A Parallel-and-Interactive Scheme for the Parameter Estimation of Michaelis-Menten Biological Systems

Shinq-Jen Wu and Anh-Tuan Ngo
Department of Electrical Engineering, Da-Yeh University, Chang-Hwa, Taiwan, ROC
Email: jen@mail.dyu.edu.tw, arcngoanhtuan@gmail.com

***Abstract*—The well-known Michaelis-Menten (MM) models give locally rich-kinetic information of proteins or metabolites. However, its parameter estimation requires increasingly large amount of experimental data and repeated modifications. In this study, we proposed a parallel-and-interactive optimization scheme for computationally MM modeling. The scheme integrates the most stochastic physiology evolution (genetic algorithm) and the less-stochastic swarm intelligence (particle swarm optimization) to ensure a flexible search. The scheme was tested with artificially time series data. Simulation results show the proposed scheme possesses good ability in global search even searching in a rather wide space (a range between 0 and 50000).**

***Index Terms*—computational intelligence, data mining, computational biology**

I. Introduction

MM systems which are described as a group of nonlinear ordinary differential equations shown in Eq. (1) are effective models for characterizing molecular biological systems and analyzing the dynamics of underlying systems. The rate change of the involved constitutes (metabolites, genes and proteins) is described as the generation flux and the consumption flux [1]:

$$\dot{x}_i = \sum_j V_{ij}^{+}(x_1,\cdots x_n) - \sum_j V_{ij}^{-}(x_1,\cdots x_n), \quad i = 1,2,...,n \tag{1}$$

where fluxes are the combination of power-law and Hill functions. It is a big challenge to perform reverse engineering from experimental time series data due to the nonlinearity and complexity. Further, MM systems possess sloppy features [2] which increase the challenge of computational approach for parameter estimation. Parameter estimation was converted into a mini-max problem through assessment error functions. Chou and Voit used dynamic flux estimation to obtain a systematic approach for the estimation of the parameters and function forms of MM systems [3].

GA, a biologically-inspired algorithm, evolves a population of individuals (chromosomes) which are encoded with bit strings or real numbers. Each individual expresses a candidate solution of optimization problems and evolves toward a better solution. Evolution starts from a population with randomly generated individuals. Individuals are stochastically selected from the population (based on their fitness) and modified through crossover and mutation to form a new population. GA is usually combined with local search or swarm techniques for convergent improvement. Nik et al. tested various in-series combinations and integrations of GA and PSO to achieve an optimal arrangement of payment inspection units in a massive network [4]. Rashidi and Sharifian recently proposed a hybrid of GA and ant colony optimization for task assignment in mobile cloud [5].

PSO, a stylized simulation for the movement of bird flock or fish school, gets optimal solutions through moving particles around a searching space according to a simple mathematical formula over particles' positions and velocity. Particles' movement is influenced by their locally best positions and is guided toward the global-best position. In this way, PSO is expected to move a swarm toward the best solution. However, prematurely convergence is the issue. Montes de Oca and coworkers proposed Frankenstein's dynamic topology to improve population diversity [6]. Taherkhani and Safabakhsh reviewed various inertial-weight adaptation strategies and proposed a stability-based adaptive inertia weight wherein the weight is determined in different dimensions for each particle [7]. Meng and coworkers took a review for various hybrids of PSO and genetic operations (selection, crossover or mutation) [8]. They proposed a crisscross search to let PSO avoid from the entrapment of local optimums of multimodal optimization problems, wherein horizontal crossover performed in the same dimension and vertical crossover operated in different dimensions. Lee and coworkers divided individuals into two groups ($p\%$ for Part 1 and $(1-p)\%$ for Part 2) according to their fitness [9]. The individuals in Part 2 were replaced by random individuals ($r\%$) and the offspring of the individuals in Part 1 through crossover and mutation ($(1-r)\%$).

Instead of combing GA and PSO in series or through integration, or enhancing the globally searching ability of PSO through topology variation or inertial weight modification, we let GA and PSO perform in parallel and communicate with each other from generation s to generations.

