Adaptive Immune Multiobjective Algorithm for LSSVM-based NARMAX model

Xia Zhou*, Sha Liu
School of Mechanical and Electrical Engineering, Jinling Institute of Technology, Nanjing, China

*Corresponding author e-mail: zhouxia@jit.edu.cn

Abstract. The LSSVM-based NARMAX model and the identification of the model are studied in the paper. By extending the immune cell subset theory and clone selection principle, a novel adaptive immune multiobjective algorithm (AIMA) is proposed. It is shown that, according to the variation of subsets in the memory set, the AIMA can produce different cell subsets adaptively, and thus regulate the related model parameters adaptively to realize the optimization. The simulation result demonstrates that with the parameters identified by the AIMA, the LSSVM-based NARMAX model has a satisfactory accuracy and it can be used to solve dynamic modeling problems.

1. Introduction

There are many ways to identify the nonlinear dynamic model. Among them, Nonlinear AutoRegressive Moving Average with eXogenous inputs (NARMAX), proposed by Billings, etc, provides a unified representation for finite realizable nonlinear systems[1]. Various models such as Hammerstein model, Wiener model, nonlinear time series model, etc, can be viewed as subsets of the NARMAX. Owing to its abilities such as high precision of approximation and fast convergence, NARMAX has been widely used in applications recently[2-4].

As an optimization problem, model identification has more than one objective in practice, where the fitting capability and the prediction capability need to be taken into account simultaneously. For this reason, adopt a multiobjective optimization method in the model identification to consider both capacities may receive a better result.

Be different from the single objective optimization, since the objectives in multiobjective optimization problems are conflict in many cases, it is a challenge to obtain a unique optimal solution to meet every objective. Thus, generally the notion of 'optimal' is transformed into 'Pareto optimal', which is also called the non-dominated optimal[5].

With the development of life science and artificial intelligence theory, immune algorithm has drawn much attention recently which is also a new branch of multiobjective optimization algorithm. On the basis of immune cell subset theory[6] and clone selection principle[7], this paper proposes a novel adaptive immune multiobjective algorithm (AIMA) for the parameter identification of dynamic model. The new algorithm divides the immune cells into multiple subsets based on the order of the input-output, and the quantity of each subset is changed adaptively according to the proportion of the optimal solutions in the Pareto set. Through this way, the AIMA can adjust the model parameters adaptively to achieve the optimization. Due to its data-driven nature, the proposed algorithm can
identify the related parameters directly without using too much prior knowledge, thus less human intervention is needed.

The rest of this paper is organized as follows: section 2 presents the improved NARMAX model and the parameter searching method in the model identification. Section 3 describes the proposed new algorithm for model parameter searching. Finally, the conclusion of the paper is outlined in section 4.

2. LSSVM-based NARMAX model

2.1 Structure of LSSVM-based NARMAX model

The general expression of the NARMAX model is [8]:

\[ y(t) = F(y(t-1), \ldots, y(t-n_y), u(t-1), \ldots, u(t-n_u), e(t-1), \ldots, e(t-n_e)) + e(t) \]  

where \( F(x) \) denotes a nonlinear function, \( u(t) \) and \( y(t) \) the input and output variables respectively, \( e(t) \) the noises. \( n_y, n_u \) and \( n_e \) the order of input, output and noise. Due to the immeasurability of the noise, a simplified NARMAX model (2) is more commonly used, which is also called NARX model [9].

\[ y(t) = F(y(t-1), \ldots, y(t-n_y), u(t-1), \ldots, u(t-n_u)) + e(t) \]  

Although various nonlinear functions can be used in the NARMAX model, for a given dynamic process, it is a challenge to select the function type. For this reason, the Least Squares Support Vector Machine (LSSVM) is employed in this paper to construct the NARMAX model. When the inputs of the LSSVM are expressed in the form of \( (y(t-1), \ldots, y(t-n_y), u(t-1), \ldots, u(t-n_u)) \), the LSSVM expression for output \( y(k) \) is naturally an extended NARMAX model.

The schematic diagram of LSSVM-based NARMAX model is shown in Figure 1, where the output of LSSVM \( y^* \) is a linear combination of intermediate nodes, and \( X^* \) represents the data for regression.

