Gene Regulatory Network Reconfiguration Model Based on Boolean Network Model

Liangzhong SHEN¹, Xiangzhen ZAN¹, Wenbin LIU²

¹Department of Information Engineering, Wenzhou Business College, Wenzhou 325000, China
²Department of Physics and Electronic Information Engineering, Wenzhou University, Wenzhou 325000, China

Abstract — We aim to improve the performance of reconstructing the algorithm for gene regulation network and propose a probabilistic Boolean network model. Firstly, the gene regulation network is studied and the weighting matrix model of gene regulation network is determined. A nonlinear mapping for genetic expression response is carried out to realize the regulation operation of genetic expression in discrete state transition. Secondly, we do a reconstruction process of gene regulation network by using probabilistic Boolean network model and analyze it based on semi-Markov decision-making process and we perform the calculation procedure of probabilistic Boolean network model. Finally, we verify the validity of the algorithm through experiments.

Keywords- probabilistic boolean network; gene regulation; network reconstruction; weighting matrix; markov decision-making

I. INTRODUCTION

The research on gene network was conducted in 1960s at first. Rarer described the features of molecular genetic system organization for controlling proaryotic organism. The other research is that Kauffman studied on the gene network dynamics through simple logic rules and put forward the simple Boolean network theory. The increase of experimental data in 1990s sped up the research of gene network and various methods for researching on gene network has been promoted, such as random model, Boolean network, logic method, threshold model and method based on differential equation, etc. After whole genome sequencing, the vast DNA serial information is revealed before the scientists, which requires people to understand him and understand how the gene expresses and how genetic expression is regulated. The expression regulation rule of these genes shall be systematically expounded on the whole genome level so as to uncover the secret of life that a genetic expression is affected by other genes. This Paper mainly adopts probabilistic Boolean network model to research on the gene regulation network reconstruction with the aim of improving the network reconstruction algorithm performance.

II. WEIGHTING MATRIX MODEL OF GENE REGULATION NETWORK

In weighting matrix model of gene regulation network, the expression value of one gene is the function of expression value of other genes. The gene expression state containing n genes shall be expressed with the vector quantity $u(t)$ in n dimensional space, $u(t)$ expresses the expression level of a gene at time $t$. A weighting matrix $W$ shows the mutual regulation effect among genes, each row of $W$ shows all regulation input of one gene and $w_i$ represents the influence of expression level of gene $j$ on gene $i$. At time $t$, the net regulation input of gene $j$ for gene $i$ is the expression level of $j$ (namely $u_j(t)$ multiplies the regulation incidence of $j$ on $i$). The total regulation input $r_i(t)$ of gene $i$ is:

$$r_i(t) = \sum_{j=1}^{n} w_{ij} u_j(t)$$

(1)

If $w_{ij}$ is positive, then gene $j$ stimulates the expression of gene $i$, if it is negative, it shows that gene $j$ inhibits the expression of gene $i$. If it is 0, it shows that gene $j$ is of no effect on gene $i$. Finally, the expression response of gene $i$ needs to take nonlinear mapping.

$$u_i(t+1) = \frac{1}{1 + e^{-(\alpha \beta + \beta)}}$$

(2)

This function is the Sigmoid function, which is commonly used in neural network, among which $\alpha$ and $\beta$ are two constants, defining the position and curvature of nonlinear mapping function curve. The expression level of gene $i$ at time $t+1$ is calculated through the above formula. In this Paper, $\alpha = 1$, $\beta = 0$, thus the Sigmoid function can be expressed as:

$$u_i(t+1) = \frac{1}{1 + e^{-(\alpha \beta)}}$$

(3)

Fig. 1 and Table 1 are the example of a gene regulation network and its corresponding weighting matrix. In this model, the regulation relation among genes can be any real number from maximum incentive value $1$ to maximum inhibitory value. The genetic expression is regulated in discrete state transition, and the expression level of all genes can be updated simultaneously.

DOI 10.5013/IJSSST.a.17.42.50 50.1 ISSN: 1473-804x online, 1473-8031 print
III. PROBABILISTIC BOOLEAN NETWORK MODEL

A. Model Description

PBN is the promotion of Boolean network, which is the class of sets constituted by Boolean network in nature. Each gene in Boolean network has only one forecasting function. Different from Boolean network, in PBN, the state of each gene has only one forecasting function. The choice probability of forecasting function \( f_i \) is \( c_{ij}^{(0)} (0 \leq c_{ij}^{(0)} \leq 1) \), \( \sum_{j=1}^{n} c_{ij}^{(0)} = 1 \), \( i = 1, 2, \ldots, n \).

Suppose that \( f_i \) is the \( j \) th probable network, then \( f_i = \{ f_{i1}^{(0)}, f_{i2}^{(1)}, \ldots, f_{in}^{(n)} \} \) \( (1 \leq j \leq l(i), i = 1, 2, \ldots, n) \). The choice probability of this network is:

\[
P_i = \prod_{j=1}^{l(i)} c_{ij}^{(0)} (j = 1, 2, \ldots, N)
\]

Among them, \( N = \prod_{i=1}^{n} l(i) \) is the number of Boolean network. The transition probability between the state of \((x_1, \ldots, x_n)\) and \((x'_1, \ldots, x'_n)\) can be acquired by the following formula:

\[
P((x_1, \ldots, x_n) \rightarrow (x'_1, \ldots, x'_n))
\]

\[
= \sum_{k=1}^{2^n} \prod_{j=1}^{n} P_{jk}(I - [f_{ik}^{(0)}(x_1, x_2, \ldots, x_n) - x'_j])
\]

For convenience of express, the binary number can be converted to decimalism through the following formula:

\[
z_i = 1 + \sum_{j=1}^{n} 2^{n-j} x_j(t)
\]

Because \( x(t) \) is taken value in \([0,1]^n\), the value of \( z_i \) is from 1 to \( 2^n \). Because \( x(t) \) to \( z_i \) is corresponding, the expression in decimalism is equivalent to the expression in components.

