STUDIA INFORMATICA 2011
Volume 32 Number 2A (96)

Jerzy MARTYNA
Uniwersytet Jagielloński, Instytut Informatyki

WAVELET SUPPORT VECTOR MACHINES AND MULTI-ELITIST PARTICLE SWARM OPTIMIZATION FOR TIME SERIES FORECASTING

Summary. In this paper, we present a new method for time series forecasting based on wavelet support vector machines (WSVM). To better represent any curve in $L^2(R^n)$ space (quadratic continuous integral space), we used a new kernel function. This function is the wavelet function. The SVM with wavelet kernel function is referred to as a wavelet SVM. In order to determine the optimal parameter of the WSVM, the multi-elitist particle swarm optimization (PSO) was used. Computational results demonstrate the effectiveness of the proposed method over the traditional methods.

Keywords: wavelet support vector machine, multi-elitist particle swarm optimization, time series forecasting

FALKOWE MASZYNY WEKTORÓW WSPIERAJĄCYCH ORAZ WIELOELITARNA OPTYMALIZACJA CZĄSTEK ROJÓW DLA PRZEWIDYWANIA SZEREGÓW CZASOWYCH

Streszczenie. W artykule przedstawiono nową wersję falkowej maszyny wektorów wspierających (ang. *wavelet support vector machines*, WSVM), którą zastosowano do przewidywania wielowymiarowych szeregów czasowych. Do wyznaczenia optymalnych parametrów falkowej maszyny wektorów wspierających użyto wieloelitarnej optymalizacji rojów cząstek (ang. *multi-elitist particle swarm optimization*, MEPSO). Efektywność uzyskanych wyników obliczeniowych została porównana z rezultatami tradycyjnych metod przewidywania wielowymiarowych szeregów czasowych.

Słowa kluczowe: falkowa maszyna wektorów wspierających, wieloelitarna optymalizacja cząstek rojów, przewidywanie szeregów czasowych

1. Introduction

The forecasting is a dynamic process in which some parameters are predicted for the given time duration. This process is characterized by many factors, such as an uncertain characteristics, fuzziness of some numbers, etc. In other words, the forecasting is an estimation of the relationship by means of the defining the mathematical model of the process and, further, by means of the development of computer and estimation techniques, some estimated data put forward for further analysis and evaluation.

The accuracy of the traditional methods of forecasting, such as the time series forecasting [3], high-order time series forecasting [13] or the regression analysis [15], is limited. Therefore, the artificial intelligence methods are in use more and more often. Among others such techniques as neural networks [23], fuzzy time series [11], expert systems [18], fuzzy-neural networks [17], neural networks and genetic algorithms [1], etc., have been applied.

Recently, a new method for regression and classification was developed by V. Vapnik [21], [22]. This method, called the *support vector machine* (SVM), can find the optimal separating hyperplane between the positive and the negative examples. This hyperplane maximizes the margin between the training examples that are closely to the hyperplane and minimizes the number of the data points which are classified wrongly. The SVM method has been successfully used for high dimensional data in many problems, such as pattern recognition [1], electric load forecasting [10], etc. Compared with the traditional neural networks, the SVM can obtain a unique global optimal solution and avoid the curse of dimensionality.

In many applications of the SVM for the pattern recognition, regression analysis, etc.,the kernel function must satisfy the Mercel conditions. The Gauss function used more often as the kernel function in the SVM, does not provide satisfying results. As a result this kernel function is not a complete orthonormal base. Therefore, the SVM used at present cannot map every curve in the $L^2(R^n)$ space.

In order to overcome this weakness of the SVM, a new kernel function based on the wavelet function has been proposed. Among others, the so-called wavelet kernel function has been suggested in a number of papers. For example, the papers by Khandoker [12], Widodo [19] described some experiments with the wavelet kernel function. The authors of these papers show that the performance of the wavelet function in the SVM is much better than the traditional kernel function. Moreover, the SVM with the kernel function can reduce the generalization error. It was observed by Lu et al. [14] in the application of the linear programming support vector regression with the wavelet dynamical systems identification.

