Applied Soft Computing 13 (2013) 939–946

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

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

Evolving Takagi–Sugeno fuzzy model based on switching to neighboring models

Ahmad Kalhor [a,*], Babak N. Araabi [a,b], Caro Lucas [a,b]

[a] Control and Intelligent Processing Center of Excellence, School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran
[b] School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

ARTICLE INFO

Article history:
Received 17 August 2011
Received in revised form 25 April 2012
Accepted 18 September 2012
Available online 3 October 2012

Keywords:
Fuzzy inference systems
Evolving Takagi–Sugeno fuzzy model
Neighboring models
Split and merge
Load forecasting
Sunspot number prediction

ABSTRACT

In this paper, we propose a new online identification approach for evolving Takagi–Sugeno (TS) fuzzy models. Here, for a TS model, a certain number of models as neighboring models are defined and then the TS model switches to one of them at each stage of evolving. We define neighboring models for an in-progress (current) TS model as its fairly evolved versions, which are different with it just in two fuzzy rules. To generate neighboring models for the current model, we apply specially designed split and merge operations. By each split operation, a fuzzy rule is replaced with two rules; while by each merge operation, two fuzzy rules combine to one rule. Among neighboring models, the one with the minimum sum of squared errors – on certain time intervals – replaces the current model.

To reduce the computational load of the proposed evolving TS model, straightforward relations between outputs of neighboring models and that of current model are established. Also, to reduce the number of rules, we define and use first-order TS fuzzy models whose generated local linear models can be localized in flexible fuzzy subspaces. To demonstrate the improved performance of the proposed identification approach, the efficiency of the evolving TS model is studied in prediction of monthly sunspot number and forecast of daily electrical power consumption. The prediction and modeling results are compared with that of some important existing evolving fuzzy systems.

1. Introduction

Online identification approaches offer appropriate solutions for modeling the demands of a large variety of non-stationary processes such as many time-varying industrial processes and natural systems. Models with fixed structure cannot adapt with structural variations of such processes. In contrast, evolving models can adapt to these processes by changing their structures and parameters. During the last decade, evolving fuzzy models have been introduced and developed vastly for online modeling applications [1]. It has been shown that fuzzy models, like modular networks, can be trained to partition a global task into several simpler subtasks and allocate distinct networks to learn each subtask [2]. Particularly, the Takagi–Sugeno (TS) fuzzy model [3] has a sufficiently adaptable structure and can be easily evolved to deal with behavioral changes in the process. Actually, in a TS fuzzy model, one can easily change the parameters of local linear models (LLMs), adjust the fuzzy membership functions, and increase or decrease the number of rules, adaptively.

Two particularly influential works in the area of evolving fuzzy models are [4,5]. Kasabov proposes an adaptive online learning algorithm as a dynamic evolving neural-fuzzy inference system (DENFIS) in [4]. DENFIS is a member of the larger family of Evolving Connectionist Systems (ECoSs) [6,7]. In DENFIS, fuzzy inference rules are created by using maximum distance clustering, which is utilized in partitioning of the input space. Angelov and Filev introduce an online identification approach for the TS model in [5], where subtractive clustering along with a concept of potential is used to define the antecedent parts of the rules. In their evolving Takagi–Sugeno (ETS) learning algorithm, the rules may be replaced or increased based on what they call the 'potential' of incoming data. The parameters of consequent parts of rules are updated through recursive optimization. Modified and Extended versions of ETS approach are suggested in [8,9]. In [10], a hierarchical on-line self-organizing learning algorithm is proposed to identify a TS fuzzy model. The rules are generated according to the desired accuracy measured by output error and a boundary criterion. The rules may be pruned by observing an error reduction ratio criterion. In [11], by a learning of the structure (similar to [10]) an online identification method for a fuzzy neural network is proposed. This is applied to NARMAX time series prediction in two

* Corresponding author at: School of Electrical and Computer Engineering, University of Tehran, P.O. Box: 14395/515, Tehran 1439957131, Iran.
Tel.: +98 21 8863 0024; fax: +98 21 8863 3029; mobile: +98 912 380 8140.
E-mail addresses: akalhor@ut.ac.ir (A. Kalhor), araabi@ut.ac.ir (B.N. Araabi), lucas@ut.ac.ir (C. Lucas).

1568-4946/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2012.09.015

different feed-forward and recurrent models. An online growing and pruning strategy is proposed in [12] to identify an evolving TS model. In [13], an online clustering approach and support vector machines are utilized to generate fuzzy rules in a TS fuzzy model form. In [14], an online clustering-based learning algorithm for switching models is proposed, where parameters and their order can change while the system switches. An evolving fuzzy predictor is suggested in [15] which employs an evolving input–output clustering: clusters are added and merged adaptively; a variant of the Levenberg–Marquardt method is used in recursive nonlinear parameter estimation. In [16], to evolve a specific form of TS fuzzy model, a modified version of vector quantization is utilized through in an incremental learning algorithm: FLEXFIS. In [17], an online identification of a neuro-fuzzy model based on indirect fuzzy clustering is proposed: In this approach, spatial features of the incoming data points are not considered directly in clustering, which causes significant reduction on redundancy of the model. In [18], an online predictor model as habitually linear and transiently evolving TS fuzzy model is proposed. In this approach, the proposed model, with respect to output error, can follow the variations of a process with agility.

In many proposed online identification approaches for evolving fuzzy models, spatial features and computed output errors of incoming data points are utilized adaptively. Spatial features of the data points are utilized generally in frame of online clustering methods to change the structure of the model; and the output errors are utilized in recursive updating of the parameters of the model. Although the spatial feature of each observed data point can be exploited indirectly in all stages of the evolving of the model, each computed output error of an observed data point is utilized adaptively just for the in-progress model. As a result, the efficiency of such online identification algorithm is decreased because we practically take away the former informative data points when the structure of the model has changed.

