國 立 交 通 大 學

電機與控制工程系

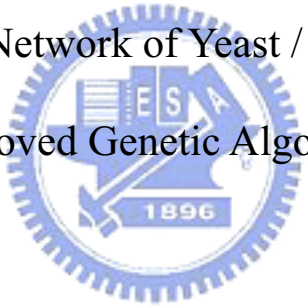
碩士論文

酵母菌與爪蟾屬青蛙蛋之基因調控網路：

改良型基因演算法

Genetic Regulatory Network of Yeast / Xenopus Frog Egg：

Improved Genetic Algorithm



研 究 生：周 家 賢

指導教授：李 祖 添 博士

吳 幸 珍 博士

中華民國九十五年七月

酵母菌與爪蟾屬青蛙蛋之基因調控網路：改良型基因演算法

Genetic Regulatory Network of Yeast / Xenopus Frog Egg： Improved Genetic Algorithm

研 究 生：周家賢　　Student：Chia-Hsien Chou

指導教授：李祖添　博士　　Advisor：Dr. Tsu-Tian Lee

吳幸珍　博士　　Dr. Shinq-Jen Wu

國 立 交 通 大 學

電機與控制工程系

碩士論文

A Thesis

Submitted to Department of Electrical and Control Engineering

College of Electrical and Computer Engineering

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Electrical and Control Engineering

July 2006

Hsinchu, Taiwan, Republic of China

中華民國九十五年七月

酵母菌與爪蟾屬青蛙蛋之基因調控網路：改良型基因演算法

學生：周家賢　　指導教授：李祖添　博士
吳幸珍　博士

國立交通大學電機與控制工程系（研究所）碩士班

摘　要

利用改良型基因演算法，針對酵母菌與爪蟾屬青蛙蛋分別做基因調控網路 modified power-law model 與 S-system 建模。藉由酵母菌實驗和爪蟾屬 Michaelis-Menten model 所得到的時間點資料集做訓練，以獲取最佳化的參數，因為 modified power-law model 和 S-system 可以清楚地描述基因在生成或消耗反應時是催化作用還是抑制作用，鑒於兩者的特色，對酵母菌的細胞週期與在爪蟾屬青蛙蛋細胞週期中的有絲分裂控制，能明瞭其基因的反應情形。具有遷移作用的改良型基因演算法不但可以做到全域搜尋，還可以精英主義的概念取得最佳的個體。最終得到的基因調控網路可以提供給生物學家在酵母菌和爪蟾屬青蛙蛋細胞週期方面做更深入的研究。

Genetic Regulatory Network of Yeast / Xenopus Frog Egg：Improved Genetic Algorithm

Student：Chia-Hsien Chou　　Advisor：Dr. Tsu-Tian Lee

Dr. Shinq-Jen Wu

Department (Institute) of Electrical and Control Engineering

National Chiao Tung University

Abstract

An improved genetic algorithm is proposed to achieve gene regulatory network modeling of Xenopus frog egg in S-system and yeast in modified power-law model respectively. Via the time-course datasets from experiment of yeast and Michaelis-Menten model of Xenopus, the optimal parameters are learned. The modified power-law model and S-system can clearly describe activative and inhibitory interaction between genes as generating and consuming process. We concern cell cycle of yeast and the mitotic control in cell cycle of Xenopus frog egg to realize gene reactions. The proposed improved genetic algorithm can achieve global search with migration and keep the best individual with elitism operation. The generated gene regulatory networks can provide biological researchers for further experiments in yeast and Xenopus frog egg cell cycle.

Acknowledgements

首先我要感謝李祖添校長和吳幸珍老師，在這兩年給我的諄諄教誨，在研究方面不吝惜地教授我一些觀念跟方法，使我獲益良多，學生不勝感激，也很感謝中正大學化工系的王逢盛教授與柯誌龍學長給予我對生物資訊研究的一些指教，還有吳政道學長跟我一起從事國科會計畫研究，也給予我很多的鼓勵與教導，非常感謝他們，協助我兩年的研究生涯。

在課業上，實驗室的學長王楀鴻，張捷，林桓宇，林洋霆，吳昆奇，吳尚鋒以及徐斌桓幫助我解決在修課時所遇到的問題，還有其他實驗室的同學林彥宏，張昭銘，張英德在修課時大家互相討論幫忙，十分感謝。此外，更有賴許多朋友的精神支持，甚或實質協助，我才能順利完成學業，掛一漏萬，可能尚有一些未提及的朋友們，在此均一併致謝。

最後要感謝我的父母周龍元與柯英雲，我的姊姊周怡姍，我的弟弟周家弘，還有我的女友蔡孟君，若不是他們長久以來的支持，不可能有我今天的小小成果。

Contents

List of tables

List of figures

Chapter 1

Introduction

1.1 Research Background

As the rapid development in cDNA microarray technologies, time-course gene dataset becomes available day by day. Hence, the construction of gene regulatory networks and signal transduction cascades for complicated biological systems has come of age. In order to approximate biological behavior for controlling metabolic/biological reaction, more and more experiments are set up for achievement of quantitative control. After post-genomic era, new scientific and technological methods on the biotechnology such as microarray technology are developed to bring massive biological knowledgeable dataset. Now system biologists are trying to describe biochemical phenomenon via mathematical model. With the mathematical model, we can realize the detailed genes-genes interaction, simulate the gene regulatory network and predict gene behavior.

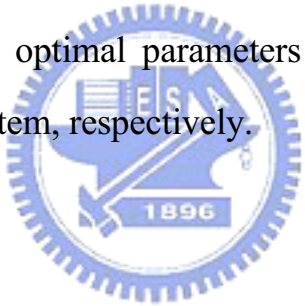
1.2 Literature Discussion

Numerous models are proposed to describe the gene network such as Boolean network, Bayesian network, Michaelis-Menten model, and S-system. Boolean network is to reconstruct gene regulatory network via Boolean function and

express gene relationship in graphical way [1, 2], which distinguish gene states to be INPUT and OUTPUT. At any time points, the state values of chosen INPUT genes are set to be 1 and 0 for non-chosen genes; the states values are given in the similar way as OUTPUT. Further, Bayesian network can also the probabilistic relationships of genes [2, 3]; joint probability distributions among genes are calculated to construct the graphical model. Michaelis-Menten model is nonlinear differential equations to describe the metabolic concentration in the biological system [4]. S-system is another nonlinear differential type expressed in power-law formalism [5, 6]. S-system describes gene regulation not only in mathematical description but also can further express into graphical form to show the activatory and inhibitory operation directly. Each equation is composed by synthesis and degradation flux; and the activation and inhibition relationship are shown in positive and negative kinetic order, respectively. In this paper, we shall develop the general S-system and another reformed nonlinear differential system, which is a modified power-low model modified from equation in [7, 8], to find out the gene regulatory network of yeast cell cycle and Xenopus cell cycle M phase control from microarray dataset, respectively.

However, the construction of such a highly nonlinear equation is a tough work. Chen and the authors first reform the nonlinear differential equation into linear form and then resolve it via linear algebra [7]. In these years, some researchers are devoted to infer gene regulatory network with various intelligent computation technologies such as hybrid differential evolution, genetic algorithm, genetic programming, ..., etc. Wang use Hybrid differential evolution and genetic algorithm to obtain the global optimal solution for highly nonlinear system and

various biochemical system [9, 10]. Kikuchi and coauthors use a genetic algorithm to transform parameters into individuals first and solve optimal parameters via evolution procedure [11]. Sakamoto and coauthors use genetic programming to develop the gene regulatory network in a tree form [12]. In this work, we shall adopt improved genetic algorithm [13] to infer the gene regulatory networks of yeast cell cycle in modified power-law model and Xenopus frog egg cell cycle in S-system. Improved evolutionary direction operator (IEDO), migration operation and elitism are combined into genetic program for global optimal, fast and best-optional searching. The input/output datasets, yeast cell cycle dataset [14], generated from Michaelis-Menten model of mitotic control in Xenopus frog eggs [15, 16], are used to train the genetic networks for searching the optimal parameters of the corresponding modified power-law model and S-system, respectively.



1.3 Content Organization

This paper is organized as follows: the biological systems, yeast cell cycle and cell cycle M phase control model of the Xenopus frog egg, are described in Chapter 2. Improved genetic algorithm is shown in Chapter 3. Chapter 4 shows the modeling and simulation results. Chapter 5 is the conclusion.

Chapter 2

Biological System

2.1 Yeast cell cycle

Research in cell cycle is very important not only for realizing cell reproduction but also for realizing cancer development. Figure 2.1 is the cell cycle that includes four phases (G1→S→G2→M). Cell grows up in G1 phase, produces RNA and synthesizes protein. During S phase, DNA is duplicated to produce two similar daughter cells. The cell continues to grow and produce protein and prepares to enter M phase during G2 phase. As DNA replication is completed, the cell enters M phase and divides.

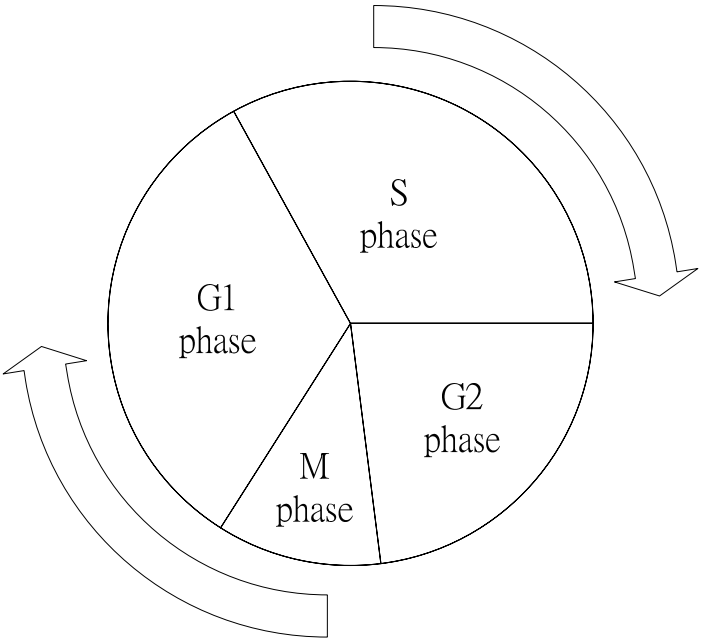


Figure 2.1 The four phases of cell cycle.

Table 2.1 The genes of the Yeast cell cycle.