The main goal of this paper is to construct a hybrid system devoted to the forecasting of the time series. To avoid the weakness of the traditional SVM method, we suggest the wavelet kernel function in the SVM. In order to optimize the SVM we propose a new method of optimization called a *multi-elitist particle swarm optimization* (MEPSO). A built this way hybrid intelligent system which can predict the time series, is a novel proposal in the field of the forecasting method implementation.

The rest of the paper is as follows. Firstly, the kernel function and their properties are presented in section 2. Afterwards, the model of the wavelet support vector machine (WSVM) is presented in section 3. The proposed optimization and the forecasting scheme are given in section 4. In section 5 we give some computational results. Finally, the paper ends with a conclusion and a direction of further work.

2. Kernels and their properties

Let points $(x, y_1)(x_2, y_2)\ldots(x_n, y_n)$ be a set of data which are independently and randomly generated from an unknow function. n denotes the total number of examples, $x_i \in R^n$ is the input and $y_i \in R$ are the target output data. Kernel functions of the form $k(x_1, x_2) = \varphi(x_1) \cdot \varphi(x_2)$, where is an inner product and φ is in general a nonlinear mapping from the input space X onto the feature space Z. In fact, the kernel function k is directly defined, φ and the feature Z is derived from its definition. The SVM introduced by V. Vapnik [21], [22] finds a hyperplane in a space different from that of the input data x. It can be said that this hyperplane in a space is induced by kernel k.

In order to guarantee the existence of a feature space the Mercer's Thorem [16] for kernel $k(x, x')$ should be satisfied.

The Mercer's theorem (Mercer, 1909)
Let $K(x, x')$ be a continuous symmetric kernel in the closed interval $a \le x \le b$ and likewise for x'. Kernel $K(x, x')$ can be expanded in the series

$$K(x, x') = \sum_{i=1}^{\infty} \lambda_i \varphi_i(x) \varphi_i(x') \quad (1)$$

with positive coefficients, $\lambda_i > 0$, for all i.

For this expansion to be valid and for its absolute and uniform coverage, it is necessary and sufficient that condition

$$\int_a^b \int_a^b K(x, x') \Psi(x) \Psi(x') dx dx' \ge 0 \quad (2)$$

should hold for all $\Psi(x)$ for which

$$\int_a^b \Psi^2(x)dx < \infty \tag{3}$$

Functions φ_i are called eigenfunctions and λ_i are called eigenvalues. The fact that all of the eigenvalues are positive means that kernel $K(x,x')$ is a positive definite [10].

We can also give the above-given condition in the time-frequency terms, namely.

Theorem 1 (Smola and Schölkopf, 1998)
The horizontal floating function is an allowable support vector's kernel function if and only if the Fourier transformation of $K(x)$ needs to satisfy the condition as follows:

$$F_x(\omega) = \frac{1}{\sqrt[n]{2\pi}} \int_0^b e^{-j\omega_x} K(x)dx \geq 0 \tag{4}$$

Let the wavelet function Ψ satisfy conditions $\Psi(x) \in L^2(R) \cap L^2(R)$ and $\overline{\Psi}(x) = 0$, where $\overline{\Psi}$ is the Fourier transformation. The wavelet function group is given by

$$\Psi_{a,m}(x) = \sqrt{a}\Psi\left(\frac{x-m}{a}\right)dx \tag{5}$$

where $a, a > 0,$ is the so-called scaling parameter, $m, m \in R,$ is the horizontal floating coefficient, $\Psi(x)$ is the "mother wavelet". Parameters m and a are called the parameters of translation and dilatation, respectively. The wavelet transform $f(x)$ is defined as

$$W(a,m) = \sqrt{a} \int_{-\infty}^{\infty} f(x)\Psi^{\bullet}\left(\frac{x-m}{a}\right)dx \tag{6}$$

where $\Psi^{\bullet}(x)$ is a complex conjugation of $\Psi(x)$.
The original signal can be obtained by classical inversion formula for $f(x)$, namely

$$f(x) = \frac{1}{C_\Psi} \int_{-\infty}^{+\infty} W(a,m)\Psi_{a,m}(x)\frac{da}{a^2}dm \tag{7}$$

where

$$C_\Psi = \int_{-\infty}^{+\infty} \frac{|\overline{\Psi}(w)|^2}{|w|} dw < \infty \tag{8}$$

