Robert SAŁAT, Michał AWTONIUK, Krzysztof KORPYSZ

Warsaw University of Life Sciences, Department of Fundamental Engineering

Black-Box system identification by means of Support Vector Regression and Imperialist Competitive Algorithm

***Abstract**. The paper proposes to use SVR network for system identification by means of the Black-Box method. Selection of the optimal network parameters as well as the selection of optimal set of regressors by the Imperialist Competitive Algorithm have been proposed. The accuracy of models built was compared to ARX model for a linear object and NARX model for a non-linear one. The results show that the use of SVR in Black-Box identification can be an useful and efficient alternative for models based on NARX and ARX structures.*

***Streszczenie.** W artykule zaproponowano użycie sieci SVR do identyfikacji systemów metodą Black-Box. Do doboru optymalnych parametrów sieci i wyboru optymalnego zestawu regresorów, zaproponowano użycie algorytmu ewolucyjnego Imperialist Competitive Algorithm. Dokładność zbudowanych modeli porównano z modelem typu ARX, dla obiektu liniowego, oraz NARX, dla nieliniowego. Otrzymane wyniki wskazują, że wykorzystanie SVR w identyfikacji Black-Box może być użyteczną i efektywną alternatywą dla modeli o strukturze ARX oraz NARX. (**Identyfikacja systemów metodą Black-Box za pomocą Support Vector Regression i Imperialist Competitive Algorithm**)*

Keywords: identification, black-box, support vector machine, evolutionary algorithms.
Słowa kluczowe: identyfikacja, black-box, support vector machine, algorytmy ewolucyjne.

Introduction

Good knowledge of behavior of the system in the control process is essential to provide high quality and achieve low production costs. Objects (plants) identification (considered as components of the system) consists in construction of mathematical description of object dynamics on the basis of measured data and forecasting the behavior of the object for new input data. Mathematical models resulting from the identification are used for the purpose of controlling, simulation tests and forecasting.

However, it is difficult to find mathematical model which is capable of accurately mapping a dynamic behaviour of the system. Most of the entity industrial systems are non-linear, resulting in very complex problems in identification and control. Usually in industry there are built linear models around the set points. These models have been well tested and described in the literature. Now among scientists there is a tendency to use and test non-linear models of systems [1].

If we have a priori knowledge and know laws of physics that occur in the system, than we can use identification on the basis of White-Box. But if the knowledge is incomplete, i.e. a model is created on the basis of the laws of physics but a number of parameters must be estimated from measurement data, then there is a Grey-Box model. However, if we don't have sufficient knowledge about the system or if the system is very complex, it is convenient to use Black-Box type identification i.e. identification based only on measured input and output signals [2]. Black-Box type identification may concern both linear systems and non-linear systems. There are many algorithms which use Black-Box method, e.g. gradient-based methods, which, however, do not guarantee the achievement of the global minimum. In the recent years, several authors [3,4,5] suggested using the method of artificial intelligence in Black-Box identification. This suggestion resulted in a large number of works, in which different techniques were presented, such as: Artificial Neural Networks (ANN), Wavelets, Fuzzy-Neural Networks (FNN) and others [5]. Models created on the basis of identification can be used both as simulators (e.g. device detection) and models for use in control systems. Among neural networks used in Black-Box identification the most frequently reported were Multi Layer Perceptron (MLP), Fuzzy-Neural Networks (FNN), Radial Basis Function (RBF), Runge-Kutta (RK), Digital Recurrent Network (DRN). Few authors used Support Vector Machines (SVM) as an identification tool. Usually the authors also ignore the rules of the selection of regressors and parameters both neural networks and SVM.
In this paper the use of Support Vector Regression (SVR) [6] was proposed for the identification of linear and non-linear systems. New approach proposed in the paper (according to authors who are not known in world literature) is the use of an evolutionary algorithm, Imperialist Competitive Algorithm (ICA), as a tool for both selection of parameters of SVR network and a tool for selecting an optimal set of regressors. The results of using SVR network with ICA were compared with ARX and NARX models which are available in Matlab System Identification Toolbox R2010b and are often used by many authors.

