Journal of Fundamental and Applied Sciences

Research Article

ISSN 1112-9867

Available online at http://www.jfas.info

DISCRETE-TIME SYSTEM IDENTIFICATION BASED ON NOVEL INFORMATION CRITERION USING GENETIC ALGORITHM

M. F. Abd Samad* and A. R. Mohd Nasir

Faculty of Mechanical Engineering, Universiti Teknikal Malaysia Melaka, 76100 Hang Tuah Jaya, Durian Tunggal, Malacca, Malaysia

Published online: 24 November 201

ABSTRACT

Model structure selection is a problem in system identification which addresses selecting an adequate model i.e. a model that has a good balance between parsimony and accuracy in approximating a dynamic system. Parameter magnitude-based information criterion 2 (PMIC2), as a novel information criterion, is used alongside Akaike information criterion (AIC). Genetic algorithm (GA) as a popular search method, is used for selecting a model structure. The advantage of using GA is in reduction of computational burden. This paper investigates the identification of dynamic system in the form of NARX (Non-linear AutoRegressive with eXogenous input) model based on PMIC2 and AIC using GA. This shall be tested using computational software on a number of simulated systems. As a conclusion, PMIC2 is able to select optimum model structure better than AIC.

Keywords: Akaike information criterion; genetic algorithm; model structure selection; parameter-magnitude information criterion; search method

Author Correspondence, e-mail: mdfahmi@utem.edu.my

doi: http://dx.doi.org/10.4314/jfas.v9i7s.54

1. INTRODUCTION

System identification is the field of approximating dynamic system models using experimental data [1]. Its basic idea is to compare the time dependent responses of the actual system to the identified model based on a performance function, hereby referred to as information criterion, giving a measure of how well the model response fits the system response [2]. The procedure of identification can be divided into several distinctive steps. These steps are data acquisition, model structure selection, parameter estimation and model validity tests [3]. Fig. 1 shows the flow of system identification. An identification procedure typically consists of estimating the parameters of different models, and next selecting the optimal model complexity within that set. Increasing the model complexity will decrease the systematic errors, however, at the same time the model variability increases [4].

The model structure selection stage is crucial in determining the form of model structure suitable to explain the problem at hand. When selecting a model structure, two considerations need to be evaluated. One is model accuracy and the other one is model parsimony known as variance and bias: $f(J) = Var(J) + Bias(J)$[5]. Hence, it is not a good idea to select the model with the smallest variance within the set because it will continue to decrease when more parameters are added.

In certain cases, there can be numerous choices of model structures to be considered. Even with the most effective information criterion, testing all model structures require a lot of time. This is why a suitable search method that optimizes the process becomes necessary.

There are many search methods introduced and one amongst them is genetic algorithm (GA). GA is recognized as a stochastic global search method. GA was developed by J. H. Holland in the 1960s at Ann Arbor, Michigan. Earlier, GA was called reproductive plan using genetic operators [6]. In his book 'Adaptation in Natural and Artificial Systems' which marks the emergence of GA, Holland explained the need for an adaptive system that is applicable to various types of problems with changing environment [7]. GA has since been studied and developed for many applications [8-10]. Application of GA has been widely used for optimal solution search through a probabilistically guided optimisation process which simulates genetic evolution. It is different compared with classical optimisation algorithm where GA is

not guided by local derivatives in search process.

