

<<链接分析>>

图书基本信息

书名：<<链接分析>>

13位ISBN编号：9787564112790

10位ISBN编号：7564112794

出版时间：2009-1

出版时间：东南大学出版社

作者：迈克·塞沃尔

页数：225

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<链接分析>>

前言

网络链接分析研究可以追溯到20世纪90年代中期。

1995年, Brazilian Marcia J. Bossy首次提出可将信息技术应用于因特网。

1996年, Larson在《万维网的文献计量: 网络空间结构初探》一文中明确将信息技术从文献计量学移植到网络中。

早期的链接分析研究同时出现在几个学科中, 包括计算机科学领域中的搜索引擎开发, 数学领域中的结构和复杂性分析。

而在1997年, Almind和Ingwersen提出了“网络计量学(webometrics)”一词, 旨在定量分析网络现象。

此后, 链接分析便成了网络计量学的主要研究内容之一。

1998年, Google的创始人Brin和Page公开了PageRank算法的核心部分, 这一链接分析算法, 作为Google的核心技术之一, 支撑着Google在商业上取得了巨大的成功, 同时, 也彰显了链接分析研究的魅力。

在信息科学研究领域, 按照Mike Fhelwall的定义, 链接分析就是采用并改进现有的信息技术, 借助文档之间的相互关联, 对文档自身的特征进行深入分析。

链接分析涉及的文档包括四个层面: 页面、目录、域名、站点。

在理论方面, 链接分析与文献计量学中的引文分析有高度相似性。

链接分析的内容主要包括: Web结构研究: 将文档视为节点, 将文档之间的链接视为连线, 便可将Web理解为一张抽象的结构图(有向图), 图中的节点与连线的属性都值得深入研究; 链接增长规律研究: 链接的建立不是随机的(“均匀链接”), 而是服从某种规律的, Web环境中, 小世界现象已不再适用, 不同类型的页面间的链接增长规律不同, 纯粹的幂定律模型(“优先链接”)已难以概括这些规律; 链接分类研究: 链接可以传达信息, 因此, 判断创建链接的动机可用于判断学术网络上信息交流的类型; 链接分析算法: Pagerank算法、HITS算法等链接分析算法应用于搜索引擎检索结果排序, 极大地提高了检索效率, 网络在发展, 这些算法也在不断的更新; 链接分析工具研究: 搜索引擎与网络爬虫一直是链接分析研究中获取数据的主要工具, 但二者都有自身的缺陷, 如何合理地使用链接分析工具以提高链接分析数据的有效性则一直是链接分析研究中讨论的主要问题。

链接分析研究中常用的工具与方法包括: 搜索引擎、网络档案、网络爬虫、链接数据库、数据清理技术、网络空间分析法、虚拟民族志方法、社会网络分析法、网络可视化方法等。

链接分析可应用于多个领域, 包括: 搜索引擎与网站设计: 链接分析算法用于对检索结果排序可有效提高检索效率, 而网站根据搜索引擎的工作原理和排序算法改进网站结构、提高网站内容质量以增加其可见度; 网站健康度检查: 通过分析网站上的人链数与出链数、入链页面类型与出链页面类型等数据判断网站在网络中的影响力及健康状况; 知识挖掘: 从链接分析算法与链接网络图中可以挖掘出网页、网站的潜在属性及潜在关联, 以获得新知识, 例如, 可借助Pajek构建知识地图, 将链接形成的网络关系可视化为一张二维图, 从图中可以明显地判断出节点的重要程度与节点间关联的紧密程度。

<<链接分析>>

内容概要

链接分析，源于对Web结构中超链接的多维分析。

当前其应用主要体现在网络信息检索、网络计量学、数据挖掘、Web结构建模等方面。

作为Google的核心技术之一，链接分析算法应用已经显现出巨大的商业价值。

本书作为英国信息科学专家迈克·塞沃尔(Mike Thelwall)教授的最新著作，从情报学的视角系统阐述了链接分析的理论、方法与应用，具体分六部分：概述、网络结构背景、学术型链接分析、链接分析的应用、链接分析的丁具和技术、总结。

从结构上看，本书逻辑严谨，条分缕析；从内容上看，本书资料翔实，通过大量的方法、工具、技术介绍及案例分析，向读者展示了整个链接分析分析研究的全貌；从表达上看，本书文字深入浅出，用较简洁的语言捕述了较复杂的理论与技术。

本书可供情报学、图书馆学、信息资源管理、信息管理与信息系统、计算机科学与技术等专业的师生课堂使用，同时，其翔实的资料可作为国内情报学、计算机科学、传播学、社会学等领域从事应用统计研究时的重要参考。

<<链接分析>>

书籍目录

第一部分 理论 第1章 前言 第2章 网络爬虫与搜索引擎 第3章 链接统计的理论基础 第4章 对链接数的解释：随机样本与相关性第二部分 Web结构 第5章 Web图中的链接结构 第6章 Web的内容结构第三部分 学术链接 第7章 大学：链接类型 第8章 大学：链接模型 第9章 大学：国际链接 第10章 院系和学科 第11章 期刊和论文第四部分 应用 第12章 搜索引擎与网站设计 第13章 西班牙大学网站健康度检验 第14章 链向大学网站的个人网页 第15章 学术网络 第16章 商业网站第五部分 工具和方法 第17章 商业搜索引擎和网络档案的使用 第18章 个人爬虫 第19章 数据清理 第20章 大学在线链接数据库 第21章 嵌入式链接分析方法 第22章 社会网络分析 第23章 网络可视化 第24章 学术链接指标第六部分 总结 第25章 总结 第26章 术语表附录：SocSciBot使用指南

<<链接分析>>

章节摘录

第1章 前言 目标 · 介绍本书的内容和结构，以及一些关键术语。

· 介绍信息科学研究方法——链接分析。

链接分析 链接分析在许多领域中有着广泛的应用，如计算机技术、理论物理、信息科学、传播学以及社会学等。

之所以能有这样广泛的应用，一方面是因为网络的重要性；另一方面是因为人们普遍认为：从网页之间的超链接中能够提取各种有用的信息。

这种认识主要源于一些相关的因素： Google的巨大成功，主要是利用一种基于链接的算法来判断网页的相关度； 期刊引用、社会人际关系等类似现象； 网络用户每天都面对各种用于研究、或用于商业、或用于娱乐的链接。

在本书中，笔者的主要目的是向新读者介绍什么是信息科学视角的链接分析。

之后，读者们就能够评价现有的研究，甚至从事自己的研究项目、形成自己的研究方法。

在本书中，笔者深信信息科学的方法对于其他学科的研究人员同样有着广泛的实用价值，尤其是那些对在线分析感兴趣的社会学家。

在研究过程中，如果将所有类型的链接分析都纳入考虑范围，必将一无所得，因为有些领域要求详细的数学算法，而另一些领域则只需要定性的分析。

本书中至少有一半的内容是研究学术网络或学术交流的，因此，读者们也可以对学术交流有一些深入的了解。

<<链接分析>>

编辑推荐

《链接分析：信息科学的研究方法》可供情报学、图书馆学、信息资源管理、信息管理与信息系统、计算机科学与技术等专业的师生课堂使用，同时，其翔实的资料可作为国内情报学、计算机科学、传播学、社会学等领域从事应用统计研究时的重要参考。

<<链接分析>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>