Here we use the following notation: C_ψ is a constant with respect to $\Psi(x)$ and $\overline{\Psi}(w) = \int \Psi(x)e^{-jwx}dx$.

The multi-dimensional wavelet function is defined as follows

$$\Psi_i(x) = \prod_{i=1}^{l} l\Psi(x_i), \quad x \in R^{l\times d} \tag{9}$$

where x is a column vector with a d dimension.

The horizontal floating kernel function for the scaling parameter of the wavelet, $a_i, a > 0$, is as follows

$$K(x,x') = \prod_{i=1}^{l} l\Psi\left(\frac{x_i - x_i'}{a_i}\right) \tag{10}$$

It is obvious that the wavelet kernel function must satisfy the condition of Theorem 1. To the wavelet kernel which satisfies this condition belongs to the Mexican hat wavelet. The Mexican hat wavelet is defined as follows [2]:

$$\Psi(x) = \frac{2}{\sqrt{3}}\pi^{-\frac{1}{4}}\left(1 - \frac{x^2}{a}\right)\exp\left(-\frac{x^2}{a}\right),\ x \in R \tag{11}$$

Theorem 2 The kernel of the Mexican hat
The Mexican hat wavelet is defined as

$$K(x,x') = \prod_{i=1}^{l} \frac{2}{\sqrt{3}}\pi^{-\frac{1}{4}}\left(1 - \frac{\|x_i - x_i'\|}{a}\right)^2 \exp\left(-\frac{\|x_i - x'_i\|^2}{2a^2}\right) \tag{12}$$

and this kernel is an allowable support vector kernel function.

3. Wavelet support vector machines

In order to build the wavelet support vector machine we can combine the wavelet kernel function with the SVM. For a set of data points given above, we now define the wavelet support vector machine (WSVM) as

$$\text{minimize} \quad \tau(w, \xi^\bullet, \varepsilon) = \frac{1}{2}\|w\|^2 + C \cdot \frac{1}{l}\sum_{i=1}^{l}(\xi_i + \xi_i^\bullet) \tag{13}$$

$$\text{subject to} \quad (w \cdot x_i + b) - y_i \le \varepsilon + \xi_i \tag{14}$$

$$y_i - (w \cdot x_i + b) \le \varepsilon + \xi_i^\bullet \tag{15}$$

$$\xi_i^\bullet \ge 0,\ \xi \ge 0,\ \varepsilon \ge 0,\ b \in R \tag{16}$$

where ξ_i and $\xi_i^\bullet$ are the slack variables corresponding to the size of the excess positive and negative deviation, respectively. w and x_i are a column vector with a d dimension, $C > 0$ is a penalty factor.

The problem given by Eq. (13) is a quadratic programming (QP) problem. By introducing Lagrangian multipliers, a Lagrangian function can be defined as

$$L(w,b,\alpha^{(\bullet)},\beta,\xi^{(\bullet)},\varepsilon,\eta^{(\bullet)}) = \frac{1}{2}\|w\|^2 + C\cdot\varepsilon + \frac{C}{l}\sum_{i=1}^{l}(\xi_i+\xi_i^{\bullet}) - \beta\cdot\varepsilon - \sum_{i=1}^{l}(\eta_i\xi_i + \eta_i^{\bullet}\xi_i^{\bullet}) - \sum_{i=1}^{l}\alpha_i(\varepsilon+\xi_i+y_i-w\cdot x_i-b) - \sum_{i=1}^{l}\alpha_i^{\bullet}(\varepsilon+\xi_i^{\bullet}-y_i-w\cdot x_i+b) \quad (17)$$