Bases of Black-Box identification

Most of real systems are non-linear, so in order to achieve an acceptable accuracy of the modelling, it is desirable to use non-liner models. Only non-linear Black-Box identification bases will be described [2]. Assuming that *u(t)* is an discrete input of non-linear system, and *y(t)* is an output, data set can be defined as:

$$\begin{aligned}\mathbf{u} &= \left[u(1),u(2),\ldots,u(t)\right] \\ \mathbf{y} &= \left[y(1),y(2),\ldots,y(t)\right]\end{aligned} \tag{1}$$

where: $u(1),\ldots,\ u(t),\ y(1),\ldots,\ y(t)$ – samples sequentially in time.

The idea of Black-Box identification is to find the connection between past observations i.e. $[u(t-1),..,\ u(t-n_u),\ y(t-1),\ldots,\ y(t-n_y)]$ and future $y(t)$.
The primary goal is estimating non-linear function

$$\begin{aligned}&y(t) = f\left(\mathbf{x}(t),\boldsymbol{\theta}\right) + e(t) = \\ &f\Big(\big[u(t-1),\ldots,u(t-n_u),\\ &y(t-1),\ldots,y(t-n_y)\big],\boldsymbol{\theta}\Big) + e(t)\end{aligned} \tag{2}$$

where: $f(\cdot)$ – unknown basis function, $\boldsymbol{x}(t)$ – regression vector, n_u – the number of past observations samples of input, n_y – the number of past observations samples of output, $\boldsymbol{\theta}$ – vector of associated parameters, $e(t)$ – noise which value should be as low as possible.

The elements of vector $\boldsymbol{x}(t)$ are called regressors and depending on the choice of the vector of regressors we can build a variety of models [2]:

- Nonlinear Finite Impulse Response (NFIR), where $\mathbf{x}(t)=[u(t-1),\dots, u(t-n_u)]$
- Nonlinear AutoRegressive with eXogenous inputs (NARX), where $\mathbf{x}(t)=[u(t-1),\dots, u(t-n_u), y(t-1),\dots, y(t-n_y)]$
- Nonlinear AutoRegressive Moving Average with eXogenous inputs (NARMAX), where $\mathbf{x}(t)=[u(t-1),\dots, u(t-n_u), y(t-1),\dots, y(t-n_y), e(t-1),\dots, e(t-n_e)]$
- Nonlinear Output Error (NOE), where $\mathbf{x}(t)=[u(t-1),\dots, u(t-n_u), y_m(t-1),\dots, y_m(t-n_y)]$
- Nonlinear Box-Jenkins (NBJ), where the vector of regressors takes all the above types.

The problem of identification of each model comes down to finding the number of regressors i.e. the number of past observations samples and non-linear function $f(\cdot)$ so as to obtain the model best most fitting to the object. Finding a suitable number of regressors is very important, as too many numbers of the regressors may lead both to creating a too complex model and variables duplication. A large number of irrelevant regressors can cause overfitting which results in a poor generalization of the model.

The Bases of Support Vector Regression (SVR) in problem identification

Since the Support Vector Machines (SVMs) and Support Vector Regression (SVR) are often presented in the literature [6,7,8], only a basic presentation of SVR in problem identification was presented. In order to present the algorithm of SVR, it must first be introduced on the projection function F of the regressors space $\Re$ to a hypothetical feature space $\aleph$, and assume that the training set is as follows

$$(\mathbf{x}_1, y_1),\dots,(\mathbf{x}_n, y_n) \in \Re^N \times \Re \tag{3}$$

where: $\boldsymbol{x}_n$ – input vector composed of regressors for NARX model, y_n – oputput.
By defining ε – insensitive loss function as

$$|y - f(\mathbf{x})|_\varepsilon = max\{0, |y - f(\mathbf{x})| - \varepsilon\} \tag{4}$$

procedure can be treated as a quadratic programming (QP), and the estimation function can be expressed in the "standard" SVR as

$$f(\mathbf{x}) = (\mathbf{w}\cdot F(\mathbf{x})) + b, \quad \mathbf{w},\mathbf{x}\in\Re^n, b\in\Re \tag{5}$$