Manuscript received January 1, 2017; revised June 27, 2017.

doi: 10.18178/lnit.5.1.39-43

II. Parallel-and-Interactive Schemes for Michaelis-Menten Modeling

GA searches the global optimal point by evolutionary operators, while PSO's particles find the better result by moving in a searching space. GA's evolutionary process always takes more time (generations) to get improvement. In the case of the same generations, GA's role is much less than PSO's role in searching better solutions. To solve this issue, we need to strengthen GA's performance to a comparatively searching ability as PSO and then take an interactive commutation. Therefore, we let GA runs twice or more times for a run of PSO, a parallel process called nxGA-PA algorithm. Fig. 1 is the proposed parallel-and-interactive scheme which integrates GA and PSO in a parallel way and share (exchange) the best individual to achieve a flexible search. The GA runs n times when PSO finishes a run such that a competitive search for GA and PSO is achieved.

The heuristic crossover and non-uniform mutation are used in this study [10], [11]. The offspring for a parents $x^{(1)}=\left(x_1^{(1)},x_2^{(1)},...,x_n^{(1)}\right)$ and $x^{(2)}=\left(x_1^{(2)},x_2^{(2)},...,x_n^{(2)}\right)$ is $y=(y_1,y_2,...,y_n)$ with $y_i=u(x_i^{(2)}-x_i^{(1)})+x_i^{(2)}$ where u is a uniformly distributed random number in the interval [0,1], and the parent $x^{(2)}$ has fitness value which is not worse than that of the parent $x^{(1)}$. If the generated offspring lies outside a feasible region a new random number u is generated to generate another offspring. This process is repeated up to k times as required. After k times attempts, if the crossover operation fails to generate a feasible offspring then a random point in the feasible region is used to replace the infeasible offspring [12].

Michalewicz's non-uniform mutation is one of the most widely used mutation operators for real-coded GA [11]. For a point $x^t=\left(x_1^t,x_2^t,...,x_n^t\right)$, the respective mutation point is

$$x_i^{t+1}=\begin{cases}x_i^t+\Delta\left(t,x_i^u-x_i^t\right), & \text{if } r\le 0.5,\\ x_i^t-\Delta\left(t,x_i^t-x_i^l\right), & \text{otherwise.}\end{cases} \tag{2}$$

The function $\Delta(t,y)$ takes value in the interval:

$$\Delta(t,y)=y\left(1-u^{\left(1-\frac{1}{T}\right)^b}\right) \tag{3}$$

where r, u are random values between 0 and 1, x_i^l, x_i^u are the lower and upper bounds of the ith component, T denotes the maximum number of generations and b determine the strength of the mutation operator. The mutation uniformly searches over the entire space in the early generation and becomes locally searching the space (closer to its descendants) in the later generation [11], [12].

