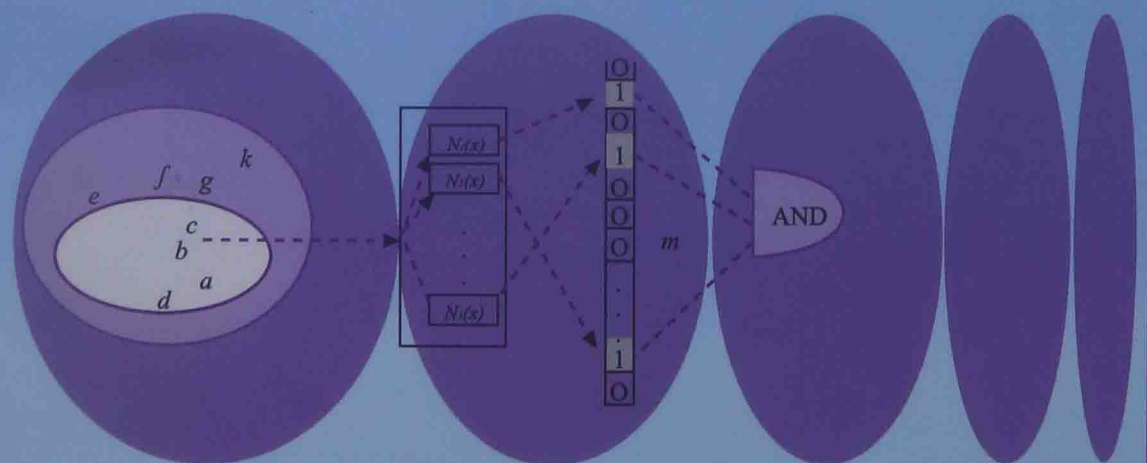


多布鲁姆过滤器

查询算法与应用研究

田小梅 谢鲲 著

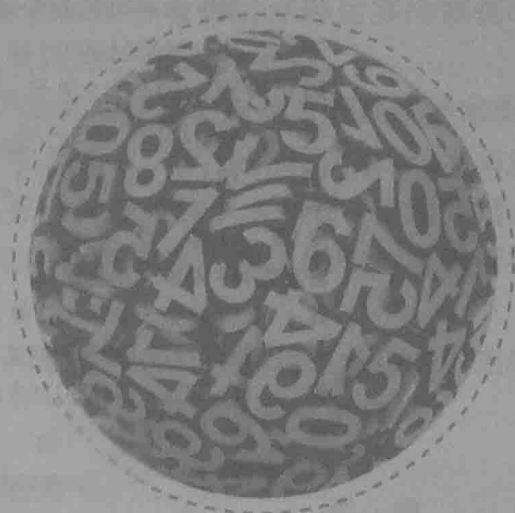
Duobulumu Guolüqi
Chaxun Suanfa Yu Yingyong Yanjiu



多布鲁姆过滤器 查询算法与应用研究

田小梅 谢鲲 著

Duobulumu Guolüqi
Chaxun Suanfa Yu Yingyong Yanjiu



湖南大学出版社

内 容 简 介

本书对多布鲁姆过滤器查询算法从理论分析和实际应用两个方面进行了深入的研究。首先分析网络高速发展给多布鲁姆过滤器查询算法带来的机遇,指出多布鲁姆过滤器查询算法研究的重要意义。然后,概括了多布鲁姆过滤器查询算法的研究现状和多布鲁姆过滤器查询算法目前的主要研究成果。考虑到单布鲁姆过滤器查询算法在解决分布式数据分发及数据同步等问题时不能完全胜任,本书提出了使用多个布鲁姆过滤器结构进行查询的数个多布鲁姆过滤器查询算法,如双布鲁姆过滤器直接查询算法、多标准布鲁姆过滤器代数运算查询算法、多计数布鲁姆过滤器代数运算查询算法,使用多标准布鲁姆过滤器进行查询的数据调和算法、使用多计数布鲁姆过滤器运算的数据调和算法及基于多标准布鲁姆过滤器运算的副本一致性维护算法等。

图书在版编目(CIP)数据

多布鲁姆过滤器查询算法与应用研究/田小梅,谢鲲著. —长沙:
湖南大学出版社, 2014. 8

ISBN 978 - 7 - 5667 - 0727 - 7

I. ①多… II. ①田… ②谢… III. ①计算机方法—研究
IV. ①TP301. 6

中国版本图书馆CIP数据核字(2014)第191204号



多布鲁姆过滤器查询算法与应用研究

DUOBULUMU GUOLVBUOYERJIAN SUANFA YU YINGYONG YANJIU

作 者: 田小梅 谢 鲲 著

责任编辑: 全 健 尚楠欣 责任印制: 陈 燕

印 装: 虎彩印艺股份有限公司

开 本: 787×1092 16开 印张: 9 字数: 219千

版 次: 2014年8月第1版 印次: 2014年8月第1次印刷

书 号: ISBN 978 - 7 - 5667 - 0727 - 7/O · 90

定 价: 22.00元

出 版 人: 雷 鸣

出版发行: 湖南大学出版社

社 址: 湖南·长沙·岳麓山 邮 编: 410082

电 话: 0731-88822559(发行部), 88821335(编辑室), 88821006(出版部)

传 真: 0731-88649312(发行部), 88822264(总编室)

网 址: <http://www.hnupress.com>

电子邮箱: presswanghj@hnu.deu.cn

版权所有, 盗版必究

湖南大学版图书凡有印装差错, 请与发行部联系

前 言

布鲁姆过滤器是一种表示集合的空间高效的有损数据结构,支持快速的数据成员查询,能有效地过滤不属于集合的成员。布鲁姆过滤器被广泛应用于数据库、网络和分布式系统,它在需要共享现有数据信息的分布式应用系统中有巨大的应用潜力。针对布鲁姆过滤器算法和应用的研究已被越来越多的研究团体所重视,涌现出了大量布鲁姆过滤器算法的变种及相关应用的研究论文,而且这种快速发展的势头还将持续下去,必定会出现更多布鲁姆过滤器算法的相关变种及应用研究。