Gene	Description
CDC28	Catalytic subunit of the main cell cycle cyclin-dependent kinase
CLN3	role in cell cycle START; involved in G(sub)1 size control; G1/S-specific cyclin, interacts with Cdc28p protein kinase to control events at START
SWI4	Involved in cell cycle dependent gene expression; both Swi4p and Swi6p are required for the in vivo protection of the SCB sequences at any cell cycle stage
SWI6	Involved in cell cycle dependent gene expression
MBP1	transcription factor
FUS3	Required for the arrest of cells in G(sub)1 in response to pheromone and cell fusion during conjugation
FAR1	Inhibitor of Cdc28p/Cln1p and Cdc28p/Cln2p complexes involved in cell cycle arrest for mating; Factor arrest protein
CLN1	G1 cyclin; role in cell cycle START
CLN2	G1 cyclin; role in cell cycle START
SIC1	P40 inhibitor of Cdc28p-Clb5 protein kinase complex
CLB5	role in DNA replication during S phase; additional functional role in formation of mitotic spindles along with Clb3 and Clb4; B-type cyclin involved in S-phase initiation
CLB6	role in DNA replication during S phase; B-type cyclin involved in S-phase initiation
CDC6	Protein involved in initiation of DNA replication; Protein that regulates initiation of DNA replication through binding to origins of replication at the end of mitosis, directing the assembly of MCM proteins and the pre-replication complex
CDC20	Cell Division Cycle; Required for onset of anaphase; adaptor for APC
GRR1	F box protein with several leucine rich repeats
CDC4	Init. of DNA synthesis & spindle pole body separation; dispensable for both mitotic and meiotic spindle pole body dupl.; essential for mitotic but not premeiotic DNA synth.; wt levels of synaptonemal complexes and intragenic recombination

The yeast cell cycle gene expression data is collected by Spellman [14]. The dataset were covered six experimental conditions (CLN2; CLN3; ALPH, CDC15, CDC28 and ELU). We use the gene time-course data from experimental condition CDC28, which contains 24 time points. We concentrate in G1 and S phase and set up our dataset from the sub network of yeast cell cycle pathway. The descriptions and datasets are available at http://cellcycle-www.stanford.edu. This dataset in Table 2.1 involves 16 genes.

2.2 M phase control of Xenopus frog egg

Michaelis-Menten model are concerned to describe the mitotic control in cell-cycle of Xenopus frog egg [15, 16],

$$\dot{x}_1 = k_1 - k_2 x_1 - k_3 x_1, \tag{1}$$

$$\dot{x}_2 = k_{pp} x_5 - \left(k_{wee} + k_{cak} + k_2\right) x_2 + k_{25} x_3 + k_3 x_1, \tag{2}$$

$$\dot{x}_3 = k_{wee} x_2 - \left(k_{25} + k_{cak} + k_2\right) x_3 + k_{pp} x_4, \tag{3}$$

$$\dot{x}_4 = k_{wee} x_5 - \left(k_{pp} + k_{25} + k_2\right) x_4 + k_{cak} x_3, \tag{4}$$

$$\dot{x}_5 = k_{cak} x_2 - \left(k_{pp} + k_{wee} + k_2\right) x_5 + k_{25} x_4, \tag{5}$$

$$\dot{x}_6 = \frac{k_a x_5 \left(1 - x_6\right)}{1 + K_a - x_6} - \frac{k_b x_6}{K_b + x_6}, \tag{6}$$

$$\dot{x}_7 = \frac{k_e x_5 \left(1 - x_7\right)}{1 + K_e - x_7} - \frac{k_f x_7}{K_f + x_7}, \tag{7}$$

$$\dot{x}_8 = \frac{k_g x_5 \left(1 - x_8\right)}{1 + K_g - x_8} - \frac{k_h x_8}{K_h + x_8}, \tag{8}$$

$$\dot{x}_9 = \frac{k_c x_8 \left(1 - x_9\right)}{1 + K_c - x_9} - \frac{k_d x_9}{K_d + x_9}, \tag{9}$$

with

$$k_2 = V_2' + x_9\left(V_2'' - V_2'\right),$$
$$k_{wee} = V_{wee}'' + x_7\left(V_{wee}' - V_{wee}''\right),$$
$$k_{25} = V_{25}' + x_6\left(V_{25}'' - V_{25}'\right),$$

where x_i, i = 1, 2, ···, 9, are the concentrations or activities of cyclin, unphosphorylated cyclin-Cdc2, Tyr-15 phosphorylated cyclin-Cdc2, doubly phosphorylated cyclin-Cdc2, Thr-161 phosphorylated cyclin-Cdc2 activated by four enzymes Cdc25, Wee1, IE and APC, respectively. Eqs. (1) ~ (4) describe four phosphorylation states of the cyclin-Cdc2 dimer. The Thr-161 phosphorylated cyclin-Cdc2 represents the M phase promoting factor (MPF). The concentration of MPF can control the phases during cell cycle. For instance, high concentration of MPF can make the cell to divide to two child cells.

The numerical solution of Michaelis-Menten model is shown as Figure 2.2 ~ 2.5. The parameter values K_a=0.1, K_b=1.0, K_c=0.01, K_d=1.0, K_e=0.1, K_f=1.0, K_g=0.01, K_h=0.01, k_a=2.0, k_b=0.1, k_c=0.13, k_d=0.13, k_e=2.0, k_f=0.1, k_g=2.0, k_h=0.15, k_1=0.01, k_3=0.5, k_{cak}=0.64, k_{pp}=0.004, $V_{25}' = 0.017$, $V_{25}'' = 0.17$, $V_2' = 0.005$, $V_2'' = 0.25$, $V_{wee}' = 0.01$, $V_{wee}'' = 0.1$ from [15, 16].

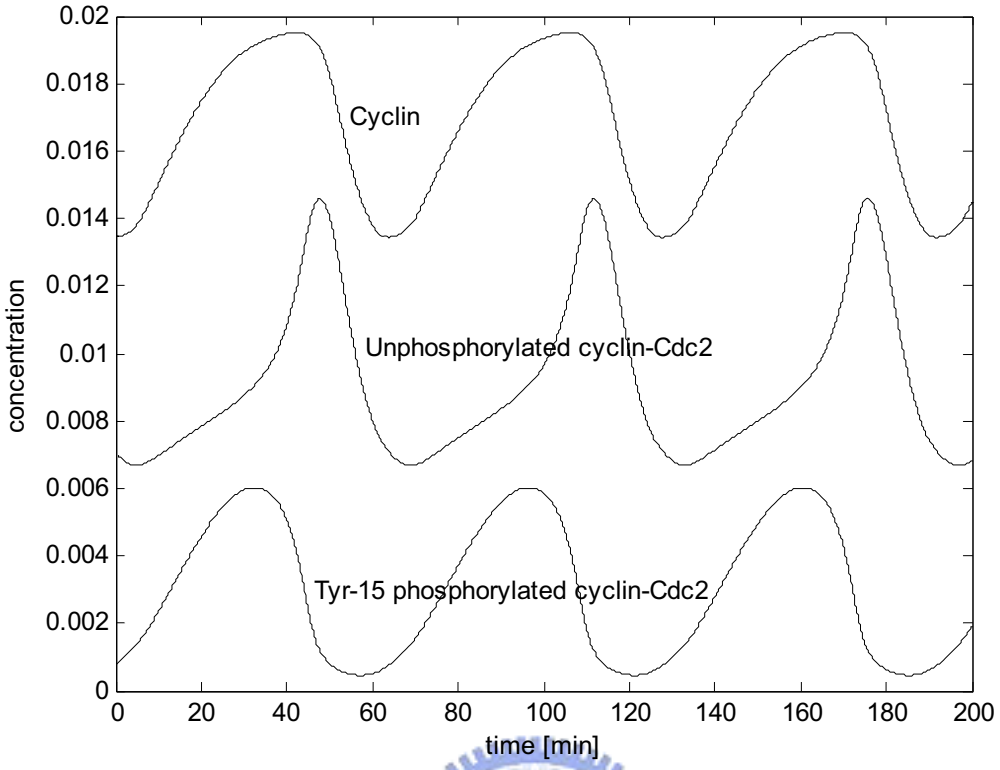


Figure 2.2 The concentrations of cyclin, unphosphorylated cyclin-Cdc2 and Tyr-15 phosphorylated cyclin-Cdc2.

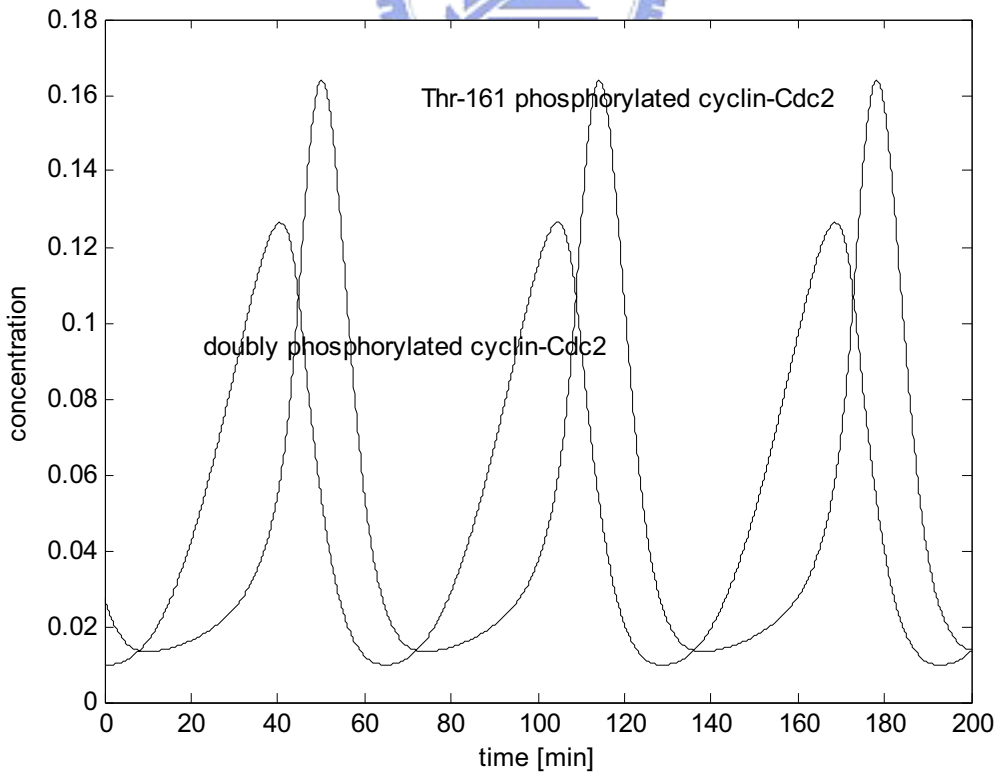


Figure 2.3 The concentrations of doubly phosphorylated cyclin-Cdc2 and Thr-161 phosphorylated cyclin-Cdc2.