The inputs of the LSSVM \( X_i, (i=1 \ldots k) \), the weight \( \alpha_i, (i=1 \ldots k) \) and the threshold \( b \) are determined by LSSVM algorithm automatically. Thus by using the LSSVM, the issue of dynamic modeling can be simplified and the problem left is to choose the following parameters:

1. The input order \( n_u \) and the output order \( n_y \) of the model;
2. the form and parameters of the LSSVM core function \( K(X_i, X^*) \). If the RBF function is adopted as the core function, there is only one parameter left, the width \( \sigma^2 \);
3. the trade-off parameter \( \gamma \).

![Figure 1. Schematic diagram of the LSSVM-based NARMAX model.](image)

2.2 Precision objectives of dynamic modeling

The commonly used precision objectives of dynamic modeling include the residual variance (VAR) and the long term prediction error (LTPE) [9]. The VAR reflects the model's fitting capability and the LTPE reflects the model's predictive capability.

In order to evaluate the predictive capability of the data-driven model better, the LTPE is redefined in a new way that only the test data are taken into account. Suppose there are \( N \) groups of dynamic
data, we set the first $N_1$ groups as training data, the remaining $(N-N_1)$ groups as testing data, and if the biggest order value of input and output is $n_{\text{max}}$, the VAR and LTPE value can be calculated as formula (3) and (4).

\[
\text{VAR} = \frac{1}{N_1-n_{\text{max}}} \sum_{i=n_{\text{max}}+1}^{N_1} (y(i) - \hat{y}(i))^2
\]

\[
\text{LTPE} = \frac{1}{N-N_1} \sum_{i=N_1+1}^{N} (y(i) - \hat{y}(i))^2
\]

where $y(i)$ represents the real output value, $\hat{y}(i)$ represents the estimated output value based on the past real input and output values. $\hat{y}(i)$ represents the estimated output value based on the past real inputs and predictive outputs.

2.3 Study of choosing model parameters

For the dynamic model, the predictive output is calculated based on the information at the previous time sample. Thus, the order of input-output will produce a significant impact on the precision of the model. In order to study the order searching method for the LSSVM-based NARMAX model, tests on an existing model are performed to observe the influence of different input-output order values. Then the key of order searching method is summarized from the analysis of the results.

A model with input order 3 and output order 2 is chosen for the test experiment, which is shown as following:

\[
y(t) = -0.4y(t-1) + 0.3y(t-2)^2 + 0.5u(t-1)u(t-3) + 0.4u(t-2)
\]

The input sequence is a zero-mean, evenly distributed, random signal with magnitude 1. 200 groups of data are generated, in which the first 150 are used for training, and the other 50 are used for validation. Then, by using the iterative searching method proposed in[10], the parameters of the corresponding LSSVM model are determined under given input-output order values, and the resulting accuracy of the model is shown in Table 1. The results are selected from the multiobjective Pareto optimal solutions under the preference of low LTPE value.

The Table 1 shows that, the VAR is not suited for selecting the order because the best VAR value is found at incorrect order $n_u=8$, $n_y=2$. On the other hand, it also shows that the LTPE and max_PE can be used as the optimization objectives because the objective values are best if the order values are correct.

However, Table 1 also shows that, for larger order values, both the LTPE and max_PE objectives are just a little worse than the objectives at correct order, which means that although using the LTPE or max_PE objectives in the order searching can easily avoid the incorrect smaller order values, it is possible to find a incorrect larger order values. To further investigate the objective values of LTPE and max PE, another test is performed for the case $n_u=3$, $n_y=2$. The result in Table 2 shows that there are conflicts between the LTPE and max_PE values, i.e. a better LTPE objective is corresponding to a worse max_PE objective.

| $n_u$ | $n_y$ | VAR    | LTPE   | max PE   |
|------|------|--------|--------|----------|
| 2    | 1    | 0.014797 | 0.01026 | 0.24629  |
| 2    | 2    | 0.005535 | 0.005721 | 0.2118   |
| 2    | 5    | 0.0020735 | 0.0074825 | 0.21471  |
| 2    | 8    | 0.0081472 | 0.014832 | 0.29654  |
| 3    | 1    | 0.00033907 | 0.001853 | 0.13351  |
| 5    | 1    | 9.9995e-005 | 0.00064245 | 0.070924 |
| 8    | 1    | 2.4225e-005 | 0.0014504 | 0.11472  |
| 3    | 2    | 2.3588e-005 | 0.00016237 | 0.044251 |
| 4    | 2    | 2.1884e-005 | 0.00024522 | 0.054824 |
| 6    | 2    | 1.8183e-005 | 0.00052991 | 0.078482 |
| 8    | 2    | 1.2821e-005 | 0.00077609 | 0.088735 |
| 3    | 3    | 3.2093e-005 | 0.00040917 | 0.065546 |
| 3    | 7    | 5.6616e-005 | 0.001124 | 0.10603  |
| 3    | 10   | 4.7076e-005 | 0.0017735 | 0.12838  |
| 6    | 5    | 4.3168e-005 | 0.0011575 | 0.11112  |
| 5    | 9    | 4.2941e-005 | 0.0022778 | 0.1461   |
For these reasons, a novel Adaptive Immune Multiobjective Algorithm (AIMA) is proposed in this paper for model parameter identification. Both the LTPE and max_PE objectives are taken into account in the optimization to determine the input-output order values, and the VAR value is used as a constraint to prevent the over-learning, that if LTPE value is smaller than the VAR value, the solution would be eliminated for punishment. Based on immune cell subset theory and clone selection principle, the AIMA can adaptively produce cell subsets for different order values, thus the large order dominant situation can be avoid and optimal small order is possible to appear.

