Journal of Theoretical and Applied Information Technology
30th November 2015. Vol.81. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A HYBRID OF OPTIMIZATION METHOD FOR MULTI-OBJECTIVE CONSTRAINT OPTIMIZATION OF BIOCHEMICAL SYSTEM PRODUCTION

[1,2]MOHD ARFIAN ISMAIL, [2]SAFAAI DERIS, [2]MOHD SABERI MOHAMAD, [2]MOHD ADHAM ISA, [2]AFNIZANFAIZAL ABDULLAH, [2]MUHAMMAD AKMAL REMLI, [2,3]SHAYMA MUSTAFA MOHI-ALDEEN

[1]Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang, Lebuh Raya Tun Razak, Gambang, Kuantan, Pahang, Malaysia.
[2]Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia.
[3]College of Computer Science and Mathematics, University of Mosul, 41002, Mosul, Iraq.
E-mail: marfian2@live.utm.my, safaai@utm.my, saberi@utm.my, mohdadham@utm.my, afnizanfaizal@utm.my, akmalmuhd@gmail.com, shaima_mustafa76@yahoo.com

ABSTRACT

In this paper, an advance method for multi-objective constraint optimization method of biochemical system production was proposed and discussed in detail. The proposed method combines Newton method, Strength Pareto Evolutionary Algorithm (SPEA) and Cooperative Co-evolutionary Algorithm (CCA). The main objective of the proposed method was to improve the desired production and at the same time to reduce the total of component concentrations involved in producing the best result. The proposed method starts with Newton method by treating the biochemical system as a non-linear equations system. Then, Genetic Algorithm (GA) in SPEA and CCA were used to represent the variables in non-linear equations system into multiple sub-chromosomes. The used of GA was to improve the desired production while CCA to reduce the total of component concentrations involved. The effectiveness of the proposed method was evaluated using two benchmark biochemical systems and the experimental results showed that the proposed method was able to generate the highest results compare to other existing works.

Keywords: *Newton Method, Strength Pareto Evolutionary Algorithm, Genetic Algorithm, Cooperative Co-evolutionary Algorithm, Biochemical System*

1. INTRODUCTION

In solving a non-linear equations system, the process of finding all the solutions for all equations in the system is required. This process is a hard task due to the nature of the equation which usually in the form of nondeterministic polynomial equation. Moreover, the complexity structure of the equations in the system that contain with many number of equations and variable, which contributed to the difficulty in solving the system [1] . Nowadays, many applications domain use non-linear equations system in real-life problem such as in chemistry domain, for example the optimization of the desired production in a biochemical system. This can be achieved by represent the biochemical system as a non-linear equations system. This is due to the knowledge of biotechnology which enable the biochemical system can be represented in mathematical model that known as ordinary differential equation (ODE).

The optimization of biochemical system production can be considered as biotechnological process which involves the fine-tuning process in the interest to improve the desired production. Besides the production, the total of component (chemical) concentrations involved also need to be considered, thus it makes two objectives need to be considered in the same time. This lead to the multi-objective problem. In addition, some constraints occurred in fine-tuning process in order to ensure a continuous optimal operation [2] . The constraints are steady state condition constraint and component concentration constraint. Therefore, the optimization process lead to the multi-objective

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

constraint optimization of biochemical system production.

The multi-objective constraint optimization of biochemical system production is about the optimization process that considers multiple objectives and constrained by several constraints. In multi-objective problem, there is no single optimum solution, but there is a set of alternative solutions. These alternative solutions are considered as optimal solution when all objectives are being considered. In general, these alternatives solution is known as Pareto optimal solution.

Recently, many works have been published in the optimization of the biochemical system production. Most of the works tends to use mathematical programming approach due to its flexibility. Example of works that use this approach including linear programming [3], geometric programming [4] and mixed inter non-linear programming [5]. The mathematical programming approach can be defined as a mathematical representation model that aimed of programing the best possible allocation of limited resources [6]. In general, this approach used mathematical model and routines in it operations [3]. Using this approach has some limitations. Usually, this approach totally depend on its initial solution which can cause the convergence problem and some error might occurred if the initial solution is not defined correctly [7], [8]. In addition, this approach requires expert knowledge in defining the decision variables and constraint, and understanding of the model and programming, which can lead to the unreliable result if the knowledge is no accurate and incomplete [8].

In order to overcome the limitations in mathematical programming approach in optimization of biochemical system production, this paper used iterative approach. The iterative approach is a procedure that generates a sequence of solution for a better solution of optimization problem. This approach use stochastic operator on a pool of candidate solution. This approach offer several advantages compare to mathematical programming approach. The iterative approach does not depends on expert knowledge because this approach uses stochastic operator, which the search direction is determine by random element [7]. Besides that, the iterative is more efficient and robust that mathematical programming approach [9].

In optimization process, the biochemical system is treated as a non-linear equations system. This is due the ODE model that represent the biochemical system where the structure biochemical system contain many equations which form a non-linear equations system [1], [10]–[12]. There are many methods in iterative approach that can be applied in solving a non-linear equations system such as numerical method, evolutionary algorithm method and swarm intelligence method. Among all methods, it has been found that Newton method, which is a numerical method, is the most popular method that is frequently used in solving a non-linear equations system. [1], [13]–[17]. In addition, Newton method is very simple and easy to apply in solving a non-linear equations system [18] and has faster convergence speed [19], [20]. Due to that, this paper used Newton method in solving the non-linear equations system.

Newton method is an iterative method that used to find an optimum point to real-value roots. Using Newton method for optimization process of biochemical system production is a good choice. This is because the biochemical system can be viewed as a non-linear equations system and Newton method can be utilized in solving the system. But, applying Newton method alone is not sufficient because Newton method only deal with biochemical system. Therefore, something is needed to deal with another part, which is multi-objective problem. There are many methods in solving multi-objective problems and the famous method is by using evolutionary algorithm (EA), especially GA [21], [22]. Within EA, many methods have been proposed such as Non-dominated Sorting Genetic Algorithm, Strength Pareto Evolutionary Algorithm (SPEA), Multi-objective Genetic Algorithm, Fast Non-dominated Sorting Genetic Algorithm and Niched Pareto Genetic Algorithm. Among all these methods, this paper uses SPEA due to its performance compare to other methods [23]–[25].