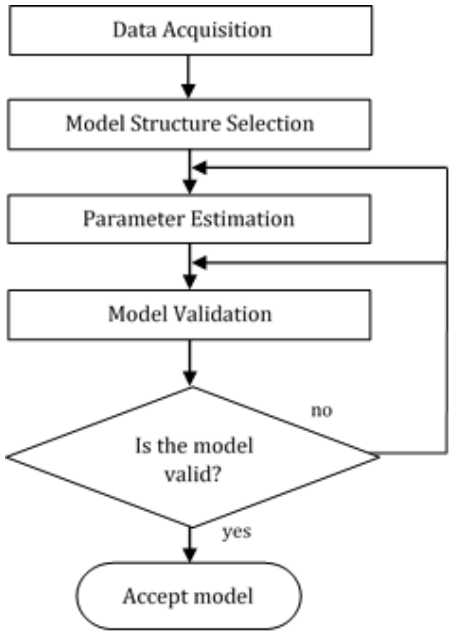

Fig.1. Flow Chart of System Identification

In this paper, the effectiveness of parameter magnitude-based information criterion 2 (PMIC2) is investigated, alongside Akaike information criterion (AIC) [11], using GA as search method, by testing on four simulated dynamic models in the form of difference equations model. These models are nonlinear autoregressive models with exogenous input (NARX) [5]. The benefit of using simulated models is the presence of an opportunity to compare the final model directly with the true model. The contribution of this paper is on application of PMIC2 on nonlinear systems that has many model choices such that genetic algorithm is incorporated within it. The paper also compares its performance to a widely known criterion i.e. AIC. The paper does not consider the model validation stage as all the results are compared directly to the true (known) model.

The next sections are as follows: Section 2 explains about information criterion; Section 3 explains the genetic algorithm; Section 4 explains the simulated models; Section 5 provides results and discussion and lastly Section 6 concludes the paper along with recommendation of future works.

2. INFORMATION CRITERION

Model complexity selection is the sub-problem of model selection [12]. Parsimony, working hypotheses, and strength of evidence are three principles that regulate the ability to make inferences [13]. Often, in order to deal with the bias-variance trade-off, the loss function or information criterion is augmented with a penalty term intended to guide the search for the "optimal" relationship penalizing undesired regressors, where regressors refer to possible terms and variables identified from model order and linearity specifications. Regularized estimation has been widely applied also in the context of system identification [14]. Several strategies have been proposed to avoid over-parameterization while utilizing all the data for training the model. The most popular strategy is to minimize a theoretically derived formula or criterion, which includes a goodness-of-fit index and a penalty factor for model complexity [15]. System identification can be framed as an optimization problem:

$$\hat{\theta} = \arg\min_{\theta} J_F(\theta, D_N) \quad (1)$$

where J_F (θ, D_N) measure how well the model described by parameter θ describes the measured data. A widely used variation of the estimation criterion includes a so-called 'regularization term' in the loss function to be minimized, that is:

$$\hat{\theta} = \arg\min_{\theta} J_F(\theta, D_N) + J_R(\theta, n) \quad (2)$$

In this case, θ is estimated by trading-off the data fitting term J_F (θ, D_N) and the regularization term J_R (θ,n) which act as a penalty to penalize certain parameters vector θ which describe 'unlikely' systems [14].

In today's literature, various types of models are proposed for system modelling such as linear autoregressive with exogenous input (ARX) model and nonlinear autoregressive with exogenous input (NARX) model [5].

The PMIC2 is developed from the approach of using parameter magnitude information in information criterion [16,17]. It includes a bias term or known as penalty function. It is written as follows:

$$PMIC2 = \sum_n (y(t) - \hat{y}(t))^2 + \sum_j \frac{1}{\theta_j} \quad (3)$$

where, θ_j is the magnitude of parameter in the model and j is the number of parameter. The other information criterion that has been widely used is Akaike information criterion (AIC) [11]. Estimation of the Kullback-Leibler information is the key to deriving the AIC, which was the first model selection criterion to gain widespread acceptance [15]. AIC is written as:

$$AIC = n\, ln\, \frac{RSS}{n} + 2p \quad (4)$$

where n is the number of observations, RSS stands for residual sum of squares, RSS/n is the maximized value of the likelihood function for the estimated model and p is the number of parameters in the statistical model [11]. RSS can be defined in formula below:

$$RSS = \sum_{t=k}^{N} \varepsilon^2(t) = \sum_{t=k}^{N} (y(t) - \hat{y}(t))^2 \quad (5)$$

where $\varepsilon(t)$ is the residual; $\hat{y}(t)$ and $y(t)$ are the k-step-ahead predicted output and actual output value at time t, respectively; and N is the number of data. The k-step-ahead prediction is used when the value of k depends on the output's smallest lag order in the selected model structure, which in turn depends on the variables selected by the search method.

