

【国家自然科学基金项目】

无结构对等网络中 的搜索算法与安全机制

WUJIEGOU DUIDENG WANGLUO ZHONG
DE SOUSUO SUANFA YU ANQUAN JIZHI

秦志光 罗绪成 马新新 编著



电子科技大学出版社

【国家自然科学基金项目】

无结构对等网络中的搜索算法与安全机制

WUJIEGOU DUIDENG WANGLUO ZHONG
DE SOUSUO SUANFA YU ANQUAN JIZHI

秦志光 罗绪成 马新新 编著

国家自然科学基金项目

无结构对等网络中的搜索算法与安全机制

秦志光 罗绪成 马新新 编著

地址：电子科技大学出版社（成都市一环路东一段151号电子信息产业开发区）

邮编：610021

责任编辑：曾志

封面设计：曾志

网址：www.wuestp.com.cn

电子邮箱：westcp@westcp.com.cn

发行：新华书店

印刷：电子科技大学出版社

版次：2008年12月第1版

印次：2008年12月第1次印刷

开本：787mm×1092mm

字数：120千字

定价：28.00元

版权所有 侵权必究

本书定价：28.00元/册，零售：25.00元/册，邮费：2.00元/册

本书定价：28.00元/册，零售：25.00元/册，邮费：2.00元/册



电子科技大学出版社

图书在版编目 (CIP) 数据

无结构对等网络中的搜索算法与安全机制/秦志光, 罗绪成,
马新新编著. —成都: 电子科技大学出版社, 2008. 12

国家自然科学基金项目

ISBN 978-7-5647-0036-2

I. 无… II. ①秦…②罗…③马… III. 计算机网络—安
全技术—研究 IV. TP393.08

中国版本图书馆 CIP 数据核字 (2008) 第 206719 号

国家自然科学基金项目

无结构对等网络中的搜索算法与安全机制

秦志光 罗绪成 马新新 编著

出 版: 电子科技大学出版社 (成都市一环路东一段 159 号电子信息产业大厦
邮编: 610051)
策划编辑: 曾 艺
责任编辑: 曾 艺
主 页: www.uestcp.com.cn
电子邮箱: uestcp@uestcp.com.cn
发 行: 新华书店经销
印 刷: 郫县犀浦印刷厂
成品尺寸: 170mm×240mm 印张 7.5 字数 150 千字
版 次: 2008 年 12 月第一版
印 次: 2008 年 12 月第一次印刷
书 号: ISBN 978-7-5647-0036-2
定 价: 29.00 元

■ 版权所有 侵权必究 ■

- ◆ 本社发行部电话: 028-83202463; 本社邮购电话: 028-83208003。
- ◆ 本书如有缺页、破损、装订错误, 请寄回印刷厂调换。

前 言

对等网络是近年来发展起来的一种具有较高扩展性的分布式系统结构，构建在应用层上的逻辑网络，因此也称之为重叠网络，是按照一定的逻辑拓扑结构将系统中的节点互联起来，通过寻址路由使得系统中任意两个节点直接通信的一种组织方法。网络中节点在功能上具有同等的地位，当节点动态加入或退出时，能够保证节点间的互联性依旧具有很好的灵活性和扩展性。相对于传统的集中式客户/服务器(C/S)模式，P2P 模式弱化了服务器担负中心角色的作用，充分利用网络终端节点的计算处理能力和资源，中央服务器承担的计算任务分布于网络终端节点完成，每个节点既可请求服务，也可提供服务。节点之间不必通过服务器就可以直接交换资源或提供、请求服务，实现系统中任意两个节点之间的直接通信，将互联网的集中管理模式引向分散管理模式。目前 P2P 技术的应用主要包括资源文件的共享和存储如 BT，分布式数据存储如 OceanStore，北大的 Maze，大规模并行计算如 SETHome 等。P2P 技术在得到广泛应用的同时，其自身的特性如动态性、匿名性等造成网络中节点行为的不可信，共享资源内容的不可靠，成为阻碍 P2P 技术应用进一步发展的障碍，而对这些存在问题的探究是 P2P 安全研究的热点方向。