where $\alpha_i^{(\bullet)} = \{\alpha_1,\ldots,\alpha_i,\alpha_1^{\bullet},\ldots,\alpha_i^{\bullet}\}$, β, $\eta_i^{\bullet} = \{\eta_1,\ldots,\eta_l,\eta_1^{\bullet},\ldots,\eta_l^{\bullet}\} \geq 0,\ i=1,\ldots,l1$ are the Lagrangian multipliers. By differentiating the Lagrangian function (17) with respect to $w,b,\varepsilon,\xi_i^{\bullet}$ and by using Karush-Kuhn-Tucker (TKK) conditions, we can obtain the corresponding dual form of function (13), namely

$$\text{maximie}\quad W(\alpha,\alpha^{\bullet}) = -\frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i^{\bullet}-\alpha_i)(\alpha_j-\alpha_j^{\bullet})K(x_i,x_j) + \sum_{i=1}^{l}(\alpha_i^{\bullet}-\alpha_i)y_i \quad (18)$$

$$\text{subject to}\quad 0 \leq \alpha_i,\ \alpha_i^{\bullet} \leq C \quad (19)$$

$$\sum_{i=1}^{l}(\alpha_i^{\bullet}-\alpha_i) = 0 \quad (20)$$

$$\sum_{i=1}^{l}(\alpha_i+\alpha_i^{\bullet}) \leq C \quad (21)$$

Further for the construction the QP problem of the WSSM and the solution this problem, we can obtain parameters $\alpha_i^{\bullet}$. Parameter b can be obtained after selecting $\alpha_j,\alpha_j \in [0,\frac{l}{C}),$ and $\alpha_k^{\bullet},\ \alpha_k^{\bullet} \in [0,\frac{l}{C})$, namely

$$b = \frac{1}{2}\left[y_j + y_k - \left(\sum_{i=1}^{l}(\alpha_i^{\bullet}-\alpha_i)K(x_i,x_j) + \sum_{i=1}^{l}(\alpha_i^{\bullet}-\alpha_i)K(x_i,x_k)\right)\right] \quad (22)$$

The regression function of WSVM can be given by

$$y = \sum_{i=1}^{l}(\alpha_i-\alpha_i^{*})\left[\prod_{j=1}^{l}\left(\Psi\frac{x^j-x_i^j}{a_i}\right)+b\right] \quad (23)$$

4. The proposed optimization algorithm

The determination of the unknown parameters of the WSVM is a multivariable optimization process in a continuous space. We propose a modification of the classical particle swarm optimization algorithm here.

We recall that the basic PSO algorithm was introduced by Eberhart and Shi (2001) [8] and Kennedy (2002) [5]. The convergence of the classical PSO is obtained with the use of a small inertia weight ω or a constriction coefficient. However, the searching process may have a poor local minimum. Therefore, here we suggest a multi-elitist strategy for searching the global best of the PSO. This modification of the PSO called MEPSO was at first proposed by Deb et al. (2002) [6]. Let β be a growth rate for each particle. If the defined fitness value of a particle of the k-th iteration is higher than that of a particle of the $(k-1)$-th iteration, the β will be increased. Thus, we move the local best of all particles which has a higher fitness value than the global best into the candidate area. As a result we obtain the replacement of the global best by the local best with the highest growth rate β. In other words, the current global best given by the fitness value is always higher than the old global best.

4.1. The MEPSO algorithm

The multi-elitist particle swarm optimization (MEPSO) algorithm works as follows: The swarm consists of n particles. Each particle has a position $X_i = (x_{i1}, \ldots, x_{id})$, a velocity $V_i = (v_{i1} \ldots, v_{id})$ and moves through a d dimensional serach space. According to the MEPSO, each particle moves towards the best previous position and towards the best particle pg in the swarm. We assume that the best previously visited position of the i-th particle gives the best fitness value, namely $p_i = (p_{i1}, \ldots, p_{id})$. Also, the best previously visited position of the swarm gives the best fitness as $p_g = (p_{g_1}, \ldots, p_{g_d})$.

