A New Approach to Dynamic Fuzzy Modeling of Genetic Regulatory Networks

Yonghui Sun*, Gang Feng, *Fellow, IEEE*, and Jinde Cao, *Senior Member, IEEE*

***Abstract*—In this paper, the dynamic fuzzy modeling approach is applied for modeling genetic regulatory networks from gene expression data. The parameters of the dynamic fuzzy model and the optimal number of fuzzy rules for the fuzzy gene network can be obtained via the proposed modeling approach from the measured gene expression data. One of the main features of the proposed approach is that the prior qualitative knowledge on the network structure can be easily incorporated in the proposed identification algorithm, so that the faster learning convergence of the algorithm can be achieved. Two sets of data, one the synthetic data, and the other the experimental SOS DNA repair network data with structural knowledge, have been used to validate the proposed modeling approach. It is shown that the proposed approach is effective in modeling genetic regulatory networks.**

***Index Terms*—Fuzzy clusters, fuzzy modeling, gene expression data, genetic regulatory networks, SOS DNA repair networks.**

I. Introduction

ONE OF THE major challenges in the postgenome era is to understand how interactions among the molecules in a cell and to determine its form and function. Methodologies to identify and analyze the complex biological networks that regulate metabolism are thus needed [1]–[3]. Furthermore, with the advance in gene microarray technologies [8], the expression levels of huge amount of genes over time can be measured simultaneously, which make it possible to investigate the complex biological processes at the molecular level, such as discriminating diseases types, testing drug effects, etc. [9]. "Reverse engineering," concerning the reconstruction and identification of genetic systems through gene expression data, is the central issue in computational molecular biology and has become one of the research focuses in the last few years [1]–[7].

Genetic regulatory networks (GRN), often abbreviated as genetic networks, seek to describe how genes or groups of genes interact with each other and to identify the complex regulatory mechanisms that control the activities of genes in living cells [5]. These networks can help us understand the intricate interactions of multiple genes under different environmental conditions. It is desirable that the constructed gene interaction models should be able to provide biologists with a range of hypotheses explaining the experimental results and suggesting optimal designs for further experiments.

In the past decades, variety of models and approaches have been developed to study the genetic networks, including Boolean networks [10], [11], linear and nonlinear differential equation models [12], [13], piecewise-linear models [14]–[16], dynamic Bayesian networks [17], [18], and fuzzy logic approaches [23], [24], etc. Boolean networks are computationally simple and do not depend on the precise experimental data [10], [11], thus they are suitable for handling both the complexity of biological networks and the challenge of generating and comparing multiple hypothetical networks. However, Boolean networks are *static* methods that cannot accurately describe the dynamic behaviors of biological systems [25]. Linear and nonlinear differential equation models provide a quantitative method to describe the continuous dynamical behaviors of genetic regulatory networks [12], [13]. However, this method has been proved to be computationally expensive and very sensitive to imprecise data [26]. Piecewise-linear models provide a qualitative approach to the analysis of genetic regulatory networks [14]–[16]. This approach has been implemented in the computational tool Genetic Network Analyzer (GNA) [15], and used for the analysis of variety of bacterial regulatory networks [16]. This approach has been shown to be able to represent the global, qualitative properties of the dynamics of gene systems. Bayesian networks, although attractive due to their ability to deal with the stochastic aspects of gene expression and noisy measurements, have ignored the dynamical aspects of gene regulation [19]. For other approaches frequently used, we refer the readers to [20]–[22] and the references therein.

Fuzzy logic [27], which provides a mathematical framework for linguistic modeling, has also been used to describe biological systems [23]. Fuzzy clustering methods have been frequently used to explore gene expression profiles for insight into both regulation and function, under the assumption that genes belong to multiple clusters and participate in multiple pathways. In [28], fuzzy k-means method was used to measure the relationship between a gene and its clusters through the membership values. In [24], by using the exhaustive search approach, the authors derived a linear fuzzy gene network model from microarray data. In [29], the fuzzy logic systems and recurrent neural networks were combined together to extract information from time-series gene expression data. In [30], the authors used

Manuscript received September 13, 2009; accepted August 29, 2010. Date of current version February 02, 2011. The work was supported in part by a grant from the Research Grants Council of the Hong Kong Special Administrative Region of China under Project CityU 113708 and in part by the City University of Hong Kong under Grant 9360131. *Asterisk indicates corresponding author.*

*Y. Sun is with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong, China (e-mail: sunyonghui168@gmail.com).

G. Feng is with the Department of Manufacturing Engineering and Engineering Management, City University of Hong Kong, Hong Kong, China, and also with Nanjing University of Science and Technology (e-mail: megfeng@cityu.edu.hk).

J. Cao is with the Department of Mathematics, Southeast University, Nanjing 210096, China (e-mail: jdcao@seu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNB.2010.2082559

a new multiscale fuzzy clustering method to model gene expression networks, which allows genes to interact between regulatory pathways and across different conditions at different levels of detail. In [31], the authors presented a hybrid approach of combining data mining and fuzzy modeling methods to build and analyze the biomolecular network of a cell process. In [32], the fuzzy approximation approach was used to investigate the robust stabilization problem of the nonlinear stochastic gene networks.

Furthermore, in [33], the authors have applied rule-based simulation techniques to model λ phage growth control, they have performed simulations to study the decision between the lytic and lysogenic pathway for wild-type λ phage and several single and double mutants. This example was also used by Shimada *et al.* [34] with a novel aspect of their work being that the regulation process is considered on two levels of abstraction. In addition, some attempts have also been made to integrate symbolic and numerical knowledge into a single formalism, for instance in the genetic grammar underlying the Metabolica system developed by Hofestädt and his colleagues [35], [36]. However, to the best of the authors' knowledge, up to now, the dynamic fuzzy system approach has never been used to model the gene expression data, which can reflect the nonlinear dynamic features of gene networks more precisely.

On the other hand, more and more useful information for gene networks could be extracted from the literature or discovered through ChIP-on-chip experiments. The priori biological knowledge for gene networks is typically qualitative, for instance, whether one gene affects another genes or not, or whether the effect is positive or negative. These kinds of information can be seen as the *structural knowledge* of genetic networks, and it is desirable to take into account these kinds of knowledge when modeling gene networks. In this way, the process of the identification can be speeded up and the obtained model can better represent gene regulatory networks.

In this paper, the dynamic fuzzy modeling approach incorporating *structural knowledge* will be used to model genetic networks from gene expression data. This method derives information on the gene interactions in a highly interpretable form (fuzzy rules) and takes into account the dynamical aspects of gene regulation, and thus is able to reveal significantly more biological relationships among genes and their products.

The rest of this paper is organized as follows. In Section II, the dynamic fuzzy gene network model is proposed, some necessary knowledge and the problem description are also given. In Section III, a generalized fuzzy clustering approach incorporating the structural knowledge is introduced to derive the parameters of local models of the membership functions and the optimal number of fuzzy rules, then the least square method is employed to estimate the parameters of the fuzzy gene network. In Section IV, our method is validated by applying it to two sets of gene expression data, and some comparisons with the existing approaches are also made. Finally, this paper concludes with some remarks and the possible future extensions.

II. Dynamic Fuzzy Model for Genetic Networks

DNA microarray technology has made a great progress in understanding the interactions among genes. The measured gene expression levels are often in the discrete-time format, which can be seen as the *observations* from system viewpoint. Many approaches have been developed to modeling the dynamic features of genetic networks based on those observed data.