通常情况下,使用布鲁姆过滤器的一般场景是:将集合 S 表示到布鲁姆过滤器这一精简结构中,在需要查询元素是否属于集合 S 时,使用布鲁姆过滤器而不是集合 S 本身进行集合成员查询,节约存储空间及提高查询的时间效率。目前大多数有关布鲁姆过滤器的扩展算法及应用研究主要是针对单个布鲁姆过滤器结构进行的。本书的研究工作则是考察如何使用多个布鲁姆过滤器结构进行相关查询,并将之用于分布式内容分发系统、数据同步系统、P2P 系统副本维护等领域。

本书对多布鲁姆过滤器查询算法从理论分析和实际应用两个方面进行了创新性的、深入的研究。首先分析网络高速发展给多布鲁姆过滤器查询算法带来的机遇,指出多布鲁姆过滤器查询算法研究的重要意义。然后,概括了多布鲁姆过滤器查询算法的研究现状和多布鲁姆过滤器查询算法目前的主要研究成果。考虑到单布鲁姆过滤器查询算法在解决分布式数据分发及数据同步等问题时不能完全胜任,作者提出了使用多个布鲁姆过滤器结构进行查询的数个多布鲁姆过滤器查询算法,如双布鲁姆过滤器直接查询算法、多标准布鲁姆过滤器代数运算查询算法、多计数布鲁姆过滤器代数运算查询算法、使用多个标准布鲁姆过滤器进行查询的数据调和算法、使用多计数布鲁姆过滤器运算的数据调和算法及基于多标准布鲁姆过滤器并运算的副本一致性维护算法。

显然,本书的研究工作不仅对布鲁姆过滤器查询算法理论研究具有重要的学术价值,而且对布鲁姆过滤器查询算法的应用具有重要的实践意义。从研究现状来看,本书应该是中国国内第一本系统研究多布鲁姆过滤器查询算法与应用的学术专著。正因为这是一个新的研究成果,更显出这本书在当前多布鲁姆过滤器查询算法理论研究和应用研究领域的代表性。

当然,这本书也还存在有很多不足,还有许多问题需作进一步探讨。希望它的出版能够起到抛砖引玉的效果,吸引更多志同道合的学者们参与到该研究领域中来。

作 者

2014 年 6 月 15 日

目 次

第 1 章 绪 论

1.1 海量数据的表示与查询	1
1.2 处理海量数据的布鲁姆过滤器	1
1.3 分布式系统中的多布鲁姆过滤器查询算法	2
1.4 多布鲁姆过滤器查询算法的研究现状	3
1.5 本书主要工作	4
1.6 本书结构与章节安排	6

第 2 章 多布鲁姆过滤器查询算法概述

2.1 布鲁姆过滤器查询算法	8
2.2 多布鲁姆过滤器查询算法的主要研究工作	22
2.3 现有多布鲁姆过滤器查询算法存在的待解决的问题	24
2.4 小 结	25

第 3 章 双布鲁姆过滤器直接查询算法

3.1 引 言	27
3.2 双布鲁姆过滤器直接查询算法的定义	27
3.3 双布鲁姆过滤器直接查询算法的性能分析	33
3.4 单布鲁姆过滤器或多布鲁姆过滤器直接查询算法的性能实验	38
3.5 双布鲁姆过滤器直接查询算法的应用探讨	40
3.6 小 结	41

第 4 章 多标准布鲁姆过滤器代数运算

4.1 引 言	43
4.2 标准布鲁姆过滤器形式化表示和相关定义	43
4.3 标准布鲁姆过滤器代数运算	44
4.4 标准布鲁姆过滤器代数运算和集合查询的关系	46
4.5 小 结	55

第 5 章 多计数布鲁姆过滤器代数运算

5.1	引言	56
5.2	布鲁姆过滤器查询算法的假阴性问题	57
5.3	计数布鲁姆过滤器代数运算的定义	58
5.4	计数布鲁姆过滤器代数运算和集合运算的关系	61
5.5	性能分析与模拟实验	67
5.6	计数布鲁姆过滤器代数运算的应用探讨	79
5.7	小结	80

第 6 章 基于多标准布鲁姆过滤器运算的精确集合调和算法

6.1	引言	82
6.2	集合调和	83
6.3	特征多项式插值调和法和 BFESR 法	85
6.4	BFESR 算法详细设计	88
6.5	算法比较	92
6.6	P2P 环境下的仿真实验评估	95
6.7	小结	98

第 7 章 基于多计数布鲁姆过滤器运算的精确集合调和算法

7.1	引言	99
7.2	计数布鲁姆过滤器减运算的查询性能	99
7.3	基于多计数布鲁姆过滤器运算的精确集合调和算法	100
7.4	各集合调和算法的性能分析与实验比较	102
7.5	CBFESR 的应用探讨与仿真实验	105
7.6	小结	110

第 8 章 基于多标准布鲁姆过滤器运算的 P2P 副本一致性维护算法

8.1	引言	111
8.2	P2P 副本更新相关工作	112
8.3	问题描述和相关定义	113
8.4	节点轨迹标签 P2P 副本一致性维护算法	115
8.5	算法理论分析	119
8.6	仿真实验和算法性能评价	120
8.7	小结	123

参考文献	124
------	-----

后 记	135
-----	-----

第1章 绪论

1.1 海量数据的表示与查询

随着物联网^[1]、下一代互联网^[2]、无线传感器网络^[3]、云计算^[4]等信息技术的发展,计算机网络在现代社会中的作用越来越举足轻重,已经成为人类社会重要的信息基础设施,并强力渗透到政治、经济、社会、军事、科技、文化、教育等各个领域。

计算机网络的广泛普及尤其是互联网的快速发展,促进了海量数据(mass data)的产生,例如海量的交互数据(互联网世界中人与人之间的交互信息、位置信息等)、海量的业务数据(企业内部的经营交易信息、物联网上的商品、物流信息等)^[5]等。根据中国互联网络信息中心(CNNIC)2014年1月16日发布的《第33次中国互联网络发展状况统计报告》^[6],截至2013年12月底,中国网民数量达到6.18亿,互联网普及率为45.8%。伴随着互联网的高速发展,社交网络逐渐成熟,移动带宽迅速提升,云计算、物联网应用更加丰富,更多的传感设备、移动终端接入到网络,由此产生的数据及增长速度迅速攀升,把人类社会带入一个以“PB”(Petabyte,1Petabyte=1 024TB)为计量单位的海量数据信息的时代。如Facebook用户数从2002年到2009年增加了20倍达到4.3亿,预计10年后社交网站用户将达到40亿。又如,国际数据公司(IDC)的研究结果表明,2011年全球产生的数据量高达1.82ZB(Zettabyte,1Zettabyte=1 024EB),相当于全球每人产生200GB以上的数据。作为对比,到2012年为止,人类生产的所有印刷材料的数据量是200PB,全人类历史上说过的所有话的数据量大约是5EB(Exabyte,1Exabyte=1 024PB)^[7]。

