Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Computation Intelligent for Eukaryotic Cell-Cycle Gene Network

Shinq-Jen Wu [1], **Cheng-Tao Wu and Tsu-Tian Lee***

Department of Electrical Engineering Da-Yeh University, Chang-Hwa, Taiwan, R.O.C
*Department of Electrical and Control Engineering National Chiao-Tung University Hsin-Chu, Tainwan, R.O.C.
[1] jen@cn.nctu.edu.tw, dau@cn.nctu.edu.tw, ttlee@cn.nctu.edu.tw

Abstract - ***C****omputational intelligent approaches is adopted to construct the S-system of eukaryotic cell cycle for further analysis of genetic regulatory networks. A highly nonlinear power-law differential equation is constructed to describe the transcriptional regulation of gene network from the time-courses dataset. Global artificial algorithm, based on hybrid differential evolution, can achieve global optimization for the highly nonlinear differential gene network modeling. The constructed gene regulatory networks will be a reference for researchers to realize the inhibitory and activatory operator for genes synthesis and decomposition in Eukaryotic cell cycle .*

Index terms: cell cycle, S system, parameter estimation, hybrid differential evolution.

1 Introduction

As well known, cancer is regarded to be from abnormal cell propagation. Therefore, the research in cell cycle is critical for realizing and further controlling the cell's growth and generation. Cell cycle includes four phases (G1--S--G2--M): G1 (Gap1) phase is the cell growth and protein synthesis; DNA molecules are duplicated in S (Synthesis) phase; cell keeps growing, and protein synthesizes and prepares for mitosis in G2 (Gap2) phase; cell cycle enters into M (Mitosis) phase as DNA replication is completed, and thereafter, mother cell divides to produce two daughter cells.

Cyclin-dependent protein kinases (Cdks) is the central molecules in cell cycle network [1, 2, 3]. During the whole cell cycle, Cdk molecules keeps constant concentration but have different activation as combined with various cyclins. In G1 state, Cdk activity is low due to cyclin partners missing. At Start, it's activity rises dramatically since cyclin synthesis induces. This phenomena keeps for throughout S, G2 and M phases. At Finish, APC (anaphase-promoting complex) is activated; cdc20 and cdh1 included in APC are to recognize specific target proteins. Cyclin/Cdk can activate Cdc20 and inhibit Cdh1. In G1 state, Cdh1/APC activity is high but Cyclin/Cdk activity is low. In S-G2-M state, cyclin/Cdk activity is high but Cdh1/APC activity is low.

Recently, researches in inferring gene regulatory network via various intelligent computation approaches have come of age. Kikuchi and the coauthors propose a genetic algorithm [4] to transform parameters into chromosomes and resolve the optimal parameters through evolution procedure. Sakamoto and coauthors use genetic programming to develop the gene regulatory network into a tree form [5]. Kimura and coauthors adopt a cooperative co-evolutionary algorithm to model the genetic network [6]. Genetic programming and least mean square were combined to identify the gene network by Ando and other authors [7]. Noman and Iba use gradual optimization strategy implemented by differential evolution to infer the modelling [8]. Wang use hybrid differential evolution (HDE) to obtain global optimal solution of highly nonlinear problems [9, 10], and further, apply to solve the optimization problems of biochemical systems. HDE evolution method can compensate the general evolution method by migration and speeding-up operation. We here adopt this technology to infer the synthesis and degration flux of genes in eukaryotic cell cycle.

This paper is organized as follows: Section 2 describe the biochemical system of cell cycle, and the constructed S-system raw model, which is written in non-linear parameterized power-law differential equation to denote transcription regulation of gene network. Section 3 is the adopted global optimization technology. Section 4 is the conclusion and future work.

2 Eukaryotic Cell-Cycle

We now consider the following biochemical pathway of an eukaryotic cell cycle.

Acknowledgement: This research is supported by the National Science Council of the R.O.C. under Grant NSC 94-2213-E-009 -124 -

[1] To whom all correspondence should be addressed.

1-4244-0033-3/06/$20.00 ©2006 IEEE.