In this paper, the measured gene expression data is to be modeled as a T-S fuzzy model with m fuzzy rules

$$\begin{aligned} &R^l: \text{IF } z_1(k) \text{ is } M_1^l \text{ and } \cdots z_r(k) \text{ is } M_r^l, \text{ THEN} \\ &y(k+1) = \hat{A}_l(q^{-1})y(k) + \hat{B}_l(q^{-1})u(k) + \hat{d}_l, \ \ l = 1, \ldots, m \end{aligned} \tag{1}$$

where R^l denotes the lth fuzzy inference rule, m the number of inference rules, M_r^l are fuzzy sets, $y(k) \in \mathbb{R}^p$ the system output variables, $u(k) \in \mathbb{R}^g$ the system input variables, $z(k) \in \mathbb{R}^r$ the premise measurable variables, which are defined as $z(k) = (y^{\mathrm{T}}(k), y^{\mathrm{T}}(k-1), \ldots, y^{\mathrm{T}}(k-n_y+1), u^{\mathrm{T}}(k), u^{\mathrm{T}}(k-1), \ldots, u^{\mathrm{T}}(k-n_u+1))^{\mathrm{T}}$, $(\hat{A}_l, \hat{B}_l, \hat{d}_l)$ is the lth local model with the shift operator q^{-1} defined by $q^{-1}y(k) = y(k-1)$, and

$$\begin{aligned} \hat{A}_l(q^{-1}) &= \hat{A}_{l1} + \hat{A}_{l2}q^{-1} + \cdots + \hat{A}_{ln_y}q^{-n_y+1}, \\ \hat{B}_l(q^{-1}) &= \hat{B}_{l1} + \hat{B}_{l2}q^{-1} + \cdots + \hat{B}_{ln_u}q^{-n_u+1}, \\ \hat{A}_{lk} &= [\hat{a}_{lkij}]_{p\times p}, \\ \hat{B}_{lk} &= [\hat{b}_{lkij}]_{p\times g}, \quad l = 1, \ldots, m. \end{aligned}$$

The dynamical fuzzy gene network (1) represents the dynamic process of biological forces regulating gene networks. The output $y(k)$ denotes the gene expression levels from microarray, which stands for mRNA concentrations and/or protein concentrations. The input $u(k)$ can be any external stimuli that influence gene regulation, such as the substances like drugs, proteins, RNAs, or expression levels of other genes.

Remark 1: In fact, in the real biological systems, there exist much more biology information for the ith gene, such as the self-degradation or dilution rate of itself and the regulatory relations with other genes. In this study, all the information is lumped into one matrix $\hat{A}_l(q^{-1})$ for simplicity.

It is worth noting that the local fuzzy model (1) only represents the properties of the gene network in a local region. By using a center-average defuzzifier, product inference and singleton fuzzifier, the fuzzy model (1) can be expressed as follows:

$$\begin{aligned} y(k+1) = \hat{A}(q^{-1}, \mu(z))y(k) + \hat{B}(q^{-1}, \mu(z))u(k) \\ + \hat{d}(\mu(z)) \end{aligned} \tag{2}$$

where

$$\begin{aligned} \hat{A}(q^{-1}, \mu(z)) &= \sum_{l=1}^{m} \mu_l \hat{A}_l(q^{-l}), \\ \hat{B}(q^{-1}, \mu(z)) &= \sum_{l=1}^{m} \mu_l \hat{B}_l(q^{-l}), \\ \mu(z) &= (\mu_1, \ldots, \mu_m). \end{aligned}$$

Remark 2: Some linear gene network models have been developed from the gene expression data [37]–[39]. However, these linear models can only describe the simple relationships among the genes and their products. The proposed fuzzy genetic network (2) can be used to represent the inherent nonlinear

relationship among the genes. It can be easily observed that the proposed model is more general than those linear models [37]–[39].

It is worth pointing out that sometimes some qualitative information for gene networks is known. Fox example, if a pair of genes have no interaction, then the corresponding entry in matrix $\hat{A}$ should be zero. Similarly, if an input has no influence on a gene that is represented as a state, then the corresponding entry in $\hat{B}$ should be zero.

Take the SOS DNA repair networks for example, it follows from the literature [44] that the relationship between the observation and the input is known, and thus the matrices $\hat{A}$ and $\hat{B}$ in fuzzy gene network (2) should have the following constrained structure:

$$\hat{A}=\begin{bmatrix} a_{11} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 & 0 & 0 & 0 \\ a_{31} & 0 & a_{33} & 0 & 0 & 0 & 0 \\ a_{41} & 0 & 0 & a_{44} & 0 & 0 & 0 \\ a_{51} & 0 & 0 & 0 & a_{55} & 0 & 0 \\ a_{61} & 0 & 0 & 0 & 0 & a_{66} & 0 \\ a_{71} & 0 & 0 & 0 & 0 & 0 & a_{77} \end{bmatrix},$$

$$\hat{B}=\begin{bmatrix} b_1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3}$$

Thus, in this paper, this useful prior knowledge for the gene network will be incorporated into the fuzzy modeling approach. In other words, some structural constraints will be enforced during the fuzzy modeling process, which can reduce the number of the parameters to be estimated and avoid the overfitting problem in modeling the gene network from the gene expression data.

In the following section, a generalized fuzzy clustering method incorporating the prior knowledge will be developed to determine the optimal number of fuzzy rules and the corresponding parameters of membership functions.

III. Identification of Fuzzy Genetic Networks

In this section, a generalized fuzzy clustering algorithm with *structural knowledge* is proposed to model gene networks from gene expression data.

Firstly, we introduce the following function:

$$\begin{aligned} h(\mu,\bar{z},\alpha)=&w_1\left[\sum_{k=1}^{N}\sum_{l=1}^{m}\mu_l(k)^{\omega}\|z(k)-\bar{z}_l\|^2\right] \\ &+w_2\left[\sum_{k=1}^{N}\sum_{l=1}^{m}\mu_l(k)^{\omega}\|e_l(k)\|^2\right], \\ \text{subject to}\quad &\sum_{l=1}^{m}\mu_l(k)=1\quad \text{for all } k \end{aligned} \tag{4}$$

where m is the number of fuzzy rules, N is the sampling data number; $\bar{z}=[\bar{z}_1,\bar{z}_2,\ldots,\bar{z}_m]$ is an m-tuple of mean prototypes, $\|z(k)-\bar{z}_l\|$ is the distance of the feature point $z(k)$ to the mean prototype $\bar{z}_l$; w_1 and w_2 are the weighting factors and ω is used to control the shape of the membership functions.

Considering the prior structural knowledge of gene networks, the predicting equation of the local linear models in the given criterion (4) should be in the following form:

$$\hat{y}^{\mathrm{T}}(k+1)=\phi^{\mathrm{T}}(k)[\alpha^l\circ M], \tag{5}$$

where $\phi(k)=[1\ z^{\mathrm{T}}(k)]^{\mathrm{T}}$, $\alpha^l=[\alpha_1^l,\ldots,\alpha_p^l],\alpha_i^l=[\alpha_{i1}^l,\ldots,\alpha_{i(1+r)}^l]^{\mathrm{T}}$; M is a constraining matrix of α^l, if an entry of α^l is constrained, the corresponding entry of M is 0, otherwise, it will be 1. The notation $\circ$ represents element-wise product, also known as the Hadamard product. Correspondingly, the predicting error in (4) is taken as $e_l(k)=y(k)-\hat{y}(k)$.

Remark 3: Most of existing results assume that the inferred gene network is a fully connected network, thus a large number of parameters need to be estimated. However, as pointed out in the last section, if the prior structural knowledge is taken into account, we can enforce some constraints on the network structure, so that the number of parameters to be estimated will be reduced.

Remark 4: Furthermore, if the prior knowledge of the gene network is not known, the proposed algorithm is also applicable. In this case, we should estimate all the parameters of the gene networks without structural knowledge. In this case, all the elements of the constraining matrix M are 1, that is, the predicting equation of the local linear models should be $\hat{y}^{\mathrm{T}}(k+1)=\phi^{\mathrm{T}}(k)\alpha^l$.

By taking the derivative of the given criterion (4) with respect to the parameters μ, $\bar{z}$ and the constrained parameter $\alpha^l\circ M$, respectively, the following theorem could be obtained readily.