where: $\mathbf{w}$ – weight vector, b – bias.
Then, problem is converted to the minimization of

$$\begin{aligned} &min\ \frac{1}{p}\sum_{i=1}^{p}(|y_i - f(\mathbf{x}_i)| - \varepsilon) = \\ &min\ \frac{1}{p}\sum_{i=1}^{p}(|y_i - \mathbf{w}\cdot F(\mathbf{x}_i) + b| - \varepsilon) \end{aligned} \tag{6}$$

where: p – the number of training data pairs.
Introducing slack variables ζ_i, ζ_i^*, the above an optimization problem can be formulated as

$$\min_{\mathbf{w},\xi,\xi_i^*}\ \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{p}\xi_i + C\sum_{i=1}^{p}\xi_i^*, \tag{7}$$

with the following functional and boundary constraints:

$$\begin{aligned} &y_i - \mathbf{w}\cdot F(\mathbf{x}_i) \le \varepsilon + \xi_i \\ &\mathbf{w}\cdot F(\mathbf{x}_i) - y_i \le \varepsilon + \xi_i^* \\ &\xi_i, \xi_i^* \ge 0 \end{aligned} \tag{8}$$

where: C, ε – user specified constant.

The optimization problem can be relatively easily solved by the introduction the Lagrange function and the formulation of so-called the dual problem regarded to Lagrange multipliers α and the introduction of the kernel term

$$k(\mathbf{x}_i, \mathbf{x}_j) = F^T(\mathbf{x}_i)F(\mathbf{x}_j), \tag{9}$$

defined in accordance with Mercer's theorem [9].
After solving the dual problem we obtain

$$f(\mathbf{x}) = \sum_{i=1}^{K}(\alpha_i^* - \alpha_i)k(\mathbf{x}, \mathbf{x}_i) + b \tag{10}$$

where: K – the number of so-called support vectors (SV).

The Bases of Imperialist Competitive Algorithm (ICA)

Described algorithm belongs to the group of evolutionary optimization methods. ICA algorithm developers were inspired by socio-political phenomena related to the territorial expansion of the great powers in the late nineteenth and early twentieth century [10].

The first stage of the algorithm is the random selection of a population of solutions called countries. Vector in size $1xn$ is meant as country, where n is the number of parameters among which the solution is sought. Each country can act as either colony or imperialist. The role which it plays depends on its strength, which is calculated using cost function defined by the programmer. Imperialist together with the belonging colonies create an empire.

The most important part of the algorithm is that individual empires compete with each other. As a result of competition the weakest empires lose their colonies for stronger powers. The competition between empires leads to the movement of colonies in the direction of imperialist, under whose control they currently are. If as a result of this movement the cost function value of one of the colonies is lower than the value, which is assigned to the imperialist's state, then the colony will take its place. The objective of the algorithm is to create the dominant empire with colonies highly concentrated on one state – the imperialist. More detailed description of the mechanism of the ICA can be found in [11].

Plants accepted for identification

Black-Box identification method using SVR and ICA has been tested on linear and non-liner systems. As an example of linear system (plant I) was taken thermal plant defined by the equation (difficult plant in control)

$$\begin{aligned} &y(t) = 1{,}77y(t-1) - 0{,}81y(t-2) + \\ &0{,}08u(t-15) + 0{,}07u(t-16) \end{aligned} \tag{11}$$

As an example of non-linear system (plant II) was taken servo model of the robot manipulator [12] defined by the differential equation

$$\begin{aligned} &y(t) = 0{,}2y^2(t-2) + 0{,}2y(t-1) + 0{,}25y(t-2) + \\ &0{,}25u(t-1) + 0{,}45\,sin\left[0{,}5\left(y(t-1) + y(t-2)\right)\right] \\ &cos\left[0{,}5\left(y(t-1) + y(t-2)\right)\right] \end{aligned} \tag{12}$$

In both cases, the sampling frequency was 2 Hz.