PB,EB乃至ZB数量级的海量数据给网络中数据信息的表示、存储、管理、定位和查询带来了新的挑战。如此海量的数据如何高效地得以表示,如何在海量数据中快速地查询到所需的数据信息等,这是网络设计者和网络服务提供商以及网络用户共同关注的问题。

数据信息的表示和查找是计算机网络和大多数应用程序的核心,是网络系统和分布式系统实现资源共享功能的基础。研究高效的数据表示及快速查询算法成为提升网络软件体系结构和进行大规模高效数据管理的关键。

1.2 处理海量数据的布鲁姆过滤器

布鲁姆过滤器(Bloom filter, BF)使用位向量简洁地表示数据集合,每个集合元素只需占用几个比特,支持数据元素的哈希查找,可快速判断一个元素是否属于这个集合。布

布鲁姆过滤器查询算法的空间效率和时间效率均较高。同时,布鲁姆过滤器作为一种表示数据集合的有损数据结构,存在误判的可能:某个不属于数据集合的元素可能被误判为属于该数据集合,这种现象称假阳性误判^[8],当然这种假阳性误判概率可以通过调整参数将其控制在很小的范围内。因此,在允许稍许误判的情况下,布鲁姆过滤器查询算法成为高效表示海量数据及快速查询海量数据的首选方案。

海量数据是发展趋势,布鲁姆过滤器查询算法作为一类较为成熟的支持高效存储与快速检索的技术,已被广泛应用于分布式数据库系统、分布式文件系统、P2P 网络系统等各类分布式系统中的海量数据处理^[9]。

1.3 分布式系统中的多布鲁姆过滤器查询算法

分布式系统在处理海量数据时,需要处理多个文件或集合,当将布鲁姆过滤器算法应用于这类系统时,往往需要对多个布鲁姆过滤器结构进行处理,如何高效地在多个布鲁姆过滤器结构上完成数据的查询,成为一个具有挑战性的课题领域。如 P2P 网络中的分布式数据分发、数据同步、多关键字检索、资源路由等应用领域,系统涉及多个布鲁姆过滤器结构的使用,但现有技术通常是针对单个布鲁姆过滤器结构进行查询,而不是针对多个布鲁姆过滤器结构进行查询。因此,探讨使用多个布鲁姆过滤器进行查询的算法即多布鲁姆过滤器查询算法,并将多布鲁姆过滤器查询算法用于内容分发、数据同步、副本维护等应用领域,成为本书的研究内容。

单布鲁姆过滤器查询算法使用单个布鲁姆过滤器结构进行查询,由于单个布鲁姆过滤器结构包含的信息仅限于单个集合的压缩信息,在需要使用多个集合的布鲁姆过滤器结构查询集合运算后的集合元素时,就遇到了限制。此时,可以考虑使用多布鲁姆过滤器查询算法来完成查询。下面从四个方面阐述多布鲁姆过滤器查询算法在分布式系统中的应用场景。

1. 内容分发网络

布鲁姆过滤器算法用于分布式文件分发、分布式存储系统等数据密集型应用中的内容分发,目的在于提高数据的存储空间效率及传输时间效率等^[10,11]。如在使用基于内容的发布/订阅范式和 P2P 技术的分布式数据分发系统中,使用布鲁姆过滤器用于减少数据存储空间及消息交换复杂度^[10]。又如,由于 P2P 系统中网络异构和网络波动(churn)的原因,使得大型数据文件分发时,各对等节点已接收内容存在着显著的差异,此时,各节点可以使用布鲁姆过滤器表示各自的已接收内容,以方便其他对等节点利用布鲁姆过滤器信息找出自身缺失的数据内容,并从目标节点处下载缺失内容,这一过程称为知情内容分发^[11]。在这些应用实例中,使用的是单个布鲁姆过滤器的查询算法,由于单布鲁姆过滤器查询算法固有的查询特性,如果不使用纠删码^[12]等辅助技术,算法将不能找出所有自身缺失的数据集合,即内容分发时有可能得不到完整的数据集合。因此,本书从多布鲁姆过滤器查询算法的角度寻求解决此类问题的方案。

2. 移动 P2P 网络中的数据同步

移动 P2P 计算领域中通常有两类数据同步方案:慢同步和快同步。慢同步是在移动

终端和台式电脑之间传输所有数据以同步的技术,由于终端间的实际差异往往远小于实际数据规模,因此该方法在带宽占用和时延方面效率较低。为改善带宽效率和时延效率,文献[13]等提出了两种基于 CPISync 的快同步技术方案:确定性方案和概率性方案。确定性方案是在确知数据集合的对称差规模的情况下使用的,但是,在移动计算领域,由于数据集合的对称差规模无法事先了解,只能使用概率性方案。通常概率性方案不能在单轮消息交换过程中完成数据同步,必须试探性地逐轮增加插值点个数,这样就可能造成多轮数据交换导致数据同步的时延过长。由于使用两个布鲁姆过滤器可估算交集的规模,也就能估算出对称差规模,本书的工作内容也包括使用多布鲁姆过滤器查询算法来完成数据同步,以提高这类概率性方案的同步时延效率。

3. 多关键字检索

当前的结构化 P2P 网络只支持单个关键字的精确匹配搜索。可是,在很多情况下,人们并不能给出准确描述拟搜索对象的关键字,而只能给出拟搜索对象的多个大致特征描述词,即需要进行模糊的多关键字搜索。对于结构化 P2P 网络系统来说,系统中的 DHT 算法通常只适合单关键字的准确匹配,如果要支持多关键字查找的功能,还要引入新的方法,如文献[14]在基于 DHT 的 P2P 搜索引擎中使用了分布式倒排索引技术和布鲁姆过滤器算法,以实现多关键字检索。但是,在资源搜索效率等方面还有进一步改进的空间。譬如,可以使用多个布鲁姆过滤器的并、交运算进一步减少 P2P 系统多关键字搜索时“或查询”、“与查询”的信息传输量,提高检索效率及使用多布鲁姆过滤器的减运算来支持“减查询”(sub query)等。