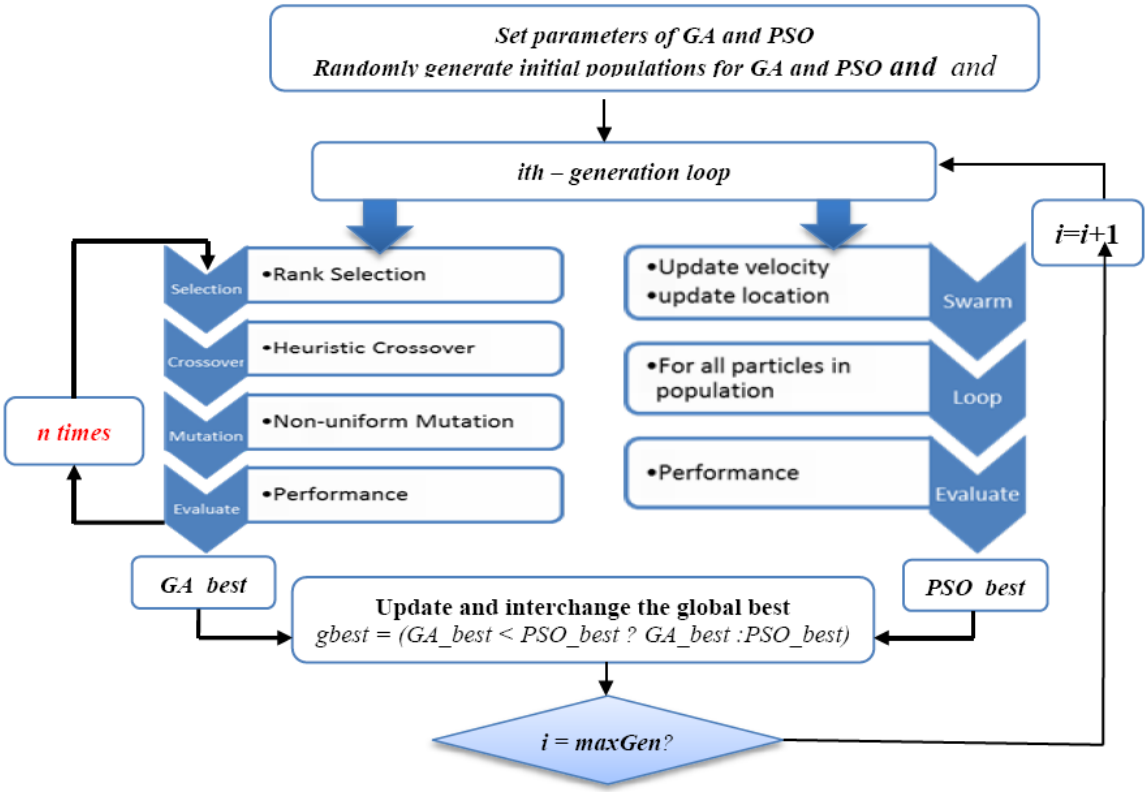


Figure 1. The scheme of parallel-and-interactive nGA-PSO algorithm.

Each particle in PSO occupies a point in a D-dimensional space (a potential solution vector). When particles fly in a space there are three forces upon them. The first one is the inertia to let particles maintain their current directions and velocities. The second force pushes each particle towards its best position in the past (personal best, pbest). The third one forces all particles

toward the best-so-far position of the swarm (global best, gbest) [13].

A. Michaelis-Menten-Based Biological Systems

Three different MM-based biological systems were used to test the proposed algorithm. These systems vary in structures, in the number of interactions and in the number of kinetic parameters. The reversible pathway [14] possesses two dependent variables and eight kinetic parameters. The metabolic pathway with branch points [15] has four dependent variables and eight parameters. The four-state kinetic system [1] has four dynamic variables and nine kinetic parameters.

The reversible pathway

The system shown in Fig. 2 is a reversible bio-reaction pathway [14].

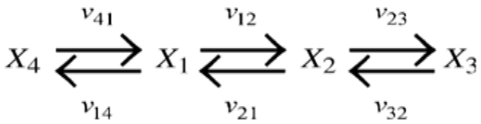


Figure 2. A reversible pathway [14].

This system has two independent variables (X_3 and X_4) and two dependent variables (X_1 and X_2). There are eighteen kinetic parameters to be estimated. The governing equations are described as

$$\dot{X}_1 = v_{41} + v_{21} - (v_{12} + v_{14}),$$

$$\dot{X}_2 = v_{12} + v_{32} - (v_{21} + v_{23}),$$

$$v_{41} = \frac{V_{41}^{\max} X_4}{1 + X_4 / K_{41} + X_1 / K_{-41}},$$

$$v_{12} = \frac{V_{12}^{\max} X_1}{1 + X_1 / K_{12} + X_2 / K_{-12}},$$

$$v_{23} = \frac{V_{23}^{\max} X_2}{1 + X_2 / K_{23} + X_3 / K_{-23}},$$

$$v_{14} = \frac{V_{14}^{\max} X_1}{1 + X_4 / K_{14} + X_1 / K_{-14}},$$

$$v_{21} = \frac{V_{21}^{\max} X_2}{1 + X_1 / K_{21} + X_2 / K_{-21}},$$

$$v_{32} = \frac{V_{32}^{\max} X_3}{1 + X_2 / K_{32} + X_3 / K_{-32}}. \quad (4)$$