3. GENETIC ALGORITHM

Genetic algorithm mimics evolution process. It conforms to the metaphor of natural biological evolution by application of the principle of survival of the fittest. In the context of system identification, it begins with genetic encoding of potential input-output relationship solutions of a system into chromosomes. A chromosome is therefore a string of code that represents a solution. Each position in the string is referred as gene. In a binary-represented GA, the variables and terms of a discrete-time system are represented by the genes of the chromosome as bit 1 for existence and bit 0 for omission [18]. The encoding is followed by selection process. The selection process refers to the process of selecting chromosomes from a set of

chromosomes (called population) to go through genetic operations. These genetic operators are crossover and mutation. The crossover operator operates by 'mating' two chromosome so that parts of their structures are exchanged. There are many varieties of crossover type e.g. single-point crossover, double-point crossover and uniform crossover. On the other hand, the mutation operator changes parts of the chromosome structure depending on specific rule e.g. whether it is coded as binary or real numbers.

After the genetic operations, a new population of different individual chromosomes is created. Based on a specific objective function, these chromosomes are assigned fitness values. In the context of system identification, such objective function refer to information criterion. Assuming a minimization of information criterion, chromosomes that has small criterion are given high fitness value while those that have big criterion are given 0 fitness value. As the process is repeated back to selection stage, only those chromosomes with high fitness are selected. Through the search process, population individuals with weak fitness are removed and a population with strong fitness is identified and maintained. This will ensure that better offsprings are produced from the parents. Fig. 2 shows the flow of genetic algorithm. The process in GA is stable and robust and can identify global optimal solution of a problem. In Duong and Stubberud, GA is used in estimating the parameters of difference functions in the form of ARMAX (Auto-Regressive Moving Average with eXogenous input) and a cubic non-linear function [19]. Table 1 shows six parameters in GA which define the characteristics of its process.

It had been shown in past researches that GA enables quick evaluation of model structures. However, the particular issue of parsimony of model structure still arises. This issue is addressed by considering the significance of the terms and variables. Various approaches are available such as the use of locally linear and cross-bilinear models [20], penalty functions [21], and information criteria such as Akaike's information criterion and the B-information criterion [22]. It allows the complexity of the structure to be reduced, gaining a parsimonious model of acceptable accuracy level. The penalty function is a straightforward approach when searching global solution with a GA [21].

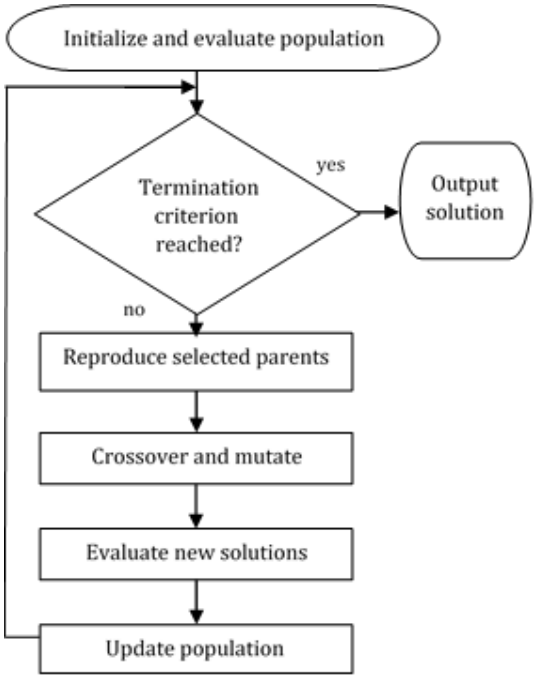

Fig.2. Flow chart of genetic algorithm

Table 1. Parameters and characteristic in GA

Parameters	Characteristics
Population size	Free parameter which trades off coverage of the search space against the time required to compute the next generation.
Length of chromosome	The length of chromosome is generally equal to the number of variables to be optimised and the lags of every variable encoded in binary. A chromosome represents a combination of genes.
Crossover probability	This probability controls the frequency at which the crossover occurs for every chromosome in the search process. This is a number between (0, l) which is determined according to the sensitivity of the variables of the search process. The crossover probability is chosen small for systems with sensitive variables. For every crossover operation, a random function generates a random number which is compared with the crossover probability. If it is less than or equal to the crossover probability then the crossover operation takes place.
Mutation	This probability controls the frequency at which mutation occurs for every

probability	gene of a chromosome in the search process. The mutation operations will be determined by a random function which generates a number between 0 and 1. If the random number is less than or equal to mutation probability then the mutation operation occurs. The selection of the mutation probability is dependent on the sensitivity of the objective function to the variables. In the case of binary coding, the mutation determines the number of bits on which mutation will be carried out.
Information criterion	Within the context of system identification, this is the main evaluation function based on which the fitness of each member of the new generation is determined. The members identified as 'fit' survive and enter a mating pool to reproduce the next generation