4. 资源路由

资源路由(resource routing)的通用框架^[15]如下:将网络看成一棵有根树,网络中节点拥有各自的资源。请求节点发出资源请求消息,并向根部路由。每个节点的资源路由表包括一个统一资源列表、一个本地节点的单独资源列表和每个孩子节点的单独资源列表,统一资源列表中包含本地节点资源和通过其孩子节点可达资源的总和。当节点接收到资源请求消息,首先检查统一资源列表是否存在该资源的查找路径。如果得到的结果是肯定的,则检查单独列表,以发现持有资源的节点;如果统一资源列表中查询不到请求资源,则将资源请求消息向树根方向传递。当然,资源路由不仅限于树型网络拓扑结构使用,也可用于其他结构的网络,如无结构 P2P 网络、结构化 P2P 网络或混合 P2P 网络等。尤其是资源列表使用标准布鲁姆过滤器表示时,资源路由表中的统一资源列表可以使用多个标准布鲁姆过滤器的并运算来完成计算^[16],非常方便且高效。

可见,对多布鲁姆过滤器查询算法及其应用开展研究,不仅可提升分布式系统的检索性能、数据分发与同步效率、路由速率,进一步拓宽布鲁姆过滤器查询算法的应用领域,而且可促进 IT 技术的更进一步发展。

1.4 多布鲁姆过滤器查询算法的研究现状

布鲁姆过滤器算法自 1970 年由 Bloom 提出以来,被广泛用于各类计算机系统,如数据库系统、文件处理系统、网络系统等,以表示海量数据集,提高查询处理速度。早期,主

要应用于数据字典^[8,17~20]、数据库操作和文件操作^[21~29]方面,随着网络技术的发展,对于布鲁姆过滤器查询算法的研究重点越来越集中于网络系统及分布式系统的应用研究^[9],如路由查找^[30~33]、数据包检测及分类^[34~38]、网络监测与测量^[39~45]、关键字检索^[14,46,47]、P2P 路由协议^[16,48]、内容分发^[11]、数据同步^[13,49~51]、Ad hoc 网络信息共享^[52]、基因组测序^[53,54]、网络安全^[55~63]、分布式缓存与存储系统^[64~67]及云计算^[29,68]等。

尽管布鲁姆过滤器的算法与应用研究得到了很多研究人员的重视,针对不同的应用场景产生了布鲁姆过滤器的各种扩展算法,拓展了布鲁姆过滤器算法的应用范围,但是目前的研究主要是针对单布鲁姆过滤器查询算法及其应用的研究,对于多布鲁姆过滤器查询算法的研究工作相对较少,应用范围也相对较窄,主要应用是将多布鲁姆过滤器查询算法用于环路检测^[69]、数据排重^[70]、数据包分类^[71,72]、IP 查找^[73,74]、关键字检索^[75]、资源路由^[16]等问题。现有的这些多布鲁姆过滤器查询算法的研究,主要是针对具体的应用环境做的工作,理论工作并不多;而且,在将多布鲁姆过滤器查询算法应用到具体问题时,通常使用的是多布鲁姆过滤器的并运算(即“或”运算)或多个布鲁姆过滤器的并行查询操作,对于其他运算的应用讨论较少,因此,在多布鲁姆过滤器查询算法领域还存在不少的研究空间。

1.5 本书主要工作

本书研究工作针对的主要问题是:在无法获知集合的并集、交集、补集、差集或对称差,也无法得到并集、交集、补集、差集或对称差对应的布鲁姆过滤器表示时,高效地使用两个或多个集合的布鲁姆过滤器表示来查询元素是否属于集合的并集、交集、补集、差集或对称差的成员,即研究多布鲁姆过滤器查询算法的性能及将多布鲁姆过滤器查询算法用于解决分布式系统中的文件分发、数据同步、副本维护等问题。本书工作具体包括以下几个方面的内容。

(1)探讨了双布鲁姆过滤器直接查询法进行并集、交集、补集、差集或对称差的成员查询时的假阳性及假阴性问题。

双布鲁姆过滤器直接查询法直接使用两个集合的布鲁姆过滤器表示来查询元素是否属于集合的并集、交集、补集、差集或对称差的成员。理论分析和实验结果表明,双布鲁姆过滤器直接查询法能够较好地支持集合的并集、交集、补集、差集及对称差的成员查询问题,其中双布鲁姆过滤器直接查询法查询并集及交集不会产生假阴性,仅有少量假阳性的存在,在很多使用布鲁姆过滤器的场合中,少量假阳性的存在对算法的影响通常较小,其应用场合不会受到影响。而使用双布鲁姆过滤器直接查询法进行补集、差集及对称差查询则除存在少量假阳性外,还存在少量假阴性,即部分本来属于补集、差集或对称差的元素被判为不属于补集、差集或对称差。通常,假阴性问题是在应用布鲁姆过滤器解决实际问题时应该尽量避免的,此时可以采用其他技术予以弥补,如知情内容分发系统^[11]中使用纠删码技术来恢复被遗漏的差集元素。

(2)研究了标准布鲁姆过滤器的代数运算和集合运算的一致性关系,及使用标准布鲁姆过滤器代数运算进行集合成员查询的性能问题。

使用双布鲁姆过滤器直接查询法进行并集、交集、补集、差集或对称差的成员查询时,

每次完成成员查询都需对两个或多个布鲁姆过滤器向量进行查询,存储多个布鲁姆过滤器向量比存储单个布鲁姆过滤器向量会增加算法的空间复杂度。标准布鲁姆过滤器表示的是集合的近似结构,集合能进行并、交、补、减等运算,那么标准布鲁姆过滤器是否能支持并、交、补、减、异或等运算?如果能,是否能用于快速查询集合的并集、交集、补集、差集及对称差中的元素?

