ϵ) used in this study are shown in Table 1. Meanwhile, the searching range for ARIMA parameters is between -100 and 100 .

The optimum parameter for SVR model obtained by PSO is $(C, \gamma, \epsilon) = (60.1924, 0.0625, 0.01)$. SVR model was developed using the optimum parameter which is then used to forecast the property crime rates using the training data and test data. After that, the residual which represents the difference between the actual value and prediction value was calculated. The residual of the training data was used to build the ARIMA model. Based on ACF and PACF, the appropriate model for the residuals is ARIMA (2,0,0), represented as

$$\hat{e}_t = a_0 + a_1 e_{t-1} + a_2 e_{t-2}, \quad (23)$$

where \hat{e}_t is the forecast for period t , e_{t-1} and e_{t-2} are the residuals at time lags $t - 1$ and $t - 2$, and a_0, a_1, a_2 are the coefficients to be estimated. Since the data varies around zero, the coefficient a_0 is not required. Therefore, the coefficients to be estimated are a_1 and a_2 . PSO was applied to estimate the value of the coefficients. The i th particle is represented by the two-dimensional vectors, $X_i = (x_{i1}, x_{i2})$ and $V_i = (v_{i1}, v_{i2})$, where the first and second dimensions of the vectors refer to a_1 and a_2 , respectively. After using the PSO, the coefficients were estimated and the ARIMA model is shown as

$$\hat{e}_t = 0.3445e_{t-1} - 0.1916e_{t-2}. \quad (24)$$

Next, the ARIMA model obtained is used to forecast the residual from the SVR model. The forecast results from SVR and ARIMA models are combined, to represent forecast results of the proposed hybrid model.

6. Analysis of the Results

Table 2 and Figure 2 show the comparison of actual values and forecast values of property crime rates for test data set. PSOSVR_PSOARIMA, PSOSVR, and PSOSVR_ARIMA were found to predict closer to the actual value and have a similar pattern with the actual data. Meanwhile, ARIMA and PSOARIMA show unsatisfactory forecasting performance with the predicted values slightly higher than the actual value. Figure 3 shows box plot of forecast errors for testing data set. Based on the box plot, there is an outlier in the error of

TABLE 2: Comparison of actual value and forecast value of property crime rates.

Year	Actual value of property crime rates	Forecast value of property crime rates				
		Individual models		Hybrid models		
		ARIMA	PSOARIMA	PSOSVR	PSOSVR ARIMA	PSOSVR PSOARIMA
2005	3431.5	3450.674	3446.673	3432.788	3415.503	3427.515
2006	3334.5	3383.523	3378.434	3346.533	3345.65	3346.024
2007	3263.5	3277.728	3279.746	3255.626	3250.134	3251.728
2008	3211.5	3230.002	3225.111	3188.747	3200.334	3193.765
2009	3036.1	3187.335	3189.071	3030.249	3037.259	3036.579

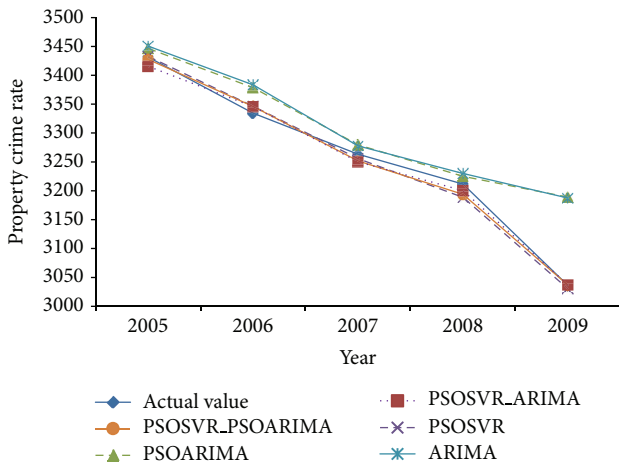


FIGURE 2: Forecasting of test data set.