4. SIMULATION SETUP

In this simulation, four NARX models are simulated using computer simulation software MATLAB. All models are denoted as Model 1, Model 2, Model 3 and Model 4 and, in the search simulation, each model is assumed to have d.c. level (constant). The following are the models written as linear regression models, its specifications, number of correct regressors and number of possible regressors:

Model 1:

$$y(t) = 0.5y(t-1) + 0.3u(t-2) + 0.3y(t-1)u(t-1) + 0.5u(t-1)^2 + e(t)$$

Specification: nonlinearity, l=2, assumed maximum output order, n_y=1, assumed maximum input order, n_u=3

Number of correct regressor = 4 out of 15

Number of possible model = 32767

Model 2:

$$y(t) = 0.1y(t-3) + 0.3u(t-1) + 0.3y(t-2)u(t-1) + 0.3u(t-1)^2 + e(t)$$

Specification: l=2, n_y=3, n_u=1

Number of correct regressor = 4 out of 15

Number of possible model = 32767

Model 3:

$$y(t) = 0.1y(t-2) + 0.2u(t-3) + 0.1y(t-1)^2 - 0.2y(t-1)y(t-2) + 0.2u(t-2)u(t-3) + e(t)$$

Specification: l=2, n_y=2, n_u=3

Number of correct regressor = 5 out of 21

Number of possible model = 2097151

Model 4:

$$y(t) = 0.1y(t-2) + 0.1u(t-1) + 0.1y(t-1)^2 - 0.2y(t-3)^2 + 0.2u(t-1)u(t-2) + e(t)$$

Specification: l=2, n_y=3, n_u=2

Number of correct regressor = 5 out of 21

Number of possible model = 2097151

The input $u(t)$ is generated from a random uniform distribution in the interval [-1, 1] to represent white signal, while noise $e(t)$ is generated from a random uniform distribution [-0.01, 0.01] to represent white noise. Five hundred data points are generated from all models. The specification of the algorithm is fixed for all models where the population size, *popsize* is set to 500, the maximum generation is 100, the mutation probability, p_m = 0.01 and the crossover probability, p_c = 0.6. This paper uses roulette-wheel selection, single-point crossover and binary bit mutation.

5. RESULTS AND DISCUSSION

Tables 2 to 5 presents the simulation results made using genetic algorithm, with PMIC2 and AIC as information criterion, for all models. The selected models are based on final selection made in genetic algorithm. The simulated model is denoted as S.M.

Table 2. Results on Model 1

Regressor	S.M.	PMIC2	AIC
d.c.			-0.01
$y(t\text{-}1)$	0.5	0.5	0.5
$u(t\text{-}1)$			
$u(t\text{-}2)$	0.3	0.3	0.29
$u(t\text{-}3)$			
$y(t\text{-}1)^2$			
$y(t\text{-}1)u(t\text{-}1)$	0.3	0.3	0.3
$y(t\text{-}1)u(t\text{-}2)$			0.01
$y(t\text{-}1)u(t\text{-}3)$			
$u(t\text{-}1)^2$	0.5	0.5	0.5
$u(t\text{-}1)u(t\text{-}2)$			0.03
$u(t\text{-}1)u(t\text{-}3)$			
$u(t\text{-}2)^2$			
$u(t\text{-}2)u(t\text{-}3)$			
$u(t\text{-}3)^2$			