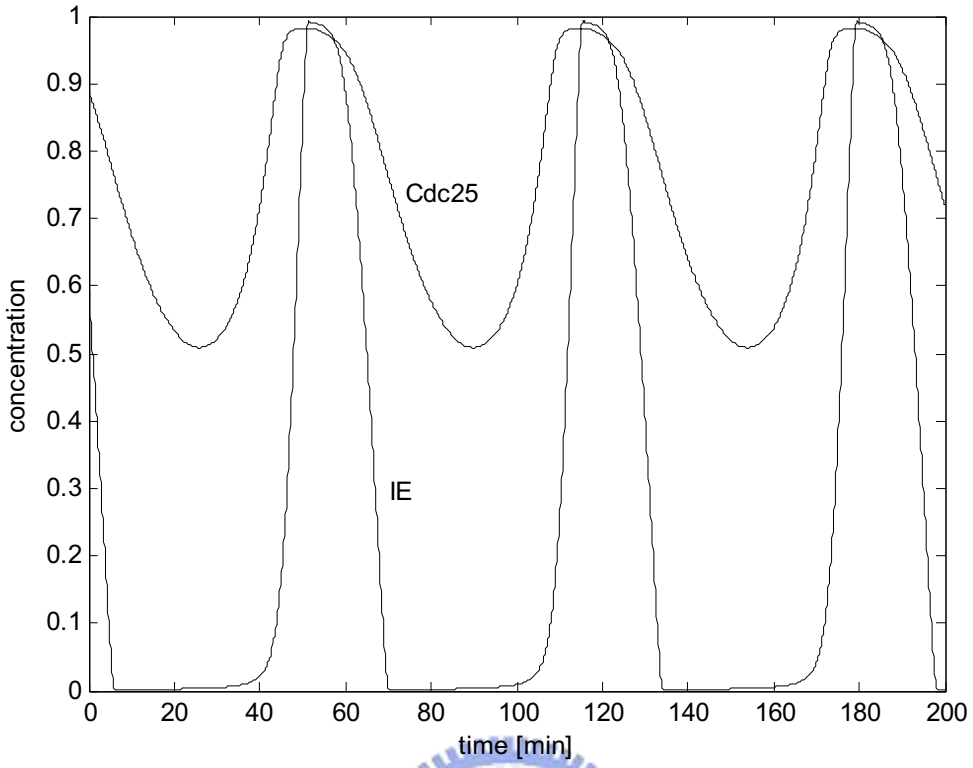


Figure 2.4 The concentrations of Cdc25 and IE.

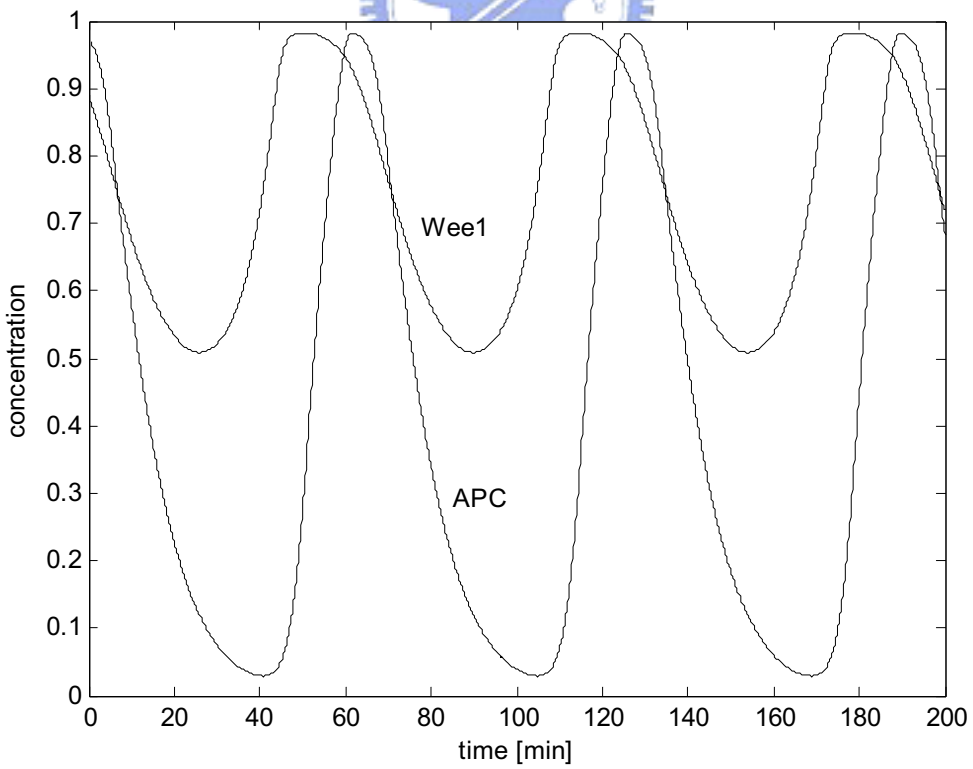


Figure 2.5 The concentrations of Wee1 and APC

Chapter 3

Improved Genetic Algorithm

3.1 Introduction

Based on general genetic algorithm, improved genetic algorithm includes improved evolutionary direction operator to speed the searching; migration operator to escape from bogging down into local solution and elitism to keeping the best to be passed down always. The flow chart is shown in Figure 3.1.

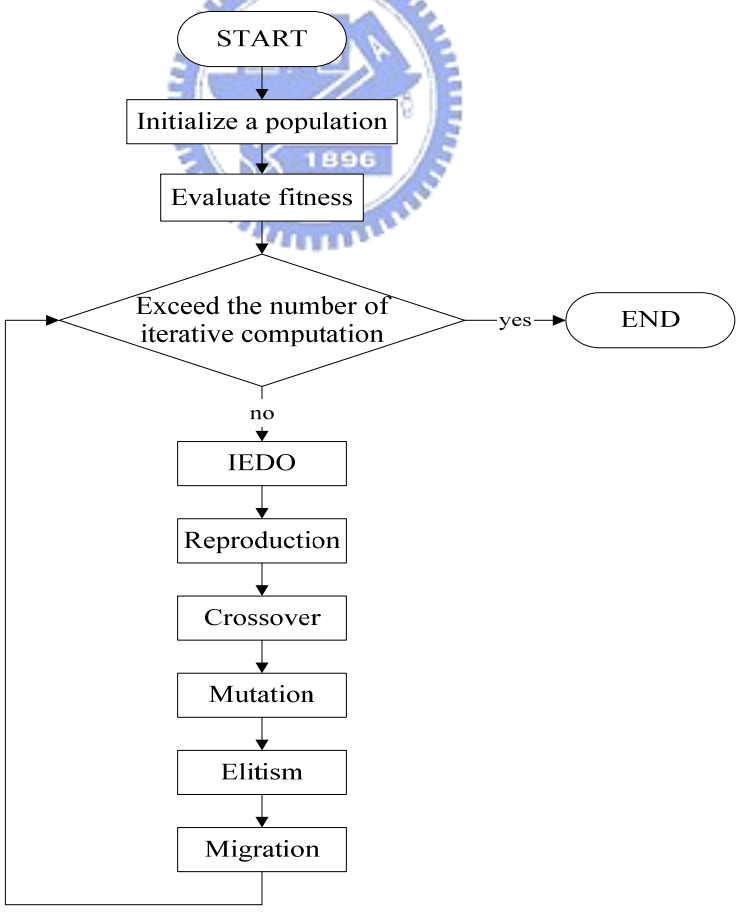


Figure 3.1 Flow chart for improved genetic algorithm.

The algorithm initializes a population first. The definition of population is shown in Figure 3.2. The chromosome in individual represents parameter. And then, we evaluate the fitness value of each individual in the population and keep the better individuals and eliminates the worse individuals though evolution procedures.

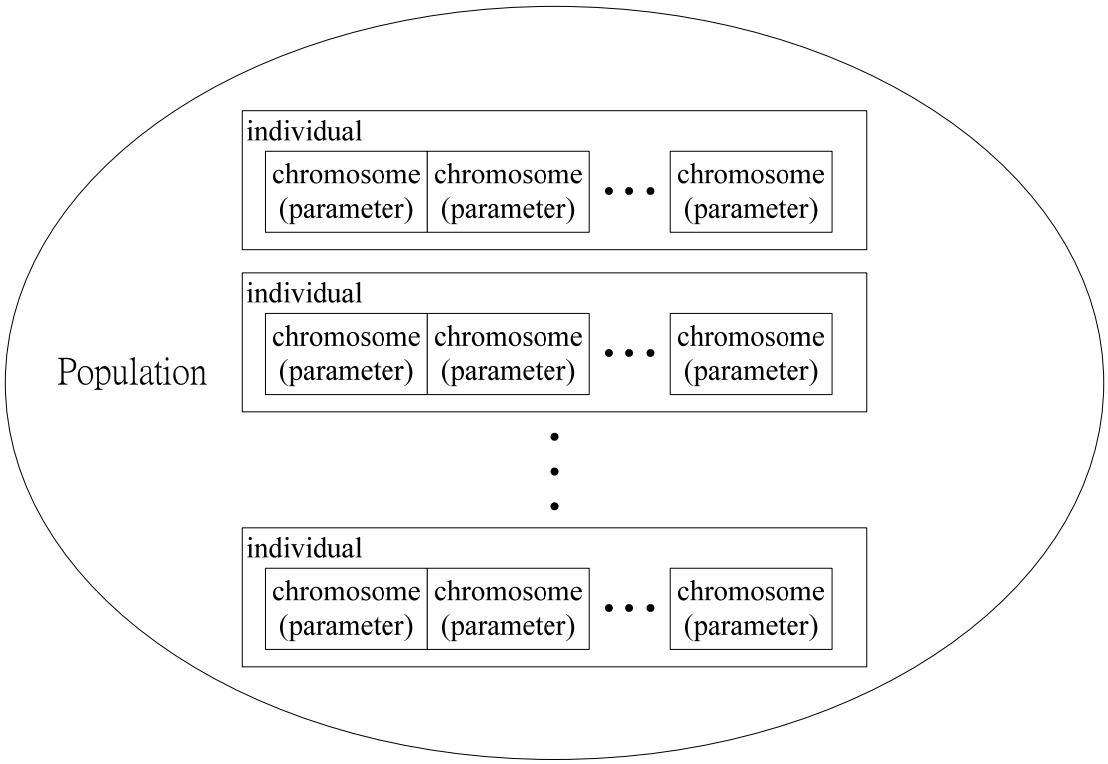


Figure 3.2 Population definition.

3.2 IEDO

Three preferred fitness value and its associated individuals are chosen to decide the evolutionary direction in the population. Improved evolutionary direction operator (IEDO) has the ability for both local and global search synchronously. We can use the IEDO operator to search quickly and converge toward the global

optimal solution. Besides, the IEDO operator can avoid bogging down in the local optimal solution, which conventional GA is usually stuck into.