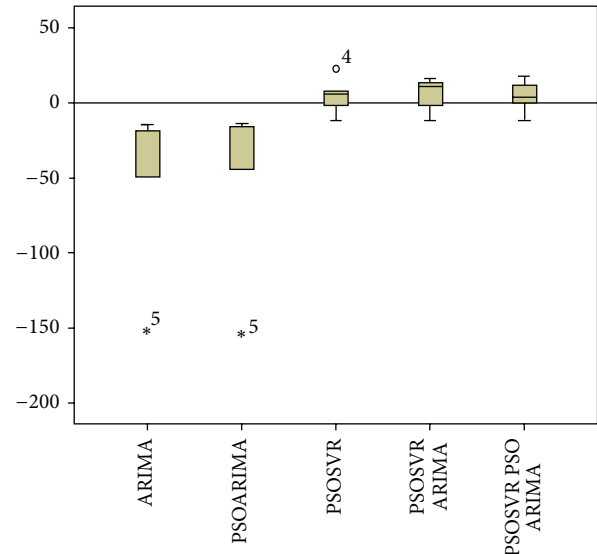


FIGURE 3: Forecasting Errors of test data set.

the PSOSVR model, while for ARIMA and PSOARIMA models, respectively, there is an extreme error value. The hybrid models, PSOSVR_PSOARIMA and PSOSVR_ARIMA, show better forecasting performance than the individual models, with no outliers or extreme error values. The proposed hybrid model was found to have the least forecast errors with a median closer to zero as compared to other forecasting models.

Table 3 shows the RMSE, MSE, MAPE, and MAD of the proposed hybrid model in comparison to the individual models, ARIMA, PSOARIMA, PSOSVR, and hybrid model, PSOSVR_ARIMA. The linear models, ARIMA and PSOARIMA, show poor performance with a relatively large error values in comparison to nonlinear model, PSOSVR and hybrid models. On the other hand, the nonlinear model, PSOSVR, shows a comparable performance to hybrid models with a slight difference in error values. The proposed hybrid model is found to have smaller errors than the individual models and PSOSVR_ARIMA. In addition, PSO is found to be able to improve the performance of the ARIMA model in the hybridization of SVR and ARIMA model. The results have shown that the use of PSO on ARIMA model is capable of improving the performance of the hybrid model PSOSVR_ARIMA. Therefore, PSOARIMA is suitable to be applied on small data set.

Table 4 shows the results of paired samples *t*-test. Paired samples *t*-test is used to compare the actual data with the

TABLE 3: Comparison of errors.

Model	RMSE	MSE	MAPE	MAD
ARIMA	72.371	5237.568	1.604	50.433
PSOARIMA	72.125	5201.967	1.544	48.387
PSOSVR	12.332	152.077	0.308	9.960
PSOSVR_ARIMA	11.704	136.980	0.319	10.568
PSOSVR_PSOARIMA	10.973	120.408	0.278	9.099

five forecasting models that have been developed, namely, ARIMA, PSOARIMA, PSOSVR, PSOSVR_ARIMA, and PSOSVR_PSOARIMA. The test is to ensure that there is no statistically significant difference of means between the actual data and the forecast values from the forecasting models. The results show that between the actual data and the forecast values from the forecasting models, the *P* value > 0.05 (0.124, 0.145, 0.463, 0.333, and 0.443). The difference between the upper and lower values for 95% interval is range between negatives and positives values. The results indicate that the hypothesis test fails to reject the null hypothesis. This implies that there is no statistically significant difference in the means between the actual data and the forecasting models. However, PSOSVR_PSOARIMA shows the smallest mean, standard deviation, and standard error mean as compared to other

TABLE 4: Paired samples test.

	Mean	Std. deviation	Std. error mean	Paired differences		t	df	Sig. (2-tailed)
				95% confidence interval of the difference				
				Lower	Upper			
Pair 1 PCR-ARIMA	-50.43259	58.03148	25.95247	-122.48819	21.62301	-1.943	4	0.124
Pair 2 PCR-PSOARIMA	-48.38690	59.74264	26.74264	-122.63637	25.86257	-1.809	4	0.145
Pair 3 PCR-PSOSVR	4.63141	5.71462	5.71462	-11.23491	20.49773	0.810	4	0.463
Pair 4 PCR-PSOSVR_ARIMA	5.64408	5.12651	5.12651	-8.58939	19.87756	1.101	4	0.333
Pair 5 PCR-PSOSVR_PSOARIMA	4.29777	5.04821	5.04821	-9.71831	18.31385	0.851	4	0.443

Legend: Pair 1: actual data and ARIMA; Pair 2: actual data and PSOARIMA; Pair 3: actual data and PSOSVR; Pair 4: actual data and PSOSVR_ARIMA; Pair 5: actual data and PSOSVR_PSOARIMA.

forecasting models. Therefore, the assumption can be made that PSOSVR_PSOARIMA is able to give better forecasting performance in comparison to other forecasting models.