By incorporating the MEPSO for the previous velocity of the particle, we can update the velocity and the particle position by means of using the equations:

$$V_{id}(k+1) = \omega \cdot V_{id}(k) + C_1\phi_1(P_{id}^l - X_{id}(k) + C_2\phi_2(p_{gd} - X_{id}(k)) \quad (24)$$

$$X_{id}(k+1) = X_{id}(k) + V_{id}(k+1) \quad (25)$$

where ϕ_1 and ϕ_2 are random positive number

The MEPSO algorithm is terminated with a maximal number of generations or best particle positions of the entire swarm. When the swarm cannot be improved after a given number of generations, the obtained values are accepted as the best.

The pseudocode of the MEPSO algorithm

```
procedure MEPSO_algorithm;
begin
  for t = 1 to t_max do
    if t < t_max then
      for j = 1 to N do {swarm size is equal to N}
        if the fitness value of particle_j in the t-th time-step > that of
particle_j in (t-1)-th
           time-step then   β_j = β_j + 1;
```

```
            endif;
               update local best_j ;
            if  the fitness of local best_j > that of global best now then
               choose local best_j put into candidate list;
            endif;;
         endfor;
           calculate β  of every candidate and record the candidate of β_max ;
           update the global best to become the candidate of   β_max ;
        else
            update  the global best to become the particle of highest fitness val-
ue.
        endif;
     endfor;
   end;
```

4.2. The Fitness Function

In order to evaluate the accuracy of a forecasting an appropriate fitness function can be used. We designed it as follows:

$$fitness = \frac{1}{l}\sum_{i=1}^{l}\left(\frac{\overline{y_i} - y_i}{y_i}\right) \tag{26}$$

where $\overline{y_i}$ denotes the forecasting value of the selected sample, y_i is the original date of the selected sample, l is the size of the selected sample. The proposed MEPSO algorithm is used for determining the parameters of the WSVM. The different parameters of the WSVM are adapted for the sake of forecasting the time series. The most adequate WSVM with the optima parameters is used in the final forecasting.

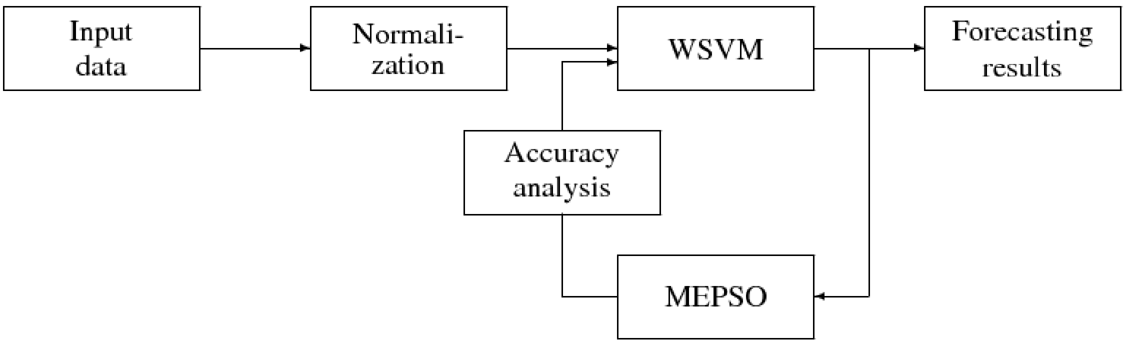

Fig. 1. Flowchart of the WSVM –based forecasting algorithm for the time series forecasting by means of the SVM and the MEPSO methods
Rys. 1. Schemat algorytmu opartego na WSVM dla przewidywania szeregów czasowych, przy użyciu metod SVM i MEPSO

The MEPSO algorithm is terminated with a maximal number of generations or best particle positions of the entire swarm. When the swarm cannot be improved after a given number of generations, the obtained values are accepted as the best.

