

第一章 总结与展望

本章作为全文的最后一章，将对本文的所述内容进行总结，并对下一步的研究工作给出一些意见。

1.1 本文总结

主要工作如下：

根据绪论

1. 本文首先介绍了下图的基本定义，存储方法，图查询类型和一些经典的查询算法。
2. 随后，我们对于图精确搜索提出了一种基于二次哈希开链法的搜索算法，有效避免了传统哈希方法冲突频发的问题，加速了整个查询过程。本文中，我们完整得介绍了这种算法，从数据库建库，到查询剪枝，直最后的子图同构检测，并通过实验证明了此算法确实可以加快整个查询过程。
3. 然后，对于图相似性搜索我们提出了一种基于节点相似度的搜索算法。本算法与 **G-Hash** 算法大致相同，但重新定义了其核心部分——图相似度度量方法，使得编码复杂度大大降低，但同时效率又不低于 **G-Hash** 算法。本文中我们除了详细介绍了此算法原理，还给出了算法设计类图，并通过实验证明了此算法完全符合我们预期的目标——运行效率不低于 **G-Hash**，甚至在某些情况下略好于 **G-Hash**，但是编码难度大大下降。

综上所述，本文对子图搜索方面做了一定基础性的研究，在认真研究了前人的经典算法基础上，进行了一些新的探索。

1.2 进一步的工作

但是，图数据查询仍是图数据管理中的一个重要领域，其中仍有一些问题需要继续去研究与改进。在未来的研究过程中，可以从以下几个方面入手：

1. 目前，时时刻刻产生着大量的图数据信息，如何在图数据库中有效存储，在内存中以何种结构存储图数据，如何用固定的磁盘页面存储不同规模尺寸的图数据，又如何对图数据进行压缩表示，这些基本的物理存储问题将直接决定 I/O 操作的用时。而在实际中对图的存取又是非常频繁的，所以如果能解决存储问题，那么必然能提高整体查询效率。
2. 除了物理存储，逻辑索引显得更为重要，如何高效索引，如何进行维护和更新，这些都是亟待解决的问题。特别是图索引的维护与更新，由于目前大量算法都处于理论阶段而并非实际运用，所以都没有涉及动态维护与更新。但是随着图数据越来越为重要，实际中对图的应用也日益增多。如果能解决这个问题，那么对图从实验室走到实际运用中将会是个强助力。
3. 图查询的计算复杂度通常能达到指数级及以上，单线程的运算速度就成了图查询速度的瓶颈。因此考虑利用 CPU/GPU 异构并行或者多处理器并行，多机并行等也是一个不错的研究课题。

参考文献