Theorem 1: Given the weighting factors w_1, w_2 and ω, the necessary conditions for minimizing the given function $h(\mu,\ \bar{z},\ \alpha)$ can be expressed as follows:

$$\begin{aligned} \bar{z}_l&=\frac{\sum_{k=1}^{N}\mu_l(k)^{\omega}z(k)}{\sum_{k=1}^{N}\mu_l(k)^{\omega}}, \\ \alpha^l\circ M&=[\Phi^{\mathrm{T}}D_l\Phi]^{-1}\Phi^{\mathrm{T}}D_lY\circ M, \\ Y&=[y(1)\ y(2)\ \ldots\ y(N)]^{\mathrm{T}}, \\ \Phi&=[\phi(1)\ \phi(2)\ \ldots\ \phi(N)]^{\mathrm{T}}, \\ D_l&=\mathrm{diag}[\mu_l(k)^{\omega}]_{N\times N}, \\ \mu_l(k)&=\left\{\sum_{j=1}^{m}\frac{[w_1\|z(k)-\bar{z}_l\|^2+w_2\|e_l(k)\|^2]^{\sigma}}{[w_1\|z(k)-\bar{z}_j\|^2+w_2\|e_j(k)\|^2]^{\sigma}}\right\}^{-1}, \\ \sigma&=\frac{1}{\omega-1},\quad l=1,2,\ldots,m. \end{aligned}$$

Proof: The proof is similar to that in [40] and thus omitted. However, it is noted that the structural knowledge is taken into account, and in this case the prediction equation is in the constrained form (5).

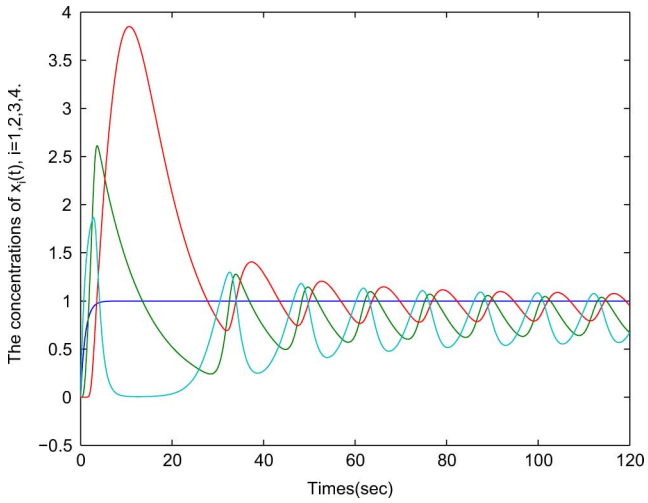

Fig. 1. Gene concentrations $x_i,\ i = 1, \ldots, 4$.

TABLE I
PARAMETERS OF EVERY RULE

Rule m	Centers $\bar{Z}_L$	Matrix θ^L
R^1	$\bar{Z}_1 = [1.3286, 1.4309, 2.4168, 2.3556]$	$\theta^1 = \begin{bmatrix} -0.4277 \\ 11.1693 \\ -8.9474 \\ -9.2529 \\ 9.3268 \end{bmatrix}$
R^2	$\bar{Z}_2 = [1.0176, 0.9791, 0.5874, 0.5527]$	$\theta^2 = \begin{bmatrix} 0.0170 \\ 1.7585 \\ -0.7619 \\ -0.0145 \\ 0.0019 \end{bmatrix}$
R^3	$\bar{Z}_3 = [0.5006, 0.5170, 0.9178, 0.9346]$	$\theta^3 = \begin{bmatrix} -0.0018 \\ 1.9582 \\ -0.9538 \\ -0.5307 \\ 0.5227 \end{bmatrix}$
R^4	$\bar{Z}_4 = [0.5522, 0.5177, 0.4560, 0.4612]$	$\theta^4 = \begin{bmatrix} 0.0013 \\ 1.8128 \\ -0.7931 \\ -0.7875 \\ 0.7640 \end{bmatrix}$
R^5	$\bar{Z}_5 = [1.0876, 1.1074, 1.0754, 1.0417]$	$\theta^5 = \begin{bmatrix} 0.0396 \\ 1.9733 \\ -1.0165 \\ 0.0825 \\ -0.0822 \end{bmatrix}$
R^6	$\bar{Z}_6 = [0.0895, 0.0858, 0.6682, 0.6878]$	$\theta^6 = \begin{bmatrix} 0.0003 \\ 1.9867 \\ -0.9823 \\ -0.0367 \\ 0.0356 \end{bmatrix}$
R^7	$\bar{Z}_7 = [0.0871, 0.1036, 1.7100, 1.7445]$	$\theta^7 = \begin{bmatrix} -0.0010 \\ 2.8646 \\ -1.6931 \\ -15.5725 \\ 15.2644 \end{bmatrix}$

Based on the above theorem, the following input space clustering algorithm can be obtained.

A. Fuzzy Input Space Clustering Algorithm

Remark 5: It is noted that, in this paper, only the qualitative information for the network structure is considered. However, other information, such as the cluster centers of the measured gene expression data might be also available in some cases. In this case, there is no need to estimate the cluster centers of the gene expression data, and this can greatly improve the convergence rate of the identification algorithm.

It is worth noting that the solutions in Theorem 1 are not suitable for the candidates for the BSMFs since the membership functions should be used to represent the space partitions only. In other words, the estimation errors should be removed from the membership functions. Thus the membership functions could be chosen as

$$\mu_l(z, \bar{z}_l, \sigma_l) = \left[\sum_{j=1}^{m} \frac{\|z - \bar{z}_l\|^{2\sigma}}{\|z - \bar{z}_j\|^{2\sigma}} \right]^{-1}. \tag{6}$$

Via Algorithm 1, the parameters of the membership functions can be obtained if the number of fuzzy rules m is known *a priori*. In order to find the optimal number of fuzzy rules, we introduce the following criterion:

$$h_z(m) = w_1 \left\{ \sum_{k=1}^{N} \sum_{l=1}^{m} \mu_l(k)^{\omega} [\|z(k) - \bar{z}_l\|^2 - \|\bar{z}_l - \bar{z}\|^2] \right\} + w_2 \left\{ \sum_{k=1}^{N} \sum_{l=1}^{m} \mu_l(k)^{\omega} [\|e_l(k)\|^2 - \bar{e}^2] \right\} \tag{7}$$

where $\bar{z} = 1/N \sum_{k=1}^{N} z(k)$, $\bar{e} = 1/N \sum_{k=1}^{N} \|y^{\mathrm{T}}(k) - \phi^{\mathrm{T}}(k-1)\alpha \circ M\|^2$, $\alpha \circ M = [(\Phi^{\mathrm{T}}\Phi)^{-1}\Phi^{\mathrm{T}}Y] \circ M$.

Algorithm 1 Fuzzy Input Space Clustering Algorithm

1: Given the weighting factors w_1, w_2 and ω, pick a termination threshold $\epsilon > 0$ and initial membership functions $\mu_l^{(0)}$, satisfying $\sum_{l=1}^{m} \mu_l^{(0)} = 1$.

2: **if** $\mathcal{I}_k = 0$

$\mu^{(j)} \rightarrow \mu^{(j+1)}$ according to Theorem 1;

else $\mu_i^{(j+1)}(k) = 0,\ \forall\, i \in \bar{\mathcal{I}}_k,\ \sum_{i \in \mathcal{I}_k} \mu_i^{(j+1)}(k) = 1,$

where

$\mathcal{I}_k \triangleq \{1 \le i \le m;\ w_1\|z(k) - \bar{z}_i\|^2 + w_2\|e_i(k)\|^2 = 0\}$,
$\bar{\mathcal{I}}_k \triangleq \{1, 2, \ldots, m\} - \mathcal{I}_k$;

3: **if** $\|\mu^{(j+1)} - \mu^{(j)}\| \le \epsilon$, **stop**;

else go to Step 2.

end if

It can be seen from (7) that the first term describes the information of the variance of the data in a mean cluster and the variance of the mean clusters themselves, the second term describes the variance of the gene expression levels in the equation cluster and the variance of the equation clusters themselves. Therefore, the optimal clustering is considered to minimize the variance in each cluster and to maximize the variance among the clusters. As the fuzzy rule number m increases, the given criterion (7) will reach a minimum, and the optimal number of fuzzy rules m can thus be determined. We then have the following identification algorithm for the rule number.