Based on model evaluation that has been carried out, we can conclude that the individual models, ARIMA, PSOARIMA, and PSOSVR, are not sufficient to model the property crime rates. The hybrid models are more suitable to be employed as forecasting model for property crime rates, where the proposed hybrid model shows the best forecasting performance. These results indicate that there are linear and nonlinear components in property crime rates. Therefore, the use of linear and nonlinear models individually is not sufficient for modeling the property crime rates. However, the optimal parameters are very important for ensuring the accuracy of SVR models. The use of PSO has facilitated the searching process for the optimal parameters of SVR model, thus able to produce more accurate models. Results of this study also demonstrated that the use of PSO in estimating the parameters of ARIMA model was able to improve the accuracy of the ARIMA model for small data sets.

7. Conclusions

This paper proposes a time series model for crime rates forecasting. The proposed model is a hybrid model that combines the nonlinear model, SVR, with linear model, ARIMA. The proposed model was used to predict the property crime rates. The PSO is used to estimate the parameters of SVR and ARIMA models. Economic indicators are used as inputs to the proposed model. The economic indicators used are gross domestic product, unemployment rate, and consumer price index. The experimental results have found that the proposed model, PSOSVR_PSOARIMA, can produce smaller errors as compared to the individual models and hybrid model, PSOSVR_ARIMA. In conclusion, it can be concluded that the proposed hybrid model is an acceptable model to be applied in the crime rates forecasting.

Acknowledgments

This work is supported by Ministry of Higher Education (MOHE) under Research University Grant (03H72). Authors would like to thank Research Management Centre (RMC),

Universiti Teknologi Malaysia, for the research activities and *Soft Computing Research Group* (SCRG) for the support and motivation in making this study a success.

References

- [1] L. Qu, Y. Chen, and Z. Liu, "Time series forecasting model with error correction by structure adaptive RBF neural network," in *Proceedings of the 6th World Congress on Intelligent Control and Automation (WCICA '06)*, pp. 6831–6835, June 2006.
- [2] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.
- [3] D. K. Wedding and K. J. Cios, "Time series forecasting by combining RBF networks, certainty factors, and the Box-Jenkins model," *Neurocomputing*, vol. 10, no. 2, pp. 149–168, 1996.
- [4] J. V. Hansen and R. D. Nelson, "Time-series analysis with neural networks and ARIMA-neural network hybrids," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 15, no. 3, pp. 315–330, 2003.
- [5] I. Rojas, O. Valenzuela, F. Rojas et al., "Soft-computing techniques and ARMA model for time series prediction," *Neurocomputing*, vol. 71, no. 4–6, pp. 519–537, 2008.
- [6] D. Zeng, J. Xu, J. Gu, L. Liu, and G. Xu, "Short term traffic flow prediction using hybrid ARIMA and ANN models," in *Proceedings of the Workshop on Power Electronics and Intelligent Transportation System (PEITS '08)*, pp. 621–625, August 2008.
- [7] R. Sallehuddin, S. M. Shamsuddin, and S. Z. M. Hashim, "Hybridization model of linear and nonlinear time series data for forecasting," in *Proceedings of the 2nd Asia International Conference on Modelling and Simulation (AMS '08)*, pp. 597–602, May 2008.
- [8] L. A. Díaz-Robles, J. C. Ortega, J. S. Fu et al., "A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: the case of Temuco, Chile," *Atmospheric Environment*, vol. 42, no. 35, pp. 8331–8340, 2008.
- [9] Y. Xiaojian and Z. Jiaping, "A comparison of hybrid ARMA-Elman models with single models for forecasting interest rates," in *Proceedings of the 2nd International Symposium on Intelligent Information Technology Application (IITA '08)*, pp. 985–989, December 2008.
- [10] R. Sallehuddin and S. M. H. Shamsuddin, "Hybrid grey relational artificial neural network and auto regressive integrated moving average model for forecasting time-series data," *Applied Artificial Intelligence*, vol. 23, no. 5, pp. 443–486, 2009.