After completing all fitness values of individuals, we choose preferred fitness values denote F_b, F_s, F_t and its associated individuals denote I_b, I_s, I_t. And then, the new individual denote I_{ideo} is calculated as

$$I_{iedo} = I_b + r_1 * D_1 * (I_b - I_s) + r_2 * D_2 * (I_b - I_t), \quad (10)$$

$$I_{iedo} = max(min(I_{iedo}, I_{max}), I_{min}), \quad (11)$$

where r_1 and r_2 are two random numbers, r_1, $r_2 \in [0, 1]$; D_1 and D_2 are the magnitude of two evolutionary directions to be 1; $I_{\max}$ and $I_{\min}$ are the upper and lower bound, respectively. The new fitness value F_{new} of the I_{ideo} is calculated; if the F_{new} is better than one of the three preferred fitness values, it would be replaced that.

3.3 Reproduction

The probability of reproduction directly depends on the fitness of the individuals. The individual with better fitness has high probability of reproduction. In contrast, the individual with worse fitness has low probability of reproduction. The rule of reproduction is shown in Figure 3.3.

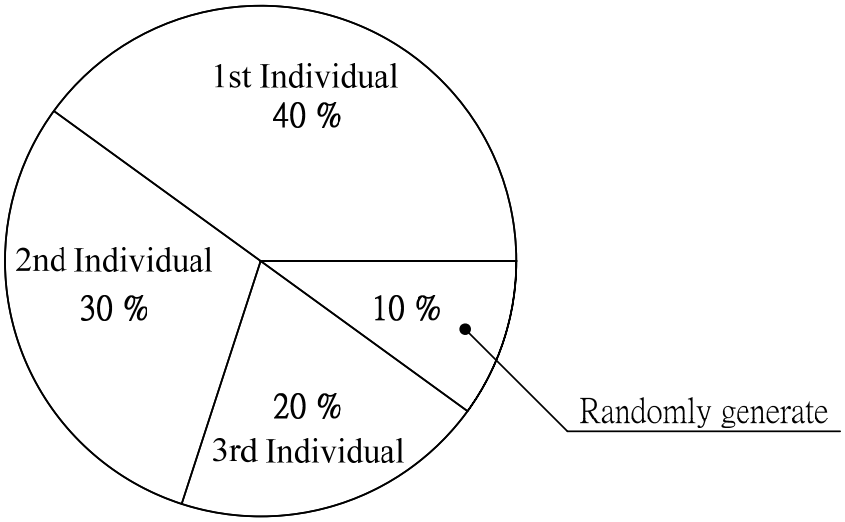


Figure 3.3 Rule of reproduction.

3.4 Crossover

Two-point crossover operation shown in Figure 3.4 is adopted and operates according to crossover probability, which involves selection of two crossover cut-points randomly and then exchanges the chosen two cut-points genes of parent individuals to generate two child individuals. And further, randomly select one of the new individuals to replace father or mother individuals.

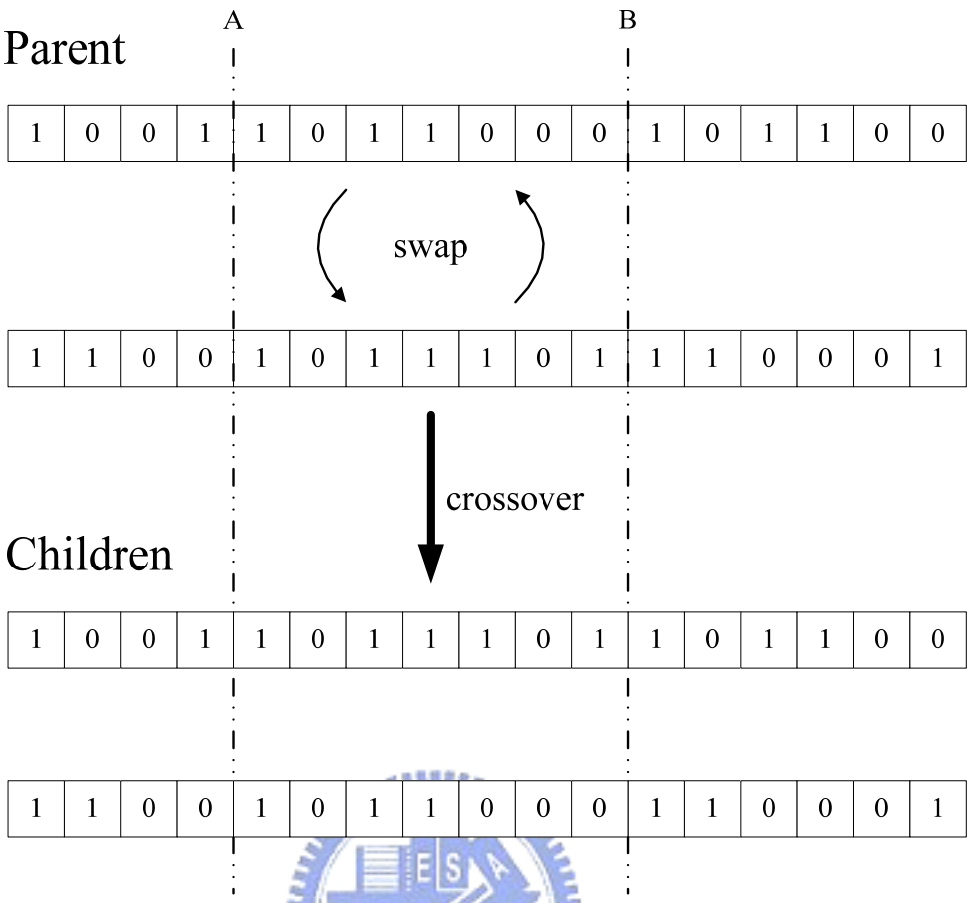


Figure 3.4 Crossover operation.

3.5 Probability Mutation and Elitism

Different from the conventional GA to choose only one gene in individual randomly for mutation operation, all genes in individuals are chosen and their mutation probability are assigned by the designer for exchanging their original values in Figure 3.5. This will bring excessive diversity in population and hence may fail to converge to temperately optimal solution. Therefore, we adopt elitism operator to decrease this effect. Elitism operator is to keep the best individual to survive for each generation and hence to ensure good characteristic to pass down

always.

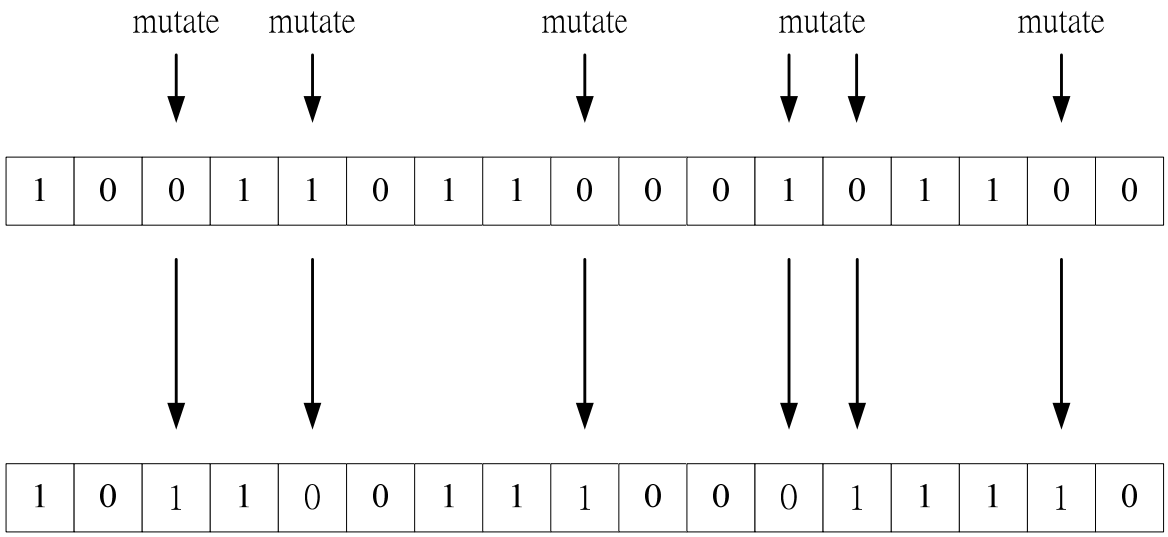


Figure 3.5 Probability mutation operation.

3.5 Migration

To wider the search space, a migration operator is done to get a new and diverse population. The degree of population diversity η is to check if the migration should be performed.

$$temp_{ij} = \begin{cases} 0, \text{ if } \left| \frac{x_{ij} - x_{bj}}{x_{bj}} \right| < \varepsilon_2, \\ 1, \text{ otherwise} \end{cases} \tag{12}$$

$$\eta = \sum_{i=1}^{NP-1} \sum_{j=1}^{Dim\_I} temp_{ij} \Big/ Dim\_I \times (NP-1), \tag{13}$$

where $\varepsilon_2 \in [0,1]$ is the tolerance of the real-valued gene diversity; x_{ij} and x_{bj} are

respectively the *j*-th chromosome in the *i*-th individual and the best individual; NP is the number of individual; *Dim_I* is the dimension of individual; η is in the range between 0 and 1. $\varepsilon_1 \in [0,1]$ is a tolerance-threshold of population diversity for migration; if η is small than ε_1, migration operate to generate a new chromosome as follows.

$$x_{ij} = \begin{cases} x_{bj} + r_2 \times \left(x_{j,\min} - x_{bj} \right), & \text{if } \dfrac{x_{bj} - x_{j,\min}}{x_{j,\max} - x_{j,\min}} > r_1 \\ x_{bj} + r_2 \times \left(x_{j,\max} - x_{bj} \right), & \text{ortherwise} \end{cases}, \qquad (14)$$

where $x_{j,\max}$ and $x_{j,\min}$ are the upper and lower bound of the *j*-th chromosome, respectively. The r_1 and r_2 are two random numbers, $r_1, r_2 \in [0, 1]$.

3.6 Fitness

Every individual is evaluated by its fitness value, which keeps the better individuals and eliminates the worse individuals. We adopt two different methods to evaluate fitness. The fitness of an individual in yeast cell cycle is defined as

$$fitness = \sum_{i=1}^{n} \sum_{j=1}^{N\text{-}1} \left(X_{ei}\left(t_0 + j\Delta t\right) - X_i\left(t_0 + j\Delta t\right) \right)^2 + \sum kinetic\ order, \qquad (15)$$

where the $X_i(t)$ is approximate value of the *i*-th variable at time *t*, $X_{ei}(t)$ is original time-course data at time *t*, *N* is number of total time points.