Within SPEA, the GA is used to represents the components in biochemical system. This is intent to fine-tuning the component concentration in order to improve the desired production and minimize the total of component concentrations involved. However, several issues arise when dealing with large biochemical system that contain with many components. This can make the representation of the solution is encoded in different type of value, this become complex, takes time in evaluate the solution and effect the optimization performance. Hence, a method is needed in order to overcome these issues. Applying the Cooperative Co-evolutionary Algorithm (CCA) is a good choice

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

where CAA has an ability to dividing a single solution into multiple sub-solutions [26], [27].

In this paper, a hybrid method that combines Newton method, SPEA and CCA is proposed and discussed in detail. The aim of the proposed method is to improve the biochemical system production and simultaneously reduce the total of component concentrations involved. The proposed method starts with Newton method that treated the biochemical system as a non-linear equations system and solving it. Then, SPEA and CCA is used for multi-objective optimization process where GA in SPEA for improve the desired production while CCA for minimize the total of component concentrations involved.

The rest of this paper is structured as follows. In the following section, the problem formulation is presented where the modelling of the biochemical system section and the model formulation of multi-objective constraint optimization section are described. Then, the discussion of the proposed method and two case studies, which are the optimization of ethanol production in *Saccharomyces cerevisiae* (*S.cerevisiae*) pathway and the optimization *tryptophan* (*trp*) of biosynthesis in *Escherichia coli* (*E.coli*) are presented. Following that, the results and discussion section before a conclusion is made.

2. MODELLING OF BIOCHEMICAL SYSTEM

The biochemical system can be modelled using ODE. Within ODE, there are two representative types that are usually used, which are S-system and generalized mass action (GMA) model. This paper focused on GMA model because of it advantages in representing the nonlinearity of biochemical system and the performance of GMA model [5], [28]. The GMA model has the following form:

$$\frac{dX}{dt} = Sv(x) \tag{1}$$

where S is the stoichiometric matrix of the system and $v(x)$ denote the vector that contains the reaction rate. The $v(x)$ is represented by the power-law function and has the following form [29]:

$$v_i = \gamma_i \prod_j x_j^{f_{ij}} \tag{2}$$

where the coefficient γ_i is the rate constant for v_i and the coefficient f_{ij} is the kinetic order. These coefficients are derived from the Taylor series in the logarithmic space around a steady state and can be define as follows [29]:

$$\gamma_i = |v_i|_0 \tag{3}$$

$$f_{ij} = \left| \frac{\delta v_i}{\delta x_j} \frac{x_j}{v_j} \right| \tag{4}$$

3. MODEL FORMULATION OF MULTI-OBJECTIVE CONSTRAINT OPTIMIZATION

The multi-objective constraint optimization of biochemical system production involves the fine-tuning (optimization) process of variables of non-linear equations system in order to improve the desired production and at the same time reduce the total of component concentrations involved. In the optimization process, there are two constraints the must be followed, namely the steady state constraint and component concentration constraint.

The steady state constraint is a condition where all the GMA models are equal to zero. This is due to the all components are in static value [2]. Therefore, the GMA model (Equation 1) become as follows:

$$\frac{dX_n}{dt} = \left[sv(x)_1, \ldots sv(x)_n \right] = 0 \tag{5}$$

This situation leads in solving a non-linear equations system. For this reason, the optimization process of multi-objective constraint optimization of biochemical system production can be considered as solving a non-linear equations system.

The fine-tuning process of variables in non-linear equations system cannot be performed indiscriminately. This is because the variables have their own constraint in order to ensure the component concentration in biochemical system remains within specific range. The purpose of this constraint is to maintain the survival of the cell. As a result, the multi-objective constraint optimization of the biochemical system production can be formulated as follows:

$$max\, F_1(v) \tag{6}$$

$$min\, F_2\left(\sum_{j=1}^{n} x_j \right) \tag{7}$$

subjected to:

$$Sv(x)_i = 0, \qquad i = 1,2,3,\ldots,n \tag{8}$$

$$x_j^L \le x_j \le x_j^U \qquad j = 1,2,3\ldots,m \tag{9}$$

where Equation 6 is the biochemical system production, Equation 7 is the total of component concentrations involved, Equation 8 is the GMA model in steady state constraint and Equation 9 is the component concentration constraint.

Journal of Theoretical and Applied Information Technology
30th November 2015. Vol.81. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. THE PROPOSED METHOD

In this section, the proposed method is presented and discussed in detail. In this method, Newton method is used to deal with non-linear equations system, SPEA is used for multi-objective problem, where the GA in SPEA is used for optimization process, and CCA is used to divide the chromosome (the representation of the solution) into multiple sub-chromosomes (sub-solution) where all sub-chromosomes are in their own sub-population. The number of sub-chromosomes is same as the number of variables in non-linear equations system that needs to be fine-tuned. In this method, a concept that represents the solution is introduced and known as cooperative chromosome, where the cooperative chromosome is formed from multiple sub-chromosomes. Figure 1 shows the proposed method in pseudo code format while Figure 2 illustrates the flowchart form. Next is the detail description of the proposed method.

Step 1: Initialize the first generation. This step is about generating randomly the initial *m* sub-chromosomes in *m* sub-populations and creates an empty external population. The external population is used to keep the chromosome (complete solution) that satisfies the constraints. All sub-chromosomes are evolving in their own sub-population. Sub-chromosome is in binary format. This is can be reviewed in line 3 to 4 in Figure 1.

Step 2: Form the cooperative chromosome. In this step, the complete solution is formed by combining all sub-chromosomes from all sub-populations. The complete solution is known as cooperative chromosome. A sub-chromosome is selected based on the fitness value where the lowest fitness value is selected and the chosen sub-chromosome is known as representative. The fitness value refers to the value that is represented by sub-chromosome. This process is known as sub-chromosome evaluation. The objective of the sub-chromosome evaluation is to minimize the total of component concentrations involved value by allowing all representatives that have the lowest fitness value from every sub-population to combine with each other. This step can be found in line 5 to 12 and 17 to 24 in Figure 1. Line 5 to 12 is occurring when generation is 0 while line 17 to 24 when iteration process takes place. The formation of cooperative chromosome is depicted in Figure 3.