The algorithms were implemented in Matlab environment. A parallel combination 2GA-PSO with a maximum of 50,000 iterations is used to get the parameters. The searching range is set from 0 to 10,000. Fig. 3 shows the simulation results wherein the solid curves denote the true system and the estimated profiles are shown by the dotted and circle points. The initial condition for Case 1 (the upper figure) is $X_1 = 4, X_2 = 3$ in an experimental environment $X_3 = 9, X_4 = 5$. The initial condition for Case 2 (the down figure) is $X_1 = 14, X_2 = 10$ in the environment $X_3 = 2, X_4 = 4.8$. The sampling time is set at 0.05 seconds for Case 1 and 0.1 seconds for Case 2. Simulation results show that the proposed scheme is able to get a nearly perfect prediction of the dynamic behavior for both cases.

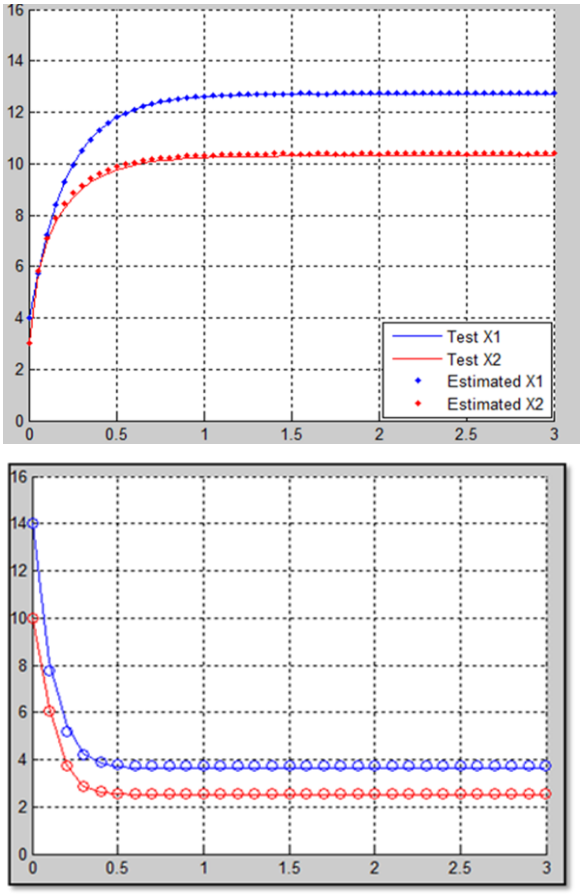


Figure 3. Simulation results for the reversible pathway tested with different initial conditions and experimental environments.

The metabolic pathway with branch points

Fig. 4 is a glyoxalase pathway with branch pints. By using X_1, X_2, X_3 and X_4 to, respectively, denote the concentration of hemithioacetal (HTA), methylglyoxal (MG), S-D-lactoylglutathione (SDLTHS), trypanothione (T(SH)2), D-glyceralde-hyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP), we have the following MM model:

$$\dot{X}_1 = v_3 - v_4 - v_5,$$

$$\dot{X}_2 = v_2 + v_1 + v_4 - v_3,$$

$$\dot{X}_3 = v_5 - v_6,$$

$$\dot{X}_4 = v_6 + v_4 - v_3, \quad (5)$$

wherein X_5 and X_6 are independent variables and the rate equations $v_1 = K_1 X_5$, $v_2 = K_2 X_6$, $v_3 = K_3 X_2 X_4$, $v_4 = K_4 X_1$, $v_5 = \frac{V_5 X_1}{K_{m5} + X_1}$, $v_6 = \frac{V_6 X_3}{K_{m6} + X_3}$.

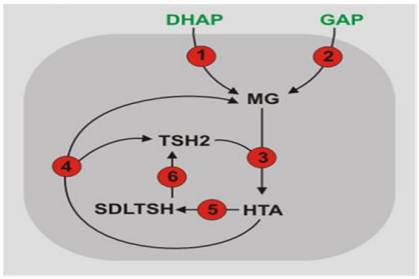


Figure 4. metabolic pathway with branch points [15].
(Source: http://jjj.biochem.sun.ac.za/database/silva/index.html)

A 3GA-PSO parallel scheme is used in the parameter estimations of the metabolic system. There are eight parameters to be estimated. To examine the searching ability we tested the scheme with two searching ranges: a small range of [0, 10] (Case 1) and a large range of [0, 10000] (Case 2). The same sampling time (0.04 seconds) and initial conditions ([0.14, 0, 0, 0.9]) with a fixed environmental condition (X_5=0.072 and X_6 =0.16) were used in both cases. Fig. 5 shows the simulation results. A perfect match for the prediction of the proposed method and the true dynamic behavior is observed.