本书通过对 P2P 领域应用技术和安全问题的介绍，并基于作者近年在该领域的研究体会和收获，旨在对将涉足该领域进行研究的初学者，以及在该领域的同行提供一个相互交流学习的机会，为彼此进一步深入该领域的研究奠定基础。本书内容主要分为 P2P 系统的体系结构和 P2P 系统的安全问题两部分。在体系结构方面，主要介绍无结构化 P2P 系统、结构化 P2P 系统和混合式 P2P 系统；在 P2P 系统的安全方面，通过从 P2P 安全问题存在的本质是节点间无明确的信任关系这点入手对 P2P 系统的信任和信誉机制进行分析、描述和研究。书中所介绍、分析和描述的模型、算法、协议主要基于国内外已发表的论文成果，所介绍的应用系统具有代表性，选择时考虑了所选对象的影响力和实际应用价值。

由于篇幅和时间的限制，本书所介绍内容的深度和广度都是有限的。对等计算是一个蓬勃发展的领域，新的算法、协议、模型不断出现，应用领域不断拓宽，我们期待着为读者奉献更高级的对等计算技术。

目 录

第一章 对等网络的研究概况	1
1.1 对等网络的发展历程	1
1.2 对等网络的技术优势	2
1.3 对等网络的体系结构	4
1.4 对等搜索的分类	5
1.4.1 无结构 P2P	5
1.4.2 结构化 P2P	6
1.4.3 混合结构 P2P	8
第二章 资源广告与兴趣感知的缓存	12
2.1 缓存策略概述	12
2.2 缓存策略的相关研究	13
2.3 IAC 的设计	13
2.3.1 共享负载特征	13
2.3.2 基于 GOSSIP 的资源广告	15
2.3.3 兴趣感知的缓存机制	19
2.4 分析 IAC	21
2.5 性能评估	21
2.6 小结	22
第三章 资源广告传播	23
3.1 穷尽搜索概述	23
3.2 复制策略的相关研究	25
3.3 穷尽搜索概率模型	26
3.4 RPXS 的设计	27
3.4.1 RPXS 的结构	28
3.4.2 节点数量评估	28

3.4.3 随机节点子集选取	30
3.4.4 副本部署	31
3.5 性能评估	34
3.6 小结	37
第四章 对等网安全问题	38
4.1 Gnutella 协议概述	38
4.2 Gnutella 面临的安全威胁和安全隐患 ^[69]	40
4.3 针对 Gnutella 的攻击类型	41
4.4 对等网安全面临的挑战 ^[74, 75]	42
4.5 信任和信誉机制提出的动机和背景	44
4.6 信任与信誉机制研究现状及意义	45
第五章 无结构化对等网信任和信誉机制	48
5.1 信任和信誉机制涉及的基本概念	48
5.2 信任和信誉机制的研究对象	50
5.3 信任和信誉机制的特性分析	51
5.4 信任和信誉机制面临的问题	51
5.5 节点信任和信誉机制研究内容的提出	52
5.5.1 信任和信誉计算模型	52
5.5.2 信任和信誉信息的安全传输	53
5.5.3 信任和信誉信息的安全存储	53
5.6 小结	54
第六章 节点的信任和信誉模型	55
6.1 主要研究内容概述	55
6.2 信任和信誉计算模型所涉及的相关问题	56
6.2.1 计算模型的构成要素	56
6.2.2 计算模型结果分析	57
6.2.3 计算模型结果的度量及信任等级划分	57
6.2.4 计算结果预测节点行为的数学描述	58