本书第4章探讨了标准布鲁姆过滤器的布尔代数,定义了标准布鲁姆过滤器的并、交、异或、补、减等代数运算,全面探讨标准布鲁姆过滤器代数运算和集合查询的关系。证明了布鲁姆过滤器的并、交仍然可以支持集合元素查询,而其他运算由于引入过多的假阴性误判而不再支持集合元素查询。

(3)研究了计数布鲁姆过滤器的代数运算和集合运算的一致性关系,及使用计数布鲁姆过滤器代数运算进行集成员查询的性能问题。

多标准布鲁姆过滤器代数运算的运算对象是标准布鲁姆过滤器等使用位向量结构的布鲁姆过滤器,而且其中的补运算、减运算及对称差运算在进行补集、差集及对称差查询时,存在假阴性,为解决这一问题,笔者从多个计数布鲁姆过滤器向量运算的角度寻求解决方法。我们对其进行理论分析和仿真实验,结果表明,计数布鲁姆过滤器的并、交、补、减、异或运算产生的新过滤器依然保持计数布鲁姆过滤器的特征,支持元素的删除操作,不会出现假阴性,能用于集合的并集、交集、补集、差集及对称差的成员查询。与双布鲁姆过滤器直接查询法相比,使用计数布鲁姆过滤器代数运算后的过滤器进行补集、差集及对称差成员查询,不存在前述假阴性问题,空间效率能提高一倍,时间效率亦能显著地得到改善。计数布鲁姆过滤器代数运算的使用有利于进一步扩展计数布鲁姆过滤器的应用范围。

(4)提出了对远程集合的对称差规模估算准确程度高的基于布鲁姆过滤器的概率性方案——准交集查询法。

与目前现有的远程对称差规模估算方法——内积法相比,准交集查询法的对称差规模估算精度比内积法高出较多,而且,内积法各布鲁姆过滤器向量空间大小的设置需相同,对于较小的集合规模,会浪费较多的存储空间;而准交集查询法则无此限制,各节点可以根据自己的集合规模选择较合适的布鲁姆过滤器向量空间大小,提高空间的存储效率。

(5)提出了基于多标准布鲁姆过滤器运算的精确集合调和的方法——BFESR。

目前已有的集合调和的方法中,近似集合调和的方法实现简单,仅需单轮消息交换,但却无法得到全部差集元素;而精确集合调和的方法,如特征多项式插值精确集合调和法^[76,77](characteristic polynomial interpolation-based synchronization, CPISync) 尽管能得到所有差集元素,但却存在算法复杂,需多轮消息交换的问题。基于多标准布鲁姆过滤器运算的精确集合调和既具有精确集合调和能得到全部差集元素的优点,也具有近似集合调和仅需单轮消息交换的优点。该方法首先使用标准布鲁姆过滤器的内积运算或准交集查询法估算出对称差规模;然后以逐轮增加的求值点和特征多项式值作为特征多项式插值算法的输入,重复调用插值算法,直至确认成功;最后进行因式分解得到差集元素,进而获取并集完成调和。通常,BFESR调和过程中,调用一次特征多项式插值算法即能成功。与已有的试探法进行比较,BFESR调和时间和消息交换轮数降低非常明显,尤其是使用准

交集查询法估算对称差规模的 BFESR 方法,其调和效率更高,通常情况下仅需单轮消息交换即可找出全部差集元素,完成准确的数据集合调和。

(6) 提出基于多计数布鲁姆过滤器运算的精确集合调和方法——CBFESR。

由于 BFESR 算法中使用的标准布鲁姆过滤器不支持集合元素的动态更新,若用于更新频繁的 P2P 网络等分布式系统则需要定时重建标准布鲁姆过滤器,这样会增加系统实现的负担及难度。因此,为解决 BFESR 调和算法的这一应用局限性,笔者提出了一种基于多计数布鲁姆过滤器运算的集合调和方法,该方法将集合用计数布鲁姆过滤器表示,利用计数布鲁姆过滤器减运算得到的新过滤器,查询并获得集合中的差集元素,再用差集和自身集合进行集合并运算,完成集合调和。CBFESR 方法仅需单轮消息交换即能找出全部差集元素,实现准确调和,而且计算简单,能支持集合元素的动态删除,适合于更新频繁的系统环境。

(7) 基于布鲁姆过滤器的轨迹标签无结构 P2P 副本一致性维护算法。

由于以往的研究多认为 P2P 系统资源是静态资源,所以在 P2P 副本一致性维护方面的研究开展不多。已经存在的无结构 P2P 网络副本一致性维护方法,要么传输消息的冗余量大,要么需要额外构造表现网络的拓扑信息结构来辅助更新消息传输。参数确定困难,方法实现受一定限制,判断更新消息是否冗余是在消息报文传播后,不能在传输的源头进行控制,不便于 P2P 系统的扩展。

针对目前无结构 P2P 副本一致性维护方法对已经更新过的节点没有记忆性而产生大量的消息冗余这一缺陷,提出一种基于布鲁姆过滤器的节点轨迹标签无结构 P2P 副本一致性维护算法。该方法从更改传播消息报文的角度出发,在更新消息传送过程中,将每轮传输的目标节点集添加到消息报文头部做记录,通过记录已经获得更新的节点,来阻止消息报文在已传播的节点中再次传播。对已经获得更新的节点进行记录,消息传输时,通过对邻居节点进行地址查询检查,完成对冗余消息的提前检测,减少消息的冗余传输。同时为了减少附加到报文中的节点地址的长度,用布鲁姆过滤器直接表示地址链表的轨迹标签,利用布鲁姆过滤器这种简洁的数据结构表示地址链表,有效减少了附加到消息报文的地址信息长度,同时利用布鲁姆过滤器的代数运算,简化传输节点的冗余判断。新的 P2P 副本一致性维护方法可大大降低冗余消息数目,提高 P2P 系统的可扩展性,副本节点网络连通性越强,消息数目和传输带宽减少越明显。

这是多布鲁姆过滤器查询算法在 P2P 副本一致性维护中的应用实例。实例中的算法思想可以进一步应用到:无结构 P2P 副本维护、无线网络资源副本维护、复杂网络病毒免疫传输维护等其他自组织网络或分布式存储系统资源一致性维护中或其他的需要记录 Trace 的场合。

1.6 本书结构与章节安排

本书共分八章,具体内容安排如下:

第 1 章,绪论。阐述高速发展的网络所产生的海量数据给数据的表示和查询带来的新的挑战,指出多布鲁姆过滤器查询算法研究的重要意义,阐述本书的主要研究工作。

第 2 章,多布鲁姆过滤器查询算法概述。首先详细介绍标准布鲁姆过滤器的工作原

理与性能分析,接着简要介绍布鲁姆过滤器查询算法的典型扩展算法,阐述了多布鲁姆过滤器查询算法的应用及主要研究成果,分析了多布鲁姆过滤器查询算法目前存在的待解决的问题,指出了本书的研究方向。