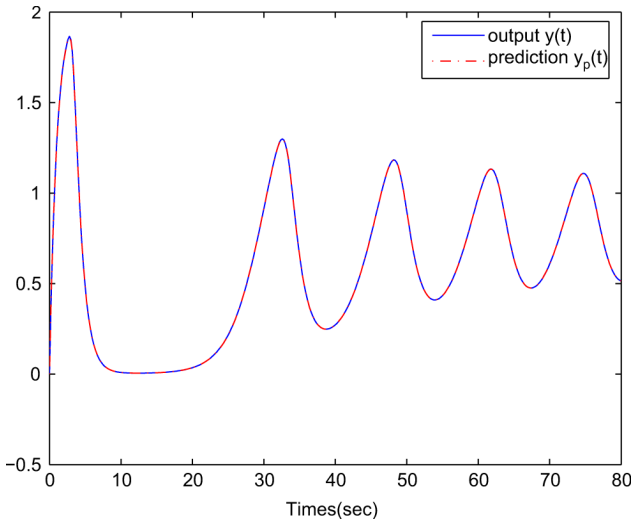

Fig. 2. Fitting curve for training data.

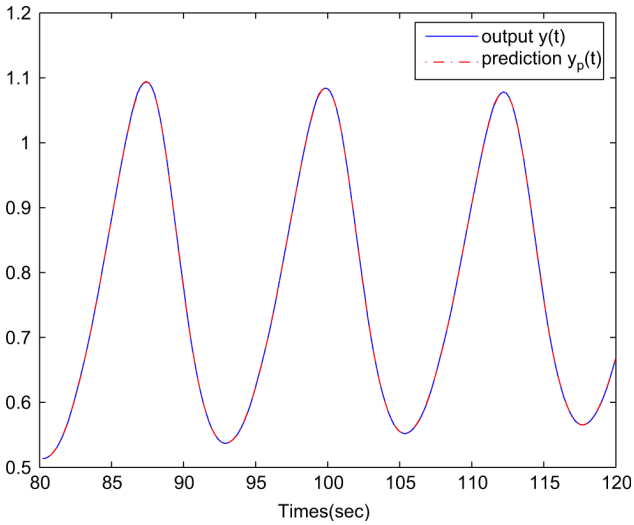

Fig. 3. Fitting curve for testing data.

B. Fuzzy Rule Number Identification Algorithm

Algorithm 2 Fuzzy Rule Number Identification Algorithm

1: Given a stopping criterion $\bar{\epsilon} > 0$ and set $m = 1$.

2: **Do** the **Algorithm 1** and obtain the BSMFs.

3: **if** $\|h_z(m+1) - h_z(m)\| < \bar{\epsilon}$, **stop**;

else $m = m + 1$, go to Step 2.

end if

Remark 6: If the cluster centers of the gene expression data are given *a priori* from knowledge of biologists, there is no need to use the above fuzzy rule number identification algorithm to find the optimal number of fuzzy rules. We will only need to use the Algorithm 1 to get the suitable parameters of the membership functions.

By taking into account the prior knowledge of the gene network structure, we can get the modified fuzzy clustering algorithm. After applying it on the measured gene expression data,

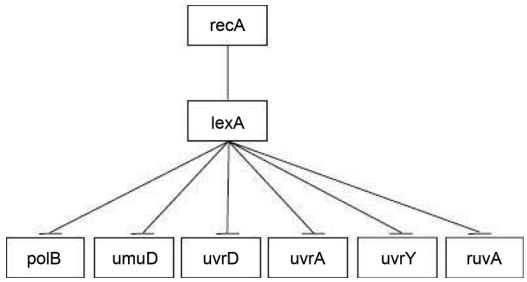

Fig. 4. Diagram of the SOS DNA repair networks [39].

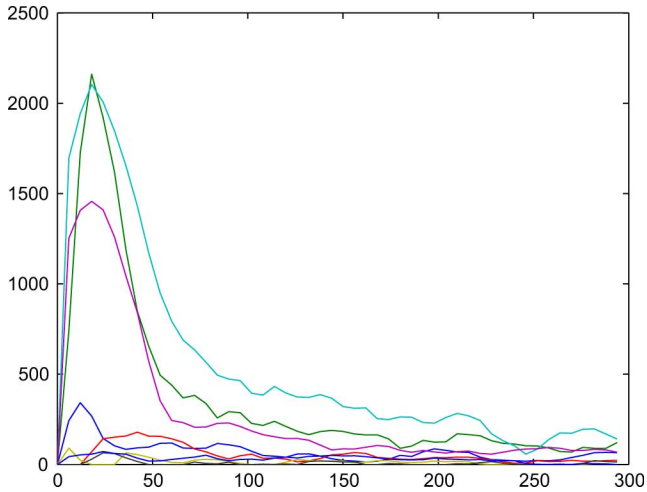

Fig. 5. The measured gene expression profiles from Exp. 1.

the parameters of the membership functions and the optimal number of fuzzy rules m for the gene network can be derived.

Next, we will show the identification of local models of the dynamical fuzzy genetic network (2) by using the least square method.

C. Identification of Local Models

Defining

$$\tilde{\phi}(k) = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_m \end{bmatrix} \otimes \phi(k) = \begin{bmatrix} \mu_1 \phi(k) \\ \vdots \\ \mu_m \phi(k) \end{bmatrix},$$

$$\Theta = \begin{bmatrix} \theta^1 \\ \vdots \\ \theta^m \end{bmatrix},$$

where $\otimes$ is the Kronecker product, one can rewrite the fuzzy model as

$$y^{\mathrm{T}}(k+1) = \tilde{\phi}^{\mathrm{T}}(k) \cdot \Theta. \tag{8}$$

When the data set is persistent exciting in the sense that the inverse defined in (9) exists, it is straightforward to apply the least square method to identify its local dynamic model parameters, θ^l, $l = 1, 2, \ldots, m$, that is,

$$\Theta = (\tilde{\Phi}^{\mathrm{T}} \tilde{\Phi})^{-1} \tilde{\Phi}^{\mathrm{T}} \tilde{Y}, \tag{9}$$

TABLE II
PARAMETERS OF RULE 1 FOR THE SOS GENE NETWORKS

Rules	Parameters
R^1	$\hat{A}_{11} = \begin{bmatrix} 0.4454 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2065 & 0.9736 & 0 & 0 & 0 & 0 & 0 \\ 0.1231 & 0 & 0.7821 & 0 & 0 & 0 & 0 \\ 0.0881 & 0 & 0 & 1.1759 & 0 & 0 & 0 \\ 0.1673 & 0 & 0 & 0 & 1.1653 & 0 & 0 \\ 0.0097 & 0 & 0 & 0 & 0 & 1.0083 & 0 \\ -0.0678 & 0 & 0 & 0 & 0 & 0 & 0.6013 \end{bmatrix}$
R^1	$\hat{A}_{12} = \begin{bmatrix} -0.1848 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.3858 & -0.2862 & 0 & 0 & 0 & 0 & 0 \\ -0.0940 & 0 & -0.0245 & 0 & 0 & 0 & 0 \\ -0.0412 & 0 & 0 & -0.4351 & 0 & 0 & 0 \\ 0.0607 & 0 & 0 & 0 & -0.3931 & 0 & 0 \\ 0.0030 & 0 & 0 & 0 & 0 & -0.5714 & 0 \\ 0.0633 & 0 & 0 & 0 & 0 & 0 & -0.0849 \end{bmatrix}$
R^1	$\hat{B}_{11} = \begin{bmatrix} 0.3144 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\hat{B}_{12} = \begin{bmatrix} 0.0391 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $\hat{d}_1 = \begin{bmatrix} 6.6918 \\ 5.5222 \\ 3.0671 \\ 4.4287 \\ -26.6810 \\ 5.6263 \\ 3.1307 \end{bmatrix}$

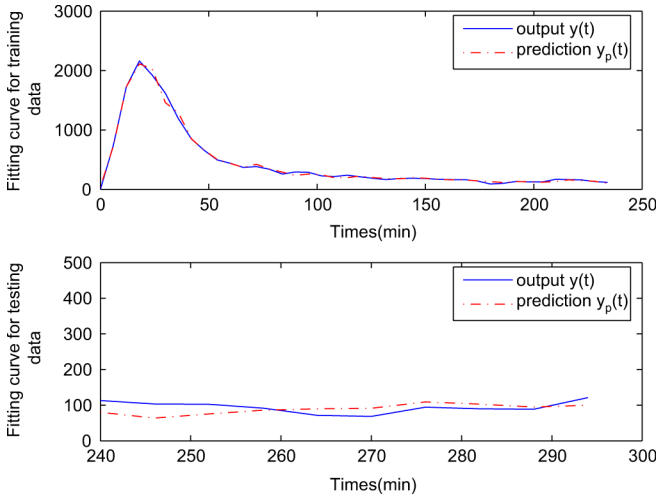

Fig. 6. Fitting curves for the gene *lexA*.

