BioSystems 162 (2017) 81–89

Contents lists available at ScienceDirect

BioSystems

journal homepage: www.elsevier.com/locate/biosystems

An improved hybrid of particle swarm optimization and the gravitational search algorithm to produce a kinetic parameter estimation of aspartate biochemical pathways

Ahmad Muhaimin Ismail [a], Mohd Saberi Mohamad [c,g,h,*], Hairudin Abdul Majid [a], Khairul Hamimah Abas [b], Safaai Deris [c,g,h], Nazar Zaki [d], Siti Zaiton Mohd Hashim [e], Zuwairie Ibrahim [f], Muhammad Akmal Remli [a]

[a] Artificial Intelligence and Bioinformatics Research Group, Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia
[b] Department of Control and Mechatronic Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia,
[c] Faculty of Creative Technology and Heritage, Universiti Malaysia Kelantan, Bachok, Kelantan, Malaysia,
[d] College of Information Technology, United Arab Emirate University, Al Ain, United Arab Emirates
[e] Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia
[f] Faculty of Electrical and Electronic Engineering, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia,
[g] Center For Computing and Informatics, Universiti Malaysia Kelantan, City Campus, Pengkalan Chepa, 16100 Kota Bharu, Kelantan, Malaysia
[h] Institute For Artificial Intelligence and Big Data, Universiti Malaysia Kelantan, City Campus, Pengkalan Chepa, 16100 Kota Bharu, Kelantan, Malaysia

ARTICLE INFO

Article history:
Received 6 October 2016
Received in revised form 23 June 2017
Accepted 21 September 2017
Available online 23 September 2017

Keyword:
Parameter estimation
Biochemical pathway
Particle swarm optimization
Gravitational search algorithm
Artificial intelligence
Bioinformatics
Metabolic engineering

ABSTRACT

Mathematical modelling is fundamental to understand the dynamic behavior and regulation of the biochemical metabolisms and pathways that are found in biological systems. Pathways are used to describe complex processes that involve many parameters. It is important to have an accurate and complete set of parameters that describe the characteristics of a given model. However, measuring these parameters is typically difficult and even impossible in some cases. Furthermore, the experimental data are often incomplete and also suffer from experimental noise. These shortcomings make it challenging to identify the best-fit parameters that can represent the actual biological processes involved in biological systems. Computational approaches are required to estimate these parameters. The estimation is converted into multimodal optimization problems that require a global optimization algorithm that can avoid local solutions. These local solutions can lead to a bad fit when calibrating with a model. Although the model itself can potentially match a set of experimental data, a high-performance estimation algorithm is required to improve the quality of the solutions.

This paper describes an improved hybrid of particle swarm optimization and the gravitational search algorithm (IPSOGSA) to improve the efficiency of a global optimum (the best set of kinetic parameter values) search. The findings suggest that the proposed algorithm is capable of narrowing down the search space by exploiting the feasible solution areas. Hence, the proposed algorithm is able to achieve a near-optimal set of parameters at a fast convergence speed. The proposed algorithm was tested and evaluated based on two aspartate pathways that were obtained from the BioModels Database. The results show that the proposed algorithm outperformed other standard optimization algorithms in terms of accuracy and near-optimal kinetic parameter estimation. Nevertheless, the proposed algorithm is only expected to work well in small scale systems. In addition, the results of this study can be used to estimate kinetic parameter values in the stage of model selection for different experimental conditions.

1. Introduction

Explaining the complex network biological processes that are characterized by dynamic behavior is one of the main issues in systems biology (Lillacci and Khammash, 2010; Raue et al., 2015). Pathways are used to describe the relationship between

* Corresponding author.
E-mail address: saberi@umk.edu.my (M.S. Mohamad).

http://dx.doi.org/10.1016/j.biosystems.2017.09.013
0303-2647/

parameters as a means of understanding the complex processes that are involved in biological systems, and mathematical models are commonly used to describe dynamic biological processes. Mathematical models can generate and predict the outcomes of the experimental hypotheses that can be employed to analyze the processes. The application of these models has enabled the construction of metabolic pathways. This phenomenon opens up opportunities for the optimization of metabolite productions in metabolic pathways (Ismail et al., 2015). Normally, these models are based on time-derivative expressions, especially ordinary differential equations (ODE), that describe a change in a state or a quantity of interest over time.

Generally, these models consist of a set of parameters that describe the physical properties of a dynamic system like the rate of reactions. Measuring these parameters is usually difficult and even impossible in some cases (Fernández Slezak et al., 2010). The parameters are often predicted based on fitting the estimated data of the model output with experimental time-series data. The goal of this fitting process is to minimize the errors between these two sets of data by adjusting the parameter values of the model (Rodriguez-Fernandez et al., 2006a). However, these experimental data are often incomplete and suffer from experimental noise (Villaverde et al., 2015). This drawback makes it challenging to find the best-fit parameters that adequately represent the actual biological processes involved. It is crucial that the best parameter values for the biochemical models are estimated and obtained by refining the model parameter values (Schilling et al., 2016). These parameter values are usually identified and measured through costly and time-consuming wet-lab experiments (Tashkova et al., 2011). Alternatively, these parameters can also be estimated using computational approaches. Thus, the estimation of the parameters can be converted into multimodal optimization problems, and global optimization algorithms are required to avoid local solutions (Banga, 2008; Sun et al., 2012). These local solutions can lead to poorly fitting data that the model itself can potentially match accurately with a set of experimental data.