In this paper, we propose an online identification approach which allows the utilization of the incoming data points and their corresponding output errors to train, concurrently, the current model and a model which will be used later as an evolved version. To this end, for a TS model, a certain number of models as neighboring models are defined and updated in parallel with the current one. Then, at each stage of evolving, the main TS model switches to one of the neighboring models.

The neighboring models – as fairly evolved versions of the current model – are generated by using specially designed split and merge operations. Actually, by each split operation, a fuzzy rule is replaced with two, and by each merge operation, two fuzzy rules combine to one. At each stage of evolving – which is held at certain times, the current model switches to one of the neighboring models which has the minimum of sum of squared errors (SSE).

In an online identification approach, it is important to reduce the computational load and redundancy of the model. To this end, in the proposed approach, we compute the output of neighboring models easily by using straightforward relations between them and the current model. This proposed ***ETS*** model which is based on switching to ***n***eighboring models is named in the rest of the paper as the 'ETS-N' model.

Also, utilizing membership functions which represent flexible validation regions for LLMs is an effective way to reduce the number of fuzzy rules in a first-order TS model [18,19]. Here, we utilize a modified structure of our former introduced TS models in [18,19] to reduce the number of the fuzzy rules as well as the number of the neighboring models; in spite of this, the main contribution, the learning method, and evolving strategy of the model are different, extensively.

The rest of the paper is organized as follows: Section 2 describes the ETS-N model and its online identification algorithm. Section 3 is devoted to case studies. There, the performance of the proposed ETS-N model is studied and compared with some other online identification approaches in prediction of the time series of monthly sunspot number and in a load forecasting problem. The paper is concluded in Section 4.

2. ETS-N fuzzy model

In this section, the structure of the ETS-N model is introduced; then updating of the linear models of ETS-N model, the split and merge operations (as tools for reconstructing) and the neighboring models are explained. At the end, the proposed online identification algorithm with an explanation about its computational complexity is presented.

2.1. The structure of the ETS-N fuzzy model

A fuzzy rule of the ETS-N model is described as follows:

$$\text{if } \mathbf{x} \text{ is } A_j \text{ then } y^j = \mathbf{z}^T\mathbf{w}_j \quad j = 1, 2, \ldots, M \tag{1}$$

where $\mathbf{x} \in \mathbf{R}^n$ is the input vector and $\mathbf{z}^T = [1\ \mathbf{x}^T]$, y^j is the local output of jth LLM, and $\mathbf{w}_j \in \mathbf{R}^{n+1}$ is the linear parameters vector of jth LLM (more generally, linear parameters of jth rule). The validity function of the jth rule is A_j, which is a fuzzy subset of the input space $\mathbf{R}^n$. The scalar output of model is computed as follows:

$$\hat{y} = \sum_{j=1}^{M} y^j \phi_j(\mathbf{x}),\ \phi_j(x) = \frac{A_j(\mathbf{x})}{S_M}$$
$$S_M = \sum_{h=1}^{M} A_j(\mathbf{x}) \tag{2}$$

where $\phi_j(\mathbf{x})$ denotes the membership function of normalized A_j. The A_j is defined as the sum of N_j standard Gaussian functions:

$$A_j(\mathbf{x}) = \sum_{i=1}^{N_i} A_{ji}(\mathbf{x})$$
$$A_{ji}(\mathbf{x}) = \exp\left(-\frac{1}{2}(x - c_{ji})^T \mathbf{S}_{ji}^{-1}(\mathbf{x} - \mathbf{c}_{ji})\right) \tag{3}$$

where $\mathbf{c}_{ji} \in \mathbf{R}^n$ and $\mathbf{S}_{ji} \in \mathbf{R}^{n\times n}$ denote the mean and the diagonal covariance matrix of Gaussian function A_{ji}, respectively. A typical contour of A_{ji}, as a standard Gaussian function, is a hyper-ellipse, therefore, A_{ji} can be defined by the center ($\mathbf{c}_{ji}$) and the diameter vector ($\mathbf{L}_{ji}$) of a contour:

$$A_{ji}(\mathbf{x}) = \exp\left(-\sum_{k=1}^{n}\left(\frac{\gamma(\mathbf{x}(k) - \mathbf{c}_{ji}(k))}{\mathbf{L}_{ji}(k)}\right)^2\right) \tag{4}$$

It is easy to see that $(\gamma^{-1}\mathbf{L}_{ji}(k))^2$ equals kth diagonal element of diagonal covariance matrix $\mathbf{S}_{ji}$, where γ^{-1} is an interpolation coefficient for membership functions, chosen typically 1/3 to make a reasonable interpolation [19]. Now, each ($\mathbf{L}_{ji}$, $\mathbf{c}_{ji}$) pair defines a standard hyper-rectangle (block) b_{ji}, centered at $\mathbf{c}_{ji}$, A three-dimensional scheme for a standard hyper-rectangle is shown in Fig. 1.

where x_1, x_2 and x_3 denote the coordinate axes of the input space. A block-based cluster (cluster) C_j is defined as the union of N_j standard blocks b_{ji}; $i = 1, 2, \ldots, N_j$. Cluster $\mathbf{C}_j$ is associated with membership function $\phi_j(.)$, constituting a fuzzy partition on $\mathbf{R}^n$. Fig. 2 shows a diagram of the structure of an ETS-N model with M fuzzy rules; the overall output of the model is the sum of the weighted outputs of LLMs, and each weight function is defined through a cluster.

Definition: Two rules of an ETS-N are considered adjacent if at least two blocks of their clusters are adjacent with each other. The