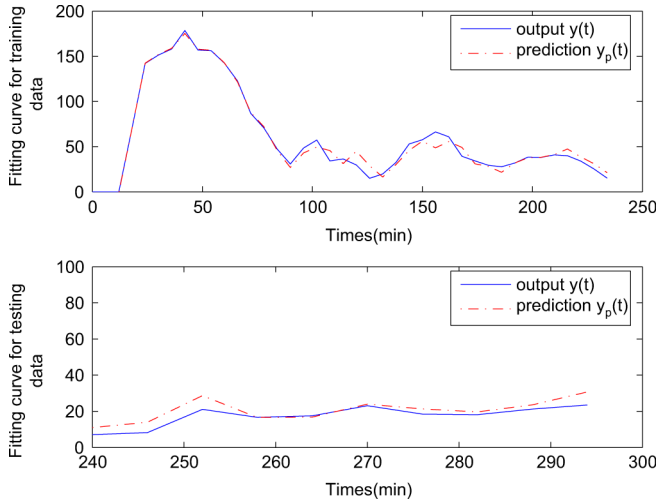

Fig. 7. Fitting curves for the gene *umuD*.

where

$$\tilde{\Phi} = \begin{bmatrix} \tilde{\phi}^{\mathrm{T}}(1) \\ \vdots \\ \tilde{\phi}^{\mathrm{T}}(N-1) \end{bmatrix}, \quad \tilde{Y} = \begin{bmatrix} y^{\mathrm{T}}(2) \\ \vdots \\ y^{\mathrm{T}}(N) \end{bmatrix}.$$

The modeling performance can be measured by the variance accounted for (VAF) index [42] given by

$$VAF = \left[1 - \frac{var(Y - \hat{Y})}{var(Y)}\right] \times 100\%, \tag{10}$$

where Y is the sequence of real plant output measurement, and $\hat{Y}$ is the fuzzy model output predictions.

Remark 7: It is noted that the fuzzy gene network (2) provides a good model to investigate the dynamics and other properties of the gene networks. It is also noted that the input-output form can be easily transformed into the state-space form [41].

IV. RESULTS AND CASE STUDIES

In order to evaluate the performance of the proposed dynamic fuzzy modeling approach, we will apply our method to two data sets, one is the synthetic data generated from a nonlinear genetic regulatory network model and the other one is the SOS DNA repair data with the known structure information.

A. Application to Synthetic Data

Our modeling method will be firstly applied to the synthetic data without structural knowledge, the data is generated from a nonlinear genetic regulatory network [43], which is described by the following differential equation:

$$\begin{aligned} \frac{dx_1}{dt} &= v_1 - k_1 x_1, \\ \frac{dx_2}{dt} &= \frac{v_2 x_4^{n_4}}{(\lambda_2 + x_4^{n_4})(\lambda_{I3} + x_3^{n_3})} - k_2 x_2, \\ \frac{dx_3}{dt} &= \frac{v_3 (x_1 x_2)^{n_{12}}}{\lambda_3 + (x_1 x_2)^{n_{12}}} - k_3 x_3, \\ \frac{dx_4}{dt} &= \frac{v_4}{\lambda_{I3} + x_3^{n_3}} - k_4 x_4, \end{aligned} \tag{11}$$

where $v_1 = 1$; $k_1 = 1$; $v_2 = 1$; $\lambda_2 = 5$; $k_2 = 0.1$; $\lambda_{I3} = 0.5$; $n_3 = 4$; $v_3 = 1$; $\lambda_3 = 5$; $k_3 = 0.1$; $v_4 = 1$; $k_4 = 1$. And we choose $n_4 = n_{12} = 4$, which are different from the values in

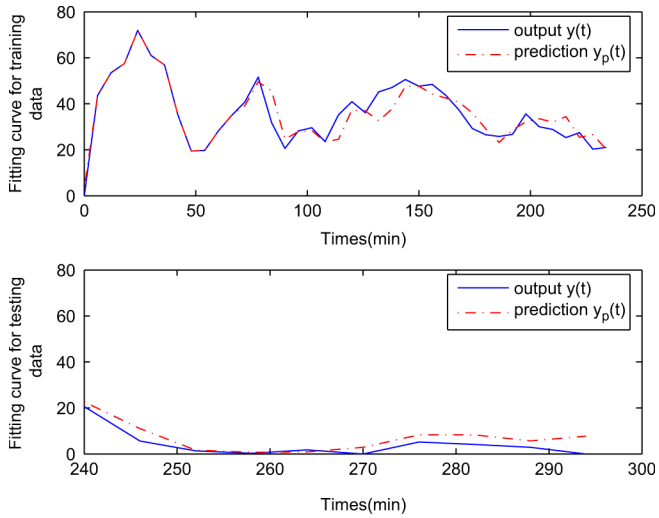

Fig. 8. Fitting curves for the gene *polB*.

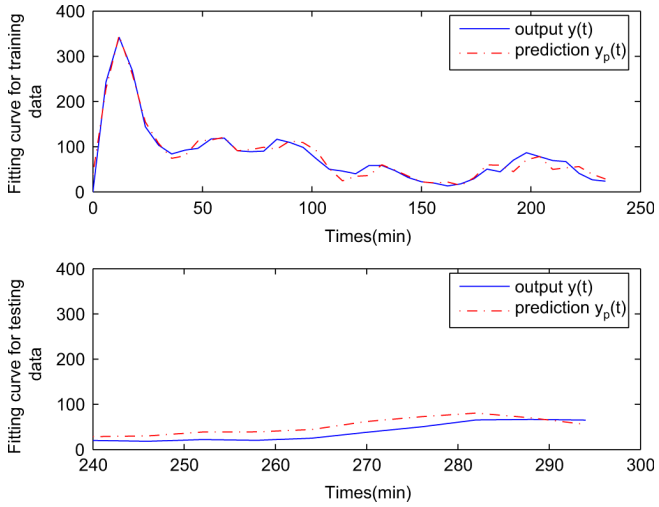

Fig. 9. Fitting curves for the gene *uvrD*.

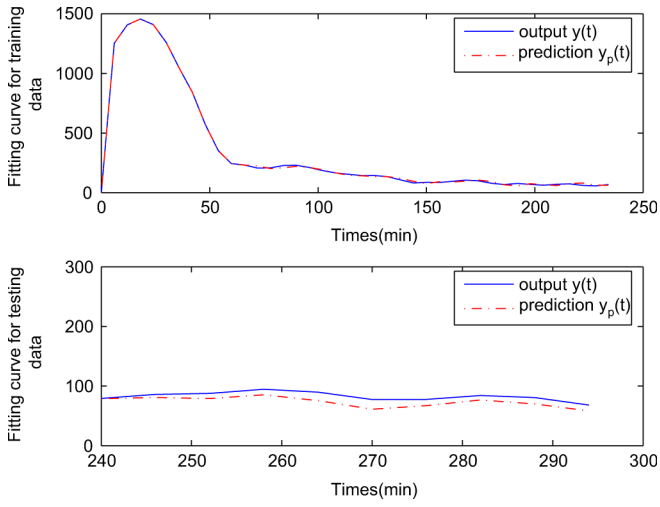

Fig. 10. Fitting curves for the gene *uvrA*.

