

第一章 基于节点距离的图相似性搜索

由前文可知，图数据能表示复杂的数据结构，在诸多领域也得到了广泛应用。从基本的生物，化学的分子结构，到交通网络，人际关系网都可以用图来建模。而对于图数据库的索引也自然成为了热点问题。

上一章我们介绍了精确子图搜索方法，但是由于真实情况下图数据库具有信息不完整，含有杂质等情形，精确搜索容易出现各种不匹配问题，难以得到我们想要的结果。同时，对于查询图的完整信息有时查询者也并不了解。所以相似性搜索的实际应用领域更加广泛。对于实际应用，相似性搜索的研究意义远超过了精确搜索。

传统的图相似性算法虽然运行效率已较为理想，但是编码上过于复杂，也无法做到对所有图数据库良好适配。所以本章我们提出了一种性能上不输于传统算法，但是实现更为简单，并且无需复杂设计即可适用于大量图数据库的基于节点距离的通用型算法。本章将首先详细介绍下我们算法的具体思路，然后给出详细的代码设计方案，最后给出对于真实数据的实验结果与分析。