6.3	信任和信誉计算模型分类	58
6.3.1	根据处理方式分类	58
6.3.2	根据搜索范围分类	59
6.4	现有信任和信誉计算模型存在的问题	61
6.5	节点行为过程具有马尔科夫性的分析	62
6.5.1	马尔科夫过程和马尔科夫链	62
6.5.2	对等网络节点状态行为的马尔科夫性	63
6.6	基于马尔科夫过程的信任和信誉计算模型	64
6.6.1	单一随机过程的马尔科夫计算模型	64
6.6.2	双随机过程的隐马尔科夫计算模型	68
6.6.3	运用信任和信誉计算模型防范 DDoS 攻击	70
6.6.4	基于 Gnutella 的对等网络 DDoS 攻击的特性分析	71
6.7	仿真实验	72
6.8	小结	77
第七章	信任和信誉值传输协议	78
7.1	Gnutella 对等网络中信息交互存在的安全威胁	79
7.2	节点信任和信誉信息的投票选举协议	79
7.2.1	节点信任和信誉信息的投票选举协议	79
7.2.2	基于节点信任和信誉信息加密的投票选举协议	80
7.2.3	增强型的投票选举协议	82
7.2.4	投票选举协议的安全性分析	83
7.3	分片、多路径投票选举协议的提出	85
7.3.1	Shamir 门限理论的概述	86
7.3.2	基于门限方案的节点信任和信誉信息的分割	87
7.3.3	节点信任和信誉信息分割与多路径算法实现	89
7.3.4	分片多路径投票选举协议的具体描述	90
7.4	多路径投票选举协议的安全性分析	93
7.5	小结	94
第八章	节点信任和信誉信息的存储策略	95
8.1	对等网络存储系统概述	95

8.2	无结构化对等网络存储技术	96
8.3	节点信任和信誉信息文件安全存储策略的提出	97
8.3.1	节点信任和信誉信息文件的路由查询算法	97
8.3.2	节点信任和信誉信息文件的冗余编码算法	101
8.4	仿真实验	108
8.5	小结	109

第一章 对等网络的研究概况

本章重点概述对等网络的研究现状，主要包括对等网络的发展历程、技术优势、体系结构和对等搜索的分类。

1.1 对等网络的发展历程

近年来，计算机行业迅猛发展，网络边界设备（如个人电脑）性能不断提高而成本却不断下降，网络基础设施的建设也日趋完善，用户的网络接入带宽不断增加。网络平台性能的提高与普及促进了新的网络应用的出现，也促进了新的互联网应用架构的出现，对等计算就是在这种条件下诞生的。对自由共享 MP3 音乐文件的需求导致了第一个 P2P 文件共享系统的出现，即 Napster，Napster 的最初版本由 Shawn Fanning 和 Parker 于 1999 年 6 月发布，其高峰期用户达到 2640 万，所产生的效应导致了对 P2P 系统研究的热潮。由于 Napster 采用集中的索引服务器来提供文件搜索服务，因此，其架构上存在性能瓶颈和单点失效的风险，而且还可能带来版权方面的法律问题，Napster 于 2001 年 7 月由于版权问题被迫关闭。

考虑到集中索引服务的风险，需要一种完全分布的对等搜索系统，Gnutella^[1]应运而生。Gnutella 是第一种构建完全分布式对等搜索网络的协议，节点随机选择邻居互相连接，最终构成一个随机图。Gnutella 采用泛洪方式传播查询请求，每个节点检查新接收到消息的生命期，如果生命期大于 0，则首先对消息生命期做减 1 操作，然后将该消息转发给所有邻居，反之则丢弃该消息。由于泛洪方式带来的消息开销呈指数增长，因此，Gnutella 的可扩展性比较差，不适合大规模应用。