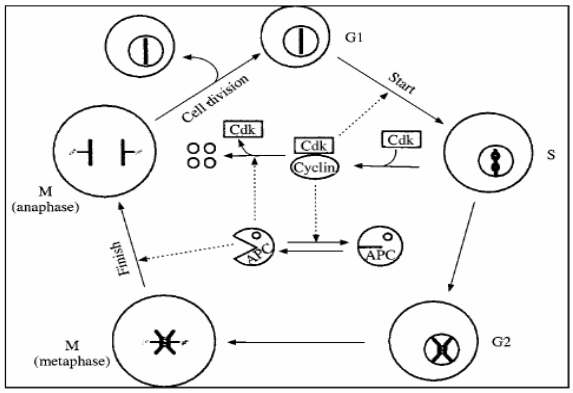

Fig.~1. Eukaryotic cell cycle [3].

This pathway can be described as

$$\frac{d[CycB]}{dt} = k_1 - (k_2^{'} + k_2^{''}[Cdh1])[CycB], \tag{1}$$

$$\frac{d[Cdh1]}{dt} = \frac{(k_3^{'} + k_3^{''}A)(1-[Cdh1])}{J_3 + 1 - [Cdh1]} - \frac{k_4 m[CycB][Cdh1]}{J_4 + [Cdh1]}, \tag{2}$$

$$\frac{d[Cdc20_T]}{dt} = k_5^{'} + k_5^{''}\frac{([CycB]m/J_5)^n}{1+([CycB]m/J_5)^n} - k_6[Cdc20_T], \tag{3}$$

$$\frac{d[Cdc20_A]}{dt} = \frac{k_7[IEP]([Cdc20_T] - [Cdc20_A])}{J_7 + [Cdc20_T] - [Cdc20_A]} - \frac{k_8[Mad][Cdc20_A]}{J_8 + [Cdc20_A]} - k_6[Cdc20_A], \tag{4}$$

$$\frac{d[IEP]}{dt} = k_9 m[CycB](1-[IEP]) - k_{10}[IEP], \tag{5}$$

$$\frac{dm}{dt} = \mu m(1 - \frac{m}{m_*}). \tag{6}$$

where $k_1 = 0.04$, $k_2^{'} = 0.04$, $k_2^{''} = 1$, $k_2^{'''} = 1$, $k_3^{'} = 1$, $k_3^{''} = 10$, $k_4^{'} = 2$, $k_4 = 35$, $k_5^{'} = 0.005$, $k_5^{''} = 0.2$, $k_6 = 0.1$, $k_7 = 1$, $k_8 = 0.5$, $k_9 = 0.1$, $k_{10} = 0.02$, $\mu = 0.01$, $J_3 = 0.04$, $J_4 = 0.04$, $J_5 = 0.3$, $n = 4$, $J_7 = 10^{-3}$, $J_8 = 10^{-3}$, $[Mad] = 1$, $m_* = 10$, $[CycB]_{threshold} = 0.1$; [CycB] is the average concentrations of cyclin B/Cdk dimmers; [Cdh1] is the average concentrations of active Cdh1/APC complexes; [$Cdc20_A$] is the concentration of "active" Cdc20; [$Cdc20_T$] is the total concentration of both active and inactive forms; k'$_i$ are rate constants, J'$_i$ are Michaelis constants, m denotes cell mass and A= [$Cdc20_A$]. Notice that cell divides, m → m/2, as [CycB] drops below 0.1 [3]. S-system is well-known for describing biochemical networks due to its easily generation for increasing genes numbers and the ability for analysis and construction [11, 12, 13]. General, the system is written into synthesis and degration flux as follows.

$$\dot{X}_i = V_i^+ - V_i^- = \alpha_i \prod_{j=1}^{n+m} X_j^{g_{ij}} - \beta \prod_{j=1}^{n+m} X_j^{h_{ij}}, \text{ for i = 1,2,…,n,}$$

where V_i^+ denotes net influx rate and V_i^- denotes net efflux rate; n is the number of dependent variables, m is the number of independent variables; α_i and β_i are rate constants, g_{ij} and h_{ij} are kinetic orders.