Input: Initial candidate solution that represents the variables in non-linear equations system.
Output: The best solution.

1. **Begin**
2. $g = 0$
3. $sp_m^g = \{sp_1^g, sp_2^g, sp_3^g, \mathrm{K}\ , sp_m^g\}$ // sp is sub-population
4. $ep = \{\}$ // ep is an empty external population
5. **For** sp_m^g
6. $sc_{mn}^g = \{sc_{m1}^g, sc_{m2}^g, sc_{m3}^g \mathrm{K}\ , sc_{mn}^g\}$ // sc is sub-chromosome
7. **For** sc_{mn}^g
8. *evaluate* sc_{mn}^g
9. **End For**
10. $rep\_sc_{mn}^g \leftarrow choose\ sc_{mn}^g$ // rep_sc is representative from sub-population
11. $coo\_c^g \leftarrow combine\ all\ rep\_sc$ // coo_c is cooperative chromosome
12. **End For**
13. *evaluate* coo_c^g
14. **While**
15. $g = g + 1$
16. **If** *not meet condition*
17. **For** sp_m^g
18. **For** sc_{mn}^{g-1}
19. $sc_{mn}^g \leftarrow reproduce\ from\ sc_{mn}^{g-1}$
20. *evaluate* sc_{mn}^g
21. **End For**
22. $rep\_sc_{mn}^g \leftarrow choose\ sc_{mn}^g$
23. $coo\_c^g \leftarrow produced\ by\ combine\ all\ rep\_sc_{mn}^g$
24. **End For**
25. *evaluate* coo_c^g
26. **End If**
27. **Else**
28. $ep \leftarrow copy\ coo\_c^g$
29. **End Else**
30. **End While**
31. **For** *ep*
32. *choose the best solution*
33. **End For**
34. *return the best solution*
35. **End**

Figure 1: The pseudo code Of The Proposed Method

Step 3: Evaluate the cooperative chromosome. This step is about the evaluation process of cooperative chromosome. Firstly, the cooperative chromosome is decoded into variables in non-linear equations system. After that, Newton method is used to solve the non-linear equations system. At this stage, two termination conditions come out; the maximum number of generations is reached and the component concentration constraints are satisfied. If these termination conditions are fulfilled, then the process moves forward to Step 6, otherwise proceed to the next step. This is can be viewed in line 13 and 25 in Figure 1. Line 13 is when generation is 0 while line 25 in the iteration process.

Step 4: Decompose the cooperative chromosome. After being tested by Newton method, the cooperative chromosome is transforming back into multiple sub-chromosomes form. Then all sub-chromosomes go back into their own sub-population for the evolution process.

Step 5: Reproduce new generation. In this step, the GA operators are used, which are selection, crossover, and mutation. By doing this, it is expected that a new generation with better quality than previous generation is produced. Then, the

Journal of Theoretical and Applied Information Technology
30th November 2015. Vol.81. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

new generation went back to Step 2. This step is happed in line 19 in Figure 1.

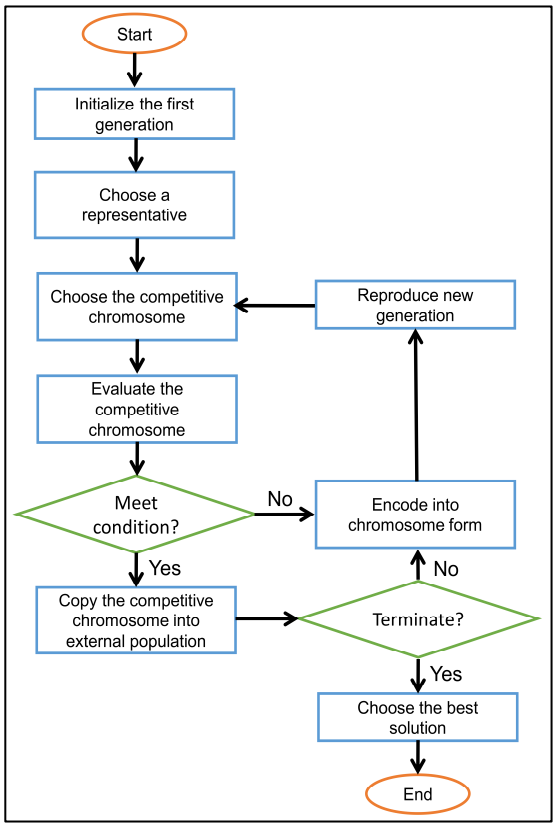

Figure 2: The Flowchart Of The Proposed Method

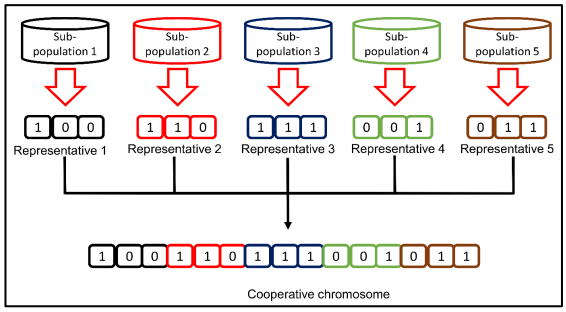

Figure 3: The Formation Of Cooperative Chromosome

Step 6: Copy the cooperative chromosome into the external population. Only cooperative chromosome that fulfills the component concentration constraint is copied into the external population. The cooperative chromosome in the external population is known as Pareto optimal. The reason for this copy process is to keep the potential best solution that found in every generation and to avoid it from being lost during the reproduction process (Step 5). Two conditions occur during the copy process; the maximum number of generations is reached and the maximum number of Pareto optimal is achieved. If all conditions are fulfilled, proceed to Step 7, or else, the process goes back to the Step 4. During the copying process, if the maximum number of Pareto optimal is achieved before the number of maximum number generation is reached, the Pareto optimal with lowest fitness value is deleted and replaced by a newly copied Pareto optimal. However, if the maximum number of generations is achieved before the maximum number of Pareto optimal is achieved, the process proceeds to the next step. This process can be found in line 28 in Figure 1.