- [11] R. Sallehuddin, *Hybridization of nonlinear and linear model for time series forecasting [Ph.D. thesis]*, Universiti Teknologi Malaysia, Johor Bahru, Malaysia, 2010.
- [12] C. H. Aladag, E. Egrioglu, and C. Kadir, "Forecasting nonlinear time series with a hybrid methodology," *Applied Mathematics Letters*, vol. 22, no. 9, pp. 1467–1470, 2009.
- [13] D. Ömer Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 4, pp. 586–594, 2010.
- [14] M. Shafie-Khah, M. P. Moghaddam, and M. K. Sheikh-El-Eslami, "Price forecasting of day-ahead electricity markets using a hybrid forecast method," *Energy Conversion and Management*, vol. 52, no. 5, pp. 2165–2169, 2011.
- [15] R. Askari Moghadam and M. Keshmirpour, "Hybrid ARIMA and neural network model for measurement estimation in energy-efficient wireless sensor networks," in *Proceedings of the International Conference on Informatics Engineering and Information Science (ICIEIS '11)*, pp. 35–48, 2011.
- [16] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models for time series forecasting," *Applied Soft Computing Journal*, vol. 11, no. 2, pp. 2664–2675, 2011.
- [17] P. F. Pai and C. S. Lin, "A hybrid ARIMA and support vector machines model in stock price forecasting," *Omega*, vol. 33, no. 6, pp. 497–505, 2005.
- [18] H. Yujun, Z. Youchan, and D. Dongxing, "Research on hybrid ARIMA and support vector machine model in short term load forecasting," in *Proceedings of the 6th International Conference on Intelligent Systems Design and Applications (ISDA '06)*, pp. 804–808, October 2006.
- [19] K. Feng and W. Xiaojuan, "Time series forecasting model with error correction by structure adaptive support vector machine," in *Proceedings of the International Conference on Computer Science and Software Engineering (CSSE '08)*, pp. 1067–1070, December 2008.
- [20] D. Y. Zhang, H. W. Song, and P. Chen, "Stock market forecasting model based on a hybrid ARMA and support vector machines," in *Proceeding of 15th Annual International Conference on Management Science and Engineering (ICMSE '08)*, pp. 1312–1317, September 2008.
- [21] L. Xiang, G. J. Tang, and C. Zhang, "Simulation of time series prediction based on hybrid support vector regression," in *Proceedings of the 4th International Conference on Natural Computation (ICNC '08)*, pp. 167–171, October 2008.
- [22] J. Che and J. Wang, "Short-term electricity prices forecasting based on support vector regression and Auto-regressive integrated moving average modeling," *Energy Conversion and Management*, vol. 51, no. 10, pp. 1911–1917, 2010.
- [23] J. H. Lo, "A study of applying ARIMA and SVM model to software reliability prediction," in *Proceedings of the International Conference on Uncertainty Reasoning and Knowledge Engineering (URKE '11)*, pp. 141–144, 2011.
- [24] K. Y. Chen and C. H. Wang, "A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan," *Expert Systems with Applications*, vol. 32, no. 1, pp. 254–264, 2007.
- [25] Y. Bao, D. Yi, T. Xiong, Z. Hu, and S. Zheng, "A comparative study on hybrid linear and nonlinear modeling framework for air passenger traffic forecasting," *Advances in Information Sciences and Service Sciences*, vol. 3, no. 5, pp. 243–254, 2011.
- [26] F. Pan, H. Zhang, and M. Xia, "A hybrid time-series forecasting model using extreme learning machines," in *Proceeding of the 2nd International Conference on Intelligent Computing Technology and Automation (ICICTA '09)*, pp. 933–936, October 2009.
- [27] K. K. Lai, L. Yu, S. Wang, and W. Huang, "Hybridizing exponential smoothing and neural network for financial time series predication," in *Proceedings of the 6th International Conference on Computational Science (ICCS '06)*, pp. 493–500, 2006.
- [28] Y. S. Lee and L. I. Tong, "Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 66–72, 2011.
- [29] J. J. Wang, J. Z. Wang, Z. G. Zhang, and S. P. Guo, "Stock index forecasting based on a hybrid model," *Omega*, vol. 40, no. 6, pp. 758–766, 2012.
- [30] M. Khashei, A. Zeinal Hamadani, and M. Bijari, "A novel hybrid classification model of artificial neural networks and multiple linear regression models," *Expert Systems With Applications*, vol. 39, no. 3, pp. 2606–2620, 2012.
- [31] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [32] Z. Ding, "Application of support vector machine regression in stock price forecasting," in *Proceeding of the International Conference on Business, Economics, and Financial Sciences, Management (BEFM '11)*, pp. 359–365, 2012.
- [33] J. Wu and L. Jin, "Daily rainfall prediction with SVR using a novel hybrid PSO-SA algorithms," in *Proceeding of the 2nd International Conference on High-Performance Networking, Computing and Communications Systems (ICHCC '11)*, pp. 508–515, 2011.
- [34] R. Liao, H. Zheng, S. Grzybowski, and L. Yang, "Particle swarm optimization-least squares support vector regression based forecasting model on dissolved gases in oil-filled power transformers," *Electric Power Systems Research*, vol. 81, no. 12, pp. 2074–2080, 2011.
- [35] Y. S. Lee and L. I. Tong, "Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 66–72, 2011.
- [36] S. Asadi, A. Tavakoli, and S. R. Hejazi, "A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization," *Expert Systems with Applications*, vol. 39, no. 5, pp. 5332–5337, 2012.
- [37] R. Rosenfeld and R. Fornango, "The impact of economic conditions on robbery and property crime: the role of consumer sentiment," *Criminology*, vol. 45, no. 4, pp. 735–769, 2007.
- [38] D. T. Altindag, "Crime and unemployment: evidence from Europe," *International Review of Law and Economics*, vol. 32, no. 1, pp. 145–157, 2012.
- [39] D. Wu and Z. Wu, "Crime, inequality and unemployment in England and Wales," *Applied Economics*, vol. 44, no. 29, pp. 3765–3775, 2012.
- [40] D. L. Yearwood and G. Koinis, "Revisiting property crime and economic conditions: an exploratory study to identify predictive indicators beyond unemployment rates," *Social Science Journal*, vol. 48, no. 1, pp. 145–158, 2011.
- [41] P. K. Narayan and R. Smyth, "Crime rates, male youth unemployment and real income in Australia: evidence from Granger causality tests," *Applied Economics*, vol. 36, no. 18, pp. 2079–2095, 2004.
- [42] K. Edmark, "Unemployment and crime: is there a connection?" *Scandinavian Journal of Economics*, vol. 107, no. 2, pp. 353–373, 2005.