4.3. The Proposed Forecasting Function

The proposed forecasting metod is described in steps as follows (see fig. 2):

1. Initialize the original data.
2. Data normalization.
3. Construct the WSVM and the QP problem of the WSVM.
4. Set the MEPSO parameters including number of swarm particles (N), swarm particie dimension (d), number of maximal iterations $(k_{\max})$, fitness value of swarm particie
5. (ς), inertia weight (ω), acceleration constans (C_1, C_2), growth rate (β), etc.
6. Optimize the parameters α_i and computer the regression coefficient coefficient b given by Eq. (22).
7. Compute the forecasting result.

5. Experimental Results

To performance evaluation of the proposed forecast method, the accuracy of prediction of the time series by means of combining theWSVM and the MEPSO was studied. We have used the script with the WSVM and the MEPSO methods which was included to Oracle Data Mining Software.

We used the following normalization:

$$\bar{y}_i^s = \frac{x_i - \min(x_i^s |_{i=1}^l)}{\max(x_i^s |_{i=1}^l) - \min(x_i^s |_{i=1}^l)}, \; s = 1, \ldots n \tag{27}$$

where i is the index of sample, x_i^s and $\bar{x}_i^s$ are the original value and the normalized value of the s sample, respectively.

For the Mexican hat wavelet kernel used in the WSVM three parameters are determined, namely

$$\begin{aligned} & v \in [0,1] \\ & a \in [0.2,\ 3] \\ & C \in \left[\frac{\max(x_{ij}) - \min(x_{ij})}{l} \times 10^{-3}, \frac{\max(x_{ij}) - \min(x_{ij})}{l} \times 10^{3} \right] \end{aligned} \tag{28}$$

We assumed that the MEPSO is convergent with the minimal value. For the minimal value of the fitness function the following values of the parameters: $C = 8700$, $v = 0.82$, $a = 95$ are obtained.

To performance evaluation of the proposed forecast metod, the accuracy of prediction of the time series by means of ARIMA metod was studied. We take into consideration the stock of the PKO B.P. noted in the Warsaw Stock Exchange [6].

The comparison of the forecasting result obtained by WSVM with the MEPSO algorithm and ARIMA metod of the PKO B.P. stock with the real value for the twelve months from the beginning of the 10 November 2008 year is given in fig. 2. It can be seen that the MEPSO can improve the global searching ability of the particie swarm optimization algorithm.

To evaluate the performance of forecasting the underlying WSVM, two performance measures are calculated, the root mean squared error (RMSE) and mean relative error (MRE). For day i to each data set, we define

$$RMSE_I = \sqrt{\frac{1}{m}\sum_{j=1}^{m}(\hat{y}_{ij} - y_{ij})^2} \tag{29}$$

$$MRE = \frac{100}{m}\sum_{j=1}^{m}\frac{|\hat{y}_{ij} - y_{ij}|}{y_{ij}} \tag{30}$$

where $\hat{y}_{ij}$ is the forecast for y_{ij}, m is the time interval.

Fig. 2. The comparison of the forecasting result obtained by WSVM and ARIMA methods of the PKO B.P. stock with the real value for the twelve months from the beginning of the 10 Nov. 2008

Rys. 2. Porównanie wyników przewidywania, uzyskane przy użyciu metod WSVM, ARIMA i rzeczywistych notowań kursu akcji PKO B.P. w ciągu 12 miesięcy, począwszy od 19 listopada 2008 r.

The Mean RMSE and Mean MRE (%) of the forecast time series are then calculated for data sets. For each forecasting methods, the fig. 3 plots the empirical cumulative distribution function (CDF) of the Mean RMSE calculated on the data sets by means of WSVM and ARIMA methods, respectively. The true data is indicated in the figures using a triangle. In both cases, our solution based on WSVM and MEPSO methods is competitively close to the true data.