Regressors selection

To define the optimal number of regressors, the impact of the number of regressors on the fit of the model to the test set was analysed. ICA algorithm was used, which had to adjust parameters of SVR network for different, predetermined input vectors, i.e. the number of samples of past observations, so as to find the best solution. In the study it was assumed that the input vector SVR has the NARX structure. The results of the analysis of all possible combinations in relation to n_u:1÷30 and n_y:1÷10 for the plant I and II are presented in Figures 1 and 2.

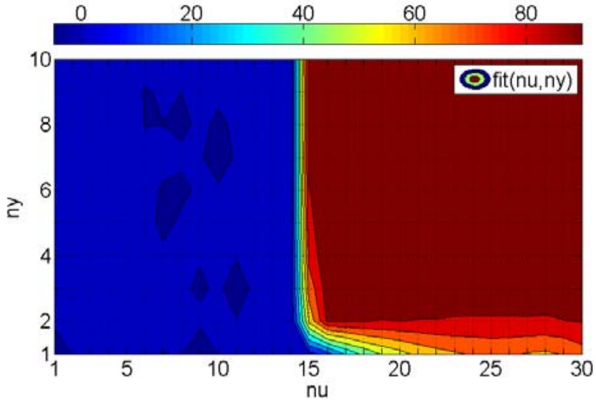

Fig.1. Impact of the number of samples of past input and output observations on the *fit* indicator for the plant I

A marked increase in the level of adjustment for the plant I is noticeable in the point (16,2). The number of regressors of the signal y should be equal to the order of a plant, which for an object described with equation 11 was 2. The number of regressors of the signal u is related to the transport delay and the order of plant. In the studied case the transport delay was 7 s, which, after taking into account the sampling frequency, gave 14 samples related to the transport delay. On the basis of the analysis the adopted regressors vector for plant I has the following form

$$[u(t-1),\ldots,u(t-16),y(t-1),y(t-2)] \tag{13}$$

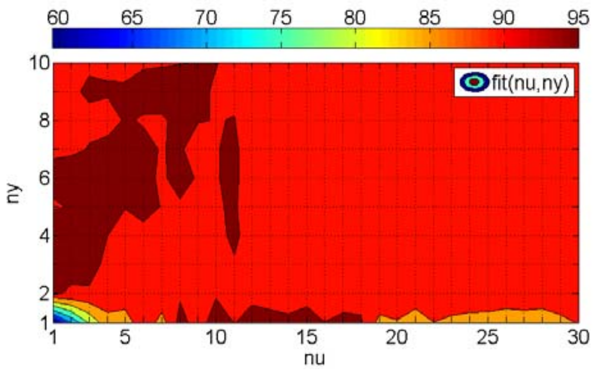

Fig.2. Impact of the number of samples of past input and output observations on the *fit* indicator for the plant II

Analysis of the impact of the number of regressors for the plant II showed that satisfactory fit was obtained for the point (1,2). Therefore a vector of regressors in this form has been adopted.

$$[u(t-1),y(t-1),y(t-2)] \tag{14}$$

Description of the SVR parameters tuning method using ICA

After selecting the optimal number of regressors i.e. the structure of input vector (equation 13 and 14) for SVR, tuning method shown in Figure 3 was performed.

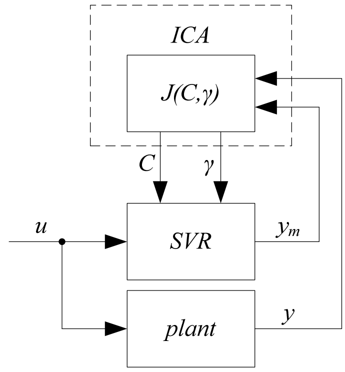

Fig.3. General scheme of SVR tuning method

The aim of ICA algorithm is to find the optimal set of parameters C and γ, which are required for proper operation of the SVR network and which ensure the best fit of the output network signal to the output signal of the identified plant (I i II). Tuning by ICS algorithm is therefore to minimize the objective function

$$J(C,\gamma)=100-fit \tag{15}$$

where: C, γ – parameters of the SVR network.