- [43] C. L. Britt, "Reconsidering the unemployment and crime relationship: variation by age group and historical period," *Journal of Quantitative Criminology*, vol. 13, no. 4, pp. 405–428, 1997.
- [44] F. Carmichael and R. Ward, "Male unemployment and crime in England and Wales," *Economics Letters*, vol. 73, no. 1, pp. 111–115, 2001.
- [45] C. F. Tang and H. H. Lean, "Will inflation increase crime rate? New evidence from bounds and modified wald tests," *Global Crime*, vol. 8, no. 4, pp. 311–323, 2007.
- [46] C. F. Tang, "The linkages among inflation, unemployment and crime rates in Malaysia," *International Journal of Economics and Management*, vol. 3, no. 1, pp. 50–61, 2009.
- [47] P. Fajnzylber, D. Lederman, and N. Loayza, "What causes violent crime?" *European Economic Review*, vol. 46, no. 7, pp. 1323–1357, 2002.
- [48] C. J. Lu, C. H. Chang, C. Y. Chen, C. C. Chiu, and T. S. Lee, "Stock index prediction: a comparison of MARS, BPN and SVR in an emerging market," in *Proceeding of the IEEE International Conference on Industrial Engineering and Engineering Management (IEEM '09)*, pp. 2343–2347, December 2009.
- [49] I. D. Lins, M. D. C. Moura, E. Zio, and E. L. Drogue, "A particle swarm-optimized support vector machine for reliability prediction," *Quality and Reliability Engineering International*, vol. 28, no. 2, pp. 141–158, 2012.
- [50] Y. Hu, C. Wu, and H. Liu, "Prediction of passenger flow on the highway based on the least square support vector machine," *Transport*, vol. 26, no. 2, pp. 197–203, 2011.
- [51] J. Wu and E. Chen, "A novel hybrid particle swarm optimization for feature selection and kernel optimization in support vector regression," in *Proceedings of the International Conference on Computational Intelligence and Security (CIS '10)*, pp. 189–194, December 2010.
- [52] N. Xin, X. Gu, H. Wu, Y. Hu, and Z. Yang, "Application of genetic algorithm-support vector regression (GA-SVR) for quantitative analysis of herbal medicines," *Journal of Chemometrics*, vol. 26, no. 7, pp. 353–360, 2012.
- [53] S. Zhao and L. Wang, "Support vector regression based on particle swarm optimization for rainfall forecasting," in *Proceeding of the 3rd International Joint Conference on Computational Sciences and Optimization (CSO '10)*, pp. 484–487, May 2010.
- [54] J. E. Hanke and D. W. Wichern, *Business Forecasting*. Prentice Hall, Englewood Cliffs, NJ, USA, 2009.
- [55] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, December 1995.
- [56] H. Wang, H. Sun, C. Li et al., "Diversity enhanced particle swarm optimization with neighborhood search," *Information Sciences*, vol. 223, pp. 119–135, 2013.
- [57] G. He and N. J. Huang, "A modified particle swarm optimization algorithm with applications," *Applied Mathematics and Computation*, vol. 219, no. 3, pp. 1053–1060, 2012.
- [58] X. Wang, J. Yang, X. Teng, W. Xia, and R. Jensen, "Feature selection based on rough sets and particle swarm optimization," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 459–471, 2007.