The fitness of an individual in Xenopus M phase control is defined as

$$fitness = \frac{\sum_{i=1}^{Dim\_I} \sum_{j=1}^{N-1} \left(X_{ei}\left(t_0 + j\Delta t\right) - X_i\left(t_0 + j\Delta t\right) \right)^2}{Dim\_I \times N}, \qquad (16)$$

where N is number of time points; $X_{ei}(t)$ and $X_i(t)$ are experiment value and estimated value of i-th reactant at time t, respectively.

Chapter 4

Modeling and Simulation Results

4.1 Yeast cell cycle

The mathematical model in yeast cell cycle is approximated from the model adopted in [7, 8].

$$\dot{X}_i(t) = G_i(t) - \lambda_i X_i(t), \quad i=1,2,\cdots,n, \tag{17}$$

where $G_i(t)$ is the transcription rate, λ_i is the self-degradation rate and n is the number of the variable, $X_i(t)$ is the concentration of the i-th gene at time t. $G_i(t)$ is a nonlinear function,

$$G_i(t) = \sum_{j=1}^{m} a_{ij} \frac{1}{1+\exp\left\{-\alpha\left[u\left(t,\beta_j,\delta\right)-\gamma\right]\right\}}, \tag{18}$$

$$u(t,\beta_j,\delta) = \begin{cases} 0 & t \le \beta_j - \delta \\ \dfrac{t-\left(\beta_j-\delta\right)}{\delta} & \beta_j - \delta \le t \le \beta_j \\ \dfrac{(\beta_j+\delta)-t}{\delta} & \beta_j \le t \le \beta_j + \delta \\ 0 & t \ge \beta_j + \delta \end{cases}. \tag{19}$$

We use a power-law function $V_i(t)$ to approximate the nonlinear function $G_i(t)$ in Eq. (18) to denote the synthesis rate.

$$V_i(t) = \lambda_i \prod_{j=1}^{n} X_j^{f_{ij}}(t), \tag{20}$$

where λ_i is the rate constant and f_{ij} is kinetic order. Further, the degradation term in Eq. (17) is replaced by $\gamma_i X_i^{k_i}(t)$ to emphasize how a gene reacts itself. The kinetic orders, f_{ij} and γ_i, can be positive or negative; positive kinetic orders indicate activating influences, but negative kinetic orders mean inhibition. In other words, the following modified power-law dynamic model is proposed.

$$\begin{aligned} \dot{X}_i(t) &= f_i(\mathbf{X}, \mathbf{P}) = V_i(t) - \gamma_i X_i^{k_i}(t) \\ &= \lambda_i \prod_{j=1}^{n} X_j^{f_{ij}}(t) - \gamma_i X_i^{k_i}(t), \quad i=1,2,\cdots,n, \end{aligned} \tag{21}$$

where n is the number of the variables; the vector **X** in Eq. (21) indicates all genes in the yeast cell cycle; the vector **P** in Eq. (21) consists the rate constants, λ_i and γ_i, and kinetic orders, f_{ij} and k_i. According to 16 genes of yeast cell cycle, there are 16 differential equations with 288 parameters.

Figure 4.1 is the pathway for the modified power-low model whose fitness is 1.1104688E-02. Black lines represent activation reaction and red lines represent inhibition reaction. The start point of the lines is the reactant and the end point is the product. For instance, the concentration of CDC28 increases rapidly as the concentrations of MBP1, CLN1, CDC6, CDC20 and GRR1 increase; however, the concentration of CDC28 decreases rapidly as the concentrations of CLN3, SWI4, FUS3, FAR1, CDC4, CLB6 andCLB6 increase.

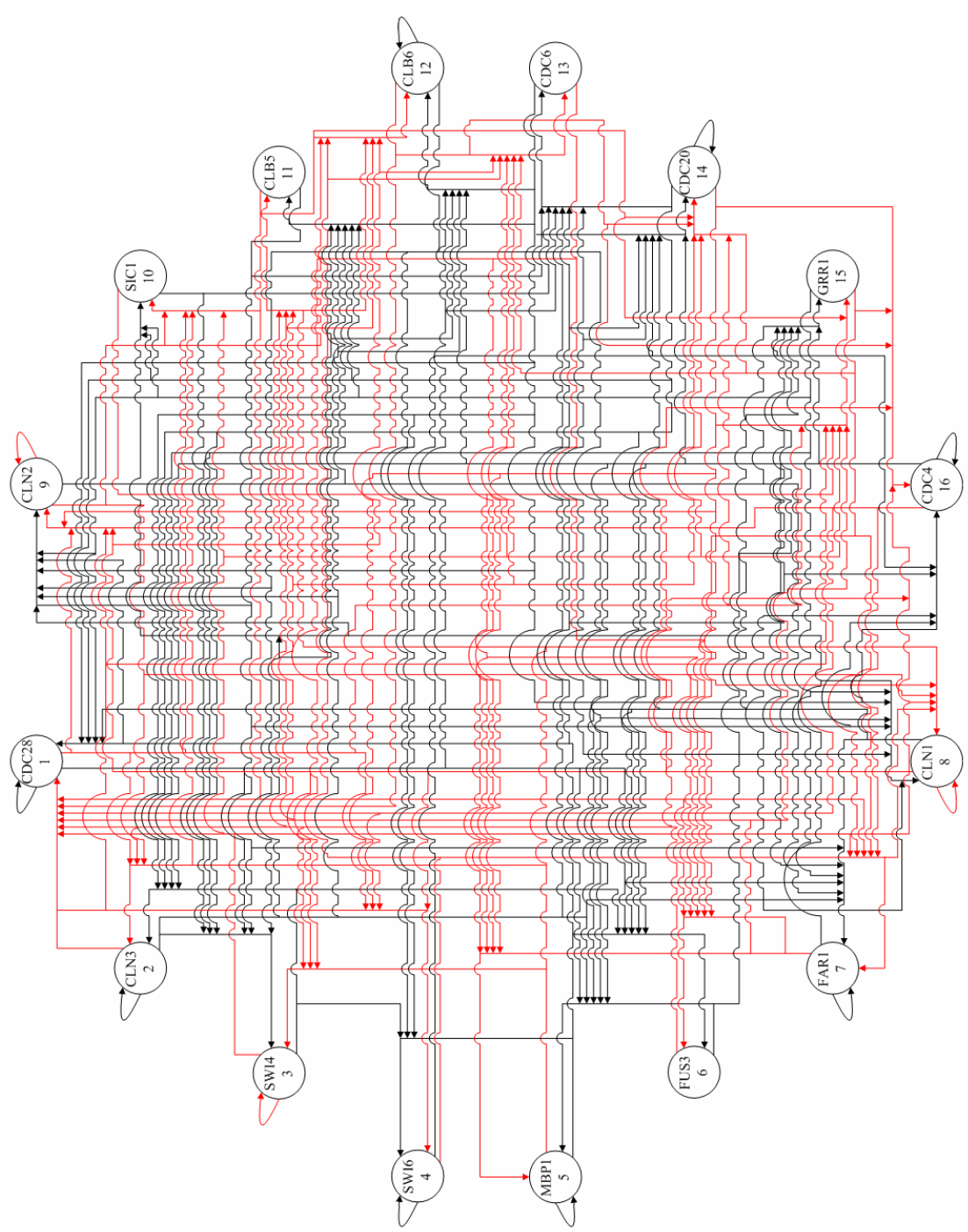


Figure 4.1 The gene regulatory network of the generated modified power-low system.

4.2 M phase control of Xenopus frog egg

We shall generate the training dataset from the Eqs. (1) ~ (9) to generate the corresponding S-system model of the frog cell cycle to further realize the gene-gene inhibitory and activatory operation for gene and enzyme synthesis and decomposition. The training dataset is shown in Figure 4.2 ~ 4.5.

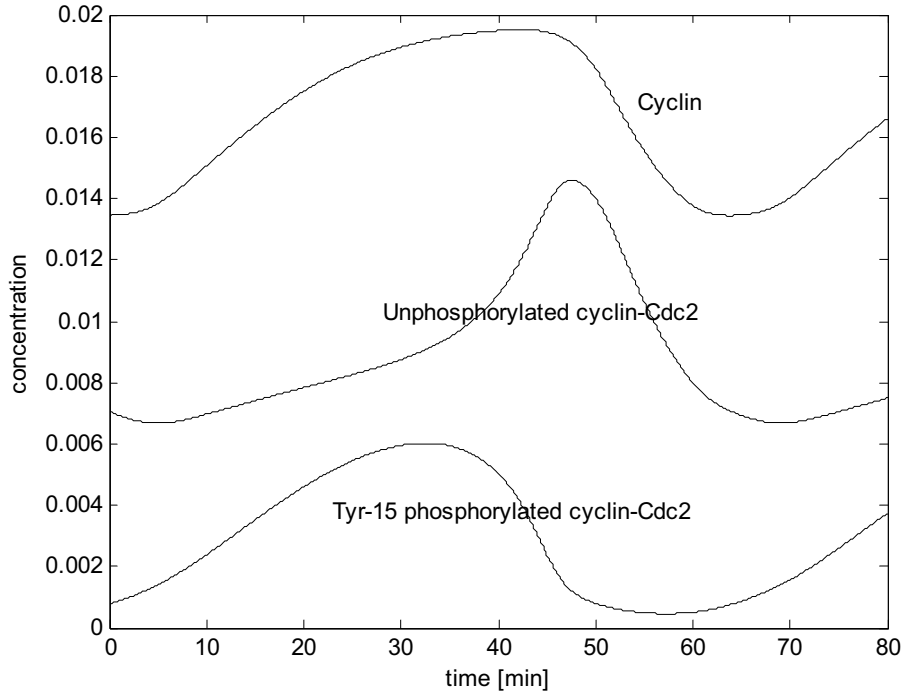


Figure 4.2 Training dataset-1.

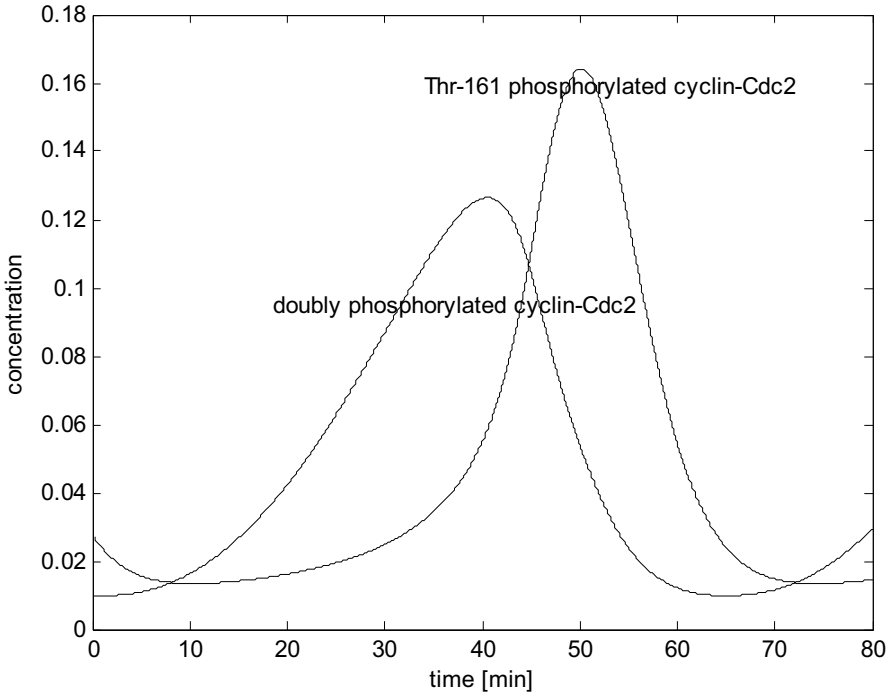


Figure 4.3 Training dataset-2.

