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

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

1Faculty of Computer Systems and Software Engineering, Universiti Malaysia Pahang, Lebuh Raya Tun Razak, Gambang, Kuantan, Pahang, Malaysia.  
2Department of Software Engineering, Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia.  
3College of Computer Science and Mathematics, University of Mosul, 41002, Mosul, Iraq.  
E-mail: marfian2@live.utm.my, safaaai@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
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 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.
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 S-system and ODE. Within ODE, there are two representative system. The fine-tuning process of variables in 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:

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

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:

$$\max F_1(v)$$

subjected to:

$$F_2(\sum_{j=1}^{n} x_j) \leq 0$$

$$x_j^i \leq x_j \leq x_j^{i+1}$$

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.

$$f_{ij} = \frac{\delta v_i \delta x_j}{\delta x_j v_j}$$

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:

$$\frac{dX}{dt} = [sv_1(x), \ldots, sv_n(x)] = 0$$

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:

$$\max F_1(v)$$

subjected to:

$$F_2(\sum_{j=1}^{n} x_j) \leq 0$$

$$x_j^i \leq x_j \leq x_j^{i+1}$$

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.
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, 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 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.

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

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:

$$\text{best} = \frac{PF_1 - PF_{best}}{2} \times 100$$  \hspace{1cm} (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 $S.\text{cerevisiae}$ and the trp biosynthesis in $E.\text{coli}$. A program based on Java was developed to test the proposed method with these biochemical systems. The program based on jMetal [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 $S.\text{cerevisiae}$ pathway. In this pathway, components were divided into two groups, which were metabolites ($X_1 - X_5$) and enzymes ($Y_1 - Y_3$). Details of the metabolites and enzymes, including their initial steady state values are presented in Table 1. This pathway has the following GMA model:
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:

\[
\text{max } F_1 = v_{PK}
\]  

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

\[
\text{min } F_2 = \sum_{j=1}^{5} X_j + \sum_{k=6}^{8} Y_k
\]

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:

\[
0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3 = 0
\]  

(11)

\[
V_i = 0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3
\]

\[
V_{HK} = 2.8632X_2 - 0.7464X_2 - 0.0243X_2
\]

\[
V_{PFK} = 0.5232X_2
\]

\[
V_{GAPD} = 8.90410^{-2}X_3 + 0.533X_3
\]

\[
V_{Gro} = 0.6159X_3
\]

\[
V_{PK} = 0.05X_3
\]

\[
V_{ATPase} = X_5X_6
\]

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:

\[
\text{max } F_1 = v_{PK}
\]  

(13)

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

\[
\text{min } F_2 = \sum_{j=1}^{5} X_j + \sum_{k=6}^{8} Y_k
\]  

(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:

\[
0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3 = 0
\]  

(11)

\[
V_i = 0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3
\]

\[
V_{HK} = 2.8632X_2 - 0.7464X_2 - 0.0243X_2
\]

\[
V_{PFK} = 0.5232X_2
\]

\[
V_{GAPD} = 8.90410^{-2}X_3 + 0.533X_3
\]

\[
V_{Gro} = 0.6159X_3
\]

\[
V_{PK} = 0.05X_3
\]

\[
V_{ATPase} = X_5X_6
\]

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:

\[
\text{max } F_1 = v_{PK}
\]  

(13)

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

\[
\text{min } F_2 = \sum_{j=1}^{5} X_j + \sum_{k=6}^{8} Y_k
\]  

(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:

\[
0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3 = 0
\]  

(11)

\[
V_i = 0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3
\]

\[
V_{HK} = 2.8632X_2 - 0.7464X_2 - 0.0243X_2
\]

\[
V_{PFK} = 0.5232X_2
\]

\[
V_{GAPD} = 8.90410^{-2}X_3 + 0.533X_3
\]

\[
V_{Gro} = 0.6159X_3
\]

\[
V_{PK} = 0.05X_3
\]

\[
V_{ATPase} = X_5X_6
\]

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:

\[
\text{max } F_1 = v_{PK}
\]  

(13)

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

\[
\text{min } F_2 = \sum_{j=1}^{5} X_j + \sum_{k=6}^{8} Y_k
\]  

(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:

\[
0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3 = 0
\]  

(11)

\[
V_i = 0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3
\]

\[
V_{HK} = 2.8632X_2 - 0.7464X_2 - 0.0243X_2
\]

\[
V_{PFK} = 0.5232X_2
\]

\[
V_{GAPD} = 8.90410^{-2}X_3 + 0.533X_3
\]

\[
V_{Gro} = 0.6159X_3
\]

\[
V_{PK} = 0.05X_3
\]

\[
V_{ATPase} = X_5X_6
\]

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:

\[
\text{max } F_1 = v_{PK}
\]  

(13)

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

\[
\text{min } F_2 = \sum_{j=1}^{5} X_j + \sum_{k=6}^{8} Y_k
\]  