Based on the M-M model in Eqs.~(1) to (6), the corresponding S-system for eukaryotic cell cycle is constructed as the following nonlinear equation with 36 parameters.

$$\dot{X}_1 = \alpha_1 - \beta_1 X_1^{h_{11}} X_2^{h_{12}},$$

$$\dot{X}_2 = \alpha_2 X_2^{g_{22}} X_4^{g_{24}} - \beta_2 X_1^{h_{21}} X_2^{h_{22}} X_4^{h_{24}} X_6^{h_{26}},$$

$$\dot{X}_3 = \alpha_3 X_1^{g_{31}} X_6^{g_{36}} - \beta_3 X_3^{h_{33}},$$

$$\dot{X}_4 = \alpha_4 X_3^{g_{43}} X_4^{g_{44}} X_5^{g_{45}} - \beta_4 X_3^{h_{43}} X_4^{h_{44}} X_5^{h_{45}},$$

$$\dot{X}_5 = \alpha_5 X_1^{g_{51}} X_6^{g_{56}} - \beta_5 X_1^{h_{51}} X_5^{h_{55}} X_6^{h_{56}},$$

$$\dot{X}_6 = \alpha_6 X_6^{g_{66}} - \beta_6 X_6^{h_{66}},$$

where X_1 denotes [CycB], X_2 denotes [Cdh1], X_3 denotes [$Cdc20_T$], X_4 denotes [$Cdc20_A$], X_5 denotes [IEP], X_6 denotes m.

3 Global Optimal Search

We shall use HDE computation intelligent technology to global research the 36 parameters of S-system in Section 2. HDE is a parallel direct search algorithm for solving global optimal problem. It can be described as the following operation.

(1) Time series data

The gene expression time-courses dataset are generated from the evolution of M-M with regards to different initial conditions.

(2) Initialization

Lower- and up-bound for each decision parameter are defined to achieve uniformly random initial populations to cover the entire search space,

$$P_i^0 = P_{\min} + I_r(P_{\max} - P_{\min}), i = 1,...,n_p,$$

where I_r is the randomly uniform distributed number.

(3) Mutation operation

Four population individuals are selected randomly in G-th generation and combined into a difference vector F_r, which denotes search direction in the solving space. Mutation factor is randomly set to be valued in 0 and 1 for

each generation, and hence, perturbed individual, based on parent individual, is generated as

$$\hat{P}_i^G = P_p^G + F_r[(P_a^G - P_b^G) + (P_c^G - P_d^G)], i = 1,...,n.$$

(4) Crossover operation

For generating mutation operation on the basis of its parent individual, the candidate individuals can converge rapidly; but the population diversity is small. The crossover operation is hence to increase the local population diversity. We shall generate high quality individuals from perturbed individuals and parent individuals by binomial crossover with crossover factor C_r setting as value between 0 and 1,

$$P_{ji}^{G+1} = \begin{cases} P_{ji}^G, if \ \ a \ \ random \ \ number > C_r, \\ \hat{P}_{ji}^G, otherwise, j = 1,...,m, i = 1,...,n. \end{cases}$$

(5) Selection and evaluation

Two selection steps are included: first step is one-to-one competition between parent and its offspring; second step is to choice the best individual in the population.

Step1: $P_i^{G+1} = MIN\left\{ f\left(P_i^G\right), f\left(P_i^{G+1}\right)\right\}, i = 1,...,n$;

Step2: $\hat{P}_b^{G+1} = MIN\left\{ f\left(P_i^{G+1}\right)\right\}, i = 1,...,n$.

(6) Acceleration operation

Acceleration operation starts to search for better solution as crossover and mutation operations can not improve optimization anymore. If the best fitness does not descends continuously from one generation to next generation, a descend method is applied to improve the convergence speed,

$$P_b^{G+1} = \begin{cases} \hat{P}_b^{G+1}, if \ f\left(\hat{P}_b^{G+1}\right) < f\left(\hat{P}_b^{G}\right), \\ P_b^N = \hat{P}_b^{G+1} - \alpha \nabla_f, otherwise. \end{cases}$$

α, the step size between 0 and 1, is determined according to the descent property; ∇_f is the gradient of the objective function.