Table 3. Results on Model 2

Regressor	S.M.	PMIC2	AIC
d.c.			-0.01
$y(t\text{-}1)$			-0.01
$y(t\text{-}2)$			
$y(t\text{-}3)$	0.1	0.1	0.09
$u(t\text{-}1)$	0.3	0.3	0.3
$y(t\text{-}1)^2$			
$y(t\text{-}1)y(t\text{-}2)$			
$y(t\text{-}1)y(t\text{-}3)$			0.3
$y(t\text{-}1)u(t\text{-}1)$			
$y(t\text{-}2)^2$			
$y(t\text{-}2)y(t\text{-}3)$			
$y(t\text{-}2)u(t\text{-}1)$		0.3	
$y(t\text{-}3)^2$	0.3		
$y(t\text{-}3)u(t\text{-}1)$			
$u(t\text{-}1)^2$	0.3	0.3	0.3

Table 4. Results on Model 3

Regressor	S.M.	PMIC2	AIC
d.c.			-0.01
$y(t\text{-}1)$			
$y(t\text{-}2)$	0.1	0.1	0.1
$u(t\text{-}1)$			
$u(t\text{-}2)$			-0.01
$u(t\text{-}3)$	0.2	0.2	0.2
$y(t\text{-}1)^2$	0.1	0.1	0.1
$y(t\text{-}1)y(t\text{-}2)$	-0.2	-0.2	-0.19
$y(t\text{-}1)u(t\text{-}1)$			
$y(t\text{-}1)u(t\text{-}2)$			
$y(t\text{-}1)u(t\text{-}3)$			0.01
$y(t\text{-}2)^2$			
$y(t\text{-}2)u(t\text{-}1)$			
$y(t\text{-}2)u(t\text{-}2)$			
$y(t\text{-}2)u(t\text{-}3)$			-0.01
$u(t\text{-}1)^2$			
$u(t\text{-}1)u(t\text{-}2)$			-0.01
$u(t\text{-}1)u(t\text{-}3)$			0.01
$u(t\text{-}2)^2$			
$u(t\text{-}2)u(t\text{-}3)$	0.2	0.2	0.19
$u(t\text{-}3)^2$			

Table 5. Results on Model 4

Regressor	S.M.	PMIC2	AIC
d.c.			-0.01
$y(t\text{-}1)$			
$y(t\text{-}2)$	0.1	0.1	0.09
$y(t\text{-}3)$			
$u(t\text{-}1)$	0.1	0.1	0.1
$u(t\text{-}2)$			
$y(t\text{-}1)^2$	0.1	0.1	0.1
$y(t\text{-}1)y(t\text{-}2)$			-0.01
$y(t\text{-}1)y(t\text{-}3)$			
$y(t\text{-}1)u(t\text{-}1)$			
$y(t\text{-}1)u(t\text{-}2)$			
$y(t\text{-}2)^2$		-0.2	
$y(t\text{-}2)y(t\text{-}3)$			
$y(t\text{-}2)u(t\text{-}1)$			0.01
$y(t\text{-}2)u(t\text{-}2)$			-0.1
$y(t\text{-}3)^2$	-0.2		-0.18
$y(t\text{-}3)u(t\text{-}1)$			
$y(t\text{-}3)u(t\text{-}2)$			
$u(t\text{-}1)^2$			
$u(t\text{-}1)u(t\text{-}2)$	0.2	0.2	0.19
$u(t\text{-}2)^2$			

From the observation, PMIC2 has selected the same model structure as simulated model in model 1 and model 3. However, for model 2 and model 4, it selected the same regressor number as simulated model with the correct magnitude of parameter each. It selected only one regressor different which is $y(t\text{-}3)^2$ instead of $y(t\text{-}2)u(t\text{-}1)$ for model 2 and $y(t\text{-}2)^2$ instead of $y(t\text{-}3)^2$ for model 4. On the other hand, AIC cannot select the right model for all models and

over-modelling occurred for all models. All the models selected by AIC have d.c. level while there is no d.c level for all simulated models.

6. CONCLUSION

From this simulation, PMIC2 proved that it performed well better than AIC in selecting a model structure as it shows that it can select the correct model for 2 models and almost select the true model for the other two (only one regressor different for each model). This was achieved without having to try all possible models i.e. using GA. GA as a search method is very encouraging when it consumes a little time in the simulation process.

7. ACKNOWLEDGEMENT

The authors would like to acknowledge all the support from Universiti Teknikal Malaysia Melaka and Ministry of Higher Education Malaysia for research grant FRGS/1/2015/TK03/UTEM/02/F000271.