[43], where $n_4 = n_{12} = 1$. These simple changes induce the oscillatory dynamics of the gene networks (11), which can be seen from Fig. 1.

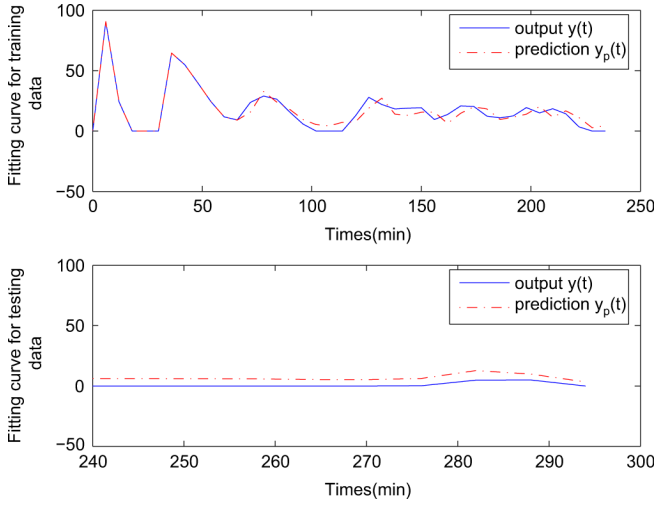

Fig. 11. Fitting curves for the gene *vurY*.

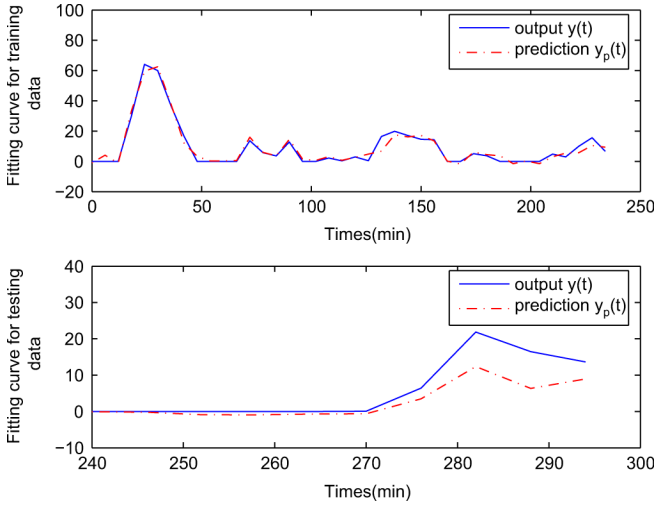

Fig. 12. Fitting curves for the gene *ruvA*.

Based on the feedback structure of this gene network, the control input can be taken as $u(t) = x_2(t)$ and the output $y(t) = x_4(t)$. In this case, the system can be seen as an SISO system.

By using the Euler scheme with a step size of $h = 0.001$ time units, the discrete time equations describing the evolution of the gene network are given by

$$\begin{aligned}
x_1(k+1) &= x_1(k) + h(v_1 - k_1 x_1(k)),\\
x_2(k+1) &= x_2(k) + h\left(\frac{v_2 x_4^{n_4}(k)}{(\lambda_2 + x_4^{n_4}(k))(\lambda_{I3} + x_3^{n_3}(k))} - k_2 x_2(k)\right),\\
x_3(k+1) &= x_3(k) + h\left(\frac{v_3 (x_1(k)x_2(k))^{n_{12}}}{\lambda_3 + (x_1(k)x_2(k))^{n_{12}}} - k_3 x_3(k)\right),\\
x_4(k+1) &= x_4(k) + h\left(\frac{v_4}{\lambda_{I3} + x_3^{n_3}(k)} - k_4 x_4(k)\right). \qquad (12)
\end{aligned}$$

Defining $T = 600h$ as the period of sampling time, the first 400 data generated from the (11) is taken as the training data and the last 200 data as the testing data.

By using the fuzzy modeling approach proposed in the last section without structural information, given stopping criteria

TABLE III
PERFORMANCE ANALYSIS OF THE SOS DATA 1

Gene	Error	Error [44]	Error [37]	Description
$uvrA$	0.0604	0.14	0.1989	Nucleotide escision repair
$lexA$	0.0703	0.10	0.2531	Transcriptional repressor
$umuD$	0.1187	0.21	0.1829	Mutagenesis repair
$polB$	0.0763	0.31	0.1419	Translesion DNA synthesis, replication fork recovery
$ruvA$	0.3924	0.22	0.9925	Double-strand break repair
$uvrD$	0.1527	0.20	0.2177	Nucleotide escision repair, recombinational repair
$uvrY$	0.2158	0.45	0.2850	SOS operon of unknown function

TABLE IV
PERFORMANCE ANALYSIS OF THE SOS DATA 2

Gene	Error	Error [44]	Error [37]	Description
$uvrA$	0.0600	0.1022	0.1610	Nucleotide escision repair
$lexA$	0.0647	0.1331	0.1466	Transcriptional repressor
$umuD$	0.0567	0.1258	0.1400	Mutagenesis repair
$polB$	0.0891	0.1755	0.3925	Translesion DNA synthesis, replication fork recovery
$ruvA$	0.1030	0.1604	0.1213	Double-strand break repair
$uvrD$	0.1740	0.2246	0.4273	Nucleotide escision repair, recombinational repair
$uvrY$	0.1883	0.3857	3.4142	SOS operon of unknown function

TABLE V
PERFORMANCE ANALYSIS OF THE SOS DATA 3

Gene	Error	Error [44]	Error [37]	Description
$uvrA$	0.0582	0.3016	0.3204	Nucleotide escision repair
$lexA$	0.0559	0.2068	0.2786	Transcriptional repressor
$umuD$	0.1005	0.1330	0.2020	Mutagenesis repair
$polB$	0.1788	0.1913	0.5847	Translesion DNA synthesis, replication fork recovery
$ruvA$	0.1462	0.1985	0.2483	Double-strand break repair
$uvrD$	0.1288	0.1780	0.2853	Nucleotide escision repair, recombinational repair
$uvrY$	0.0967	0.2432	0.3349	SOS operon of unknown function

TABLE VI
PERFORMANCE ANALYSIS OF THE SOS DATA 4

Gene	Error	Error [44]	Error [37]	Description
$uvrA$	0.0628	0.3250	0.3758	Nucleotide escision repair
$lexA$	0.0576	0.1793	0.1991	Transcriptional repressor
$umuD$	0.0573	0.1461	0.1848	Mutagenesis repair
$polB$	0.1029	0.1417	0.2165	Translesion DNA synthesis, replication fork recovery
$ruvA$	0.0201	0.0362	2.5699	Double-strand break repair
$uvrD$	0.0616	0.0862	0.1431	Nucleotide escision repair, recombinational repair
$uvrY$	0.0675	0.3060	0.2343	SOS operon of unknown function

$\epsilon = 0.4$, $\bar{\epsilon} = 0.2$, the parameters $\omega_1 = \omega_2 = 0.01$ and $\omega = 1.1$, one can obtain the optimal fuzzy rules $m = 7$ and the corresponding parameters shown in Table I with the associated VAF index value as $\mathrm{VAF} = 99.9994\%$.

The obtained dynamic fuzzy gene network for the synthetic data generated from the gene networks (11) can be expressed as follows:

$$\hat{y}(k+1) = \sum_{l=1}^{m} \mu_l(k)\phi^{\mathrm{T}}(k)\theta^l, \tag{13}$$

where $\phi(k) = [1,\ y(k),\ y(k-1),\ u(k),\ u(k-1)]^{\mathrm{T}}$, the membership functions are defined as (6) with $\sigma = 10$, the centers of the BSMFs and the other corresponding parameters are listed in Table I.

The simulation results are presented as follows: Fig. 2 is the fitting curve for the training data, while Fig. 3 is the fitting curve for the testing data. It can be observed from these simulation results that a quite satisfactory fitting result has been achieved.