2001 年 3 月，Niklas Zennström 等引入了另一种无需集中索引的 P2P 共享应用，即基于 FastTrack 协议^[2]的 Kazaa^[3]。Kazaa 是一种层次化的无结构对等网络，根据节点在网络中所担当的角色，可以把节点划分为超级节点和叶节点。超级节点之间建立随机连接，形成一个随机图，本质上，超级节点网络就是一个无结构对等网络；叶子节点连接到一个或者多个超级节点，并将共享信息发布在所连接的超级节点。在资源搜索过程中，超级节点代理叶子节点完成资源的搜索。层次

化对等网络利用了节点的异构性,并且缩小了搜索网络的规模,因此,具有较高的实用价值。

针对 Gnutella 的不可扩展性,研究人员考虑从体系结构方面来解决可扩展的大规模对等搜索。基于分布式数据结构方面的研究成果, I. Stoica 等提出了第一个分布式哈希表(DHT)——Chord^[4]之后,其他类似的 DHT 也被提了出来,如 CAN^[5]、Tapestry^[6]、Pastry^[7]、Kademlia^[8]等。DHT 的节点之间具有确定的关系,并且数据项和节点之间的关系也是确定的,正因为 DHT 的这种严格的拓扑结构要求和数据分布要求,DHT 又称为结构化对等网络。为了维护这种确定的拓扑结构,需要额外的开销,特别是在高度动态的条件下,所需的开销更多,因此 DHT 适用于相对稳定的应用场景。DHT 具有哈希表的接口,因此,使用简单方便;通常,DHT 的邻居数和查询时延均为 $O(\log N)$,其中 N 为系统中的节点总数,因此,具有较好的可扩展性;DHT 的另一个优点是能够保证搜索的成功。

当前的研究热点之一是优化的搜索算法和新型应用。利用对等网络的特点,可以构建低成本、高可用的大规模信息共享系统,如 P2PWEB、基于 P2P 的 WIKI 系统等。这些应用对资源搜索具有更高的要求,如支持复杂查询、更高的搜索成功率(无论流行资源还是稀有资源)等。在应用方面,用户还需要更高级的资源共享模式,而不是简单的免费共享。

1.2 对等网络的技术优势

对等网络具有低成本、自组织、鲁棒性、可扩展和隐私保护等技术优势^[9-11],这些特点的具体说明如下:

(1) 低成本

对于客户机/服务器架构,中央服务器需要大量投资以满足负载要求,而对等网络则无需大量投资,具有低成本的优势。首先,对等网络可以有效聚集大量空闲的资源,如存储空间、网络带宽和 CPU 计算能力等,节点之间相互协作来完成特定的功能,因此,其基础设施成本和维护成本分散到不同节点,避免了中央服务器的大笔投资。其次,对等网络构建灵活,不需要对网络基础设施做任何更改,任何中档个人电脑均可发起建立一个对等网络(如 Bittorrent)。

(2) 自组织

对等网络中的节点以“ad-hoc”的方式相互连接,构成一个松散耦合的网络系统,节点自主地加入和退出系统,自主的选择邻居节点。当节点退出系统或者发生网络故障时,相关节点更新其邻居列表,自组织为新的网络系统,节点的所

有行为均具有自治性,不需要中央节点的协调。自组织特性降低了对等系统的维护开销,并且提高了对等系统的鲁棒性。

(3) 鲁棒性

在大规模网络系统中,网络故障、节点故障的发生是不可避免的,这些故障对系统的稳定性和服务质量造成影响。在客户机/服务器模式下,服务器成为整个系统的关键节点,一旦服务器发生故障,导致整个系统的崩溃。在对等模式下,系统的服务能力是节点服务能力的聚集,系统的服务能力是分散在不同的节点,而且节点提供的服务具有冗余性,因此,部分节点和网络发生故障,对整个系统的影响不大,保证了系统的稳定性和服务的可用性。对等网络还是一个动态的自组织系统,节点和网络故障触发网络的重新布局,形成新的稳定系统。因此,对等模式具有很高的鲁棒性。