(7) Migration operation

This operation is a wide search exploration to balance the diversity decrease from acceleration operation and hence to avoid local or premature solution. A new population based on the best individual P_b^G is generated,

$$P_{ji} = \begin{cases} P_{jb} + \delta_{ji}\left(P_{j\min} - P_{jb}\right), if \ \tilde{\delta}_{ji} < \dfrac{P_{jb} - P_{j\min}}{P_{j\max} - P_{j\min}}, \\ P_{jb} + \delta_{ji}\left(P_{j\max} - P_{jb}\right), otherwise, \end{cases}$$

where $j = 1,...,m, i = 1,...,n$; a lower-bound and upper-bound of the decision parameters for the j-th gene are used; δ_{ji} and $\tilde{\delta}_{ji}$ are uniformly distributed random number between 0 and 1. Migration operation is performed only if the degree of population diversity η satisfies

$$\eta = \left\{ \sum_{\substack{i=1 \\ i \neq b}}^{n} \sum_{j=1}^{m} \eta_{ji} \right\} / \left(m\left(n-1\right)\right) < \varepsilon_1,$$

$$\eta_{ji} = \begin{cases} 0, if \left| \dfrac{P_{ji}^G - P_{jb}^G}{P_{jb}^G} \right| < \varepsilon_2; j = 1,...,m; i = 1,...,n, \\ 1, otherwise, \end{cases}$$

where η is the ratio of the total diversified genes to overall genes, except for the best individual; tolerance ε_1 and ε_2 is set to be value between 0 and 1.

The performance index is defined as the fitness of test datasets as follows. Figure 2 shows the simulation results for Eukaryotic cell cycle. The fitness value is 3.7411301E-06.

$$Fitness = \frac{\sum_i (model_i - exp_i)^2}{points \ of \ exercise}.$$

The simulation results for the generated s-system of the Eukaryotic cell are show in Figure 2. Based on the constructed model, the related biochemical reactions can be realized. For the change of concentration of cyclin B/Cdk dimmers, Cyclin B/Cdk dimmers is synthesized by cyclin B and Cdk, and the increasing in concentrations of Cdh1/APC will bring increasing degradation influence in concentration of Cycin B/Cdk.

4 Conclusion

S-system is a good candidate for gene network model due to its easier generation as gene numbers increase. However, the construction of S-system is a tough work. We here use an computation intelligent method, HDE, to infer the gene regulator network of Eukaryotic cell cycle. HDE can resolve high dimensional problem with global and diverse research direction. The inferred genetic network can provide readers to figure out the gene activatory and inhibitory relationship in cell cycle control.

Acknowledgements

The authors thank to Dr. Feng-Sheng Wang for concepts and suggestion of parameter estimation. Dr. Wang is a professor at the department of Chemical Engineering in National Chung Chen University. His researches are in the systems biology.

References

[1] Nasmyth, K., "Evolution of the cell cycle," *Philos. Trans. R. Soc. Lond. Ser. B: Biol. Sci. 349, 271-281, 1995.*

[2] Novak, B., Csikasz-Nagy, A., Gyorffy, B., Nasmyth, K. and Tyson, J. J., "Model scenarios for evolution of the eukaryotic cell cycle," *Philos. Tans. R. Soc. Lond. Ser. B: Biol. Sci. 353, 2063-2076, 1998b.*

[3] Tyson, J. J. and Novak, B., "Regulation of the Eukaryotic Cell Cycle: Molecular Antagonism, Hysteresis, and Irreversible Transitions," *J. theor. Biol., 210, 249-263, 2001.*

[4] Kikuchi S, Tominaga D, Arita M, Takahashi K, Tomita M, "Dynamic modeling of genetic networks using genetic algorithm and S-system," *Bioinformatics, 19:643-650, 2003.*

[5] Sakamoto, E. and Iba, H., "Inferring a System of Differential Equations for a Gene Regulatory Network by using Genetic Programming", *Proceedings of the Evolutionary Computation Congress*, Vol.1, p.p. 720-726, 2001.

[6] Kimura, S. and Ide, K., "Inference of S-system models of genetic networks using a cooperative co-evolutionary algorithm," *Bioinformatics, Vol. 21 no. 7 p. 1154-1163, 2005.*

[7] Ando, S., Sakamoto, E. and Iba, H., "Evolutionary Modeling Inference of Gene Network," *Information Sciences, 145(3-4):237-259, 2002.*

[8] Noman, N. and Iba, H., "Inference of Gene Regulatory Networks Using S-system and Differential Evolution," *GAECO*, p.p. 439-446, June, 2005.

