# 摘 要

大规模无标度网络创建技术是复杂网络领域研究的热点问题之一。现有的现实世界图结构不能满足特殊研究的需求。Barabási-Albert (BA)是较著名的无标度网络构造模型，在进行无标度网络创建时，现有方法在处理节点效率低，存在重复遍历等问题。本文针对以上问题展开研究，旨在提升节点处理效率，主要研究内容如下。

首先，通过深入分析现存无标度网络创建过程的缺陷，发现当前方法处理效率低的原因主要为：节点重复遍历问题。（重复节点处理，访问无用节点）

其次，提出一种多叉树节点处理策略及相应的算法，以便解决在内存中处理大规模数据重复遍历导致的重复计算的问题，该算法根据现有BA模型改善数据存储结构。根据无标度网络的特点，提出基于ROLL-TREE的索引结构，并设计了相应的三叉树搜索策略以便提高算法效率。

再次，基于三叉树索引的高效处理结果，通过进一步改善节点存储结构，从而构建更符合现实特性的无标度网络，进一步优化节点处理结果。

最后，基于不同查询，通过实验分别从结果数量、运行时间以及原子操作次数等多方面对算法高效性进一步验证。

关键词：无标度网络，索引，

# 第1章 绪 论

## 1.1 研究背景

在[网络理论](https://zh.wikipedia.org/w/index.php?title=%E7%BD%91%E7%BB%9C%E7%90%86%E8%AE%BA&action=edit&redlink=1)中，无标度网络是带有一类特性的[复杂网络](https://zh.wikipedia.org/wiki/%E5%A4%8D%E6%9D%82%E7%BD%91%E7%BB%9C)，其典型特征是在网络中的大部分节点只和很少节点连接，而有极少的节点与非常多的节点连接。这种关键的节点（称为“**枢纽**”或“**集散节点**”）的存在使得无尺度网络对意外故障有强大的承受能力，但面对协同性攻击时则显得脆弱。现实中的许多网络都带有无尺度的特性，例如[因特网](https://zh.wikipedia.org/wiki/%E4%BA%92%E8%81%94%E7%BD%91)、金融系统网络、[社会人际网络](https://zh.wikipedia.org/wiki/%E7%A4%BE%E4%BC%9A%E6%80%A7%E7%BD%91%E7%BB%9C)等等。

小世界网络模型说明了少量的随机捷径会改变网络的拓扑结构，从而出现小世界效应，无标度网络出现的主要用途是为了能够准确描述现实世界网络结构，揭示了增长性和择优机制在复杂网络演化过程中的普遍性以及幂律的重要性，且在当前多个研究领域中复杂网络属于热门研究。，为了能够顺应复杂网络创建技术的发展，越来越多的行业开始根据现实世界网络的结构特征以及数据存储方式来设计与本领域相关的语义，以便能够实现与无标度网络构造技术的接轨。

在现有研究成果的基础上，本文结合所研究无标度网络的结构特征、节点处理效率的情况，以及进行高效的节点处理研究。

## 1.2 研究现状

复杂网络的理论和应用已经有了相对成熟的发展，在学术界以及社会层面都引起了很轰动的效果，同时对于这些理论以及应用的研究也留下了广阔的发展空间，

## 1.3 研究内容

从存在的复杂网络研究内容显示，在现实网络节点分布均匀的情形下，现有的研究成果能够快速且简单的实现网络构建，但是，最新研究结果发现，在节点分布多变，社区结构属性构成复杂的情况下，现有研究成果较难满足需求，

本篇论文提出了一种基于内存的可以以轮盘形式快速构建BA无标度网络的数据结构，这种结构具有以下特性：(a)快速：比基础的单节点处理速度快1000倍；(b)精准：新提出的数据结构严格遵循基于精确的BA模型，而不是近似模型；(c)可扩展性：可以应用到其他具有“优先连接”幂律增长性的模型中。

研究在内存中快速高效的构建无标度网络构建技术可以为建立新的复杂网络模型提供参考。

## 1.4 本文结构

本文设计的内容为在内存中快速高效的构建无标度网络构建技术，重点内容为构建ROLL\_MULTI\_TREE结构、扩展的MULTI\_TREE结构，本文总体上可分5章，以下为从第2章起的结构：

第2章为相关基础知识概述，主要介绍了无标度网络的特征，复杂网络以及幂律分布的基本概念。

第3章对已有方法存在的问题进行改进优化，且将基本的ROLL\_TREE结构进行扩展，最后从理论和示例中对该算法进行了论证。

第4章，同样根据创建多叉树的策略，提出对ROLL\_MULTI\_TREE的改进算法，并在此基础上分析相应的节点分布特征以及各个结构对节点处理效率的影响。，分别从理论和示例中对本章的算法进行了论证。

第5章，首先，对本文研究中提出的ROLL\_MULTI\_TREE结构与基础的ROLL\_TREE结构处理节点的效率进行了对比，然后对文中所提出的所有算法用实验分析的形式来验证其高效性。

最后，对研究生期间的研究成果进行总结；分析当前的研究内容，提出其以后有待改进的地方；最后提出未来研究工作方向的瞻望。

# 第2章 基础知识概述

复杂网络：

对于复杂网络，钱学森先生给出了严格的定义，

小世界网络：

无标度网络：

具有

度分布：

节点的度服从幂律分布，

特点：1.增长性：有新节点加入

2.优先连接：连接的概率与已存在的节点的属性有关

a)适应度模型：引入适应度来表示节点的特性

b)局域世界演化模型：选取局域世界，优先连接概率与局域世界有关

幂律分布：各个参数的含义

复杂网络的传播：俄

三叉树效率较作者高

分析四叉树 多叉树的树高，度分布最高点变化情况