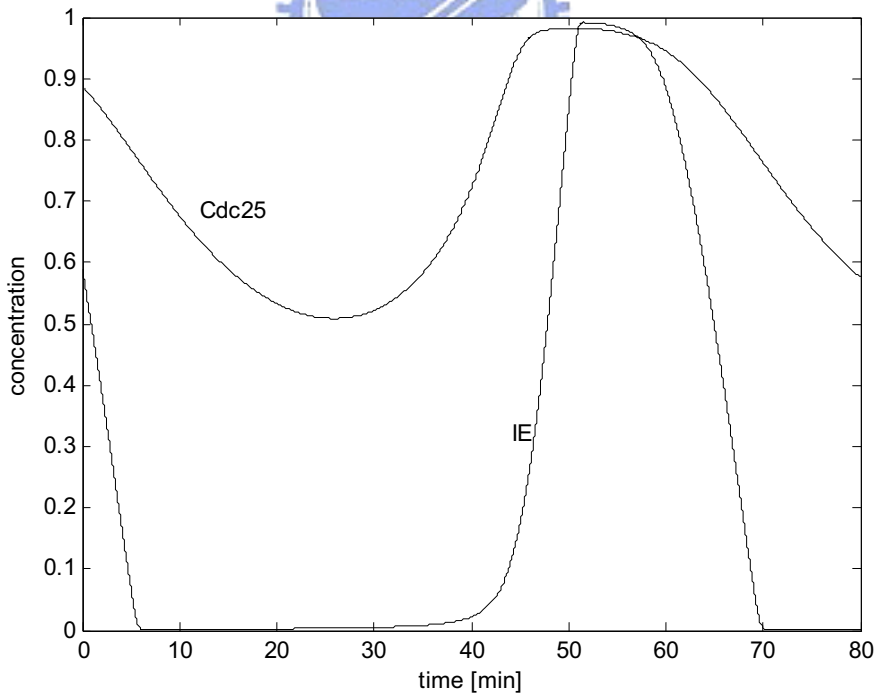


Figure 4.4 Training dataset-3.

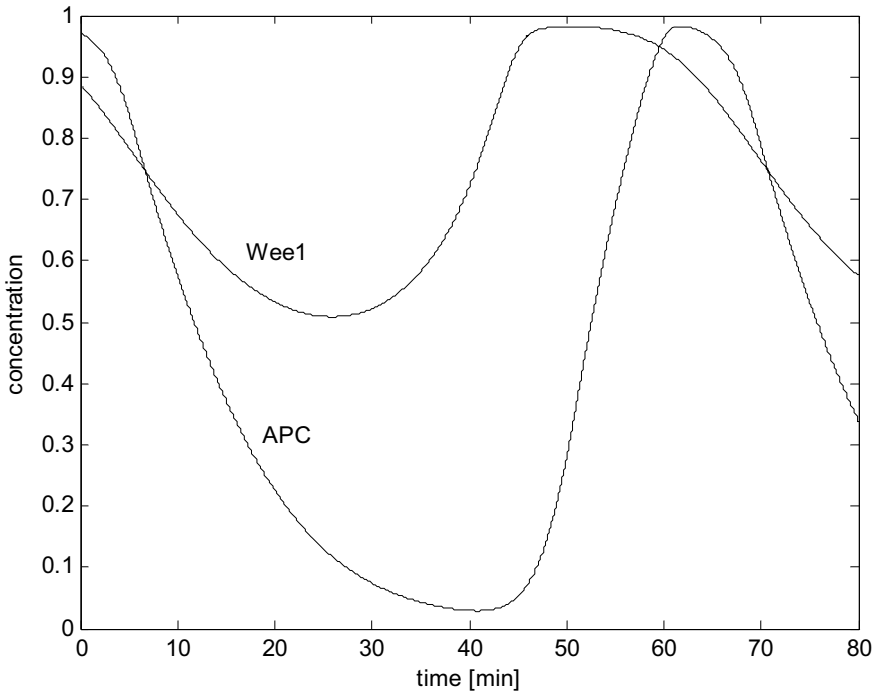


Figure 4.5 Training dataset-4.

S-system use power-law flux to describe the synergism and saturation of the biological system,

$$\dot{x}_i = \alpha_i \prod_{j=1}^{n} x_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} x_j^{h_{ij}}, \quad \text{for } i = 1,2,...,n, \tag{22}$$

where x_i is the state variable or reactant; n is the number of x_i. α_i is the production rate-constant and β_i is the degradation rate-constant; both can be positive or zero. g_{ij} and h_{ij}, are kinetic orders; their values can be positive to indicate activating influences or negative to denote inhibition. We now construct our S-system structure for Xenopus frog egg as

$$\dot{x}_1 = \alpha_1 x_1^{g_{11}} x_9^{g_{19}} - \beta_1 x_1^{h_{11}} x_9^{h_{19}} \tag{23}$$

$$\dot{x}_2 = \alpha_2 x_1^{g_{21}} x_2^{g_{22}} x_3^{g_{23}} x_5^{g_{25}} x_6^{g_{26}} x_7^{g_{27}} x_9^{g_{29}} - \beta_2 x_2^{h_{22}} x_3^{h_{23}} x_6^{h_{26}} x_7^{h_{27}} x_9^{h_{29}} \tag{24}$$

$$\dot{x}_3 = \alpha_3 x_2^{g_{32}} x_3^{g_{33}} x_4^{g_{34}} x_6^{g_{36}} x_7^{g_{37}} x_9^{g_{39}} - \beta_3 x_2^{h_{32}} x_3^{h_{33}} x_6^{h_{36}} x_7^{h_{37}} x_9^{h_{39}} \tag{25}$$

$$\dot{x}_4 = \alpha_4 x_3^{g_{43}} x_4^{g_{44}} x_5^{g_{45}} x_6^{g_{46}} x_7^{g_{47}} x_9^{g_{49}} - \beta_4 x_4^{h_{44}} x_5^{h_{45}} x_6^{h_{46}} x_7^{h_{47}} x_9^{h_{49}} \tag{26}$$

$$\dot{x}_5 = \alpha_5 x_2^{g_{52}} x_4^{g_{54}} x_5^{g_{55}} x_6^{g_{56}} x_7^{g_{57}} x_9^{g_{59}} - \beta_5 x_4^{h_{54}} x_5^{h_{55}} x_6^{h_{56}} x_7^{h_{57}} x_9^{h_{59}} \tag{27}$$

$$\dot{x}_6 = \alpha_6 x_5^{g_{65}} x_6^{g_{66}} - \beta_6 x_5^{h_{65}} x_6^{h_{66}} \tag{28}$$

$$\dot{x}_7 = \alpha_7 x_5^{g_{75}} x_7^{g_{77}} - \beta_7 x_5^{h_{75}} x_7^{h_{77}} \tag{29}$$

$$\dot{x}_8 = \alpha_8 x_5^{g_{85}} x_8^{g_{88}} - \beta_8 x_5^{h_{85}} x_8^{h_{88}} \tag{30}$$

$$\dot{x}_9 = \alpha_9 x_8^{g_{98}} x_9^{g_{99}} - \beta_9 x_8^{h_{98}} x_9^{h_{99}} \tag{31}$$

Note that since concentrations of x_1, x_2, x_3, x_4 and x_5 are too small as compared to other variables. According to Eqs. (23) ~ (31), there are 83 parameters to estimate. The scale-up operation is adopted to normalize all states variables to a computation reasonable range to improve the computation error. Another test data from Michaelis-Menten model is used to demonstrate the performance of the improved genetic algorithm program, Figure 4.6 ~ 4.14 is the simulation results with the estimated fitness 3.0766122E-08 for *N*=80,000, *Dim_I*=83. The low fitness value ensures the good fitting of the simulation results with the datasets and also guarantees the reliability of the generated S-system. From the constructed S model, we can realize the interaction between various genes in Xenopus frog egg. For instance, the concentration of x_2 increases rapidly as concentrations of x_1, x_3 and x_5 increase; the concentration of x_2 decreases rapidly as concentrations of x_7 and x_9 increase.

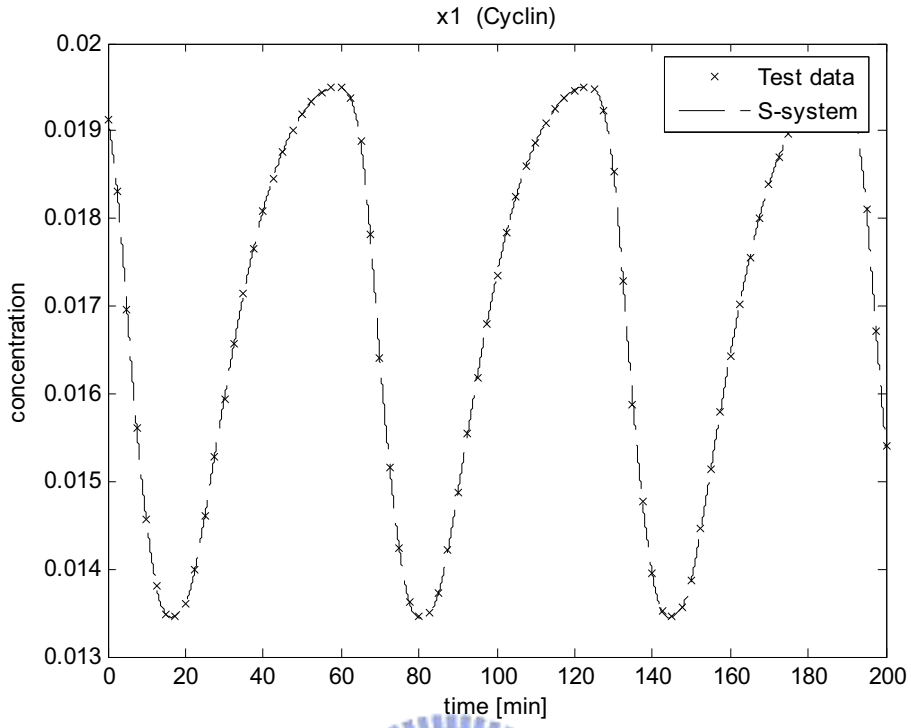


Figure 4.6 Cyclin evolution.