Table 2  The LTPE and max_PE values for \( n_x = 3, n_y = 2 \)

| No | LTPE       | max_PE     |
|----|------------|------------|
| 1  | 0.00016237 | 0.044251   |
| 2  | 0.00016815 | 0.043962   |
| 3  | 0.00017286 | 0.043673   |
| 4  | 0.00018090 | 0.042589   |
| 5  | 0.00019263 | 0.040316   |

3. The Adaptive Immune Multiobjective Algorithm (AIMA)

This section describe the AIMA on the basis of the immune cell subset theory and clone selection principle. In the AIMA, memory cells are introduced to keep the Pareto optimal solutions, and the memory cells are classified according to the order values. If there are different order values in the memory cells, the algorithm is running in the multiple sets mode, otherwise it is running in the single set mode. In the following, the definitions of the operators are given first, then the algorithm is introduced in detail. After that, the algorithm is validated by means of identifying the parameters of a known system.

3.1 Operator definition

For descriptive convenience, following notations are given first.

(1) \( C \) represents the cell, which is a candidate solution in the algorithm;
(2) \( MC \) represents the memory cell, which is used to represent a Pareto optimal solution;
(3) \( MCS \) represents the memory cell group, and \( MCS_i \) represents the \( i \)-th memory cell subset.

**Definition 1:** The dominant relationship between multiobjective optimization solutions

Suppose the optimization problem has \( m \) objectives that can be described as: \( \text{min}(f_1(X), f_2(X), \ldots, f_m(X)) \). For two candidate solutions \( C_i \) and \( C_j \), if \( \forall t \in 1 \ldots m, f_t(C_i) \leq f_t(C_j) \), together with \( \exists s \in 1 \ldots m, f_s(C_i) < f_s(C_j) \), we say that \( C_i \) dominates \( C_j \), and marked it as \( C_i \succ C_j \); if \( \exists t \in 1 \ldots m, f_t(C_i) < f_t(C_j) \), together with \( \exists s \in 1 \ldots m, s \neq t, f_s(C_i) > f_s(C_j) \), we say that \( C_i \) has nothing to do with \( C_j \), and marked it as \( C_i \prec C_j \). A nondominated solution is a solution which can dominate all the rest solutions except the solutions that have nothing to do with it[10].

**Definition 2:** Cell affinity calculating operator

In the AIMA, affinity represents the order's adaptability of each subset. The affinity value is only calculated when the algorithm is running in the multiple sets mode. The affinity value can be calculated as equation (6). The cells' affinity is equal to the reciprocal of the sum of all input-output order when they belong to the same memory cell subset, otherwise, the affinity value is set as 1.

\[
\text{aff}(x) = \begin{cases} 
1 & \text{if } x \in MCS_i \\
\frac{1}{\sum_{i \in MCS} n_i} & \text{if } x \notin MCS 
\end{cases}
\]  

(6)

**Definition 3:** Dominated number calculating operator

Dominated number calculating operator is used to calculate the number of the cells which dominate a given cell. The operator is denoted by \( \text{dom} \) and calculated by equation (7).

\[
\text{dom}(C) = \sum_j \text{sig}(C, C_j)
\]

\[
\text{sig}(C, C_j) = \begin{cases} 
1 & C_j \succ C \\
0 & \text{others}
\end{cases}
\]

(7)
Theorem 1: In a set, the cells which have smaller dominated number cannot be dominated by
the cells which have larger dominated number. Especially, the cells which are not dominated by other
cells are nondominated cells of the set.