第3章,双布鲁姆过滤器直接查询算法。它使用两个独立的布鲁姆过滤器结构直接查询集合的并集、交集、补集、差集或对称差成员,分析了双布鲁姆过滤器直接查询算法的假阳性及假阴性问题,对双布鲁姆过滤器直接查询算法的性能进行了实验验证,并探讨了双布鲁姆过滤器直接查询算法的应用范围。

第4章,标准布鲁姆过滤器代数运算。讨论了集合代数运算和标准布鲁姆过滤器代数运算的关系,得到通过标准布鲁姆过滤器代数运算直接支持集合查询的使用技巧。

第5章,多计数布鲁姆过滤器代数运算。首先定义了计数布鲁姆过滤器的并、交、减、补、异或运算,对每种代数运算作了详尽的理论分析与证明,并使用模拟实验对理论结果进行了证实。通过理论和实验比较了多计数布鲁姆过滤器代数运算和双布鲁姆过滤器直接查询算法的性能,并讨论了计数布鲁姆过滤器代数运算的应用范围。

第6章,基于多标准布鲁姆过滤器运算的精确集合调和算法。该算法首先使用布鲁姆过滤器估算出远程集合的对称差规模,然后再调用 CPISync 算法,完成精确集合调和。该方法中最重要的要估算出远程集合的对称差规模,对称差规模估算的准确程度直接影响算法的执行效率。针对现有估算方法(使用多布鲁姆过滤器的内积运算,称内积法)对称差规模的估算准确程度不够高的缺点,本书也提出了一种估算对称差规模的新方法——准交集查询法,并将之用于集合调和。实验结果证实,基于多标准布鲁姆过滤器运算的精确集合调和大多数情况下仅需1轮消息交换即可完成集合调和,调和时间降低较为显著。

第7章,基于多计数布鲁姆过滤器运算的精确集合调和算法。由于 BFESR 算法中使用的标准布鲁姆过滤器不支持集合元素的删除操作,对于数据集更新频繁的 P2P 网络等分布式系统来说不太适合应用,因此,为解决 BFESR 调和算法的这一应用局限性,本书提出了使用多计数布鲁姆过滤器运算的集合调和算法(CBFESR 算法)。首先,对使用计数布鲁姆过滤器减运算完成集合调和的可行性及优越性进行了详细的分析与论证。然后使用模拟实验和仿真实验证实基于多计数布鲁姆过滤器运算的精确集合调和算法同时具有近似集合调和和精确集合调和方法的优点,仅需单轮消息交换即能找出全部差集元素,实现准确调和。

第8章,基于多标准布鲁姆过滤器运算的 P2P 副本一致性维护算法。针对无结构 P2P 网络副本一致性维护问题,从直接更改消息报文角度出发,提出一种基于布鲁姆过滤器运算的节点轨迹标签无结构 P2P 副本一致性维护算法,通过在传输消息的报文中添加已接收更新消息的节点轨迹地址链表标签,可在消息传输源节点进行冗余判断来减少冗余消息数目。针对直接存储节点地址轨迹标签算法的消息长度随着消息传输轮数和网络度数增加,算法中地址链表轨迹标签采用布鲁姆过滤器表示,可以减少添加到报文中的轨迹长度,利用布鲁姆过滤器并运算简化传输节点的冗余判断。实验结果表明,基于布鲁姆过滤器的节点轨迹标签算法可以大大降低冗余消息数目,提高 P2P 系统的可扩展性。副本节点网络连通性越强,消息数目和传输带宽减少越明显。

第 2 章 多布鲁姆过滤器查询算法概述

2.1 布鲁姆过滤器查询算法

2.1.1 标准布鲁姆过滤器结构与操作

标准布鲁姆过滤器 (Bloom filter, BF)^[8] 结构的组成如下: 1 个长度为 m 的位向量 $\mathbf{BF}(b_1, b_2, b_3, \dots, b_m)$, 1 个由 k 个哈希函数 h_1, h_2, \dots, h_k 组成的哈希函数组和 1 个数据集 $S = \{s_1, s_2, \dots, s_n\}$ 。通常, k 个哈希函数相互独立, 取值为 $0 \sim (m-1)$ 之间的整数。在插入 S 中元素之前, 该位向量所有二进制位均为 0。集合 S 中所有数据元素插入到位向量 $\mathbf{BF}(b_1, b_2, b_3, \dots, b_m)$ 后, 该位向量结构称为 S 的布鲁姆过滤器表示, 记为 $\mathbf{BF}(S)$ 。

初始化: 在初始化位向量时, \mathbf{BF} 的每个二进制位都为 0。

元素的插入: 当将元素 x 插入 \mathbf{BF} 时, 根据哈希函数计算 x 对应的 k 个哈希地址 $h_1(x), h_2(x), \dots, h_k(x)$, 并将 \mathbf{BF} 向量的该 k 个二进制位置为 1。

成员查询: 若要查询给定的元素 y 是否属于集合 S , 则首先进行哈希函数的计算, 求出 $h_1(y), h_2(y), \dots, h_k(y)$, 根据位向量 $\mathbf{BF}(S)$ 的第 $h_1(y), h_2(y), \dots, h_k(y)$ 个二进制位是否全为 1, 判定 y 是否属于集合 S 。以图 2.1 所示的 \mathbf{BF} 位向量为例, 表示的 $S = \{x, y, z\}$, 查询 w 是否属于集合 S 时, 由于 w 对应的哈希地址中有一个二进制位为 0, 从而判定 w 不在 \mathbf{BF} 所表示的集合 S 中。而查询 z 时, z 对应的所有二进制位均为 1, 从而 z 被判定为集合 S 的成员。