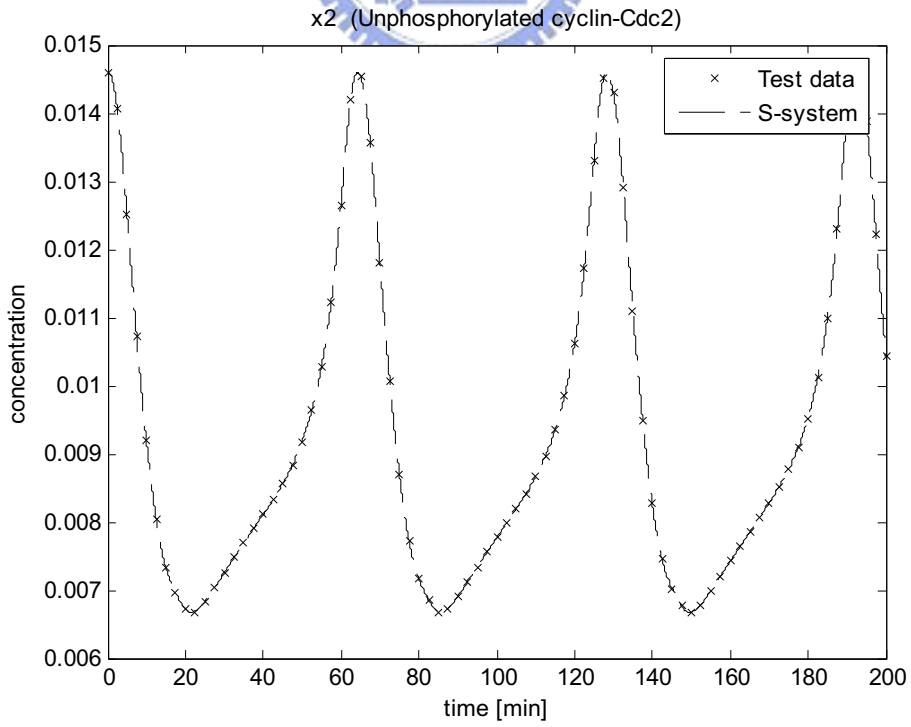


Figure 4.7 Unphosphorylated cyclin-Cdc2 evolution.

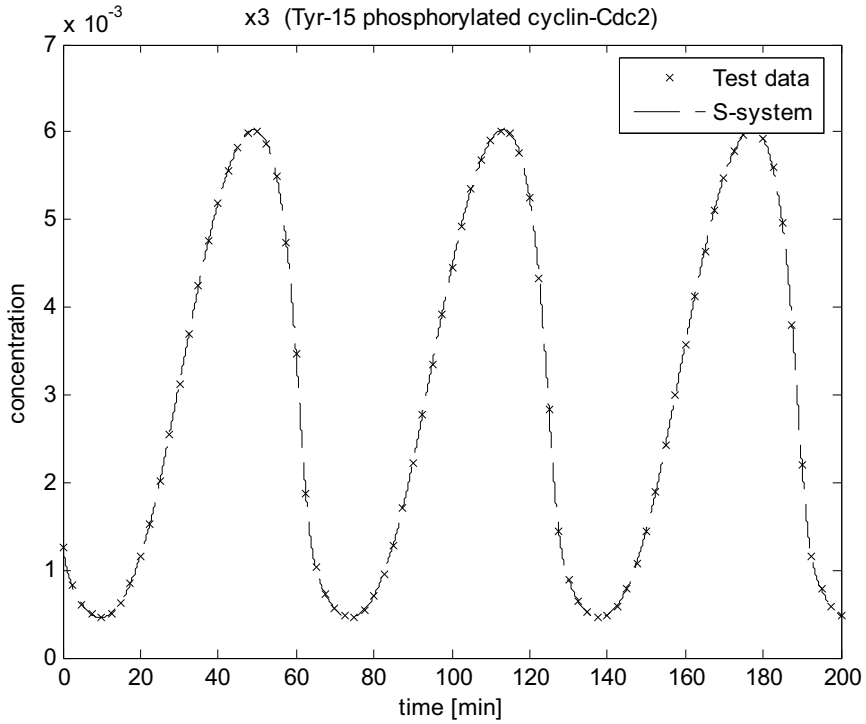


Figure 4.8 Tyr-15 phosphorylated cyclin-Cdc2 evolution.

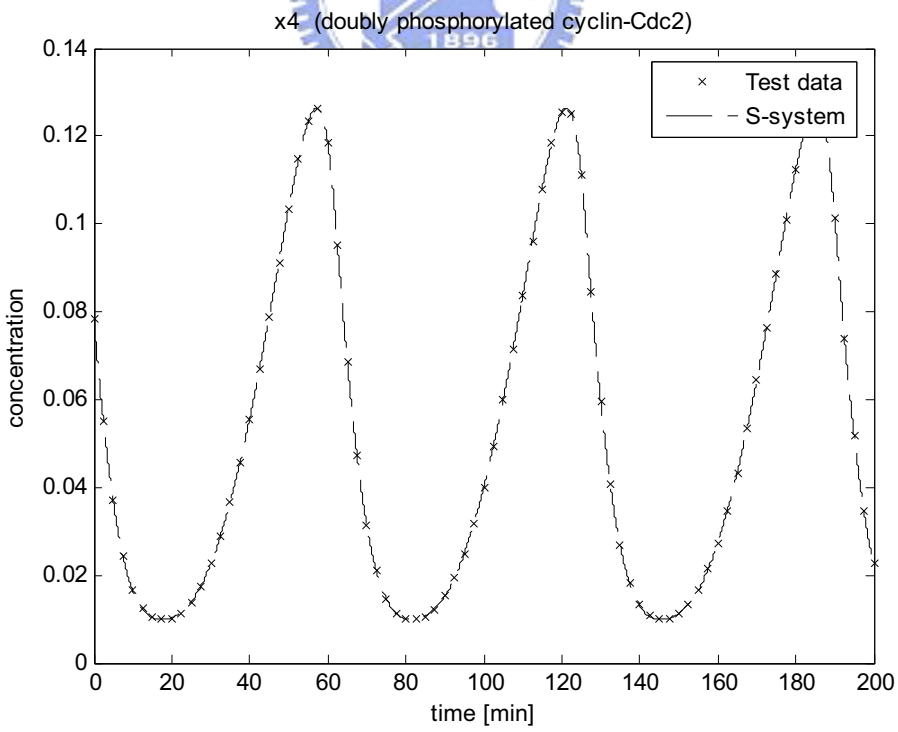


Figure 4.9 Doubly phosphorylated cyclin-Cdc2 evolution.

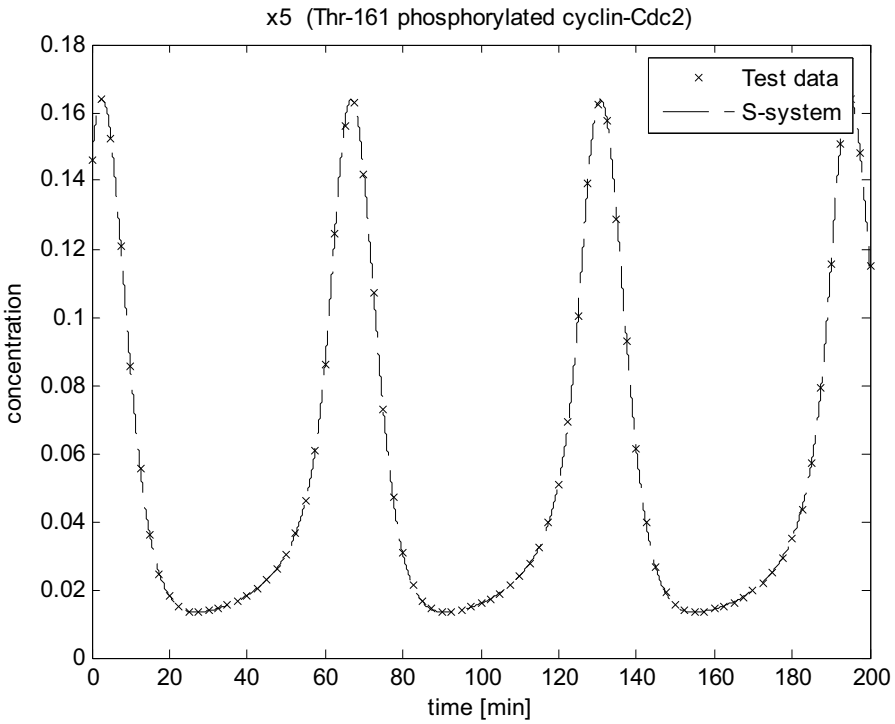


Figure 4.10 Thr-161 phosphorylated cyclin-Cdc2 evolution.

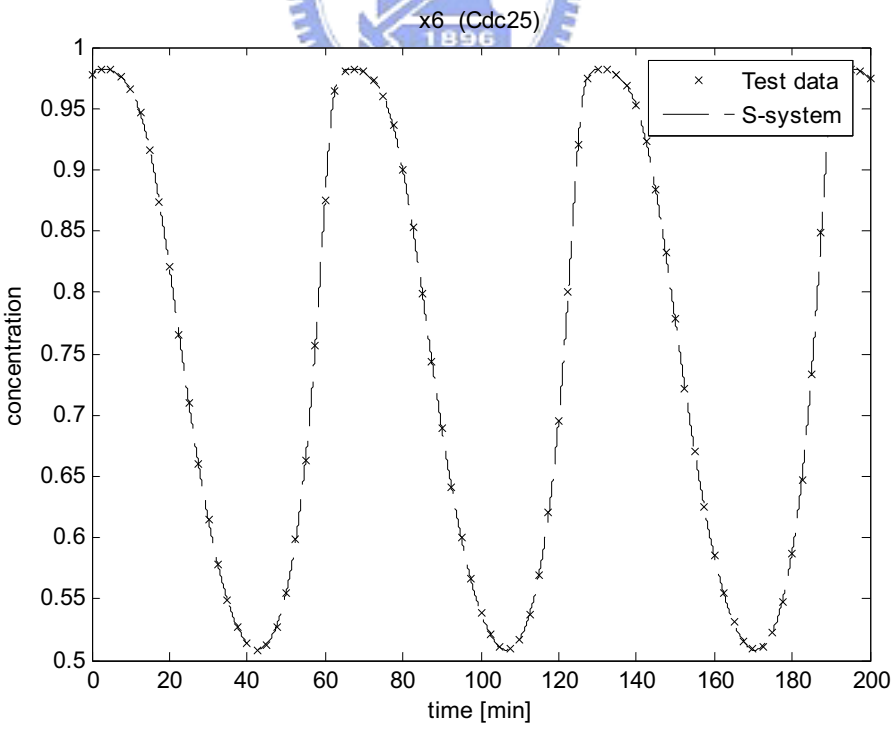


Figure 4.11 Cdc25 evolution.