Proof:
Based on the reduction to absurdity method,
suppose \( C_i \succ C_j \), and \( \text{dom}(C_i) > \text{dom}(C_j) \).
From \( C_i \succ C_j \), we can get \( \forall X, X \succ C_i \Rightarrow X \succ C_j \),
therefore, we have \( \text{dom}(C_i) \geq \text{dom}(C_j) + 1 \), which conflicts with the assumption. Following this proof,
if \( \text{dom}(C_i)=0 \), it means that there are no cells that can dominate \( C_i \), thus, \( C_i \) is the nondominated
individual in the set.

Definition 4: Clone operator
Clone means asexual reproduction, thus a group of identical cells can be descended from a single
original ancestor. Before cloning, the clone source should be selected in advance.

In order to prevent the algorithm from prematurity, minimal number of the clone source is set as
20\% of the total number. If the number of nondominated individuals and memory cells reaches the
minimal size, these cells will be selected as the clone source directly. If not, other dominated cells will
be selected to join the clone source where the cells dominated by fewer cells have the priority.

Since the algorithm may be running in the single set mode or the multiple sets mode, the clone
operators are defined respectively. When it is running in the single set mode, the order of all
individuals are identical; the numbers of clone cells are evenly distributed. When it is running in the
multiple sets mode, the clone proportions are distributed according to the cells' affinity. The lower the
affinity is, the larger the clone size is. Suppose the cells need to be cloned are composed by \((C_1, C_2, ..., C_k)\), and the modified coefficient is \( M \), the clone size of cell \( C_j \) is:

\[
M_j = \text{int} \left( \frac{M \prod_{i=1}^{k} \text{aff}(C_i)}{\text{aff}(C_j)} \right)
\]

Definition 5: Mutation operator
After clone, mutation is implemented on the cells. In order to improve the searching speed, the
mutation proportion is set as 100\%. When the algorithm is running in the single set mode, mutation
operator is implemented only on the LSSVM parameters. When the algorithm is running in the
multiple sets mode, mutation operator is implemented on both the LSSVM parameters and the input-
output order.

Definition 6: Redundant memory cell removal operator
For the single set mode, the crowding distance method proposed in [10] will be adopted as the
removal strategy. For the multiple sets mode, removal methods will be determined according to the
current iteration phases, and the cell which has the biggest difference between the LTPE objective and
the max PE objective will be deleted one by one. At the beginning of the algorithm, in order to
prevent the algorithm from prematurity, the cells in subsets which have more individuals are more
likely to be removed; after certain generations, the cells in subsets which have less individuals will be
removed to accelerate the convergence speed.

3.2 Summary of the AIMA
With the operators described in section 3.1, the whole procedure of the AIMA can be summarized in
Figure.2. As aforementioned, the algorithm can be implemented in either the single set mode or the
multiple sets mode, thus an adaptive scheme is designed to select the mode depending on the
nondominated cells of current generation and the memory cells of the previous generation. In the
algorithm, these cells are classified according to their order values and the proportion of each subset is
calculated. If the proportion of a certain subset has exceeded a given percentage, the single set mode
will be implemented and the cells of the next generation will have the identical input-output order values.
The main differences between the single set mode and the multiple sets mode are shown as following:

1. In the multiple sets mode, the affinity values are calculated to determine the distribution of cloning number, while in the single set mode, the input-output order values are fixed; thus the affinity values will not be calculated any more and every clone source is cloned equally.

2. In the single set mode, the mutation operator is only performed on the parameters of the LSSVM, while in the multiple sets mode, the mutation operator is performed on both the parameters of the LSSVM and the input-output order values.

3. In the multiple sets mode, the cells of next generation is composed of three parts: new cells selected from mutated cells in the last generation, new cells produced according to the order values of subsets in the memory set, and new cells produced according to the random generated order values and the LSSVM parameters. In the single set mode, the next generation is composed of two parts only: new cells selected from the mutated cells in the last generation, and new cells produced according to the random LSSVM parameters (have the same order values).

Figure.2 Flowchart of the AIMA.

Main steps which have not described in section 3.1 are explained briefly as follows:

1. Initialization
   At the beginning of the algorithm, the candidate cells are initialized randomly and the memory sets are initialized to be empty. The input-output order values and LSSVM parameters are randomly constructed according to their ranges.

2. The mode switching condition
   It is a trade-off between the convergence capability and the computational burden of the algorithm, therefore two conditions of the switching from multiple sets mode to single set mode are given according to the quantity of input-output order values. When the quantity is below a certain number,
the mode is switched if all nondominated cells' order values are identical and not changed for three consecutive generations. When the quantity is beyond a certain number, the mode is switched if there is a subset in the memory set whose proportion exceeds a given percentage for three consecutive generation. After the switching, the cells which have different order will be cleared for fast convergence.

