International Journal of Information Retrieval Research
Volume 12 • Issue 1

Intelligent Framework With Controlled Behavior for Gene Regulatory Network Reconstruction

Bhavana Bansal, Jaypee Institute of Information Technology, Noida, India*
Aparajita Nanda, Jaypee Institute of Information Technology, Noida, India
Anita Sahoo, Jaypee Institute of Information Technology, Noida, India

ABSTRACT

Gene regulatory networks (GRNs) are the pioneering methodology for finding new gene interactions getting insights of the biological processes using time series gene expression data. It remains a challenge to study the temporal nature of gene expression data that mimic complex non-linear dynamics of the network. In this paper, an intelligent framework of recurrent neural network (RNN) and swarm intelligence (SI)-based particle swarm optimization (PSO) with controlled behaviour has been proposed for the reconstruction of GRN from time-series gene expression data. A novel PSO algorithm enhanced by human cognition influenced by the ideology of Bhagavad Gita is employed for improved learning of RNN. RNN guided by the proposed algorithm simulates the nonlinear and dynamic gene interactions to a greater extent. The proposed method shows superior performance over traditional SI algorithms in searching biologically plausible candidate networks. The strength of the method is verified by analyzing the small artificial network and real data of Escherichia coli with improved accuracy.

KEYWORDS

Gene Regulatory Networks, Human Cognition-Based PSO, Recurrent Neural Networks, Time-Series Gene Expression Data

1. INTRODUCTION

Recent years have seen the advent of DNA microarray technology and information retrieval has been proved essential for the reconstruction of gene regulatory networks (GRNs) from temporal gene expression data. GRNs proves important for understanding many unkown biological functionalities and processes. It gives insights of the activities of genes and provide knowledge about transcriptional regulations among them (Aalto et al., 2020). GRNs is a virtual network of genes and their mutual influences, where node of the network is a gene and edges are the influence from the regulator to the target gene which either activates or suppress target gene's ability of protein formation (Morgan et al., 2019). GRNs have been successfully applied in diagnostics and contributes in identification

DOI: 10.4018/IJIRR.2022010104 *Corresponding Author

of essential genes (Xie et al., 2020). The well known issue encountered in the analysis of temporal data for GRN reconstruction problem is the curse of dimensionality (Altman & Krzywinski, 2018).

In context of the computational models used for GRNs reconstruction from time series data, researchers has adopted several methods (Delgado & Gómez-Vela, 2019; Razaghi-Moghadam & Nikoloski, 2020) such as Boolean networks, Bayesian networks (BNs), dynamic Bayesian networks (DBNs) and linear additive genetic model. Boolean networks model (Barman & Kwon, 2018) considers only two states for each gene: active and inactive. This model does not take into consideration the intermediary effects on the genes which cause information loss. Bayesian networks (BNs) model (Sanchez-Castillo et al., 2018) are graph based models forming a genetic network as a directed acyclic graphs. This model effectively handles noise, missing values and the random nature of gene expression data, however it does not take into account the dynamical nature of GRNs and the temporal aspect of the data. The limitations of BNs were overcome by dynamic Bayesian networks (DBNs) (Adabor & Acquaah-Mensah, 2019). The linear additive genetic model (Luque-Baena et al., 2014) may identify linear regulatory relationships but does not consider the non-linear behaviour of GRNs.

Motivation: Considering the limitations of these models researchers adopted recurrent neural network (RNN) for the problem of GRNs reconstruction. RNN model clearly manifested the temporal nature of gene expression data and non-linear dynamics among gene regulations which is essential for GRN reconstruction. This model has an ability to consider the feedforward and feedback loops of the genetic regulation network (Biswas & Acharyya, 2016, 2018). Time-series data is the input to the RNN model. The data contains expression levels ($x_i(t)$) of genes at consecutive time points. The gene expression level $x_i(t+1)$ of a gene (i) of the current time point ($t+1$) of an RNN.

Layer is simulated from the expression levels $x_j(t)$ of genes (j) (regulator genes) at previous time point t accompanied by the set of genetic network parameters. In terms of GRNs topology very few regulatory genes j influence a target gene i which concludes that genetic network connectivity is

Figure 1. The overview of the work flow