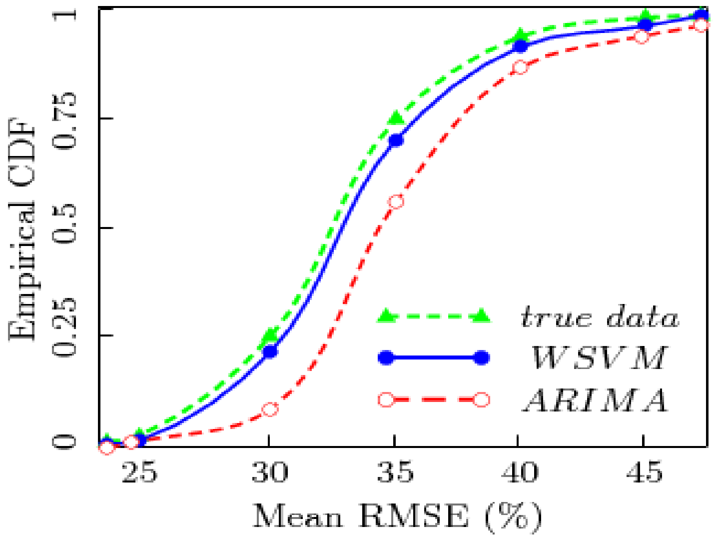

Fig. 3. The empirical cumulative distribution function (CDF) calculated by means of WSVM and ARIMA, respectively

Rys. 3. Funkcja kumulacyjnego rozkładu (CDF) w zależności od średniej wartości RMSE obliczone przy użyciu WSVM i ARIMA

6. Conclusion

The concept of combining the SVM method and the MEPSO technique into the hybrid system for forecasting the time series is very promising. A mathematical model for this system has been also developed. By means of using the computational experiment it has been shown that the forecasting of the time series can be achieved with satisfactory accuracy.

The scope of the analytical model presented in this paper is limited to one method of optimization, e.g. the MEPSO method. As a future work, we plan to develop an SVM method that will be improved by another optimization method in order to solve the formulated forecasting problem.

BIBLIOGRAPHY

1. Azadeh A., Ghaderi S. F., Tarvedian S., Saberi M.: Integration of Artificial Neural Networks and Genetic Algorithm to Predict Electrical Energy Consumption. Applied Mathematics and Computation, Vol. 186, 2007, p. 1731÷1741.
2. Białasiewicz J. T.: Wavelets and Approximations. Warszawa: WNT, 2000.

3. Box G. E. P., Jenkins G. M.: Time Series Analysis: Forecasting and Control. Holdan Day, San Francisco 1976.
4. Burges C. J.: A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery, 2, 1998, p. 1÷43.
5. Clerc M., Kennedy J.: The Particle Swarm - Explosion, Stability, and Convergence In Multidimensional Complex Space. IEEE Trans. on Evolutionary Computation, Vol. 6, No. 1, 2002, p. 58÷73.
6. Deb K., Pratap A., Agarwal S., Meyarivan T.: A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II. IEEE Trans. on Evolutionary Computation, Vol. 6, No. 2, 2002.
7. Dom Maklerski BOŚ http://bossa.pl/.
8. Eberhart R. C., Shi Y.: Particle Swarm Optimization: Developments, Applications and Resources. Proc. of IEEE Int. Conference on Evolutionary Computation, Vol. 1, 2001, p. 81÷86.
9. Haykin S.: Neural Networks: A Comprehensive Foundation. 2nd Ed., New Jersey: Prentice Hall, 1999.
10. Hong W. C.: Electric Load Forecasting by Support Vector Model. Applied Mathematical Modelling, Vol. 33, No. 6, p. 2444÷2454.
11. Hwang J. R., Chen S. M., Lee C. H.: Handling Forecasting Problems Using Fuzzy Time Series. Fuzzy Sets and systems, Vol. 100, 1998, p. 217÷228.
12. Khandoker A. H. K., Lai D. T. H., Begg R. K., Palaniswarni M.: Wavelet-Based Feature Extraction for Support Vector Machines for Screening Balance Impatients in the Elderly. IEEE Trans. on Neural Systems and Rehabilitation Engineering, Vol. 15, No. 4, 2007, p. 587÷597.
13. Lee L. W., Wang L. W., Chen S. M.: Handling Forecasting Problems Based on Two-Factor High-Order Time Series. IEEE Trans. on Fuzzy Systems, Vol. 14, No. 3, 2006, p. 468÷477.
14. Lu Z., Sun J., Butts K. R.: Linear Programming Support Vector Regression with Wavelet Kernel: A New Approach to Nonlinear Dynamical System Identification. Mathematics and Computers in Simulation, Vol. 79, 2009, p. 2051÷2063.
15. McQuarrie A. D. R., Tsai Ch.-L.: Regression and Time Series Model Selection. World Scientific, 1998.
16. Mercer J.: Functions of Positive and Negative Type and Their Connection with the Theory of Integral Equations. Philos. Trans. Roy. Soc., London, 209 (1909), p. 415÷446.
17. Piramuthu S.: Theory and Methodology - Financial Credit-Risk Evaluation with Neural and Neural-Fuzzy Systems. European Journal of Operational Research, Vol. 112, 1991, p. 310÷321.