B. Application to SOS DNA Repair Networks

The proposed fuzzy modeling method with *structural knowledge* will be applied to the real experimental data set: SOS DNA repair networks data. The SOS data set contains gene expression measurements presented by Ronen *et al.* [44] for the SOS DNA repair networks of *E. coli* bacterium. The specific study consists of four experiments under various light intensities (experiments 1 and 2: $5\ \mathrm{Jm}^{-2}$, experiments 3 and 4: $20\ \mathrm{Jm}^{-2}$). Each experiment consists of 50 points with a time period of 6 min, where eight major genes were monitored: *uvrD, lexA, umuD, recA, uvrA, uvrY, ruvA,* and *polB*. The specific data set can be downloaded from Uri Alon's homepage at http://www.weizmann.ac.il/mcb/UriAlon/Papers/SOSData in the form of four different data sets, corresponding to Exp. 1, Exp. 2, Exp. 3, and Exp. 4, respectively.

As pointed out in [44], in this SOS networks, the sole input is the gene *recA*, and the master regulator is the gene *lexA*. The gene *lexA* inhibits all the rest genes of the SOS network under normal condition. When DNA damage is sensed, the normally

suppressed genes become active, and the rest genes have no relationship or connections, the diagram of this SOS DNA repair networks can be seen from Fig. 4.

Based on the architecture of the SOS gene networks, an SIMO (one input, seven outputs) fuzzy gene network can be developed to describe each one of the genes involved in the process.

For convenience, the following mean error criterion used in [44] is employed to test the modeling performance for the given data set

$$\text{Error} = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i^{\text{measured}} - y_i^{\text{predicted}}|}{y_i^{\text{measured}}}, \qquad (14)$$

where N is the number of samples in the data set.

In this case study, firstly the Exp. 1 data set is used, Fig. 5 shows the measured gene expression profiles from Exp. 1. The first 40 points in Exp. 1 are taken as the training data, and the last 10 points for testing purpose. By using the proposed fuzzy modeling method with the known structure information, given stopping criteria $\epsilon = 0.3$, $\bar{\epsilon} = 0.5$, the parameters $\omega_1 = 1$, $\omega_2 = 8$, and $\omega = 1.36$, we can get the optimal fuzzy rule number $m = 3$, the parameters of the fuzzy Rule 1 for the SOS gene networks shown in Table II. The fitting curves for the training data and the testing data can be seen from Figs. 6–12, respectively. It can be seen from those results that the performance of the proposed fuzzy modeling approach with structural knowledge is satisfactory.

The results based on the mean error criterion (14) are shown in Table III. In [44], the authors used the the singular value decomposition method to model gene expression data, where the same Exp. 1 data set was used. In addition, the authors in [37] modeled the gene expression from time-course gene expression data in the form of linear systems. Comparison results with those in [44] and [37] are also shown in Table III, respectively. It can be observed that our approach produces much more accurate results for this SOS genetic networks than those in [44] and [37] on Exp. 1 data set.

In order to further validate the proposed approach, we also apply the fuzzy modeling method to the other three experimental data sets in comparison with the other two modeling methods, that is, the linear system modeling approach proposed in [37] and the singular value decomposition method used in [44]. The results are summarized in Tables IV–VI, respectively. It can be observed that the approach presented in this study manages to produce much more accurate results for this gene network.

In spite of the noise inherent in the experimental data, the proposed approach indeed recovers the general trends of the expression patterns for all genes, and the results also demonstrate the accuracy and efficiency of our method in successfully capturing the interactions among the considered genes.

V. Conclusions

In this paper, a novel fuzzy modeling approach, incorporating the prior knowledge of gene networks, has been proposed to construct genetic networks from gene expression data. The proposed approach has two distinctive features. One is that the prior structural knowledge on gene networks can be utilized so that the faster convergence of the identification process can be achieved. The other is that the nonlinear dynamic property inherent in most gene networks can be well captured so that the better prediction capability of the resulting fuzzy gene network can be achieved. The proposed approach has been applied to two data sets, the synthetic data from a numerical example and the real SOS DNA repair networks data with structural knowledge, respectively. It is observed from the comparison with some existing results that the dynamic fuzzy gene network provides much better modeling accuracy, and thus could be a more promising method in modeling of gene regulatory networks.

References

[1] P. D'haeseleer, S. Liang, and R. Somogyi, "Gene expression data analysis and modeling," presented at the Pac. Symp. Biocomput., Waimea, HI, Jan. 4–9, 1999.

[2] N. Holter, A. Maritan, M. Cieplak, N. Fedoroff, and J. Banavar, "Dynamic modeling of gene expression data," *Proc. Nat. Acad. Sci. USA*, vol. 98, no. 4, pp. 1693–1698, Feb. 2001.

[3] D. Angeli, J. Ferrell, and E. Sontag, "Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 7, pp. 1822–1827, Feb. 2004.

[4] P. Smolen, D. Baxter, and J. Byrne, "Mathematical modeling of gene networks," *Neuron*, vol. 26, no. 3, pp. 567–580, Jun. 2000.

[5] H. de Jong, "Modelling and simulation of genetic regulatory systems: A literature review," *J. Comp. Biol.*, vol. 9, no. 1, pp. 67–103, Jan. 2002.

[6] H. Bolouri and E. Davidson, "Modelling transcriptional regulatory networks," *BioEssay*, vol. 24, pp. 1118–1129, 2002.

[7] E. Sontag, "Molecular systems biology and control," *Eur. J. Control*, vol. 11, no. 5, pp. 1–40, 2005.

[8] D. Lockhart and E. winzeler, "Genomics, gene expression and DNA arrays," *Nature*, vol. 405, no. 6788, pp. 827–836, 2000.

[9] G. Mclachlan, K. Do, and C. Ambroise, *Analyzing Microarray Gene Expression Data*. Hoboken, NJ: Wiley, 2004.

[10] S. Liang, S. Fuhrman, and R. Somogyi, "Reveal, a general reverse engineering algorithm for infernce of genetic network architectures," in *Proc. Pac. Symp. Biocomput.*, 1998, vol. 3, pp. 18–29.

[11] T. Akutsu, S. Miyano, and S. Kuhara, "Identification of genetic networks from a small number of gene expression patterns under the Boolean network model," in *Proc. Pac. Symp. Biocomput.*, 1999, vol. 4, pp. 17–28.

[12] T. Chen, H. He, and G. Church, "Modeling gene expression with differential equations," in *Proc. Pac. Symp. Biocomput.*, 1999, vol. 4, pp. 29–40.

[13] L. Chen and K. Aihara, "Stability of genetic regulatory networks with time delay," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, pp. 602–608, May 2002.

[14] L. Glass, "Classification of biological networks by their qualitative dynamics," *J. Theor. Biol.*, vol. 54, no. 1, pp. 85–107, Oct. 1975.

[15] H. de Jong, J. Geiselmann, C. Hernandez, and M. Page, "Genetic network analyzer: Qualitative simulation of genetic regulatory networks," *Bioinformatics*, vol. 19, no. 3, pp. 336–344, 2003.

[16] H. de Jong, J. Geiselmann, G. Batt, C. Hernandez, and M. Page, "Qualitative simulation of the initiation of sporulation in Bacillus subtilis," *Bull. Math. Biol.*, vol. 66, no. 2, pp. 261–299, Mar. 2004.

[17] K. Murphy and S. Mian, *Modeling Gene Expression Data Using Dynamic Bayesian Networks*. Berkeley, CA: Comput. Sci. Div., Univ. California, 1999.

[18] B. Perrin, L. Ralaivola, A. Mazurie, S. Bottani, J. Mallet, and F. d'Alché-Buc, "Gene networks inference using dynamic Bayesian networks," *Bioinformatics*, vol. 19, no. Suppl. 2, pp. 138–148, Sep. 2003.

[19] N. Friedman, M. Linial, I. Nachman, and D. Pe'er, "Using Bayesian networks to analyze expression data," *J. Comp. Biol.*, vol. 7, no. 3–4, pp. 601–620, Aug. 2000.

[20] M. Eisen, P. Spellman, P. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proc. Nat. Acad. Sci. USA*, vol. 95, no. 25, pp. 14863–14868, Dec. 1998.