Step 7: Choose the best solution. In this step, the better solution is chosen based on its fitness value where the Pareto optimal with highest fitness value is chosen as the best solution. The last step is found in line 31 to 33 in Figure 1. The fitness function for Pareto optimal and cooperative chromosome is given as follows:

$$best_c = \left| \frac{PF_1 - PF_2}{2} \right| \times 100 \qquad (10)$$

where PF_1 is the percentage improvement of F_1 from its steady state value while PF_2 is the percentage minimization of the F_2 from its steady state value.

5. MODEL AND EXPERIMENTAL DATA

Two biochemical systems were used in this paper, which were the fermentation pathway in *S.cerevisiae* and the *trp* biosynthesis in *E.coli*. A program based on Java was developed to test the proposed method with these biochemical systems. The program based on jMetel [30] and JAMA version 1.0.3 were used. These programs can be downloaded at http://jmetal.sourceforge.net/index.html and http://math.nist.gov/javanumerics/jama/. The following sub-sections are the description of the biochemical systems.

5.1 Optimization Of Ethanol Production In *Saccharomyces Cerevisiae* Pathway

In this pathway, the proposed method is used to improve the ethanol production. Figure 4 illustrates the schematic representation of *S.cerevisiae* pathway. In this pathway, components were divided into two groups, which were metabolites (X_1 - X_5) and enzymes (Y_1 - Y_6). Details of the metabolites and enzymes, including their initial steady state values are presented in Table 1. This pathway has the following GMA model:

Journal of Theoretical and Applied Information Technology
30th November 2015. Vol.81. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$\begin{aligned}
\frac{dX_1}{dt} &= V_{in} - V_{HK} \\
\frac{dX_2}{dt} &= V_{HK} - V_{PFK} - V_{Carb} \\
\frac{dX_3}{dt} &= V_{PFK} - V_{GAPD} - 0.5V_{Gro} \\
\frac{dX_4}{dt} &= 2V_{GAPD} - V_{PK} \\
\frac{dX_5}{dt} &= 2V_{GAPD} + V_{PK} - V_{HK} - V_{Carb} - V_{PFK} - V_{ATPase}
\end{aligned} \tag{11}$$

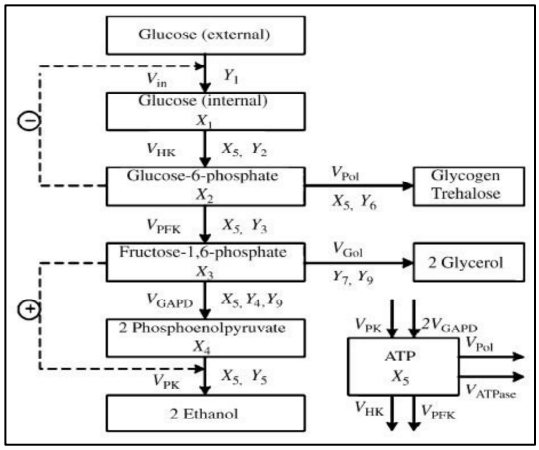

Figure 4: The Schematic Representation Of S.cerevisiae Pathway

Table 1: The Detail Of Metabolites And Enzymes In S.cerevisiae Pathway

At the initial steady state, all fluxes in the GMA model can be formulated as follows:

Metabolite	Acronym	Symbol	Initial Steady State
Glucose (internal)	Glc_{in}	X_1	0.0345
Glucose-6-phosphate	G6P	X_2	1.0110
Fructose-1,6-phosphate	FDP	X_3	9.1440
Phosphoenolpyruvate	PEP	X_4	0.0095
Adenosine triphosphate	ATP	X_5	1.1278
Enzyme	**Acronym**	**Symbol**	**Initial Steady State**
Glucose transport	V_{in}	Y_1	19.70
Hexokinase	V_{HK}	Y_2	68.50
Phosphofructo-1-kinase	V_{PFK}	Y_3	31.70
Glyceraldehyde dehydrogenase	V_{GAPD}	Y_4	49.90
Pyruvate kinase	V_{PK}	Y_5	3440.00
Polysaccharide biosynthesis	V_{Carb}	Y_6	14.31
Polyol biosynthesis	V_{Gro}	Y_7	203.00
ATPase	V_{ATPase}	Y_8	25.10

$$\begin{aligned}
V_{in} &= 0.8122X_2^{-0.2344}Y_1 \\
V_{HK} &= 2.8632X_1^{0.7464}X_5^{0.0243}Y_2 \\
V_{PFK} &= 0.5232X_2^{0.7318}X_5^{-0.3941}Y_3 \\
V_{Carb} &= 8.904\times10^{-4}X_2^{8.6107}Y_7 \\
V_{GAPD} &= 7.6092\times10^{-2}X_3^{0.6159}X_5^{0.1308}Y_4 \\
V_{Gro} &= 9.272\times10^{-2}X_3^{0.05}X_4^{0.533}X_5^{-0.0822}Y_8 \\
V_{PK} &= 9.471\times10^{-2}X_3^{0.05}X_4^{0.533}X_5^{-0.0822}Y_5 \\
V_{ATPase} &= X_5X_6
\end{aligned} \tag{12}$$

The ethanol production is given by the pyruvate kinase (v_{PK}) flux. Therefore, the first objective of the multi-objective problem for this pathway can be formulated as follows:

$$max\, F_1 = V_{PK} \tag{13}$$

For the second objective, the total component concentrations involved is formulated as follows:

$$min\, F_2 = \sum_{j=1}^{5} X_j + \sum_{j=6}^{6} Y_j \tag{14}$$

For the steady state constraint, all the equations in the GMA model are equal to zero. Hence, the GMA of this pathway (given by Equation 11) can be formulated as follows:

$$\begin{aligned}
&V_{in} - V_{HK} = 0 \\
&V_{HK} - V_{PFK} - V_{Carb} = 0 \\
&V_{PFK} - V_{GAPD} - 0.5V_{Gro} = 0 \\
&2V_{GAPD} - V_{PK} = 0 \\
&2V_{GAPD} + V_{PK} - V_{HK} - V_{Carb} - V_{PFK} - V_{ATPase} = 0
\end{aligned} \tag{15}$$

For the component concentration constraint, the metabolites constraint were set to approximately 20% from their steady state values, which were in the range of 0.8 and 1.2 [2], [31]. Meanwhile, for the enzymes constraint, the values were set in the range of 0-50 [2], [31].