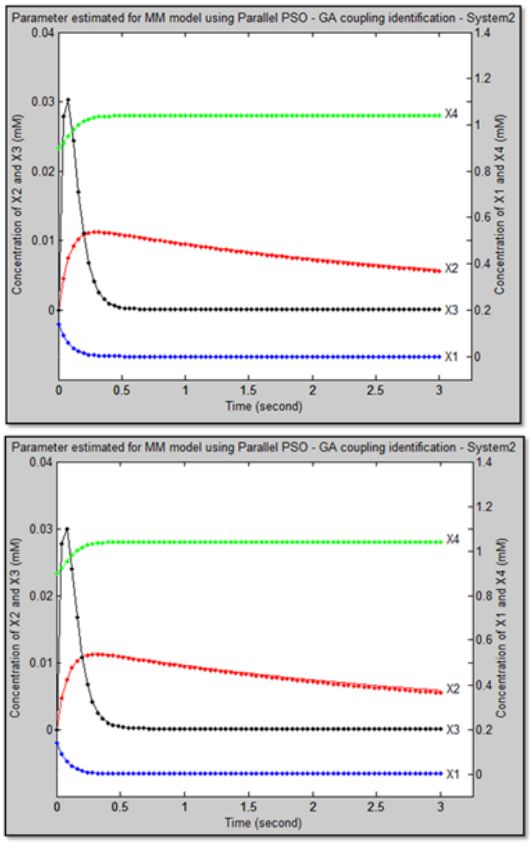


Figure 5. Simulation results for the metabolic pathway tested in a small searching space [0, 10] (Case 1, the upper figure) and a large range of [0, 10000] (Case 2, the lower figure).

Four-state metabolic pathways

A four state pathway in Fig. 6 describes an enzyme molecule E that catalyzes the reaction from the substrate S to the product P. It has two electrically distinct conformational states: E with the binding site for the substrate S exposed (denoted by SE), and E* with the binding site for the product P exposed (denoted by E*P).

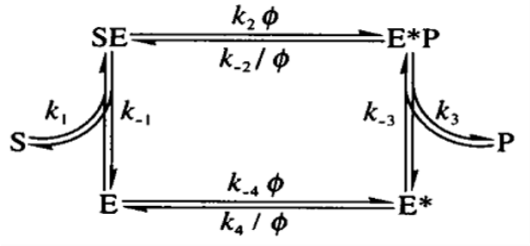


Figure 6. four-state metabolic pathways. [1]

The rate equations are shown as follows.

$$dE/dt = (k_{-1})SE + (k_4/\phi)E^* - (k_{-4}\phi + k_1 S)E,$$
$$dSE/dt = (k_1 S)E + (k_{-2}/\phi)E^* P - (k_2\phi + k_{-1})SE,$$
$$dE^*P/dt = (k_2\phi)SE + (k_{-3}P)E^* - (k_{-2}/\phi + k_3)E^* P,$$
$$dE^*/dt = (k_{-4}\phi)E + (k_3)E^* P - (k_4/\phi + k_{-3}P)E^* \quad (6)$$

We use a parallel combination 3GA-PSO to estimate the parameter of the four-state kinetic systems. We further use a very large range ([0, 50000]) to show that the global searching ability of the proposed scheme is very strong. Fig. 7 is the simulation results for searching in a range of [0, 10] (the upper figure) and [0, 50000] (the down figure). A nearly perfect prediction is achieved in both cases.