(4) 可扩展

传统的客户机/服务器模式,服务器的功能限制了其最大服务能力,从而决定了能够支持的客户机数量。设计过大的服务能力需要更多的投资,而这些能力并不一定会充分使用;相反,过小的服务能力可能导致负载过重,给后期的系统升级带来困难。与之相对,对等模式具有客户机/服务器模式不可比拟的可扩展性。对等模式中,节点既是客户机又是服务器,节点之间相互协作实现特定的功能,新节点的加入同时给系统带来了额外的资源,如存储空间、网络带宽、文件资源、CPU 计算能力等。实际对等网络的规模也说明了对等模式的可扩展性,如 Gnutella 网络的活跃节点数量通常超过 1 百万^[12]。因此,对等模式具有高可扩展性。

(5) 隐私保护

由于对等网络中的节点是自治的,除直接邻居外,其他节点无法知晓节点的加入和退出。节点之间通过相互转发消息完成系统的功能,对任意节点来说,很难区分消息的来源和目的,结合密码学手段,对等网络能够很容易地实现匿名功能^[13-15],包括发送者匿名、接收者匿名和发送者接收者互匿,从而保护用户隐私。相对而言,客户机/服务器模式很难做到匿名性和用户隐私的保护,因为中央服务器能够轻易获取客户机的信息,至少包括 IP 地址等信息。

对等网络的内在优势决定其成为构建大规模互联网应用的首选架构,其中最重要的应用之一为信息共享,其基础为对等搜索。虽然集中式搜索引擎,如 Google、Baidu、Yahoo 等能够提供信息搜索的能力,但集中式搜索引擎存在其自身的缺陷。首先,需要大量的基础设施投资,如高性能数据中心的建立;其次,集中式服务常常成为网络攻击的目标,导致服务的瘫痪,而对等网络具有内在的鲁棒性,单个或者部分节点的失败不会导致系统功能的下降;第三,由于商业原因,集中式搜索引擎存在内容审查和操纵搜索结果的可能,可能导致搜索结果具有倾向

性^[16]。综上所述,对等搜索具有集中式搜索不可比拟的优势,因此,研究对等搜索能够解决大规模信息共享的基础性问题。

1.3 对等网络的体系结构

P2P 计算的发展过程中,提出了各种各样的系统,根据不同的分类依据,这些系统可以划分为不同的种类。

(1) 系统集中程度

根据系统的集中程度,可以将 P2P 系统分为集中式(centralized)和完全分布式(decentralized)。集中式 P2P 系统采用中央服务器提供资源索引功能,基于该索引服务,节点可以快速定位所查询的资源,而资源的交换则是通过普通节点之间的直接连接进行,无需中央服务器的参与,其典型代表是 Napster 和 Maze^[17]。由于共享版权文件的问题, Napster 最终被迫关闭。Napster 虽然是一个可行的 P2P 架构,但是存在其自身的局限性:中央索引服务器存在单点失效的可能,从而导致服务不可用;中央索引服务器也是性能瓶颈,虽然可以提供更好的基础设施来解决该问题,但是在没有合适的商业模式的条件下,显然是不可行的;版权问题同样会困扰集中式的索引服务。为了克服 Napster 的不足,完全分布式的 P2P 系统取代了 Napster,成为文件共享的有效工具, Gnutella 是完全分布式 P2P 系统的代表。从严格意义上讲, Gnutella 是一种 P2P 搜索协议,本文简单统称实现 Gnutella 协议的 P2P 系统为 Gnutella 系统。在完全分布式系统中,最直接的搜索方法是传播查询请求给系统中所有节点, Gnutella 采用泛洪实现这种资源查询方式。查询节点发送查询请求给其所有邻居,邻居节点对查询消息的生命期(TTL)做减 1 操作,如果 $TTL > 0$ 则以相同的方式转发该查询请求给当前节点的邻居节点。