18. Rodionov S. N., Martin J. H.: An Expert System-Based Approach to Prediction of Year-to-Year Climatic Variations in the North Atlantic Region. International Journal of Climatology, Vol. 19, 1999, p. 951÷974.
19. Widodo A., Yang B. S.: Wavelet Support Vector Machine for Induction Machine Fault Diagnosis Based on Transient Current Signal. Expert Systems and Applications, Vol. 35, No. 1-2, 2008, p. 307÷316.
20. Wu C. H., Tzeng G. H., Lin R. H.: A Novel Hybrid Genetic Algorithm for Kernel Function and Parameter Optimization in Support Vector Regression. Expert Systems Applications, Vol. 36, No. 4, p. 4725÷4735.
21. Vapnik V. N.: The Nature of Statistical Learning. New York: Springer Verlag, 1995.
22. Vapnik V. N.: Statistical Learning Theory, New York: JohnWiley and Sons, 1998.
23. Zhang G. P., Patuwo B. E., Hu Y. M.: Forecasting with Artificial Neural Networks: The State of the Art. Int. Journal of Forecasting, Vol. 14, 1998, p. 35÷62.

Recenzenci: Prof. dr hab. inż. Jerzy Klamka
Dr inż. Jerzy Respondek

Wpłynęło do Redakcji 31 stycznia 2011 r.

Omówienie

W artykule wprowadzono falkowe maszyny wektorów wspierających (ang. *Wavelet Support Vector Machines)* dla przewidywania szeregów czasowych. Istotą tej metody jest zastosowanie falkowej funkcji jądrowej w maszynie wektorów wspierających. Jako falkową funkcję jądrową użyto falkę znaną jako „meksykański kapelusz". Wartości parametrów falkowej maszyny wektorów wspierających uzyskano dzięki oryginalnej metodzie, opartej na wieloelitarnej optymalizacji rojów cząstek (ang. *Multi-Elitist Particie Swarm Optimization*). Opracowaną metodę przewidywania szeregów czasowych zastosowano do prognozowania kursu akcji papierów wartościowych spółki KGHM S.A., notowanych na Giełdzie Papierów Wartościowych w Warszawie. Dla porównania wyników przewidywania, uzyskanych przy użyciu falkowej maszyny wektorów wspierających wraz z wieloelitarną optymalizacją rojów cząstek, obliczono dwa mierniki trafności prognoz *ex post*: standardową średnią błędów względnych (ang. *root mean squared error*, RMSE) oraz relatywną standardową średnią błędów (ang. *mean relative error*, MRE).

Address

Jerzy MARTYNA: Uniwersytet Jagielloński, Instytut Informatyki, ul. Prof. S Łojasiewicza 6, 30-348 Kraków, Polska, martyna@ii.uj.edu.pl.