

参考文献

- [1] 增材制造技术的研究现状及其发展趋势 [J]. 郭志飞,张虎. 机床与液压. 2015(05).
- [2] 增材制造(3D 打印)技术发展[J]. 卢秉恒,李涤尘. 机械制造与自动化. 2017(04).
- [3] 3D 打印技术研究现状和关键技术[J]. 张学军,唐思熠,肇恒跃,郭绍庆,李能,孙兵兵,陈冰清. 材料工程. 2016(02).
- [4] 基于 BIM 的 3D 打印技术在建筑行业的应用研究[J]. 季安康,王海飙. 科技管理研究. 2016(24).

A SINGLE SOURCE REACHABLE QUERY METHOD FOR GENE REGULATORY NETWORKS

Zhaoyuan Zhang, Ziheng Ding, Hanwen Wang, Keyi Liu, Weiyiqi Wang
School of Computer Science and Engineering, Northeastern University, China

Summary. *With the deepening of research on gene regulatory network, the construction technology of gene regulatory network has been relatively mature. However, according to the characteristics of gene regulatory network, it is still a problem to realize efficient reachable probability query and reachable path query among gene pairs on the constructed gene regulatory network. Therefore, this paper proposes an efficient and accurate single-source reachable query algorithm GRN-RQ based on gene regulatory network, in which the meaning of single-source is: the reachable query between gene pairs with both starting and ending nodes being single nodes can be realized. Algorithm GRN-RQ mainly includes three parts, respectively, an active graph pruning algorithm AP, which is applicable to different node situations, is used to reduce the size of probability graphs, a concurrent reachable query algorithm PSRA for partial instance graph generation and staged graph search, which is used to improve operational efficiency, the dual nodes optimal reachable path query algorithm DOP based on the dual path evaluation rule DPFR. Through experimental verification, the algorithm GRN-RQ proposed in this paper can not only quickly obtain the reachable probability between gene pairs and the optimal reachable path between them on the gene regulatory network, but also improve the accuracy of the reachable query result compared with the traditional method.*

Keywords. *gene regulatory network, single source reachable query, graph pruning, concurrent reachable query, optimal reachable path.*

Paper Background

As the gene regulatory network construction technology becomes more and more mature in recent years, a fast and accurate reachable query algorithm is urgently needed to be applied to the established gene regulatory network, so as to provide a reliable basis for biomedical research based on the relationship between genes. The existing relevant reachable query algorithms have some drawbacks and limitations when applied to gene regulatory networks, some of which cost too much time, some have unstable results, and others are not applicable to gene regulatory networks. Based on this, the project team proposes a reachable query algorithm suitable for gene regulatory network in detail.

Paper Innovation Scheme

The innovation scheme includes the innovation ideas of the algorithm and the specific algorithm introduction.

There are three main innovation ideas of the project algorithm. The first is to design efficient and diverse graph pruning algorithms based on gene regulatory network. The second is to use the concurrent reachable query algorithm to save time. The third is to design dual path evaluation rules to get the optimal path and suboptimal path. The specific algorithm design is as follows.

In terms of graph pruning algorithm, considering the large-scale and diverse characteristics of the gene regulatory network, a graph pruning algorithm with three processes is proposed. The algorithm can reduce the scale of the graph to a minimum without changing the reachable probability of a given two nodes, which is convenient for the subsequent calculation of the reachable probability.

In terms of reachable query algorithm, the traditional reachable query based on Monte Carlo method consumes a lot of time in generating possible world graphs, so the query efficiency is not high. In view of the shortcomings of this method, the project team proposed an innovative query algorithm for generation of possible world graphs and concurrent reachable queries, which greatly shortened the time required for reachable queries and improved query efficiency.

In terms of the reachable path query algorithm, based on the designed dual path evaluation rules, the project team used greedy algorithm to get the optimal path between two nodes on the probability graph on the one hand, and added the path probability correlation calculation to the concurrent reachable query algorithm to get the suboptimal path.

Academic Value

The pruning algorithm based on gene regulatory network is highly profitable in directed acyclic probability graph, and provides a method to reduce the scale of graph without changing the probability between nodes. Besides a concurrent reachable probability query algorithm is proposed for the first time, which greatly reduces the time and space complexity. In addition, it is the first time that path evaluation rules and path query algorithms are designed for the gene regulatory network represented by probability graph in the paper, which provides a new way to explore the relationship between gene nodes. At last, the algorithm proposed in this paper has a certain universality and can be applied to related fields involving reachable query on probability graph.

The relevant experimental results of the algorithm can also verify project team's hypothesis. In terms of the consuming time, the basic algorithm consumes 1457.625ms in average, while the PSRA-AP proposed in the paper only used 33.0125ms at average. From the comparison of the running time of the algorithm proposed in the paper and the basic random algorithm on the breast cancer gene regulatory network, it can be seen that the running time of the algorithm proposed in the paper is greatly shortened. In terms of the accuracy, the variance of the basic algorithm is 7.299×10^{-6} , on the contrary, that of the proposed algorithm is 7.26×10^{-6} , which is obtained from the 20 sets of data tested on the breast cancer gene regulatory network. It can be seen from the calculation of variance that the test data variance of the algorithm proposed in the paper is smaller, so the experimental results are more accurate.