Hence, the problem statement for the multi-objective constraint optimization for *S.cerevisiae* pathway can be formulated as follows:

$$\begin{aligned}
&max\, F_1 = V_{PK} \\
&min\, F_2 = \sum_{j=1}^{5} X_j + \sum_{k=1}^{6} Y_k
\end{aligned} \tag{16}$$

subjected to:

$$\begin{aligned}
&Sv(x)_i = 0, \qquad i = 1,2,3,4,5 \\
&X_j^{0.8} \le X_j \le X_j^{1.2} \qquad j = 1,2,3,4,5 \\
&Y_k^0 \le Y_k \le Y_k^{50} \qquad k = 1,2,3,4,5,8
\end{aligned} \tag{17}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5.2 Optimization Of *Tryptophan* Biosynthesis In *Escherichia Coli* Pathway

For *E.coli* pathway, the proposed method is used to optimize the *trp* production. Figure 5 shows the *E.coli* pathway and Table 2 give the detail of all components in this pathway. This pathway formulates the GMA model as follows:

$$\begin{aligned}\frac{dX_1}{dt} &= V_{11} - V_{12}\\ \frac{dX_2}{dt} &= V_{21} - V_{22}\\ \frac{dX_3}{dt} &= V_{31} - V_{32} - V_{33} - V_{34}\end{aligned} \tag{18}$$

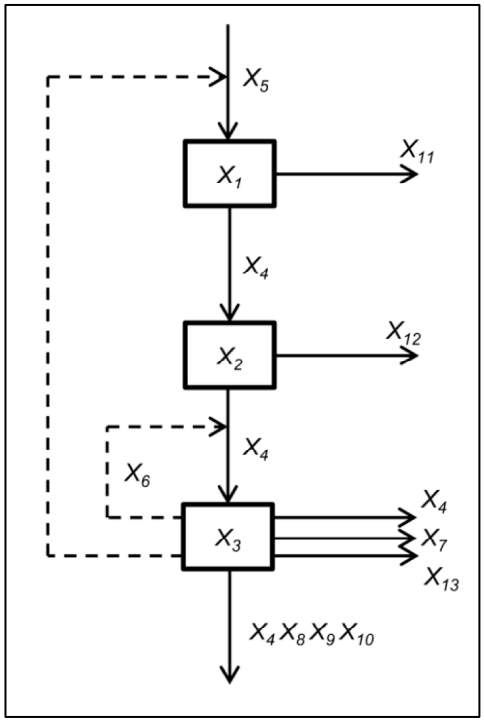

Figure 5: The Schematic Representation Of E.Coli Pathway

Table 2: The Value Of Components In E. Coli Pathway

Reaction concentration	Initial Steady State
X_1	0.184654
X_2	7.986756
X_3	1418.931944
X_4	0.00312
X_5	5
X_6	2283
X_8	430

In Equation 18, X_1 is *mRNA* concentration, X_2 is enzyme concentration and X_3 is *trp* concentration. For this pathway, the initial steady state of all fluxes can be formulated as follows:

$$\begin{aligned}V_{11} &= 0.6403X_3^{-5.87\times10^{-4}}X_5^{-0.8332}\\ V_{12} &= 1.0233X_1X_4^{0.0035}X_{11}^{0.9965}\\ V_{21} &= X_1\\ V_{22} &= 1.4854X_2X_4^{-0.1349}X_{12}^{0.8651}\\ V_{31} &= 0.5534X_2X_3^{-0.5573}X_6^{0.5573}\\ V_{32} &= X_3X_4\\ V_{33} &= 0.9942X_3^{7.0426\times10^{-4}}X_7\\ V_{34} &= 0.8925X_3^{3.5\times10^{-6}}X_4^{0.9760}X_8X_9^{-0.0240}X_{10}^{-3.5\times10^{-6}}\end{aligned} \tag{19}$$

The *trp* production of *E.coli* pathway is given by given by reaction V_{34}. This leads to the first objective for the multi-objective constraint optimization of this pathway as follows:

$$max\, F_1 = V_{34} \tag{20}$$

For the second objective, the total component concentrations involved in producing the best production can be formulated as follows:

$$min\, F_2 = \sum_{j=1}^{6} X_j + X_8 \tag{21}$$

For the steady state constraint, all the equations in the GMA model (Equation 18) are equal to zero and can be formulated as follows:

$$\begin{aligned}V_{11} - V_{12} &= 0\\ V_{21} - V_{22} &= 0\\ V_{31} - V_{32} - V_{33} - V_{34} &= 0\end{aligned} \tag{22}$$

In *E.coli* pathway, only several components were fine-tuned, which were X_1 to X_6 and X_8, while the other components used fixed value. The component concentration constraint can be formulated as follows [3], [28], [32]:

$$\begin{aligned}&X_k^{0.8} \le X_k \le X_k^{1.2} \qquad k = 1,2,3\\ &0 \le X_4 \le 0.00624\\ &4 \le X_5 \le 10\\ &500 \le X_6 \le 5000\\ &0 \le X_8 \le 1000\end{aligned} \tag{23}$$

Therefore, the problem statement for the multi-objective constraint optimization of *E.coli* pathway can be formulated as follows:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$max\ F_1 = V_{34}$$
$$min\ F_2 = \sum_{k=1}^{6} X_k + X_8 \quad (24)$$

subjected to:

$$\begin{aligned} & sv(x)_i = 0, \qquad i = 1,2,3 \\ & X_k^{0.8} \le X_k \le X_k^{1.2} \qquad k = 1,2,3 \\ & 0 \le X_4 \le 0.00624 \\ & 4 \le X_5 \le 10 \\ & 500 \le X_6 \le 5000 \\ & X_7 = 0.0022 X_5 \\ & 0 \le X_8 \le 1000 \\ & X_9 = 7.5 \\ & X_{10} = 0.005 \\ & X_{11} = 0.9 \\ & X_{12} = 0.02 \\ & X_{13} = 0 \end{aligned} \quad (25)$$

6. EXPERIMENTAL RESULTS AND DISCUSSIONS

In order to obtain the best result, several experiments were performed to determine which parameter settings can produce the highest result. Table 3 shows the list of parameter settings used to obtain the best result for all biochemical systems. For Newton method, fixed parameters were used, the maximum number of iterations was set to 50 and the tolerance was set to 10^{-6}.