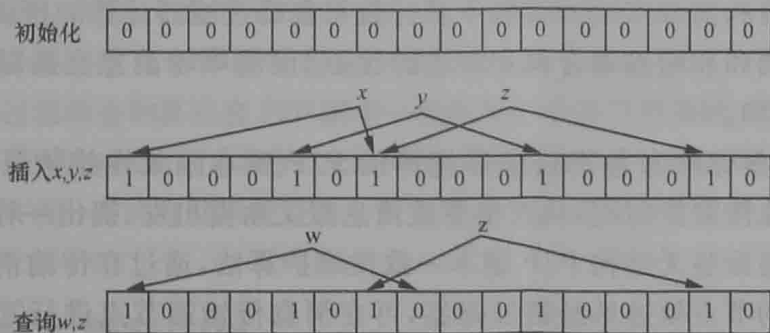


图 2.1 标准布鲁姆过滤器算法

使用 $\mathbf{BF}(S)$ 判定元素是否集合 S 的成员, 如果结果为否定的, 那么该元素必定不是集合 S 的成员; 如果结果为肯定的, 那么该元素或者确实是集合 S 的成员, 或者是一个假阳性 (false positive), 即元素原本不在集合 S 中被判为集合 S 的成员。

2.1.2 标准布鲁姆过滤器查询算法性能分析

标准布鲁姆过滤器存在误判的可能,如图 2.2 所示,不在集合 S 中的元素 g 使用标准布鲁姆过滤器进行成员查询时,被误判为集合 S 的成员,即出现了假阳性误判,这种误判概率称假阳性概率(false positive probability)。此时, g 便是一个假阳性。由标准布鲁姆过滤器的插入操作特性可知,对于集合 S 中的成员,在查询时是不会将其判为不属于集合 S 的,即不会出现假阴性(false negative)。

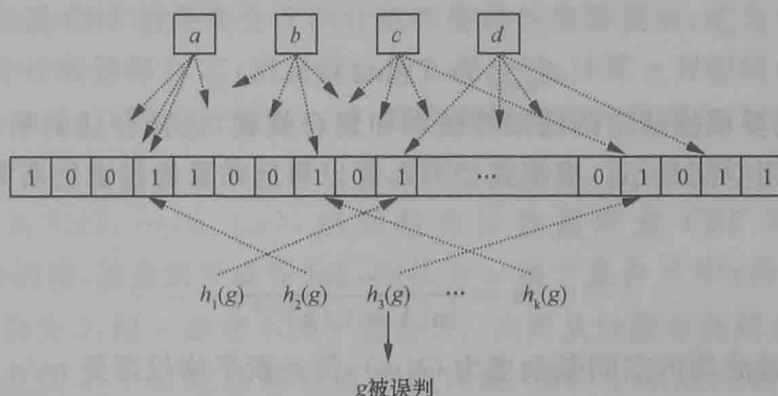


图 2.2 标准布鲁姆过滤器的假阳性误判实例

布鲁姆过滤器假阳性概率的大小直接影响查询算法的准确程度,当假阳性概率为 1% 时,说明使用布鲁姆过滤器查询算法被判为集合 S 成员的元素中有 1% 的可能是假阳性,其实它并不在 S 中,99% 的可能是它确实是集合 S 的成员。因此,设计布鲁姆过滤器时,首先应该关注的是布鲁姆过滤器的假阳性概率。

布鲁姆过滤器的假阳性概率计算公式为^[8,30]:

$$fp = (1 - p)^k = \left[1 - \left(1 - \frac{1}{m} \right)^{kn} \right]^k \approx (1 - e^{-kn/m})^k \quad (2.1)$$

由式(2.1)可看出,我们可以认为 fp 取决于两个参数: m/n 和 k , m/n 称为每元素比特数。当 m/n 取值固定时,根据式(2.1)可画出 fp 随 k 值变化的趋势图,如图 2.3 所示。

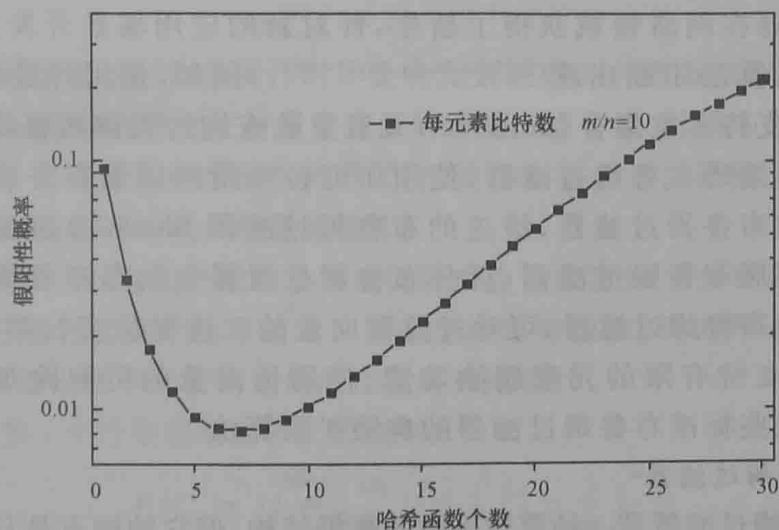


图 2.3 布鲁姆过滤器假阳性概率随哈希函数个数变化趋势

由图 2.3 可看到, k 值并不是越大越好, 也不是越小越好, 有一个最优 k 值 k_{opt} 使得布鲁姆过滤器假阳性概率 fp 最低。如, 当 $k=7$ 时, $fp_{min}=0.00819$, 即 $k_{opt}=7$ 。

理论最优 k 值 k_{opt} 的计算公式为:

$$k_{opt} = \ln 2 \times (m/n) \quad (2.2)$$

此时, 假阳性概率的最小值为:

$$fp_{min} = (1/2)^{k_{opt}} \approx (0.6185)^{m/n} \quad (2.3)$$

在实际应用中, k 值应为整数, 为减少哈希函数的计算量, 通常取式(2.2)的下整值作为次优 k 值, 即

$$k_{subopt} = \lfloor (\ln 2)(m/n) \rfloor \quad (2.4)$$

有时, 我们需要根据给定的假阳性概率和集合规模, 选择合适的哈希函数个数, 使得布鲁姆过滤器占用空间最小。根据式(2.1), 可以得出布鲁姆过滤器所需存储空间计算公式:

$$m = \frac{-kn}{\ln[1 - (fp)^{1/k}]} \quad (2.5)$$

标准布鲁姆过滤器的空间复杂度为 $O(m)$, 每元素平均仅需要 m/n 位比特, 与元素本身的大小无关, 其空间效率较高。在标准布鲁姆过滤器中, 插入与查询操作都需要计算 k 次哈希函数及访问位向量的 k 个比特, 因此它们的时间复杂度均为 $O(k)$, 与位向量空间大小 m 和布鲁姆过滤器中表示的集合规模 n 无关。而且, 由于布鲁姆过滤器所用的哈希函数通常是相互独立的, 因此, 哈希函数的计算及位向量的访问非常适合于并行实现, 从而更进一步提高布鲁姆过滤器的时间效率。