ICA algorithm parameters adopted in tuning process were:
- the number of countries 20
- the initial number of imperialists-states 5
- the number of iteration 20.

The limits of optimised parameters were:
- parameter C: 600 ÷ 80000
- parameter γ: 10^{-6} ÷ 10^{-4}

The evaluation of fitting was performed with Mean Squared Error (MSE) and *fit* indicator which is calculated according to the formula [13]

$$fit=100\%\left(1-\frac{\sqrt{(y_1-y_{m1})^2+\cdots+(y_n-y_{mn})^2}}{\sqrt{(y_1-\bar{y})^2+\cdots+(y_n-\bar{y})^2}}\right) \tag{16}$$

where: $y_{1\ldots n}$ – the sample of the plant output, $y_{m1\ldots mn}$ – the sample of the SVR output, $\bar{y}$ – the mean of the sample of the plant output.

Results

Proposed method of identification was compared with ARX model for linear plant and NARX model for non-liner plant. These models were built using System Identification Toolbox 7.4.3 of Matlab R2010b. SVR network was also implemented in Matlab R2010b environment. This network used the SMO algorithm which is available in LIBSVM toolbox [14]. The kernel function was Gaussian Radial Basis Function (RBF), RBF the satisfied by SVR kernel according to Marcer's theorem. Both on training and testing mode a fault tolerance was assumed to be ε=0.001. Training and testing data ware divided into two equal subsets. The general mode of SVR network working in test mode is shown in Figure 4.

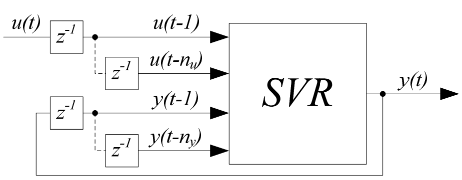

Fig.4. The general mode of action of SVR during testing

The SVR responses and plant I and II for training and testing data are shown in Figure 5 and 6.

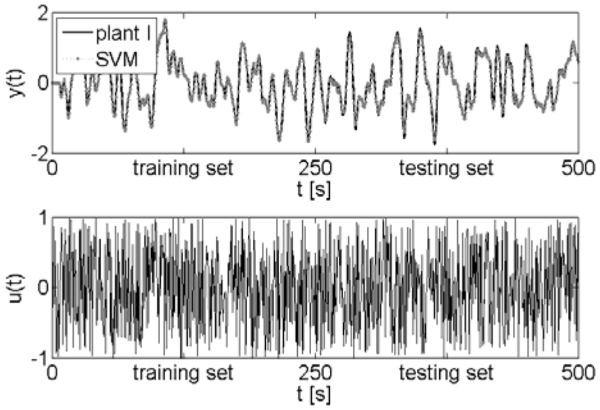

Fig.5. The responses for SVR and plant I

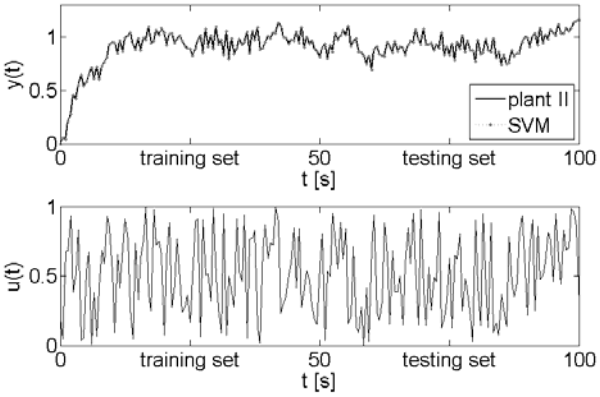

Fig.6. The responses for SVR and plant II

The values of calculated indicators with SVR network parameters are summarized in Tables 1 and 2.