Table 3: Parameter Settings In Obtaining The Best Result

Parameter	*S.cerevisiae* pathway	*E.coli* pathway
Number of sub-population	11	7
CR: Crossover rate	0.8	0.8
MR: Mutation rate	0.1	0.2
Maximum number of Pareto Optimal	100	100
Maximum number of generations	100	100

The best results obtained by the proposed in *S.cerevisiae* pathway are shown in Table 4. The table shows the best result in a single run, the average of 100 runs, standard deviation and the comparison with other existing works for all variables involved. In this pathway, the proposed method able to increase the ethanol production up to 53nM and reduce the total of component concentrations involved to 295.84nM. Several existing works were compared with the proposed method and gives in Table 4. It can be clearly seen that the performance of the proposed method is better than all existing works in term of increase the ethanol production and reduce the total of component concentrations involved.

Table 5 shows the best result achieved by the proposed method in *E.coli* pathway. The result was measured in term of the best in a single run, average and standard deviation and the comparison with other existing works. The proposed method was subjected to 100 independent runs. The proposed method able to increase the *trp* production up to 3.99nM and reduce the total of component concentrations involved to 6015.96nM. In evaluating the performance of the proposed method, the results (*trp* production and the total of component concentrations involved) in best of single run were compared with other existing works. As shown in Table 5, the *trp* production produced by the proposed method was the highest and the proposed method perform better in reducing the total of component concentrations involved compare to other existing works.

In evaluating the concept of multiple sub-chromosomes that introduced in this paper, the proposed method was compared with the single chromosome representation (without CCA). Several experiments were performed by using parameter settings as indicated in Table 3. Figure 6 and Figure 7 give the comparison of the results in *S.cerevisiae* pathway and *E.coli* pathway respectively. It can be seen clearly that all the production results from all pathways produced by the proposed method were higher compared to the results that only used single chromosome representation. This is because the proposed method introduces the optimization strategy in which the concept of candidate solution representation where it allows every variable (in non-linear equations system) to be represented separately by the sub-chromosome and evolve in their own sub-population. By doing that, it makes all variables altered and fine-tuned. This does not happen in the single chromosome representation since there are possibilities that not all variables are altered and fine-tuned as all variables are grouped together into a single chromosome. In conclusion, the candidate solution representation introduced in this paper is able to increase the performance of the proposed method in improving the production.

Journal of Theoretical and Applied Information Technology
30th November 2015. Vol.81. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

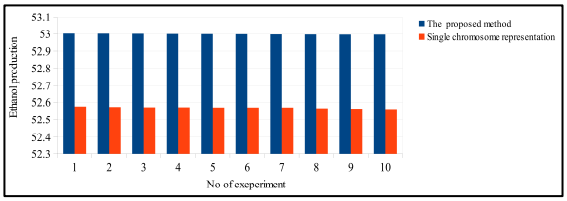

Figure 6: The Comparison Results Of Multiple Sub-Chromosomes Concept In S.cerevisiae Pathway

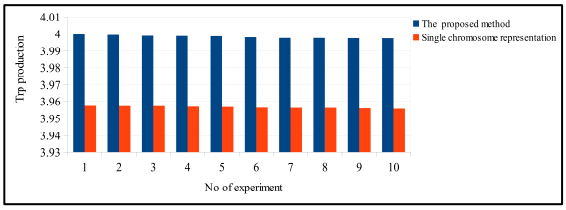

Figure 7: The Comparison Results Of The Multiple Sub-Chromosomes Concept In E.coli Pathway

Besides the concept of multiple sub-chromosomes, the sub-chromosome evaluation concept also contributed to the effectiveness of the proposed method, particularly in minimizing the total of component concentration involved. This can be achieved when a representative from each sub-population is selected in order to form the cooperative chromosome. The selection of a representative is based on the fitness value where the lowest fitness value is selected and combines with other representatives, thus ensure that the total of component concentrations involved can be minimized. In validating this strategy, several experiments were conducted where comparisons were made in choosing the representative from sub-population: choose based on fitness value; choose randomly; and choose based on the index that refers to the order sub-chromosome generated. The experiment used the parameter settings in Table 3 for *S.cerevisiae* pathway and *E.coli* pathway. The results of this experiment are given in Figure 8 and Figure 9 for *S.cerevisiae* pathway and *E.coli* pathway respectively. From both figures, it can be observed that the representative that was selected based on fitness was able to minimize the total of component concentrations involved compared to random-based and index-based. It can be clearly seen that the graph of fitness-based decreased smoothly. For the graph of random-based and index-based, sometimes the total of component concentrations involved value from the next generation is higher than the previous generation. This is due to the process of selecting the representative from all sub-populations. This process does not happen in selecting the representation in random-based and index-based,

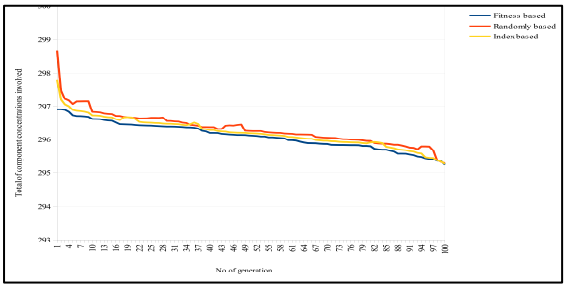

Figure 8: The Comparison Results Of The Sub-Chromosome Evaluation Concept In S.cerevisiae Pathway

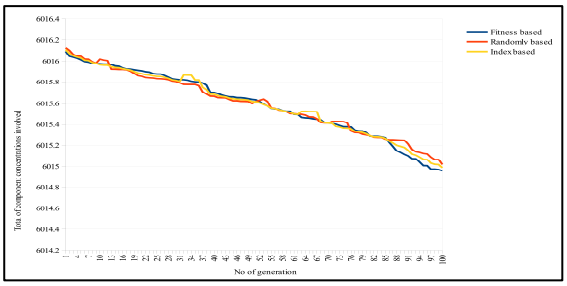

Figure 9: The Comparison Results Of The Sub-Chromosome Evaluation Concept In E.coli Pathway

and sometimes it makes the total of component concentrations involved higher in the next generation. As a conclusion, using the sub-chromosome evaluation concept is effective as it allows the representative to be selected based on the fitness value to ensure that the total of component concentrations involved can be minimized.