Global optimization algorithms employ stochastic searching strategies to identify a set of possible solutions that are randomly selected based on the given search space. Furthermore, these algorithms are widely used to estimate parameters for various biological models (Chassagnole et al., 2001; Curien and Bastien, 2009; Galazzo and Bailey, 1990; Sun, 2012). Particle swarm optimization (PSO) (Ng et al., 2013; Shi and Eberhart, 1999), the Bee algorithm (BA) (Leong et al., 2013; Pham et al., 2006), the Firefly Algorithm (FA) (Yang, 2009), differential evolution (DE) (Chong et al., 2014), scatter search (SS)(Rodriguez-Fernandez et al., 2006b), simulated annealing (SA) (Villaverde et al., 2012), and others, have already been used to estimate the parameters involved in various biological system models. The main advantages of these models are that they offer researchers the ability to find the best and easiest ways to implement solutions for high-dimensional problems. Despite these advantages, these algorithms often suffer from high computational costs as they try to obtain a global optimum within the large search space (Baker et al., 2010; Fong, 2014; Sun, 2012). In addition, the generated solutions might not represent the actual near-optimal solutions.

In multimodal optimization problems, PSO is often stuck in local optimal, which is the result of a poor global search. The standard gravitational search algorithm (GSA) also has some drawbacks; for example, it has a poor convergence if the initial population is not well generated (Kumar and Sahoo, 2014). Moreover, it often incurs a large computational cost due to the large searching space. Thus, a hybrid of particle swarm optimization and the gravitational search algorithm (PSOGSA) is proposed, which combines the social thinking (*gbest*) ability of PSO with the exploration capability of GSA. This hybrid is able to perform well in optimization problems, especially in minimization problems (Mirjalili and Hashim, 2010). Nevertheless, PSOGSA often incurs a high computational cost when obtaining the global optimum solutions. Besides, the advantages of employing PSO capability in the hybrid in terms of its rapid convergence speed are also weakened (Shanhe and Zhicheng, 2014). Standard and previous algorithms of parameter estimation that have been employed to deal with noisy data often suffer from such poor solutions, and there are typically high errors between experimental and estimated outputs. Hence, a high-performance optimization algorithm is required to maintain fast convergence frequently and improve the quality of the solutions.

This paper proposes an improved hybrid of PSOGSA (IPSOGSA). This improvement has enhanced the search for a global optimum (the best set of kinetic parameter values) by reducing the searching space and focusing the search on the high possibility of feasible solution areas. Hence, there are upsurges in the performance of the proposed algorithm with the advantage of fast convergence in obtaining the global optimum and near-optimum solutions.

The paper is structured as followed. First, we present a problem formulation on the parameter estimation in kinetic model. Then, we present the description of the proposed algorithm phases accompanied with the details on each phase. Next, the experimental setup is explained and consists of the description of model case studies, parameter setting and performance evaluation for the estimation results. We then present the result and discussion section that discuss the results and findings from this study. Finally, the paper is summarized in the conclusion section.

2. Materials and algorithms

The problems that are inherent in biological system estimation will be briefly formulated and explained in this section before the IPSOGSA and experimental setups are explained.

2.1. Problem formulated

The aim of the parameter estimation problem is to attain the near-optimal set of parameters that can minimize the differences between the estimated model output and the experimental time series data. Usually, the nonlinear least squares error function is implemented to minimize differences. Parameter estimation for biological systems can be expressed as per Eqs. (1)–(3) (Lillacci and Khammash, 2010). Where s is the compound in the biochemical system model $s(x)$, which comprises a set of parameters $x = \{x_1, x_2, .. x_n\}$ where n is the number of parameters. The reaction rate of compound s can be represented as a series ODE in the following form:

$$\frac{ds}{dt} = g(s(u,x), t), \tag{1}$$

$$s(t_0) = s(0), \tag{2}$$

$$y = h(s(u,x), t) + e, \tag{3}$$

where g and h are the nonlinear functions, t is the sampling time, and e is the generated measurement noise by random Gaussian noise $N(0,1)$, while y is the rate of reaction and $s(x)$ is the biochemical compounds with set of parameter x. On the other hand, u is the input signal to the reaction of s process.

2.2. An improved hybrid of particle swarm optimization and the gravitational search algorithm (IPSOGSA)

In PSOGSA, the standard PSO has been improved through modifying the process by which acceleration is calculated, before employing this to update the velocity and population process. As with PSO, this hybrid also carries other operations, such as initialization, update velocity, and position. This hybrid adopts the