8. REFERENCES

[1] Saleem A., Taha B, Tutunji T. and Al-Qaisia A. Identification and cascade control of servo-pneumatic system using particle swarm optimization. Simulation Modelling Practice and Theory, 2015, 52: 164-179.

[2] Alfi A. and Fateh M.M. Parameter identification based on a modified PSO applied to suspension system. Journal of Software Engineering & Applications, 2010, 3: 221-229.

[3] Keesman K.J. System Identification: An Introduction. Springer Science & Business Media, 2011.

[4] Riddef' F.D., Pintelon R., Schoukens J. and Gillikinb D.P. Modified AIC and MDL model selection criteria for short data records. IEEE Transactions on Instrumentation and Measurement, 2004, 54: 144-150.

[5] Ljung L. System Identification: Theory for the User. 2nd ed. Upper Saddle River: Prentice Hall, 1999.

[6] Holland, J. H. Genetic Algorithms and the Optimal Allocation of Trials. SIAM Journal of Computing,1973, 2(2): 88-105.

[7] Holland. J. H. Adaptation in Natural and Artificial Systems. MIT Press ed. Massachusetts

Institute of Technology, 1992.

[8] Zalzala, A. M. S. and Fleming, P. J. (Eds). Genetic algorithms in engineering systems. London, UK:The Institution of Electrical Engineers, 1997.

[9] Goldberg, D. E. Genetic algorithms in search, optimization and machine learning. Reading,Massachusetts: Addison Wesley Publishing Company, 1989.

[10] Gen, M. and Cheng, R. Genetic algorithm and engineering design. New York: John Wiley, 1997.

[11] Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 1974, 19(6).

[12] Kristinsson K. and Dumont G.A. System identification and control using genetic algorithms. IEEE Transactions on Systems, Man, and Cybernetics, 1992, 22(5): 1033-1046.

[13] Xiaoyong L., Huajing F. and Zhaoxu C. A novel cost function based on decomposing least-square support vector machine for Takagi-Sugeno fuzzy system identification. IET Control Theory & Applications,2013, 8(5): 338-347.

[14] Prando G., Pillonetto G. and Chiuso A. The role of rank penalties in linear system identification. IFAC-Papers online,2015, 48(28): 1293-1300.

[15] Xinshu X., Ying L. and Mukkamala R. A model order selection criterion with applications to cardio-respiratory-renal systems. IEEE Transactions on Biomedical Engineering, 2005, 52(3): 445-453.

[16] Samad M.F.A., Jamaluddin H., Ahmad R., Yaacob M.S. and Azad A.K.M. Effect of penalty function parameter in objective function of system identification. International Journal of Automotive and Mechanical Engineering, 2013, 7: 940-954.

[17] Samad M.F.A and Nasir A.R.M. Parameter Magnitude-Based Information Criterion for Identification of Linear Discrete-Time Model. In Proceedings of Innovative Research and Industrial Dialogue'16: 2017, 129-130.

[18] Jamaluddin, H., Abd Samad, M. F., Ahmad, R. and Yaacob, M.S., Optimum Grouping in a Modified Genetic Algorithm for Discrete-Time, Non-Linear System Identification, Proc. IMechE, Part I: J. Systems and Control Engineering, 2007,221(7): 975-989.

[19] Duong, V. and Stubberud, A. R. System identification by genetic algorithm. In IEEE Aerospace Conference Proceedings, 2002, 5: 2331–2337.

[20] Wei, H. L., Billings, S. A., and Liu, J. Term and variable selection for non-linear system identification. International Journal of Control, 2004, 77(1): 86–110.

[21] Michalewicz, Z. Genetic algorithms+data structures=evolution programs. Second edition. Berlin–Heidelberg: Springer Verlag, 1994.

[22] Veres, S. M. Structure selection of stochastic dynamic systems. OPA (Amsterdam) BV, 1991.

How to cite this article:
Abd Samad M F, Mohd Nasir A R. Discrete-time system identification based on novel information criterion using genetic algorithm. J. Fundam. Appl. Sci., 2017, *9(7S), 584-599.*