(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:

\[
0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3 = 0
\]  

(11)

\[
V_i = 0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3
\]

\[
V_{HK} = 2.8632X_2 - 0.7464X_2 - 0.0243X_2
\]

\[
V_{PFK} = 0.5232X_2
\]

\[
V_{GAPD} = 8.90410^{-2}X_3 + 0.533X_3
\]

\[
V_{Gro} = 0.6159X_3
\]

\[
V_{PK} = 0.05X_3
\]

\[
V_{ATPase} = X_5X_6
\]

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:

\[
\text{max } F_1 = v_{PK}
\]  

(13)

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

\[
\text{min } F_2 = \sum_{j=1}^{5} X_j + \sum_{k=6}^{8} Y_k
\]  

(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:

\[
0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3 = 0
\]  

(11)

\[
V_i = 0.8122X_2 - 0.5232X_2 - 8.6107X_2 + 8.90410^{-2}X_3
\]

\[
V_{HK} = 2.8632X_2 - 0.7464X_2 - 0.0243X_2
\]

\[
V_{PFK} = 0.5232X_2
\]

\[
V_{GAPD} = 8.90410^{-2}X_3 + 0.533X_3
\]

\[
V_{Gro} = 0.6159X_3
\]

\[
V_{PK} = 0.05X_3
\]

\[
V_{ATPase} = X_5X_6
\]
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:

\[
\begin{align*}
\frac{dx_1}{dt} &= V_{11} - V_{12} \\
\frac{dx_2}{dt} &= V_{21} - V_{22} \\
\frac{dx_3}{dt} &= V_{31} - V_{32} - V_{33} - V_{34}
\end{align*}
\]  

(18)

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:

\[
\begin{align*}
V_{11} &= 0.6403X_1^{1.487\times10^{-6}}X_4^{0.032} \\
V_{12} &= 1.0233X_1X_4^{0.0035}X_3^{0.0065} \\
V_{21} &= X_1 \\
V_{22} &= 1.4854X_2X_4^{-0.1349}X_3^{0.8651} \\
V_{31} &= 0.5534X_2X_4^{-0.5573}X_3^{0.5573} \\
V_{32} &= X_3X_4 \\
V_{33} &= 0.9942X_3^{0.0286}\times 10^{-6}X_4 \\
V_{34} &= 0.8925X_3^{3.5\times10^{-6}}X_4^{0.0070}X_5X_6^{0.0240}X_7^{3.5\times10^{-6}}
\end{align*}
\]  

(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:

\[
\max F_1 = V_{34}
\]  

(20)

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

\[
\min F_2 = \sum_{j=1}^{6} X_j + X_8
\]  

(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:

\[
\begin{align*}
V_{11} - V_{12} &= 0 \\
V_{21} - V_{22} &= 0 \\
V_{31} - V_{32} - V_{33} - V_{34} &= 0
\end{align*}
\]  

(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]:

\[
\begin{align*}
X_k^{0.3} &\leq X_k \leq X_k^{1.2} & k = 1, 2, 3 \\
0 &\leq X_4 \leq 0.00624 \\
4 &\leq X_5 \leq 10 \\
500 &\leq X_6 \leq 5000 \\
0 &\leq X_8 \leq 1000
\end{align*}
\]  

(23)

Therefore, the problem statement for the multi-objective constraint optimization of E.coli pathway can be formulated as follows:
\[ \max F_1 = V_{14} \]
\[ \min F_2 = \sum_{k=1}^{6} x_k + x_6 \] (24)

subjected to:
\[ s v(x_i) = 0, \quad i = 1, 2, 3 \]
\[ x_i^{\min} \leq x_i \leq x_i^{\max} \quad k = 1, 2, 3 \]
\[ 0 \leq x_4 \leq 0.00624 \]
\[ 4 \leq x_7 \leq 10 \]
\[ 500 \leq x_8 \leq 5000 \]
\[ x_9 = 0.0022x_4 \]
\[ 0 \leq x_4 \leq 1000 \]
\[ x_9 = 7.5 \]
\[ x_{10} = 0.005 \]
\[ x_{11} = 0.9 \]
\[ x_{12} = 0.02 \]
\[ x_{13} = 0 \] (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>
<thead>
<tr>
<th>Parameter</th>
<th>S. cerevisiae pathway</th>
<th>E. coli pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sub-population</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>CR: Crossover rate</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>MR: Mutation rate</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>Maximum number of Pareto Optimal</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Maximum number of generations</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

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 ability 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 compared 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 clearly seen 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.
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 \textit{S.cerevisiae} pathway and \textit{E.coli} pathway. The results of this experiment are given in Figure 8 and Figure 9 for \textit{S.cerevisiae} pathway and \textit{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, 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 \textit{S. cerevisiae} pathway and \textit{E. coli} pathway. The experimental results showed that the proposed method performed well when compare to other existing works.