7. CONCLUSION

In this paper, an advance method that hybrid Newton method, SPEA and CCA was presented and discussed in detail. The method was proposed in order to improve the biochemical system production and simultaneously reduce the total of component concentrations involved. The proposed method works by viewing a biochemical system as a non-linear equations system. Then, Newton method was used in dealing with non-linear equations system, GA in SPEA for fine-tuning the variables in non-linear equations system and CCA to reduce the total of component concentrations involved. Several experiments were performed on the benchmark biochemical system, namely *S. cerevisiae* pathway and *E. coli* pathway. The experimental results showed that the proposed method performed well when compare to other existing works.

ACKNOWLEDGEMENT

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The authors would like to express their appreciation for the support of the sponsors of the Potential Academic Staff Grant No. Q.J130000.2728.01K19 from the Universiti Teknologi Malaysia, managed by the Research Management Center, Universiti Teknologi Malaysia. This work was also supported by the Fellowship scheme under Universiti Malaysia Pahang and the Malaysian Ministry of Education. We would also like to express our gratitude to the editor and anonymous reviewers who reviewed this paper.

REFRENCES:

[1] C. Grosan and A. Abraham, "A New Approach for Solving Nonlinear", *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions*, Vol. 38, No. 3, 2008, pp. 698–714.

[2] H. Link, J. Vera, D. Weuster-Botz, N. T. Darias, and E. Franco-Lara, "Multi-objective steady state optimization of biochemical reaction networks using a constrained genetic algorithm," *Computers and Chemical Engineering*, Vol. 32, No. 8, 2008, pp. 1707–1713.

[3] J. Vera, C. Gonzalez-Alcon, A. Marin-Sanguino, and N. Torres, "Optimization of biochemical systems through mathematical programming: Methods and applications," *Computers & Operations Research*, Vol. 37, No. 8, 2010,pp. 1427–1438.

[4] G. Xu and L. Wang, "An Improved Geometric Programming Approach for Optimization of Biochemical Systems," *Journal of Applied Mathematics*, Vol. 2014, 2014.

[5] P. K. Polisetty, E. P. Gatzke, and E. O. Voit, "Yield optimization of regulated metabolic systems using deterministic branch-and-reduce methods," *Biotechnology and Bioengineering*, Vol. 99, No. 5, 2008, pp. 1154–1169.

[6] S. P. Bradley, A. C. Hax, and T. L. Magnanti, *Applied Mathematical Programming*. Reading, MA: Addison-Wesley, 1977.

[7] A. P. Mariano, C. B. B. Costa, D. de F. de Angelis, F. M. Filho, D. I. P. Atala, M. R. W. Maciel, and R. M. Filho, "Optimization strategies based on sequential quadratic programming applied for a fermentation process for butanol production," *Applied Biochemistry and Biotechnology*, Vol. 159, No. 2, 209, pp. 366–381.

[8] R. Fourer, D. M. Gay, and B. W. Kernighan, "A Modeling Language for Mathematical Programming," *Management Science*, Vol. 36, 1990, pp 519–554.

[9] E. Balsa-Canto, J. R. Banga, J. A. Egea, A. Fernandez-Villaverde, and G. M. Hijas-Liste, "Global Optimization in Systems Biology: Stochastic Methods and Their Applications," *Advances in Systems Biology*, Vol. 736, I. I. Goryanin and A. B. Goryachev, Eds. Springer New York, 2012, pp. 409–424.

[10] M. Babaei, "A general approach to approximate solutions of nonlinear differential equations using particle swarm optimization," *Applied Soft Computing*, Vol. 13, No. 7, 2013, pp. 3354–3365.

[11] Y.-Z. Luo, G.-J. Tang, and L.-N. Zhou, "Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-Newton method," *Applied Soft Computing*, Vol. 8, No. 2, 2008, pp. 1068–1073.

[12] O. H. Sendín, J. Vera, N. V. Torres, and J. R. Banga, "Model based optimization of biochemical systems using multiple objectives: a comparison of several solution strategies," *Mathematical and Computer Modelling of Dynamical Systems*, Vol. 12, No. 5, 2007, pp. 469–487.

[13] C.-S. Liu, "A modified Newton method for solving non-linear algebraic equations," *Journal of Marine Science and Technology*, Vol. 17, No. 3, 2009, pp. 238–247.

[14] Y. Mo, H. Liu, and Q. Wang, "Conjugate direction particle swarm optimization solving systems of nonlinear equations," *Computers & Mathematics with Applications*, Vol. 57, No. 11–12, 2009, pp. 1877–1882.

[15] N. Ujevic, "A method for solving nonlinear equations," *Applied Mathematics and Computation*, Vol. 174, No. 2, 2006, pp. 1416–1426.

[16] C.-Y. Ku, W. Yeih, and C.-S. Liu, "Solving Non-Linear Algebraic Equations by a Scalar Newton-homotopy Continuation Method," *International Journal of Nonlinear Sciences and Numerical Simulation*, Vol. 11, No. 6, 2010, pp. 435–450.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[17] E. G. Birgin, N. Krejić, and J. M. Martínez, "Globally Convergent Inexact Quasi-Newton Methods for Solving Nonlinear Systems," *Numerical Algorithms*, Vol. 32, No. 2–4, 2003, pp. 249–260.