[9] Lin, Y. C., Hwang, K. S. and Wang F. S., "Co-Evolutionary Hybrid Differential Evolution for Mixed-Integer Optimization Problems," *Eng. Opt*., Vol. 00, 1-20, 2001

[10] Tsai, K. Y. and Wang, F. S., "Evolutionary optimization with data collocation for reverse engineering of biological networks," Bioinformatics, vol.21 no. 7 ,1180-1188, 2005.

[11] Savageau, M.A., "Biochemical Systems Analysis: a Study of Function and Design in Molecular Biology," Addison-Wesley, Reading, Massachusetts, 1976.

[12] Voit, E.O., Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists, xii + 530 pp., Cambridge University Press, Cambridge, U.K., 2000.

[13] Veflingstad, S. R., Almeida, J. and Voit, E. O., "Priming nonlinear searches for pathway identification," Theoretical Biology and Medical Modeling, 1:8, 2004.

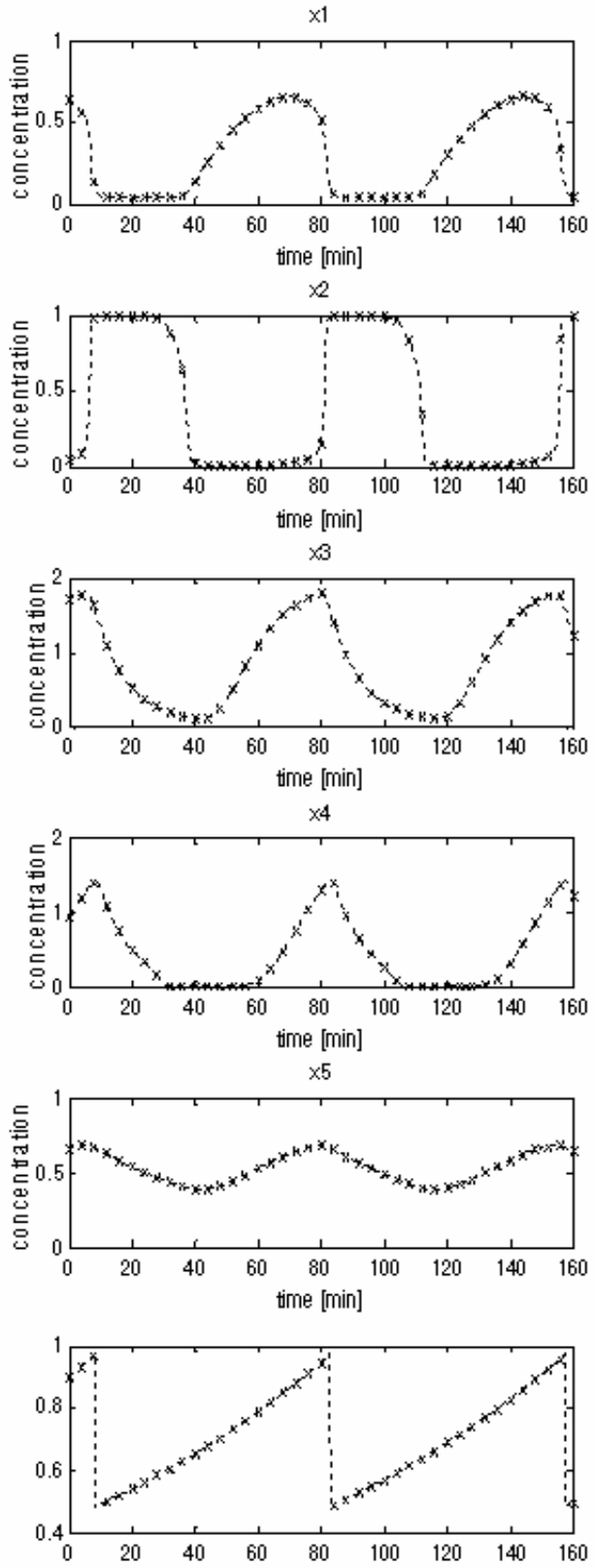

Fig.2 The simulation results for Eukaryotic cell cycle. The test data is denoted by x; estimated values are shown with dash line.