B. Analysis of Semi-Markov Decision-making Process

Firstly, the state of PBN is given, at any time \( t \geq 0 \), the state \( x(t) \) of PBN is the decimal expression of gene state vector. Therefore, the state space is formed by all probable state vectors, namely \( S = [0, 1, \ldots, 2^n] \). Given PBN, some states represent certain property of gene regulation network. For example, whether some gene expressions result in the network reaches to undesired state. According to desired and undesired state, the state space can be divided into different kinds. Suppose that \( S_c \subset S \) is the set of desired states, then \( S_c = S - S_u \) corresponds to the set of undesired states.

Secondly, introduce semi-Markov nucleus to describe the dynamic evolution of PBN. For arbitrary \( t \in R_+ \), the given control input \( a \in A(x) \), semi-Markov nucleus \( Q \) is:

\[
Q(t, y | x, a) = H(t, | x, a) p(y | x, a)
\]

In the formula, \( H(t, | x, a) \) and \( p(y | x, a) \) respectively represent the sojourn time distribution and transition probability.

Thirdly, in order to complete the description of the first objective model of semi-Markov decision-making process in PBN, an optimal criterion shall be introduced. Firstly, the definition of strategy shall be given. The strategy actually is the operation criterion of decision maker for choosing action. For all \( x \in S \), the strategy is the function \( S \) affecting \( S \times R \) and meets the condition that \( g(x, \lambda) \) belongs to \( A(x) \). In other words, function \( g(x, \lambda) \) is the action taken at the state \( x \). Suppose that \( G \) is the set of all strategies.

For arbitrary \((x, \lambda) \in R_+ \times S \times R, g \in G \) , according to famous Tulcea’s theorem, there is random process \( \{S_n, J_n, A_n, \lambda, n \geq 0\} \). Under suitable condition, corresponding to discrete time process \( \{S_n, J_n, A_n\} \), the semi-Markov decision-making process \( \{Z(t), A(t), t \in R_+\} \) can be defined as:

\[
Z(t) = J_n, A(t) = A_n, S_n \leq t < S_{n+1}, n \in R_+, n \geq 0
\]

Suppose that \( t \) represents the time of the first undesired states set \( S_u \), namely \( t = \inf \{ t : Z(t) \in S_u \} \). For arbitrary \( g \in G \) and \((x, \lambda) \in S \times R_+\), the first objective probability criterion can be defined as:

\[
D^p(x, \lambda) = P_t^p \left[ \int_0^t r(Z(t), A(t)) dt \leq \lambda \right]
\]

Here, the first objective model of semi-Markov decision-making process of PBN is completed.
The selected comparison algorithm operates individually 10 times. Table 2 is the sample average of expected matrix.

<table>
<thead>
<tr>
<th></th>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>x5</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>0.0095</td>
<td>0.0110</td>
<td>-0.0019</td>
<td>-0.0075</td>
<td>0.5213</td>
</tr>
<tr>
<td>x2</td>
<td>0.2844</td>
<td>0.0090</td>
<td>-0.0025</td>
<td>0.0005</td>
<td>0.0025</td>
</tr>
<tr>
<td>x3</td>
<td>0.0000</td>
<td>-0.5975</td>
<td>0.0125</td>
<td>0.9986</td>
<td>0.0000</td>
</tr>
<tr>
<td>x4</td>
<td>0.0140</td>
<td>0.0050</td>
<td>0.0000</td>
<td>0.6939</td>
<td>0.0021</td>
</tr>
<tr>
<td>x5</td>
<td>0.0000</td>
<td>0.0015</td>
<td>-0.8117</td>
<td>0.0017</td>
<td>-0.0003</td>
</tr>
</tbody>
</table>

IV. EXPERIMENT CONTRASTIVE ANALYSIS

In order to test the effect of the algorithm in this Paper on network reconstruction of gene regulation, this algorithm is compared with the representative genetic algorithm and the gene regulation network in Fig. 2 is ratiocinated respectively.

Figure 3 is the gene network acquired by the algorithm in this Paper and Fig. 4 is the gene network acquired by classic genetic algorithm. Through comparison, it can be concluded that compared with traditional genetic algorithm, the optimal algorithm of improved particle swarm can preferably reflect the network structure of gene regulation network and its regulation relations.

V. CONCLUSIONS

This Paper puts forward a reconstruction algorithm of gene regulation network based on probabilistic Boolean network model, designs the weighting matrix model of gene regulation network and constructs the reconstruction process of gene regulation network by using probabilistic Boolean
network model and constructs the calculation process of probabilistic Boolean network model. The experiment data verifies the validity of the algorithm. Due to the limited experimental condition, the proposed algorithm only conducts performance verification on simple network and it does not verify the reconstruction algorithm performance of large-scale gene regulation network. The next step is to focus on the research of actual system development of algorithm.

ACKNOWLEDGMENTS

Zhejiang Provincial Natural Science Foundation of China(No.R1110261,LY13F010007);

REFERENCES