Energy and Buildings 90 (2015) 76–84

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Prediction of building energy consumption using an improved real coded genetic algorithm based least squares support vector machine approach

Hyun Chul Jung [a], Jin Sung Kim [b], Hoon Heo [a,*]

[a] Dept. of Control and Instrumentation Eng., Korea University, Seoul 137-701, Republic of Korea
[b] Energy Solution Lab, Living and Energy R&D Center, LG Electronics, Seoul 153-802, Republic of Korea

ARTICLE INFO

Article history:
Received 10 April 2013
Received in revised form 9 December 2014
Accepted 16 December 2014
Available online 6 January 2015

Keywords:
Least-square support vector machines
Genetic algorithm
Golden section method
Building energy consumption
Prediction

ABSTRACT

The least-squares support vector machine (LSSVM) strategy has played a crucial role in the forecasting of building energy consumption owing to its remarkable nonlinear mapping capabilities in prediction. In order to build an effective LSSVM method, its two free parameters, the regularization parameter and the kernel parameter, must be selected carefully. However, LSSVM using a conventional real-coded genetic algorithm (RCGA) or differential evolution algorithm (DEA) for determining the aforementioned two parameters consumes excessive amounts of computation time. In this study, a novel LSSVM for effective prediction of daily building energy consumption is designed by utilizing a hybrid of the direct search optimization (DSO) algorithm and RCGA, called the DSORCGA. The proposed DSORCGA differs from the conventional RCGA in terms of the reproduction operator and the crossover operator, and is used to optimize free parameters of LSSVM for faster computation speed and higher predictive accuracy. Finally, in a MATLAB2010a environment, actual building energy consumption data are adopted to run the proposed DSORCGA-LSSVM and conventional RCGA-LSSVM and DEA-LSSVM. Further, the simulation results in the target period are compared with those of actual recorded energy consumption data, and improvement in computation time is revealed via numerical simulation.

1. Introduction

In recent years, energy conservation has received wide regional and national attention, with prediction of the energy consumption of a building becoming one of the most important individual and societal energy conservation issues. However, predicting building energy consumption is complicated because energy consumption data present nonlinear patterns caused by such factors as climate, settled population, and seasonal changes. Therefore, accurate and rapid prediction of building energy consumption is challenging.

Several prediction strategies have been adopted, such as time-series analysis [10], Fourier methods [11], and artificial neural networks methods. Because of the strong nonlinear mapping behavior of energy consumption data, some researchers have begun using artificial intelligent techniques to focus on the prediction of building energy consumption. In particular, the artificial neural network (ANN) theory is widely used in nonlinear time series prediction, which has been applied generally to predict building energy consumption [1–9]. However, conventional ANN has several drawbacks, such as the need for a large number of controlling parameters, difficulty in obtaining stable solutions, and thus lack of generalizability. Furthermore, its training procedure is time consuming and can be trapped in local minima [12].

Recently, support vector machines (SVM), developed by Vapnik [13] in 1998, have received increasing attention. SVMs are based on the structural risk minimization principle (SRM) and have been shown to be superior to the traditional empirical risk minimization principle (ERM) employed in conventional neural networks. SRM seeks to minimize the upper bound of the generalization error, which consists of the sum of training errors. In other words, SVM overcomes the over-fitting problem and achieves improved generalization performance by minimizing the structural risk, instead of the empirical risk on which previous techniques were based [14]. Further, many studies apply SVM regression to building energy consumption prediction [15,16] with the prediction accuracy outperforming that of other prediction methods. The SVM procedure of the learning problem requires quadratic programming (QP) with a linear constraint. However, the size of a matrix relevant to the QP

* Corresponding author. Tel.: +82 2 3290 3974; fax: +82 2 929 7808.
E-mail address: heo257@korea.ac.kr (H. Heo).

http://dx.doi.org/10.1016/j.enbuild.2014.12.029
0378-7788/

problem is directly proportional to the number of training points associated with the computation speed.

In order to reduce the complexity of optimization processes, Suykens and Vandewalle proposed a reformulated version of SVM, called least-squares SVM (LSSVM) [17]. LSSVM uses the square errors instead of the nonnegative errors in the optimization problem and equality constraints instead of inequality constraints, in contrast to conventional SVM. In other words, LSSVM solves a set of linear equations instead of the QP problems solved in standard SVM, thereby significantly reducing the computational time of the learning process [18]. LSSVM has been applied to various fields such as pattern recognition [19,20], system identification [21,22], and time series prediction [23].

For LSSVM, two parameters, the regularization constant and the kernel function parameter, are called free parameters and they play a crucial role in the performance of the SVM. However, inappropriate free parameters lead to over-fitting or under-fitting, and different parameter settings may also cause significant differences in the performance [24]. Thus, selecting the optimal free parameters is an important procedure in LSSVM. However, no general guidelines are available for selecting these parameters [13,25]. Moreover, it is time-consuming and difficult to select suitable parameters to obtain good performance by trial and error [26]. In order to solve these problems, genetic algorithms (GA), differential evolution algorithms (DEA), immune algorithms (IA), and other evolutionary strategies have been employed to find free parameters of LSSVM [26–29]. In particular, substantial research has been conducted on the application of GA techniques to improve the accuracy of building energy consumption prediction [26,30–32]. Pai and Hong showed that SVM with GAs are superior to other competitive prediction strategies (regression and ANNs) [33].

