Threonine Biosynthesis Pathway Simulation Using IBMDE with Parameter Estimation

Chuii Khim Chong¹, Mohd Saberi Mohamad¹*, Safaai Deris¹, Mohd Shahir Shamsir², Yee Wen Choon¹, and Lian En Chai¹

¹ Artificial Intelligence and Bioinformatics Research Group, Faculty of Computing, Universiti Teknologi Malaysia, Skudai, 81310 Johor, Malaysia
ckchong2@live.utm.my, {saberi,safaai}@utm.my, {ywchoon2,lechai2}@live.utm.my
² Department of Biological Sciences, Faculty of Biosciences and Bioengineering, Universiti Teknologi Malaysia, Skudai, 81310 Johor, Malaysia
shahir@utm.my

Abstract. When analysing a metabolic pathway through mathematical model, it is important that the significant parameters are being correctly estimated. However, this process often comes across problems such as easily being trapped in local minima, repetitive exposure to worse results during the search process, and occurrence of noisy data. Thus, an improved Bee Memory Differential Evolution algorithm (IBMDE), which is a hybrid of the Differential Evolution algorithm (DE), the Kalman Filter (KF), Artificial Bee Colony algorithm (ABC), and a memory feature is presented this paper. IBMDE is an improved estimation algorithm as previous work only utilised DE. The threonine biosynthesis pathway is the metabolic pathways used in this paper. For metabolite O-Phosphohomoserine production simulation, the IBMDE able to produce the estimated optimal kinetic parameter values with significantly reduced error rate (63.67%) and shows a faster convergence time (71.46%) compared to the Nelder Mead (NM), the Simulated Annealing (SA), the Genetic Algorithm (GA), and DE respectively. In addition, IBMDE demonstrates to be a reliable estimation algorithm.

Keywords: Parameter Estimation, Differential Evolution Algorithm, Kalman Filter, Artificial Bee Colony Algorithm, Memory feature.

1 Introduction

Systems biology manifests a biological system by a set of ordinary differential equations (ODEs) in mathematical models [1]. The essential interactions show quantitatively through ODEs with the aim in explaining the behaviour at the system level. Hand-tuning and in-vitro biochemical experiments are the main methods to retrieve the values of unknown parameters in the ODEs [2]. Under some conditions, these values are collected through estimation, and therefore, it is important that the

* Corresponding author.
estimation methods used in mathematical models are thoroughly studied [1]. Parameter estimation in system biology reduces the variance between experimental data and simulated data. It usually works as a part of a recursive process to develop mathematical models that are able to provide optimal estimated values for biological systems. Nevertheless, increasing number of unknown parameters and noisy experimental data of dynamic biochemical pathways cause most traditional estimation methods to generate inaccurate estimations [3].

Under the category of evolutionary algorithms, the Differential Evolution algorithm (DE) has been found to be the best estimation algorithm. It works to optimize a problem repetitively with the fixed objective function. The major advantages of DE are efficiency, high speed, ease of use, and simplicity [4]. It had been implemented by Moonchai et al. [5] to improve the production of bacteriocin by estimating the control parameters which were temperature and pH. Therefore, in this paper, it is implemented as the main estimation algorithm to solve the increasing number of unknown parameters. Kalman gain, $K$, value is used by the Kalman Filter (KF) in handling the noisy data through normalization. IDE (the Improved Differential Evolution algorithm) is the hybrid of these two algorithms [6].

Easily trapped in the local minimal due to faster convergence speed [4] and attempts to expose to worse results during the search process repeatedly are the disadvantages of DE. Therefore, to solve the mentioned disadvantages, IDE is then further combine with two modifications – the artificial bee colony algorithm (ABC) and a memory feature to generate the improved Bee Memory Differential Evolution algorithm (IBMDE). ABC capable to rise the probability in finding the optimal solutions by the food source possibility and this characteristic avoids the trapped in local minima [7]. The memory feature, however, capable to keep track of the best candidate ever during the search process with the extra memory named $g_{best}$ and this prevents the worse result from being exposed again.

The paper is structured into four sections, where Section 2 introduce the method implemented, IBMDE, and subsequently Section 3 with experimental setup, results, and discussion. Finally, the conclusion and future works is showed in Section 4.

## 2 Methodology

An improved estimation algorithm, the Improved Bee Memory Differential Evolution algorithm (IBMDE) is presented in this paper. Previous work [5] only used DE while IBMDE uses a hybrid of DE, KF, ABC, and memory feature for parameter estimation. The details of the IBMDE is demonstrated in Figure 1. As in conventional DE, a $n \times m$ population matrix, $P$, the initial population is produced in the initialization process. $m$ is the number of unknown parameters and $n$ is the number of generations. Each gene of the individuals in the initial population gained its own value based on Equation (1). Each gene represents a parameter value and individual, $(\text{Ind}_i)$, indicates a set of estimated parameter values (possible solution) which $i$ is the index variable where $1 \leq i \leq n$. For initialisation process, $g$ indicates the ones matrix with dimension of $n \times 1$, $I_{\text{initial}} = \{I_1, I_2, \ldots, I_m\}$ where $I$ is the initial values for each parameter, and $\text{rand}(n,m)$ indicates a $n \times m$ matrix with random values between 0 and 1.

\[
P = g \times I_{\text{initial}} \times 10^{0.5 \times \text{rand}(n \times m)}
\] (1)