[18] C.-S. Liu and S. N. Atluri, "A novel time integration method for solving a large system of non-linear algebraic equations," *Computer Modeling in Engineering and Sciences*, Vol. 31, No. 2, 2008, pp. 71–83.

[19] S. Taheri and M. Mammadov, "Solving systems of nonlinear equations using a globally convergent optimization algorithm," *Global Journal of Technology & Optimization*, Vol. 3, 2013, pp. 132–138.

[20] M. Á. Fernández and M. Moubachir, "A Newton method using exact jacobians for solving fluid–structure coupling," *Computers & Structures*, Vol. 83, No. 2–3, 2005, pp. 127–142.

[21] A. Konak, D. W. Coit, and A. E. Smith, "Multi-objective optimization using genetic algorithms: A tutorial," *Reliability Engineering & System Safety*, Vol. 91, No. 9, 2006, pp. 992–1007.

[22] P. Ngatchou, A. Zarei, and A. El-Sharkawi, "Pareto Multi Objective Optimization," in *Proceedings of the 13th International Conference on, Intelligent Systems Application to Power Systems*, 2005, pp. 84–91.

[23] S. Jiang and S. Yang, "A Fast Strength Pareto Evolutionary Algorithm Incorporating Predefined Preference Information," in *Proceedings of the 15th UK Workshop on Computational Intelligence*, 2015.

[24] W. Sheng, Y. Liu, X. Meng, and T. Zhang, "An Improved Strength Pareto Evolutionary Algorithm 2 with application to the optimization of distributed generations," *Computers & Mathematics with Applications*, Vol. 64, No. 5, 2012, pp. 944–955.

[25] E. Zitzler, K. Deb, and L. Thiele, "Comparison of Multiobjective Evolutionary Algorithms: Empirical Results," *Evolutionary Computation*, Vol. 8, No. 2, 2000, pp. 173–195.

[26] M. A. Ismail, H. Asmuni, and M. R. Othman, "The fuzzy cooperative genetic algorithm (FCoGA): The optimisation of a fuzzy model through incorporation of a cooperative coevolutionary method," *Journal of Computing*, Vol. 3, No. 11, 2011, pp. 81–90.

[27] J. Gu, M. Gu, C. Cao, and X. Gu, "A novel competitive co-evolutionary quantum genetic algorithm for stochastic job shop scheduling problem," *Computers & Operations Research*, Vol. 37, No. 5, 2010, pp. 927–937.

[28] A. Marin-Sanguino, E. O. Voit, C. Gonzalez-Alcon, and N. V Torres, "Optimization of biotechnological systems through geometric programming," *Theoretical Biology and Medical Modelling*, Vol. 4, 2007, pp. 38–54.

[29] M. A. Savageau, *Biochemical systems analysis: a study of function and design in molecular biology*. Reading, Massachusetts: Addison–Wesley, 1976.

[30] J. J. Durillo and A. J. Nebro, "jMetal: A Java framework for multi-objective optimization," *Advances in Engineering Software*, Vol. 42, No. 10, 2011, pp. 760–771.

[31] F. Rodriguez-Acosta, C. M. Regalado, and N. V Torres, "Non-linear optimization of biotechnological processes by stochastic algorithms: Application to the maximization of the production rate of ethanol, glycerol and carbohydrates by Saccharomyces cerevisiae," *Journal of Biotechnology*, Vol. 65, No. 1, 1999, pp. 15–28.

[32] G. Xu, "Steady-state optimization of biochemical systems through geometric programming," *European Journal of Operational Reseach*, Vol. 225, No. 1, 2013, pp. 12–20.

[33] M. A. Ismail, S. Deris, M. S. Mohamad, and A. Abdullah, "A Hybrid of Newton Method and Genetic Algorithm for Constrained Optimization method of the Production of Metabolic Pathway," *Life Science Journal*, Vol. 11, No. 9s, 2014, pp. 409–414.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4 : The best result and the comparison with other existing method in S.cerevisiae pathway.

Variable	Best Solution	Average	Standard deviation	Work in [31]	Work in [32]	Work in [33]
X_1	0.8281	0.9890	0.1223	1.14	1.102	1.11
X_2	0.9244	1.0035	0.1064	1.05	1.046	1.03
X_3	1.1397	0.9945	0.1158	1.15	1.141	1.13
X_4	1.1317	1.1291	0.0288	1.17	1.171	1.18
X_5	1.0900	0.9981	0.1178	1.12	1.113	1.14
Y_1	49.9873	49.9858	0.0091	49.97	50	49.99
Y_2	45.0051	45.0767	0.1743	44.77	45.953	45.83
Y_3	49.9020	49.9184	0.0175	49.89	50	49.92
Y_4	47.4848	47.4227	0.2472	47.26	47.772	47.97
Y_5	49.6307	49.5232	0.3143	48	48.366	48.30
Y_8	49.7893	49.8053	0.0339	49.75	50	49.79
F_1	**53.0002**	**52.8145**	**0.3483**	52.0843	52.5118	52.57
F_2	**295.8353**	**296.4461**	**0.4104**	295.27	297.664	297.384

Table 5 : The best result and the comparison with other existing method in E.coli pathway

Variable	Best Solution	Average	Standard deviation	Work in [28]	Work in [3]	Work in [32]	Work in [33]
X_1	1.1920	0.9753	0.138	1.1900	1.2000	1.2000	1.1100
X_2	1.1396	1.1007	0.1288	1.1480	1.1500	1.1150	1.1140
X_3	0.8	0.8	1.6×10^{-15}	0.8000	0.8000	0.8000	0.8000
X_4	0.0054	0.0054	1.2×10^{-5}	000041	0.0040	0.0054	0.00538
X_5	4.3880	4.3007	0.2943	4.0000	4.0000	4.0110	4.7540
X_6	5000	5000	0	5000	5000	5000	5000
X_8	1000	1000	0	1000	1000	1000	1000
F_1	**3.9865**	**3.9802**	**0.0085**	3.0620	3.0620	3.9460	3.9750
F_2	**6015.9596**	**6016.0232**	**0.2157**	6016.3759	6016.3759	6016.5652	6016.2168