GAs were first proposed by Holland [34], with further study by Goldberg [35]. A general GA starts with a population of individuals, which evolves into advanced populations (according to a fitness function) as a result of naturally inspired operators, such as reproduction, crossover, mutation, and inversion [36].

Recently, real-coded GA (RCGA) methods have been introduced for a wide range of applications [26–28,30–33]. An RCGA uses a real value as a parameter of the individuals in the population without using the coding and encoding procedure before calculating the fitness value [37]. Thus, an RCGA is more straightforward, faster, and more efficient than conventional binary-coded GA. Recently, Kim et al. proposed a hybrid concept called DSORCGA that consists of the direct search optimization (DSO) method and RCGA. The DSORCGA hybrid differs from conventional RCGA in terms of the reproduction and crossover operators. It has been shown that the robustness and convergence speed of DSORCGA is superior to that of conventional RCGA [38].

In this study, the DSORCGA method proposed by Kim et al. [38] is employed in the LSSVM, namely DSORCGA-LSSVM, to determine the free parameters of LSSVM in a more effective manner by optimizing free parameters simultaneously from the training data. To verify the efficiency and performance of the proposed algorithm, a fair comparison is made with conventional RCGA and DEA, which were developed for the same problem and condition. Eventually, LSSVM with DSORCGA optimization was applied to building energy consumption prediction. The superior performance of the proposed algorithm, as compared with conventional RCGA and DEA methods, is confirmed by simulation.

The remainder of this paper is organized as follows. Section 2 introduces LSSVM and then establishes the process of optimized free parameters based on LSSVM. Section 3 discusses the procedures of conventional RCGA and the proposed DSORCGA. The simulation results of applying the proposed DSORCGA-LSSVM model to building energy consumption prediction are presented in Section 4. Finally, conclusions are drawn in Section 5.

2. Least-squares support vector machines and the optimization process of free parameters

LSSVM is a technique for regression. When LSSVM is used to estimate building energy consumption, input and output variables must be chosen first. Hence, this study takes historical data of building energy consumption for the last four weekdays as the input parameters of the building energy consumption prediction model. The daily quarter-hourly weekday building energy consumption is chosen as the model's output.

Consider a training data set $\{(x_1, y_1), \ldots, (x_N, y_N)\}$ with input data $x_i \in R^N$ and output data $y_i \in R^N$, where N is the size of sample data. In order to obtain the function dependence relation, SVM maps the input space into a high-dimension feature space and constructs a linear regression as follows:

$$y = f(x) = w^T \phi(x) + b \tag{1}$$

with $\phi(\cdot): R^n \to R^{n_\phi}$, a nonlinear function that maps the input space into a so-called higher dimensional feature space; w and b are the regression parameters to be solved.

The LSSVM regression estimation involves primal and dual problems. Given the training data set $\{(x_1, y_1), \ldots, (x_N, y_N)\}$, the aim is to estimate the model Eq. (1), where f is parameterized as in Eq. (4). We can formulate the following optimization scheme to infer our parameters:

$$\begin{aligned} &\min_{w,b,e} L_P(w, e) = \frac{1}{2}||w||^2 + \frac{\gamma}{2}\sum_{i=1}^{N} e_i \\ &s.t. \quad y_i = w \cdot \phi(x_i) + b + e_i, \quad i = 1, 2, \ldots, N \end{aligned} \tag{2}$$

where error variables $e = (e_1, e_2, \ldots, e_N)^T$, $e_i \in R$, and $\gamma > 0$ is a penalty parameter also called a regularization parameter, which is included to control the bias-variance trade-off. The above statement is in fact the same formulation as is used in the case of ridge regression in the feature space defined by $\phi(\cdot)$. Note that in some cases, w becomes infinite-dimensional, and the above problem formulation cannot be used to solve the problem. Therefore, we perform the computations in another problem, named the dual problem of Lagrangian multipliers, after applying Mercer's condition. The corresponding Lagrangian of Eq. (2) is

$$L_D(w, b, e_i, \alpha) = \frac{1}{2} w^T w + \frac{\gamma}{2}\sum_{i=1}^{N} e_i - \sum_{i=1}^{N} \alpha_i (w^T \phi(x_i) + b + e_i - y_i) \tag{3}$$

Here, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T$, $\alpha_i \in R$ are Lagrangian multipliers. According to the Karush–Kuhn–Tucker (KKT) conditions, the first-order conditions for optimality are given by

$$\begin{cases} \dfrac{\partial L_D}{\partial w} = 0 \to w = \displaystyle\sum_{i=1}^{N} \alpha_i \phi(x_i) \\ \dfrac{\partial L_D}{\partial b} = 0 \to 0 = \displaystyle\sum_{i=1}^{N} \alpha_i \\ \dfrac{\partial L_D}{\partial e_i} = 0 \to \alpha_i = \gamma e_i, \quad i = 1, 2, \ldots, N \\ \dfrac{\partial L_D}{\partial \alpha_i} = 0 \to y_i = w^T \phi(x_i) + b + e_i, \quad i = 1, 2, \ldots, N \end{cases} \tag{4}$$

From Eq. (4), combining the first and the last condition yields

$$y_i = \sum_{i=1}^{N} \alpha_i \phi(x_i)^T \phi(x_i) + b + e_i \tag{5}$$