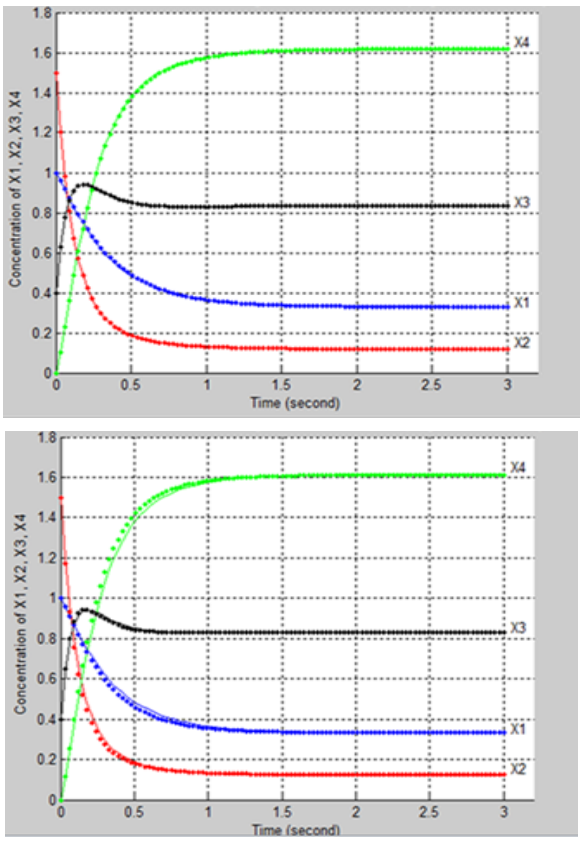


Figure 7. Simulation results for the four-state metabolic pathway tested in a small searching space [0, 10] (Case 1, the upper figure) and a very large range of [0, 50000] (Case 2, the lower figure).

III. Conclusion

GA always serve in good condition for global best due to stochastic genetic operators, while PSO can deal with the locally optimal searching due to the characteristic of foraging movement. The proposed parallel and interactive scheme let GA (evolutionary algorithms) and PSO (swarm optimizations) maintain their own searching characteristics and achieve a more effective search through information communication such that the defects of GA in slow convergence and PSO in premature convergence are compensated. The nGA-PSO scheme considers the situation that GA takes more time to achieve a comparable improvement than PSO. Simulation results for these three systems shows that the larger the dimension of system is the larger the value of n is (n=2 for two states and n=3 for four states). The technical contribution of this study is that through in-depth understanding of the essence of evolutionary and swarm algorithms we can use a simple scheme to achieve a fantastic compensation.

Acknowledgment

This research was supported by grant number MOST 105-2221-E-212-013 from the Ministry of Science and Technology of Taiwan, R.O.C.

References

[1] B. Robertson and R. D. Astumian, "Michaelis– Menten equation for an enzyme in an oscillating electric field," *Biophysical Journal*, vol. 58, pp. 969-974, 1990.

[2] R. N. Gutenkunst, J. J. Waterfall, F. P. Casey, K. S. Brown, C. R. Myers, and J. P. Sethna. (October 2007). Universally sloppy parameter sensitivities in systems biology models. *PLoS Computational Biology*. [Online]. 3(10): e189. Available: http://dx.doi.org/10.1371/journal.pcbi.0030189

[3] I. C. Chou and E. O. Voit. (July 2012). Estimation of dynamic flux profiles from metabolic time series data. *BMC Syst. Biol*. [Online]. 6(84). Available: http://bmcsystbiol.biomedcentral.com/articles/10.1186/1752-0509-6-84

[4] A. A. Nik, F. M. Nejad, and H. ZAkeri, "Hybrid PSO and GA approach for optimizing surveyed asphalt pavement inspection units in massive network," *Automation in Construction*, vol. 71, pp. 325-345, 2016.

[5] S. Rashidi and S. Sharifian, "A hybrid heuristic queue based algorithm for task assignment in mobile cloud," *Future Generation Computer Systems*, vol. 68, pp. 331-345, 2017.

[6] M. A. Montes de Oca, T. Stützle, M. Birattari, and M. Dorigo, "Frankenstein's PSO: A composite particle swarm optimization algorithm," *IEEE Trans. on Neural Networks*, vol. 13, no. 5, pp. 1120-1132, 2009.

