

Chapter 12

Application of Fireworks Algorithm in Bioinformatics

Yuchen Zhang

School of Computer Science, Shaanxi Normal University, China

Xiujuan Lei

School of Computer Science, Shaanxi Normal University, China

Ying Tan

Peking University, China

ABSTRACT

Fireworks Algorithm (FWA) has been applied to many fields in recent years, showing a strong ability to solve optimization problems. In this chapter, FWA is applied to some research hotspots in bioinformatics, such as biclustering of gene expression data, disease-gene prediction, and identification of LncRNA-protein interactions. This chapter briefly introduces some backgrounds of bioinformatics and related issues. Through corresponding bioinformatics' problems to optimization problems, some specific optimization functions are constructed and solved by the Fireworks Algorithm. The simulation results illustrate that the fireworks algorithm shows high performance and potential application value in the field of bioinformatics.

INTRODUCTION

Fireworks algorithm (FWA) (Tan & Zhu, 2010) has been applied to many research fields as a relatively new heuristic optimization algorithm proposed by Tan et al. The algorithm searches the optimal solutions of problems in solution space by simulating the explosion process of fireworks. The performance of this algorithm has also been affirmed by many scholars. For example, A power system reconstruction scheme based on FWA is proposed to minimize the power loss and voltage profile enhancement (Mohamed Imran, Kowsalya, & Systems, 2014). Milan Tuba et al. solved multilevel image threshold problem by using FWA (Tuba, Bacanin, & Alihodzic, 2015). In the large-scale non-convex economic

DOI: 10.4018/979-8-3693-3026-5.ch012

load dispatch problem, the FWA with chaotic sequence operator is applied well (Pandey et al., 2018). At the same time, FWA also derived many improved versions, such as Enhanced FWA (Zheng, Janecek, & Tan, 2013) and Adaptive FWA (Li, Zheng, & Tan, 2014). Thus, FWA is a mature optimization algorithm, and its searching process follows Markova random process.

Bioinformatics, as an interdisciplinary subject of computer science, biology and mathematics, is a hot research field in recent years. Bioinformatics is mainly concerned with the collection, processing, storage, distribution, analysis and interpretation of biological information. It combines the tools and techniques of mathematics, computer science and engineering, and biology to reveal the biological secrets contained in a large number of complex biological data. Through computational methods, sequence alignment, protein complexes mining, essential proteins identification, disease-genes prediction and the relationships exploration between proteins and non-coding RNAs can be carried out. And these problems may be converted into appropriate optimization problems, which can be solved by swarm intelligence algorithm. Therefore, this chapter intends to apply FWA to bioinformatics.

With the rapid development of high-parallel and high-throughput sequencing technology, a large number of gene expression levels have been measured at the same time, thus obtaining a large number of microarray gene expression data (Xiang, Yang, Ma, & Ding, 2003). There are a lot of useful information in the unbalanced gene expression data (Eisen, Spellman, Brown, & Botstein, 1998), such as similarity of gene expression, conditional specificity of gene expression (proliferation, differentiation and canceration of cells). Therefore, data mining and machine learning technology are used widely. (Cheng & Church, 2000) Cheng et al. first applied bi-clustering analysis to gene expression data. Yang et al. combined the method of dealing miss values with the definition of biclustering based on the CC algorithm, and proposed the Flexible Overlapped BiClustering (FLOC) algorithm to obtain multiple double clusters (Yang, Wang, Wang, & Yu, 2002). According to the greedy strategy, Angiulli et al. proposed a new biclustering algorithm RWB. This algorithm utilizes the random walk strategy (Angiulli, Cesario, & Pizzuti, 2008). At the same time, heuristic intelligent algorithms have been applied to this NP-hard problem. A biclustering algorithm based on evolutionary computation is proposed (Divina & Aguilar-Ruiz, 2006). In addition, some new evaluation criteria have also been introduced (Divina, Pontes, Giráldez, Aguilar-Ruiz, & medicine, 2012). However, the new swarm intelligence optimization algorithms proposed in recent years have been applied only a few times. Some scholars found that the cuckoo optimization algorithm can solve the problem of biclustering very well (Balamurugan, Nataraajan, & Premalatha, 2018). Lu Yin et al. combined cuckoos algorithms with genetic algorithms to mine submatrices (Yin, Qiu, & Gao, 2018). Therefore, based on FWA, this chapter intends to use it to carry out the study of biclustering of gene expression data.

Genes are known to be effective carriers of genetic material in DNA. Genes play important roles in today's disease research. Identifying diseases associated genes is one of the important tasks in the study of pathology in complex diseases. Although a lot of biomolecular knowledge and research methods (such as linkage analysis) (Easton, Bishop, Ford, & Crockford, 1993) related to diseases have been obtained at present, a lot of manpower and material resources have been required, and sometimes expected results are not sure to be obtained. Numerous studies have confirmed that genes related to the same or similar diseases, their functions are usually similar or related (Goh et al., 2007). This similarity or correlation may be a direct binding in the physical sense or belong to the same protein complex, or there may be an indirect interaction, such as participation in the same metabolic pathway or cellular process. This makes the development of disease-gene prediction algorithms based on molecular network become a hot topic in recent years. Kohler et al. proposed that applying the random walk with restart (RWR) model

29 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/application-fireworks-algorithm-bioinformatics/342530

Related Content

The Genetic Code, Hadamard Matrices, Noise Immunity, and Quantum Computers

Sergey Petoukhov and Matthew He (2010). *Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications* (pp. 115-131).

www.irma-international.org/chapter/genetic-code-hadamard-matrices-noise/37899

SPCCTDM, a Catalogue for Analysis of Therapeutic Drug Monitoring Related Contents in the Drug Prescription Information

Sven Ulrich, Pierre Baumann, Andreas Conca, Hans-Joachim Kuss, Viktoria Stieffenhofer and Christoph Hiemke (2010). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 1-11).

www.irma-international.org/article/spcctdm-catalogue-analysis-therapeutic-drug/45162

An Innovative Approach to Enhance Collaboration in the Biomedical Field

Georgia Tsiliki, Manolis Tzagarakis, Spyros Christodoulou, Sophia Kossida and Nikos Karacapilidis (2013). *International Journal of Systems Biology and Biomedical Technologies* (pp. 51-64).

www.irma-international.org/article/innovative-approach-enhance-collaboration-biomedical/78392

Emerging Web Tools and Their Applications in Bioinformatics

Shailendra Singh and Amardeep Singh (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 1769-1783).

www.irma-international.org/chapter/emerging-web-tools-their-applications/76147

Image Processing Tools for Biomedical Infrared Imaging

Gerald Schaefer and Arcangelo Merla (2010). *Biocomputation and Biomedical Informatics: Case Studies and Applications* (pp. 187-197).

www.irma-international.org/chapter/image-processing-tools-biomedical-infrared/39612