Table 1. Values of indicators for a linear plant

	fit [%]		MSE [-]	
	training set	testing set	training set	testing set
ARX	100	100	0	0
SVR C=80000 γ=1,139*10^{-5}	99,27	91,39	2,34*10^{-5}	3,04*10^{-4}

Table 2. Values of indicators for a non-linear plant

	fit [%]		MSE [-]	
	training set	testing set	training set	testing set
NARX	97,66	95,61	2,74*10^{-5}	1,83*10^{-5}
SVR C=49439 γ=5,31*10^{-5}	99,95	99,6	2,6*10^{-7}	1,52*10^{-7}

The proposed method of identification for linear plant is characterized by the lower match indicator and higher mean squared error in comparison with the classical model ARX. The proposed model of identification, though, shows a better match to NARX model for non-linear plant.

Summary

In this article were presented the capabilities of using ICA for SVR network tuning. Also the capability of using ICA and SVR in choosing the optimal set of regressors was shown. According to these results SVR network better deals with the identification of non-linear plants which is related to the non-linear nature of the SVR. The fit indicator for non-linear plant is about 4% higher with respect to NARX model.

The authors are planning to test in further works the method described in this article for creating a models of real systems in the presence of the measurement noises.

REFERENCES

[1] Juditsky A., Hjalmarsson H., Benveniste A., Delyon B., Ljung L., Sjoberg J., Zhang Q., Nonlinear black-box models in system identification: Mathematical foundations, Automatica, 31 (1995), n.12, 1725-1750

[2] Sjoberg J., Zhang Q., Ljung L., Benveniste A., Delyon B., Glorennec P.-Y., Hjalmarsson H., Juditsky A., Nonlinear black-box modeling in system identification: a unified overview, Automatica, 31 (1995), n.12, 1691-1724

[3] Gretton A., Doucet A., Herbrich R., Rayner P.W.J., Schölkopf B., Support Vector Regression for Black Box system identification, 11th IEEE Workshop on Statistical Signal Processing (2001)

[4] Rivals I., Personnaz L., Black-box modeling with State-Space Neural Networks, w: Neural Adaptive Control Technology, World Scientific (1996), 237-264

[5] Ranković V.M., Nikolić I.Z., Identification of Nonlinear Models with Feedforward Neural Network and Digital Recurrent Network, FME Transactions, 36 (2008), n.2, 87-92

[6] Smola A., Scholkopf B., A tutorial on support vector regression, Statistics and Computing, 14 (2004), n.3, 199-222

[7] Salat R., Osowski R., Support Vector Machine for soft fault location in electrical circuits, Journal of Intelligent and Fuzzy Systems, 22 (2011), n.1, 21–31

[8] Salat R., Awtoniuk M., Odwzorowanie dynamiki pracy regulatora PID zaimplementowanego w sterowniku PLC za pomocą Least Squares-Support Vector Machines, Przegląd Elektrotechniczny, 5a (2012), 227-230

[9] Vapnik V., Statistical Learning Theory, Wiley-Interscience (1998)

[10] Atashapaz Gargari E., Lucas C., Imperialist competitive algorithm: An algorithm for optimization inspired by imperialist competition, IEEE Congress on Evolutionary Computation (2007)

[11] Awtoniuk M., Dostrajanie regulatora z pominięciem eksperymentu identyfikacji obiektu przy użyciu Imperialist Competitive Algorithm, Elektronika: konstrukcje, technologie, zastosowania, 12 (2012), 46-49.

[12] Han W.-Y., Han J.-W., Lee C.-G., Development of a Self-tuning PID Controller based on Neural Network for Nonlinear Systems, 7th Mediterranean Conference on Control and Automation (1999)

[13] Huyck B., De Brabanter K., Logist F., De Brabanter J., Van Impe J., De Moor B., Identification of a pilot scale distillation column: A kernel based approach, 18th World Congress of the International Federation of Automatic Control (2011)

[14] Chang ., Lin C., LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology 2 (2011), n.3, 1-27

***Authors**: dr inż. Robert Sałat, mgr inż. Michał Awtoniuk, dr inż. Krzysztof Korpysz, Warsaw University of Life Sciences, Department of Fundamental Engineering, ul. Nowoursynowska 164, 02-787 Warszawa, E-mail: robert_salat@sggw.pl, michal_awtoniuk@sggw.pl, krzysztof_korpysz@sggw.pl.*