[7] M Taherkhani and R. Safabakhsh, "A novel stability-based adaptive inertia weight for particle swarm optimization," *Applied Soft Computing*, vol. 38, pp. 281-295, 2016.

[8] A. Meng, Z. Li, H. Yin, S. Chen, and Z. Guo, "Accelerating particle swarm optimization using crisscross search," *Information Sciences*, vol. 329, pp. 52-72, 2016.

[9] W. P. Lee, Y. T. Hsiao, and W. C. Hwang. (January 2014). Designing a parallel evolutionary algorithm for inferring gene networks on the cloud computing environment. *BMC Systems Biology*. [Online]. 8(5). pp. 876-880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar

[10] A. H. Wright, "Genetic algorithms for real parameter optimization," in *Foundations of Genetic Algorithms I*, G. J. E. Rawlins, Ed., San Mateo, CA: Morgan Kaufmann Press, 1991, pp. 205–218.

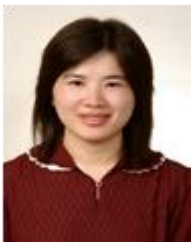
[11] K. Deep and M. Thakur, "A new mutation operator for real coded genetic algorithms," *Applied Mathematics and Computation*, vol. 193, pp. 211-230, 2007.

[12] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, New York: Springer-Verlag, 1992.

[13] J. Kennedy, "The particle swarm: Social adaptation of knowledge," presented at the IEEE International Conference on Evolutionary Computation, April 1997.

[14] E. O. Voit, *Computational Analysis of Biochemical Systems*, Cambridge, U.K.: Cambridge University Press, 2000.

[15] M. Silva, Ferreira, E. N. Antonio, A. M. Tomas, C. Cordeiro, and A. P. Freire, "Quantitative assessment of the glyoxalase pathway in Leishmania infantum as a therapeutic target by modeling and computer and simulation," *The FEBS J.*, 2005.



Shinq-Jen Wu received the B.S. degree in chemical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1986; the M.S. degree in chemical engineering from the National Tsing-Hua University, Hsinchu, Taiwan, R.O.C., in 1989; the M.S. degree in electrical engineering from the University of California, Los Angeles, CA, in 1994; and the Ph.D. degree in electrical and control engineering from the National Chiao-Tung University, Hsinchu, Taiwan, R.O.C., in 2000.

From September 1989 to July 1990, she was with the Laboratory for Simulation and Control Technology of the Chemical Engineering Division of Industrial Technology Research Institute, Hsinchu, Taiwan, R.O.C. Since then until 1991, she joined the Chemical Engineering Department, Kao-Yuan Junior College of Technology and Commerce at Kaohsiung, Taiwan, R.O.C. From 1995 to 1996, she was an engineer at the Integration Engineering Department, Macronix International Co., LTD, Hsinchu, Taiwan, R.O.C. She now join the Electrical Engineering Department, Da-Yeh University, Chang-Hwa, Taiwan, R.O.C.

She is a member of PHI-TAU-PHI Scholastic Honor Society. Her name is included in Asian Admirable Achievers, in Asia/American Who's Who, in Asia/Pacifica Who's Who, and in Marquis Who's Who in Science and Engineering/in the World/in America/in Asia. She is a Life Fellow of the International Biographical Association. She is a Scientific Adviser to the IBC Director General. She got 21st Century Award for Achievement from IBC and The Albert Einstein Award of Excellent from ABI. She is the Editor-in-Chief of Anatomy & Physiology: Current Research, and the editor of journal Advances in Fuzzy Sets and Systems, International Journal of Advanced Robotic Systems, Global Journal of Cancer Therapy and Journal of Food Science and Nutrition Therapy.

Her research interests include ergonomics-based smart car, Petri-net modeling for cancer mechanism, bio-inspired intelligent techniques for the inference of biological systems, soft-sensor for on-line tuning, optimal GA-fuzzy-neural modeling, control and tracking, and optimal fuzzy estimation/tracking/regulation.

Anh-Tuan Ngo is a graduate in electrical engineering of Da-Yeh university, changhwa, Taiwan. His research is to develop a hybrid evolutionary algorithm for the parameter identification of Michaelis-Menten models.