(2) 网络拓扑和数据分布的特征

根据网络的拓扑结构和数据分布的特征, P2P 系统可以分为三大类,即:无结构(Unstructured P2P)、结构化(Structured P2P)和混合式(Hybrid P2P)。无结构 P2P 系统中,节点之间建立随机的连接,节点和资源之间无确定的关联。 Gnutella 网络是典型的无结构 P2P 系统,所构成的网络拓扑结构是一个随机图。结构化 P2P 系统中,节点之间存在确定的关系,资源和节点之间也存在确定的关系。这类系统主要指分布式哈希表(DHT),如 Chord^[4]、CAN^[5]、Tapestry^[6]等。邻居关系是确定的,而非随机的,资源的存放也是确定的。无结构 P2P 和结构化 P2P 都是 P2P 系统发展过程中为解决特定的问题而提出的架构,两者在面临不同的资源分布和需求的条件下,分别各有其优缺点。混合架构结合了两者的优势,

在无结构 P2P 系统中构建一个分布式哈希表, 分别解决各自擅长的问題。

(3) 系统的层次

根据覆盖网络是否具有层次结构, P2P 系统可以划分为层次化的(hierarchical)和非层次化的(non-hierarchical)。完全分布式的纯 P2P 系统通常是非层次化的, 所有节点都处于同一层次, 如 Gnutella。而混合系统和部分分布式纯 P2P 系统属于层次化系统, 如 KaZaa^[3], 层次化系统充分利用了节点的异构性, 能够提供更好的可扩展性, 具有很高的路由效率。

1.4 对等搜索的分类

对等计算的典型应用之一是资源共享, 包括文件共享、存储空间共享、CPU 计算资源的共享等。在这类应用中, 其核心问题是如何发现(或者定位)所需的共享对象, 该问题统称为资源搜索问题。一般来说, 当前的资源搜索算法可以分为三类: 无结构 P2P、结构化 P2P 和混合 P2P。

1.4.1 无结构 P2P

无结构 P2P 中, 节点之间、节点和资源之间均无确定关系。从整个系统来说, 所构成的覆盖网络的拓扑结构是一个随机图, 资源在网络中随机分布, 在这样的条件下, 设计有效的资源搜索算法是一个挑战性问题。在无结构 P2P 的发展过程中, 提出了下列主要的资源搜索算法:

(1) 泛洪搜索算法

在无结构 P2P 网络中, 每个节点只具有局部网络状态信息, 即该节点的邻居信息, 无结构 P2P 网络的拓扑结构是一个随机图, 因此, 设计高效的资源搜索是一个难题。由于无结构对等网络的特点, 数据项定位只能采用盲目搜索(Blind search)策略, 其目标在于以最小的开销把查询请求发送给尽可能多的节点。泛洪搜索是最直接的盲目搜索, 对新接收到消息, 节点首先根据消息的类型做本地操作(如本地关键字匹配), 然后对消息的生命期做减 1 操作, 如果生命期大于 0, 则将该消息转发给所有邻居节点(消息来源节点除外), 反之则丢弃该消息。泛洪机制简单、容错, 适用于网络规模较小的情况。由于泛洪机制的消息量呈指数增长, 因此, 可扩展性差。在网络规模较大的情况下, 产生大量的冗余消息, 导致网络带宽的浪费, 同时造成网络拥塞, 恶化网络的性能。

(2) 随机游走

随机游走常用于解决随机网络中的各种问题, 资源搜索是其中之一。就简单

的随机游走而言,每走一步,节点选择一个随机邻居,将消息转发给该邻居。随机游走能够很好地适应网络的动态性,并且降低冗余消息的传播。由于随机游走采用串行方式传播查询请求,导致其查询时延较大,为了解决这个问题,Qin Lv等^[18]提出了 k -随机游走,初始节点同时发起 k 个并行的随机游走,每个行走者之间是独立的。 k -随机游走是在查询冗余消息和查询时延之间的一个折中,其搜索性能远远高于泛洪搜索。随机游走的另一个变体是偏好随机游走,查询消息的下一跳选择不是随机的,而是根据邻居节点的属性来做出决策,比如选择节点度最大的邻居作为下一跳。L. A. Adamic等^[19]提出了幂律图中基于偏好随机游走的搜索算法,充分利用高度节点在搜索中的重要角色,提高了大规模幂律网络中资源搜索的性能。