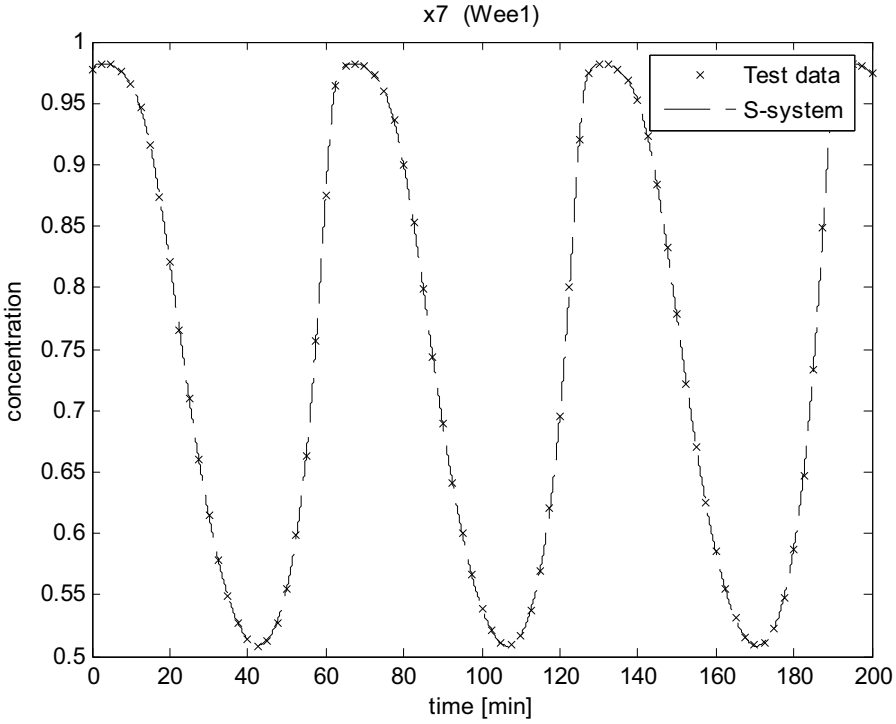


Figure 4.12 Wee1 evolution.

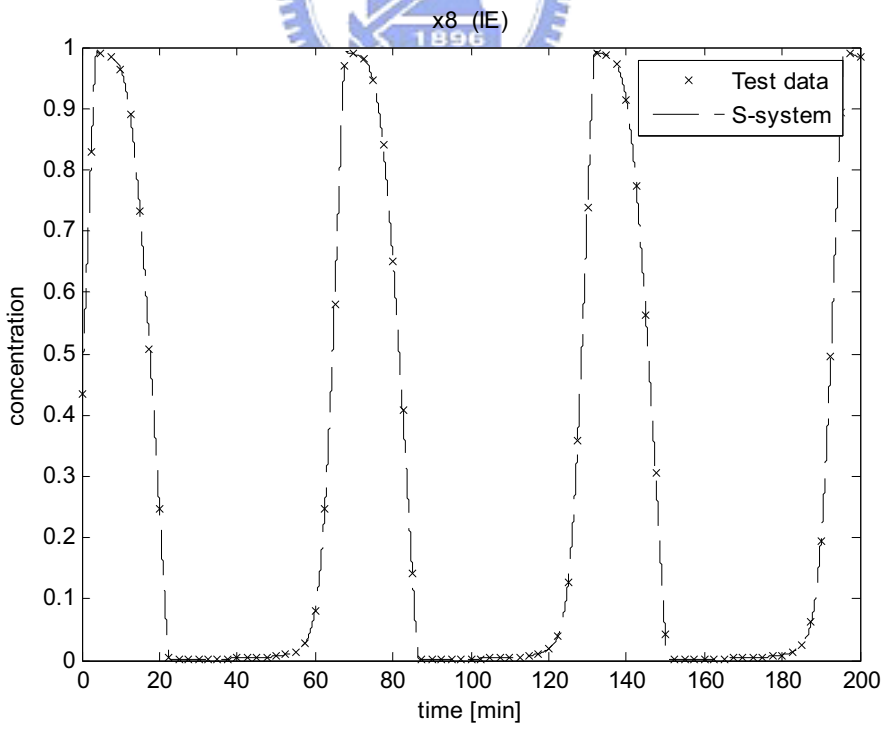


Figure 4.13 IE evolution.

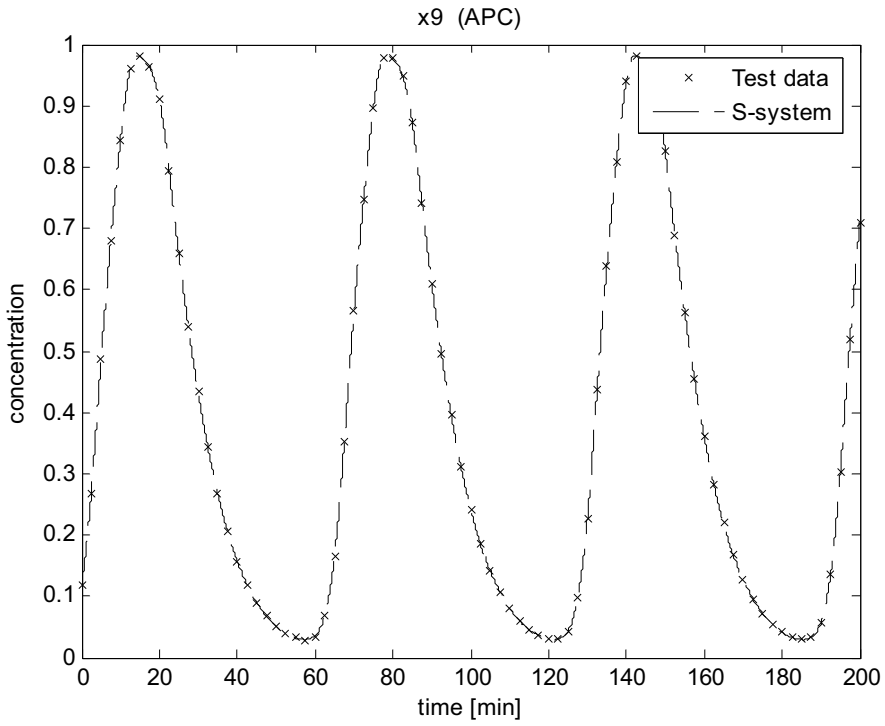


Figure 4.14 APC evolution.

Chapter 5

Conclusion

Improved genetic algorithm technique, which involves IEDO, migration operation and Elitism, is adopted to construct mathematical models for yeast and Xenopus frog egg. The gene time-course data form experiment of yeast cell cycle is used to construct gene regulatory network in modified power-law model. The training and test dataset are generated from Michaelis-Menten metabolic model of mitotic cell-cycle control of Xenopus frog egg in S-system. The two proposed gene regulatory networks reveal activatory and inhibitory operations for gene/enzyme synthesis and decomposition. Hence, the networks can provide biological researchers for further experiments in yeast and Xenopus frog egg cell cycle.

References

[1] Akutsu, T., Miyano, S. and Kuhara, S., "Identification of genetic networks from a small number of gene expression patterns under the Boolean network model", *Proceedings of the Pacific Symposium Biocomputation*, Vol. 4, p.p. 17–28, 1999.

[2] De Jong, H., "Modeling and simulation of genetic regulatory systems: a literature review", *J. Comput. Biol*, Vol. 9, p.p. 67–103, 2002.

[3] Pe'er, D., Regev, A., Elidan, G. and Fridman, N., "Inferring subnetworks from expression profiles", *Bioinformatics*, Vol. 17, p.p. 215–224, 2001.

[4] Gonick, L. and Wheelis, M., *The Cartoon Guide to Genetics*, Collins, 1991.

[5] Savageau, M.A., *Biochemical System Analysis: A Study of Function and Design in Molecular Biology*, Addison-Wesley, 1976.

[6] Voit, E.O., *Computational Analysis of Biochemical System*, Cambridge University Press, 2000.

[7] Chen, H.C., Lee, H.C., Lin, T.Y., Li, W.H. and Chen, B.S., "Quantitative characterization of the transcriptional regulatory network in the yeast cell cycle", *Bioinformatics*, Vol.20, pp. 1914-1927, 2004.

[8] Singer, R. and Penman, S., "Messenger RNA in HeLa cells: kinetics of formation and decay", *J. Mol. Biol.*, Vol.78, pp.321-328, 1973.

[9] Tsai, K.Y. and Wang, F.S., "Evolutionary Optimization with Data Collocation for Reverse Engineering of Biological Networks" , *Bioinformatics*, Vol. 21, p.p. 1180–1188, 2005.

[10] Chiou, J. P. and Wang, F. S., "A hybrid method of differential evolutionary with application to optimal control problems of a bioprocess system", *Proceedings 1998 IEEE on Evolutionary Computation Conference*, p.p. 627–632, 1998.

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

[12] 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.

[13] Chiang, C.L., Su, C.Y. and Wang, F.S., "Augmented Lagrangian Method for Evolutionary Optimization of Mixed-Integer Nonlinear Constrained Problem", *International Mathematical Journal*, Vol. 2, No. 2, p.p. 119–154, 2002.

[14] Spellman, P.T., Sherlock, G., Zhang, M.Q., Iyer, V.R., Anders, K., Eisen, M.B., Brown, P.O., Botstein, D. and Futcher, B., "Comprehensive identification of cell cycle regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization", *Mol. Bio. Cell*, Vol. 9, pp. 3273-3297, 1998.

[15] Schmidt, H. and Jacobsen E.W., "Linear systems approach to analysis of complex dynamic behaviors in biochemical networks", *IEE Syst. Biol.*, Vol. 1, No. 1, p.p. 149–158, 2004.

[16] Marlovits, G., Tyson, C.J., Novak, B. and Tyson, J.J., "Modeling M-phase control in Xenopus oocyte extracts: the surveillance mechanism for unreplicated DNA", *Biophys. Chem.*, Vol. 72, p.p. 169–184, 1998.

Vita

姓　　名：周家賢

出生日期：67年10月20日

籍　　貫：高雄市

學　　歷：高雄市立信義國小 74~80

高雄市立五福國中 80~83

國立高雄工商專校 電機工程科 83~88

國立台灣科技大學 電機工程系 88~90

國立交通大學 電機與控制工程研究所 93~95