

**课 程 实 验 报 告**

**课程名称： 大数据分析**

**专业班级：**

**学 号：**

**姓 名：**

**指导教师：**

**报告日期：**

**计算机科学与技术学院**

**目录**

[实验二 PageRank算法及其实现 1](#_Toc166675391)

[**1.1实验目的** 1](#_Toc166675392)

[**1.2 实验内容** 1](#_Toc166675393)

[**1.3 实验过程** 1](#_Toc166675394)

[1.3.1 编程思路 1](#_Toc166675395)

[1.3.2 遇到的问题及解决方式 1](#_Toc166675396)

[1.3.3 实验测试与结果分析 1](#_Toc166675397)

[**1.4 实验总结** 1](#_Toc166675398)

# 实验二 PageRank算法及其实现

## **1.1实验目的**

1、学习pagerank算法并熟悉其推导过程；

2、实现pagerank算法，理解阻尼系数的作用；

3、将pagerank算法运用于实际，并对结果进行分析。

## **1.2 实验内容**

利用实验一得到的出现次数最多前1000个的title之间的引用关系<title,<title1,…,titlek>>，由title为节点构造有向图，编写pagerank算法的代码，根据每个节点的入度计算其pagerank值，迭代直到误差小于10-8。

实验进阶版考虑加入teleport β，用以对概率转移矩阵进行修正，解决dead ends和spider trap的问题。

输出title及其对应的pagerank值。

## **1.3 实验过程**

### 1.3.1 编程思路

### 1.3.2 遇到的问题及解决方式

### 1.3.3 实验测试与结果分析

## **1.4 实验总结**