(3) 混合搜索

泛洪搜索和随机游走都有其优缺点,泛洪搜索能快速局部扩展,而随机游走能快速纵深渗透。就单独一种搜索方式而言,泛洪在扩展深度增加的条件下,其消息开销呈指数增加,但是覆盖的节点也迅速增加;随机游走在扩展深度增加的条件下,其覆盖的网络节点增加缓慢,但是消息开销小。C. Gkantsidis等^[20]提出了组合随机游走和浅层泛洪的混合搜索算法。对一个新消息,节点首先局部扩展该消息——将该消息转发给1跳或者2跳范围的所有邻居。然后,对消息的生命期做减1操作,如果该消息生命期大于0,则将该消息转发给一个随机邻居。混合搜索策略降低了搜索的消息开销和时延,同时提高了搜索的成功率。

(4) 复制策略

在数据项流行度很高的条件下,泛洪搜索的性能很好,具有低时延、高成功率的性能测度,然而,自然的副本分布无法有效提高资源搜索的性能。Edith Cohen等^[21]详细研究了无结构对等网络中的复制策略,考虑副本量和查询流行度之间的关系,提出了一致复制、正比复制和平方根复制策略。通过分析发现,一致复制和正比复制具有相同的平均搜索开销,而平方根策略具有优化的平均搜索开销。根据生日悖论的原理,R. A. Ferreira等^[22]提出了查询率无关的复制策略,该策略利用生日悖论原理确定副本的分配模型,采用随机游走方法来实现副本的部署,该策略保证了搜索的高成功率。BubbleStorm^[23]是另一种查询率无关的复制策略,相似的,其副本分配源于生日悖论,并且考虑流量成本,BubbleStorm采用随机多图上的BubbleCast实现副本的部署。

1.4.2 结构化 P2P

基于泛洪的无结构 P2P 网络存在可扩展性差的问题,该问题引出了从两个方面来提高 P2P 系统的搜索性能。一方面,改进泛洪算法或者提出新的消息传播算

法,而网络的结构特性保持不变;另一方面,提出一种新的 P2P 体系结构,因此,结构化 P2P 系统应运而生。一般来说,结构化 P2P 系统主要指分布式哈希表(DHT),在结构化 P2P 系统中,节点之间、节点与共享对象之间具有确定的关系,而且所构成的覆盖网络具有严格的拓扑结构。

分布式哈希表是一种分布式数据结构,支持数据的分布式存储和查询。DHT 是在传统哈希表数据结构的基础上演化而来,对资源的“value”,如文件名等,作哈希运算而获得资源在哈希表中的索引 $H(\text{value})$,通过这种方式将资源集合的元素存储在哈希表中。资源集合包括了整个 P2P 系统中所有节点提供的资源,这是一个动态的大规模集合,因此用某个中心服务器来存储这个集合所对应的哈希表是不现实的。如果把这个哈希表按照特定的策略进行划分,并且按照特定的策略将划分后的分区存储到特定的节点,这样就构成了一个分布式哈希表。为了对分布式哈希表进行操作,需要系统中的节点相互协作,完成分布式哈希表的维护,支持对哈希表的查询和增删操作,因此,系统中的每个节点按照特定的算法选择某些节点作为邻居,这种以系统的参与者为节点,节点之间的邻居关系为边构成的网络就成为覆盖于现有网络(如 IP 网络)之上的逻辑网络,称之为覆盖网络。