总的来说, 标准布鲁姆过滤器的优势在于空间占用较少和查询速度快, 非常适合于存储空间受到限制又能容忍少许误判的应用场合。

2.1.3 标准布鲁姆过滤器的典型扩展算法

自 Burton Bloom 在 20 世纪 70 年代提出标准布鲁姆过滤器之后, 布鲁姆过滤器被广泛用于拼写检查和数据库系统中。近一二十年, 伴随着计算机网络的普及和发展, 布鲁姆过滤器在网络领域获得了新生, 针对新的应用场景开发的各类标准布鲁姆过滤器的扩展算法不断出现^[8,30,34,56,64,73,78~99]。例如, 使用计数器向量的计数布鲁姆过滤器, 能支持多重集合(multi-set)元素重数查询的光谱布鲁姆过滤器、动态计数过滤器和空间编码布鲁姆过滤器, 使用压缩位向量的压缩布鲁姆过滤器、能表示多维元素的多维布鲁姆过滤器, 级连的布鲁姆过滤器 Bloomier 过滤器, 向量空间可动态伸展的可扩展布鲁姆过滤器、拆分布鲁姆过滤器和动态布鲁姆过滤器, 可反向列举元素的可逆布鲁姆过滤器, 可对过滤器向量的二进制位实行清零操作的广义布鲁姆过滤器, 能支持有限的元素删除功能、使用位向量的可删除布鲁姆过滤器等。下面逐一介绍这些标准布鲁姆过滤器的典型扩展算法。

1. 计数布鲁姆过滤器

标准的布鲁姆过滤器是一种简洁高效的数据结构, 但它的缺点是只支持插入和查询两种操作, 而不能有效地支持删除操作。在所表达的集合是静态集合或者集合元素只

增加不减少的时候,标准布鲁姆过滤器可以很好地工作,但是如果表达的集合成员经常变动,标准布鲁姆过滤器因为不支持删除操作,从而不适合该应用场合。计数布鲁姆过滤器^[64,95]很好地解决了上述问题。

计数布鲁姆过滤器将标准布鲁姆过滤器位向量的每一位扩展为一个 r 位的计数器,从而成为一个计数器向量。在插入元素时将对应的 k 个计数器的值分别加 1,删除元素时将对应的 k 个计数器的值分别减 1。 n 个元素的数据集合 $S = \{s_1, s_2, \dots, s_n\}$,通过 k 个哈希函数 h_1, h_2, \dots, h_k 映射到由 m 个计数器组成的计数器向量 $CBF = (C_1, C_2, C_3, \dots, C_m)$ 中。计数器向量 CBF 就是集合 S 的计数布鲁姆过滤器表示,记为 $CBF(S)$ 。在初始化时, CBF 的每个计数器都为 0。当元素 x 插入集合时,计算 x 对应的 k 个计数器索引值 $h_1(x), h_2(x), \dots, h_k(x)$ (即哈希地址),将 CBF 向量相应计数器加 1 ($C_{h_j(x)}++$, $1 \leq j \leq k$),完成元素的插入。查询给定的元素 x 是否属于集合 S ,首先计算 x 对应的 k 个计数器索引值 $h_1(x), h_2(x), \dots, h_k(x)$,然后检查计数器向量 CBF 对应 k 个计数器 ($C_{h_j(x)}$, $1 \leq j \leq k$) 的值,若全大于或等于 1,则认为 x 属于集合 S 中;否则,只要此 k 个计数器中任一计数器为 0,则 x 必定不属于集合 S 。当要从计数布鲁姆过滤器删除元素 x 时,计算出 x 对应的 k 个计数器索引值后将向量 CBF 中相应计数器进行减 1 运算 ($C_{h_j(x)}--$, $1 \leq j \leq k$) 即可。

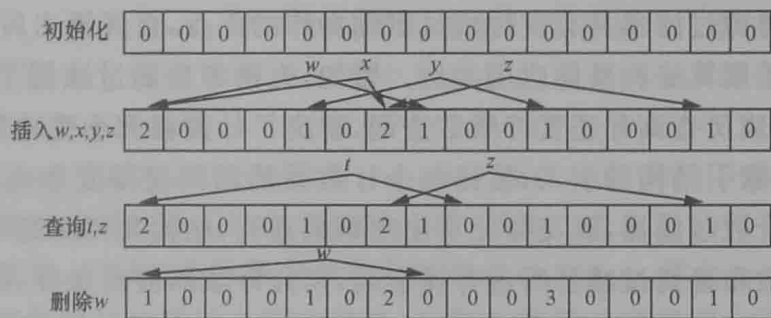


图 2.4 计数布鲁姆过滤器算法

计数布鲁姆过滤器同样存在假阳性,其假阳性概率与标准布鲁姆过滤器的相同,为:

$$fp_{m,n,k}^{CBF} = (1 - e^{-kn/m})^k \quad (2.6)$$

计数布鲁姆过滤器执行删除操作时不是将元素映射的计数器采取清零操作,而是将相应计数器减 1,从而能方便地支持集合成员的增减。或者说,计数布鲁姆过滤器通过占用 r 倍于标准布鲁姆过滤器的存储空间代价,使得布鲁姆过滤器除了支持集合成员的插入、查询操作外,还能支持成员的删除操作。

因为在元素插入时,会使计数布鲁姆过滤器对应计数器加 1,而计数器是由有限的比特位组成,因此可能会造成计数器溢出。为避免计数器溢出,应选择适当的计数器位数 r 。下面通过计算溢出概率来获得通用的计数器长度设计。

我们先计算第 i 个计数器增加 j 次的概率:

$$P(C_i = j) = \binom{nk}{j} \left(\frac{1}{m}\right)^j \left(1 - \frac{1}{m}\right)^{nk-j} \quad (2.7)$$

上式右端的表达式中,第一个分量表示从 nk 次哈希中选择 j 次的组合数,第二个分