(3) The termination condition of the algorithm
The algorithm is terminated when it has entered into the single set mode and its memory sets are not varied for consecutive two generations, or it has been implemented for five generations in this mode.

3.3 Validation of the AIMA
In this section, the AIMA is validated by means of identifying the parameters of a known system (9).

\[ y(t) = -0.6y(t-1)+0.3y(t-2)+0.4u_1(t-1)+0.2u_1(t-2)u_2(t-3)-0.5u_2(t-2)u_2(t-4) \]  

(9)

The input sequence is a 2-dimensional, zero-mean, evenly distributed, random signal with magnitude 1. In the experiment, 200 groups of data are produced, the first 150 are used for training and the other 50 are used for testing. The range of order is set as [1, 10], the range of \( \gamma \) is set as [10, 2000] and the range of \( \sigma^2 \) is set as [0.001, 1000].

10 simulations are run for the AIMA to search for the order. In each simulation, the algorithm can be switched into the single set mode, which is the condition for the termination. The fastest switching is at the seventh generation and the slowest switching is at the sixteenth generation. 9 simulations successfully find the correct order (2, 4, 3) while only one simulation fails, finding the order (3, 4, 3).

A set of feasible parameters of LSSVM and its corresponding objective values are shown in Table 3.

| parameter | value | parameter | value |
|-----------|-------|-----------|-------|
| \( n_{u1} \) | 2 | \( \sigma^2 \) | 116.3 |
| \( n_{u2} \) | 4 | VAR | 2.6482e-005 |
| \( n_y \) | 3 | LTPE | 2.1423e-004 |
| \( \gamma \) | 1852.6 | max_PE | 0.068088 |

The comparison between the real values and the LSSVM estimation values are shown in Figure 3. The results show that the AIMA can find proper order values and LSSVM parameters, furthermore, due to its data-driven nature, the LSSVM can be easily adapted to other type of system without using the preliminary knowledge.
4. Conclusions

In order to solve the dynamic modeling problem, the LSSVM-based NARMAX model and the parameter searching method for it are studied in the paper. Based on the immune cell subset theory and clone selection principle, the AIMA is presented to identify the model parameters adaptively. The simulation results show that proper order values and LSSVM parameters can be found by the AIMA without too much preliminary knowledge. Furthermore, due to its data-driven nature, the LSSVM-based NARMAX model and the parameter searching based on AIMA can be easily adapted to solve other dynamic modeling problems.

References

[1] Chen S, Billings S A. Representations of nonlinear systems: the NARMAX Model. International Journal of Control. 49(1989)1013-1032.
[2] Shuai Liu, Zhi Li, Ruilin Lin. Study of the energetic electron fluxes at geostationary orbit with a time-varying NARMAX method. 2nd International Conference on Computer, Communications, and Control Technology, 2015.
[3] Tsai J, Wang C, Kuang C, et al. A NARMAX model-based state-space self-tuning control for nonlinear stochastic hybrid systems. Applied Mathematical Modelling. 34 (2010)3030-3054.
[4] Boynton R. J, Balikhin M A, Billings S A, et al. The analysis of electron fluxes at geosynchronous orbit employing a NARMAX approach. Journal of Geophysical Research: Space Physics. 118 (2013)1500-1513.
[5] Veldhuisen D A V, Lamont G B. Multiobjective evolutionary algorithms analyzing: the state-of-the-art. IEEE Trans Evolutionary Computation8 (2010) 125-147.
[6] Mosmann TR, Sad S. The expanding universe of T-cell subsets: Th1, Th2 and the more. Immunology today. 17(1996):138-146.
[7] Leandro N. de Castro, Fernando J. Von Zuben. Learning and Optimization Using the Clonal Selection Principle. IEEE Trans Evolutionary Computation. 6 (2002)239-251.
[8] Billings S A, Chen S. Extended Model Set, Global seta and threshold model identification of severely nonlinear systems. International Journal of Control. 50(1989)1897-1923.
[9] Vazquez KR, Fonseca CM, Fleming P J. Identifying the structure of nonLinear dynamic systems using multiobjective genetic programming. IEEE Transactions on systems, man, and cybernetics-part A: system and humans. 34(2004)531-545.
[10] Deb K, Pratap A, Agarwal S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation.6(2002)182-197.