**ACKNOWLEDGEMENT**
The authors would like to express their appreciation for the support of the sponsors of the Potential Academic Staff Grant No. Q.J13000.2728.01K19 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.

REFERENCES:


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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>0.8281</td>
<td>0.9890</td>
<td>0.1223</td>
<td>1.14</td>
<td>1.102</td>
<td>1.11</td>
</tr>
<tr>
<td>$X_2$</td>
<td>0.9244</td>
<td>1.0035</td>
<td>0.1064</td>
<td>1.05</td>
<td>1.046</td>
<td>1.03</td>
</tr>
<tr>
<td>$X_3$</td>
<td>1.1397</td>
<td>0.9945</td>
<td>0.1158</td>
<td>1.15</td>
<td>1.141</td>
<td>1.13</td>
</tr>
<tr>
<td>$X_4$</td>
<td>1.1317</td>
<td>1.1291</td>
<td>0.0288</td>
<td>1.17</td>
<td>1.171</td>
<td>1.18</td>
</tr>
<tr>
<td>$X_5$</td>
<td>1.0900</td>
<td>0.9981</td>
<td>0.1178</td>
<td>1.12</td>
<td>1.113</td>
<td>1.14</td>
</tr>
<tr>
<td>$Y_1$</td>
<td>49.9873</td>
<td>49.9858</td>
<td>0.0091</td>
<td>49.97</td>
<td>50</td>
<td>49.99</td>
</tr>
<tr>
<td>$Y_2$</td>
<td>45.0051</td>
<td>45.0767</td>
<td>0.1743</td>
<td>44.77</td>
<td>45.953</td>
<td>45.83</td>
</tr>
<tr>
<td>$Y_3$</td>
<td>49.9020</td>
<td>49.9184</td>
<td>0.0175</td>
<td>49.89</td>
<td>50</td>
<td>49.92</td>
</tr>
<tr>
<td>$Y_4$</td>
<td>47.4848</td>
<td>47.4227</td>
<td>0.2472</td>
<td>47.26</td>
<td>47.772</td>
<td>47.97</td>
</tr>
<tr>
<td>$Y_5$</td>
<td>49.6307</td>
<td>49.5232</td>
<td>0.3143</td>
<td>48</td>
<td>48.366</td>
<td>48.30</td>
</tr>
<tr>
<td>$Y_6$</td>
<td>49.7893</td>
<td>49.8053</td>
<td>0.0339</td>
<td>49.75</td>
<td>50</td>
<td>49.79</td>
</tr>
<tr>
<td>$F_1$</td>
<td>53.0002</td>
<td>52.8145</td>
<td>0.3483</td>
<td>52.0845</td>
<td>52.518</td>
<td>52.57</td>
</tr>
<tr>
<td>$F_2$</td>
<td>295.8353</td>
<td>296.4461</td>
<td>0.4104</td>
<td>295.27</td>
<td>297.664</td>
<td>297.384</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1.1920</td>
<td>0.9753</td>
<td>0.138</td>
<td>1.1900</td>
<td>1.2000</td>
<td>1.2000</td>
<td>1.1100</td>
</tr>
<tr>
<td>$X_2$</td>
<td>1.1396</td>
<td>1.1007</td>
<td>0.1288</td>
<td>1.1480</td>
<td>1.1500</td>
<td>1.1150</td>
<td>1.1140</td>
</tr>
<tr>
<td>$X_3$</td>
<td>0.8</td>
<td>0.8</td>
<td>$1.6 \times 10^{-15}$</td>
<td>0.8000</td>
<td>0.8000</td>
<td>0.8000</td>
<td>0.8000</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0.0054</td>
<td>0.0054</td>
<td>$1.2 \times 10^{-10}$</td>
<td>0.00041</td>
<td>0.0040</td>
<td>0.0054</td>
<td>0.00538</td>
</tr>
<tr>
<td>$X_5$</td>
<td>4.3880</td>
<td>4.3007</td>
<td>0.2943</td>
<td>4.0000</td>
<td>4.0000</td>
<td>4.0101</td>
<td>4.7540</td>
</tr>
<tr>
<td>$X_6$</td>
<td>5000</td>
<td>5000</td>
<td>0</td>
<td>5000</td>
<td>5000</td>
<td>5000</td>
<td>5000</td>
</tr>
<tr>
<td>$X_8$</td>
<td>1000</td>
<td>1000</td>
<td>0</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
<td>1000</td>
</tr>
<tr>
<td>$F_1$</td>
<td>3.9865</td>
<td>3.9802</td>
<td>0.0085</td>
<td>3.0620</td>
<td>3.0620</td>
<td>3.9460</td>
<td>3.9750</td>
</tr>
<tr>
<td>$F_2$</td>
<td>6015.9596</td>
<td>6016.0232</td>
<td>0.2157</td>
<td>6016.3759</td>
<td>6016.3759</td>
<td>6016.5652</td>
<td>6016.2168</td>
</tr>
</tbody>
</table>