[21] L. Soinov, M. Krestyaninova, and A. Brazma, "Towards reconstruction of gene networks from expression data by supervised learning," *Genome Biol.*, vol. 4, no. 1, R6, Jan. 2003.

[22] C. Ma and C. Chan, "Inferring gene regulatory networks from expression data by discovering fuzzy dependency relationships," *IEEE. Trans. Fuzzy Syst.*, vol. 16, no. 2, pp. 455–465, Apr. 2008.
[23] P. Woolf and Y. Wang, "A fuzzy logic approach to analyzing gene expression data," *Physiol. Genomics*, vol. 3, no. 1, pp. 9–15, Jun. 2000.
[24] B. Sokhansanj, J. Fitch, J. Quong, and A. Quong, "Linear fuzzy gene network models obtained from microarray data by exhaustive search," *BMC Bioinf.*, vol. 5, no. 108, Aug. 2004.
[25] L. Glass and S. Kauffman, "The logical analysis of continuous, nonlinear biochemical control networks," *J. Theor. Biol.*, vol. 39, pp. 103–129, 1973.
[26] J. Tegner, M. Yeung, J. Hasty, and J. Collins, "Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling," *Proc. Nat. Acad. Sci. USA*, vol. 100, no. 10, pp. 5944–5949, May 2003.
[27] L. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, pp. 338–352, 1965.
[28] A. Gasch and M. Eisen, "Exploring the conditional coregulation of yeast gene expression through fuzzy k-means clustering," *Genome Biol.*, vol. 3, no. 11., Oct. 2002, RESEARCH0059.1-22.
[29] I. Maraziotis, A. Dragomir, and A. Bezerianos, "Gene networks reconstruction and time-series prediction from microarray data using recurrent neural fuzzy networks," *IET Syst. Biol.*, vol. 1, no. 1, pp. 41–50, 2007.
[30] P. Du, J. Gong, E. Wurtele, and J. Dickerson, "Modeling gene expression networks using fuzzy logic," *IEEE. Trans. Syst., Man, Cybern. B, Cybern.*, vol. 35, no. 6, pp. 1351–1359, Dec. 2005.
[31] X. Hu, B. Sokhansanj, D. Wu, and Y. Tang, "A novel approach for mining and fuzzy simulation of subnetworks from large biomolecular networks," *IEEE. Trans. Fuzzy Syst.*, vol. 15, no. 6, pp. 1219–1229, Dec. 2007.
[32] B. Chen, Y. Chang, and Y. Wang, "Robust H_∞-stabilization design in gene networks under stochastic molecular noises: Fuzzy-interpolation approach," *IEEE. Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 1, pp. 25–42, Feb. 2008.
[33] S. Meyers and P. Friedland, "Knowledge-based simulation of genetic regulation in bacteriophage lambda," *Nucleic Acids Res.*, vol. 12, no. 1, pp. 1–9, 1984.
[34] T. Shimada *et al.*, "Knowledge-based simulation of regulatory action in lambda phage," *Int. J. Artif. Intell. Tools*, vol. 4, no. 4, pp. 511–524, 1995.
[35] R. Hofestädt, "A simulation shell to model metabolic pathways," *Syst. Anal. Model. Simul.*, vol. 11, no. 4, pp. 253–262, 1993.
[36] R. Hofestädt and F. Meineke, "Interactive modelling and simulation of biochemical networks," *Comput. Biol. Med.*, vol. 25, no. 3, pp. 321–334, 1995.
[37] F. Wu, W. Zhang, and A. Kusalik, "Modeling gene expression from microarray expression data with state-space equations," in *Proc. Pac. Symp. Biocomput.*, 2004, vol. 9, pp. 581–592.
[38] Z. Wang, F. Yang, D. W. C. Ho, S. Swift, A. Tucker, and X. Liu, "Stochastic dynamic modeling of short gene expression time-series data," *IEEE. Trans. NanoBiosci.*, vol. 7, no. 1, pp. 44–55, Mar. 2008.
[39] H. Xiong and Y. Choe, "Structural systems identification of genetic regulatory networks," *Bioinformatics*, vol. 24, no. 4, pp. 553–560, Jan. 2008.
[40] S. Cao, N. Rees, and G. Feng, "Analysis and design for a class of complex control systems—Part I: Fuzzy modelling and identification," *Automatica*, vol. 33, no. 6, pp. 1017–1028, 1997.
[41] S. Cao, N. Rees, and G. Feng, "Analysis and design for a class of complex control systems-Part II: Fuzzy controller design," *Automatica*, vol. 33, no. 6, pp. 1029–1039, 1997.
[42] J. Roubos, S. Mollov, R. Babuska, and H. Verbruggen, "Fuzzy model-based predictive control using Takagi-Sugeno models," *Int. J. Approx. Reas.*, vol. 22, pp. 3–30, 1999.
[43] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach, *Systems Biology in Practice: Concepts, Implementation and Application*. Hoboken, NJ: Wiley, 2005.
[44] M. Ronen, R. Rosenberg, B. Shraiman, and U. Alon, "Assigning numbers to the arrows: Parameterizing a gene regulation network by using accurate expression kinetics," *Proc. Nat. Acad. Sci. USA*, vol. 99, no. 16, pp. 10555–10560, Aug. 2002.

Yonghui Sun received the M.S. degree from Southeast University, Nanjing, China, in 2007, and the Ph.D. degree from City University of Hong Kong, in 2010.

His research interests include systems biology, stochastic control, and fuzzy modeling and control.

Gang Feng (S'90–M'92–SM'95–F'09) received the B.Eng. and M.Eng. degrees in automatic control from Nanjing Aeronautical Institute, China, in 1982 and 1984, respectively, and the Ph.D. degree in electrical engineering from the University of Melbourne, Australia, in 1992.

He has been with City University of Hong Kong since 2000, where he is now Chair Professor of Mechatronic Engineering. He is also ChangJiang Chair Professor at Nanjing University of Science and Technology, awarded by Ministry of Education, China. He was Lecturer/Senior Lecturer at the School of Electrical Engineering, University of New South Wales, Australia, from 1992 to 1999. His current research interests include hybrid systems and control, modeling and control of energy systems, and intelligent systems & control.

Prof. Feng was awarded an Alexander von Humboldt Fellowship in 1997–1998, and the IEEE TRANSACTIONS ON FUZZY SYSTEMS Outstanding Paper Award in 2007. He is an Associate Editor of IEEE TRANSACTIONS ON FUZZY SYSTEMS and *Mechatronics*, and was an Associate Editor of IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE TRANSACTIONS ON SYSTEMS, MAN & CYBERNETICS, PART C, and *Journal of Control Theory and Applications*.

Jinde Cao (M'07–SM'07) received the B.S. degree from Anhui Normal University, Wuhu, China, the M.S. degree from Yunnan University, Kunming, China, and the Ph.D. degree from Sichuan University, Chengdu, China, all in mathematics/applied mathematics, in 1986, 1989, and 1998, respectively.

He was with Yunnan University from 1989 to 2000. Since 2000, he has been with the Department of Mathematics, Southeast University, Nanjing, China. From 2001 to 2002, he was a Postdoctoral Research Fellow with the Department of Automation and Computer-Aided Engineering, Chinese University of Hong Kong, Shatin, Hong Kong. He was a Visiting Research Fellow and a Visiting Professor with the School of Information Systems, Computing and Mathematics, Brunel University, Middlesex, U.K., from 2006 to 2008. He is the author or coauthor of more than 160 research papers and five edited books. His current research interests include nonlinear systems, neural networks, complex systems, complex networks, stability theory, and applied mathematics.

Dr. Cao was an Associate Editor of IEEE TRANSACTIONS ON NEURAL NETWORKS from 2006 to 2009. He is an Associate Editor of *Journal of the Franklin Institute*, *Mathematics and Computers in Simulation*, *Neurocomputing*, *International Journal of Differential Equations*, *Discrete Dynamics in Nature and Society*, and *Differential Equations and Dynamical Systems*. He is a reviewer of *Mathematical Reviews* and *Zentralblatt-Math*.