DHT 的核心是有效地定位资源,支持这种功能的关键在于高效的消息路由算法,当前对这个问题的解决方案关系到覆盖网络的拓扑结构和所采用的路由策略。对于大规模网络,已证明有多种模型提供较小的路由表和短的直径,这些模型包括了环状结构(Ring)、Plaxton Tree 结构、超立方体结构、De bruijn 图、蝶网等。在 P2P 系统中,每个节点具有自治的特点,节点可以动态地加入和退出系统,而且每个节点所具有的系统信息是局部的,在这种条件下,贪心路由策略是一种最好的选择。资源(或者资源指针)存储在区块中,区块又存放在相应的节点中,解决资源定位问题就需要解决有效的消息路由算法,该算法能够快速路由查询请求到负责节点。不同的节点组织模式具有不同的拓扑结构,相应也具有不同的直径,寻求具有小直径的大规模图模型是研究的一个主要问题。在极图理论中,研究了图的“度-直径”问题:图的直径为 Δ ,节点最大的出度为 d ,关于图的大小 n 的上界(即 Moor bound)^[24]为:

$$1 + d + d^2 + \cdots + d^\Delta = \frac{d^{\Delta+1} - 1}{d - 1} \quad (1-1)$$

即: $n \leq \frac{d^{\Delta+1} - 1}{d - 1}$ 。当确定 n 和 d 的条件下,可以得出 $\Delta = \Omega(\log n / \log d)$ 。而

在现有的构造中,蝶网和 De bruijn 图能够取得 $O(\log n / \log d)$ 的直径,因此是优化的解。

与无结构 P2P 相似, DHT 也有其优缺点。首先, DHT 能够在给定的空间开销和时间开销的条件下保证查询成功, 无论对于流行的数据项还是稀有数据项, DHT 都能够搜索成功; 其次, DHT 能够很好地支持“精确匹配”查询, 而对于“复杂查询”, DHT 是无法有效胜任的; 再次, DHT 对于节点的“扰动”比较敏感, 为了保证其正确的功能, 所需的维护开销较大。

1.4.3 混合结构 P2P

无结构 P2P 和结构化 P2P 都分别具有各自的优缺点。就查全率 (Recall) 而言, 结构化 P2P 具有 100% 的查全率, 而无结构 P2P 系统并不能保证这一点。对于查询时延, 在不同数据项流行度的条件下, 这两者具有较大的差异性。如果数据项流行度比较高, 无结构 P2P 在局部范围内 (如 $TTL=2$) 进行泛洪就能命中, 而结构化 P2P 的查询时延为 $O(\log N)$, 其中 N 为网络中节点的数量。如果数据项比较稀少, 则泛洪查询的效率远远低于结构化 P2P。实际系统跟踪数据的分析结果表明数据项的流行度服从 MZipf 分布^[25], 部分数据项的流行度较高, 而另一部分数据项的流行度较低。根据这种流行度分布和两种 P2P 架构的特点, 一种合理的想法是结合两种 P2P 架构来解决 P2P 系统中高效搜索的问题。

(1) 简单混合 P2P

虽然各种 P2P 设计层出不穷, 但是哪一种是最好的设计呢? 研究人员和产业界均没有这方面的共识。在无结构 P2P 网络中, 查询以泛洪方式传播给邻居节点, 通过查询消息的 TTL 限制传播的范围, 如 Gnutella 网络和 KaZaa 等, 这种设计无法保证很高的查全率。在结构化 P2P 网络中, 将倒排索引存储在分布式哈希表上, 基于关键词的倒排索引示例如下:

<“对等网络”: (doc3, doc4, doc5, doc6 ...) >

<“资源搜索”: (doc1, doc3...) >

资源搜索时, 查询请求被路由到索引关键词的节点, 从而获得针对每个关键词的查询结果, 这些结果的交集构成此次资源搜索的结果, 结构化 P2P 网络保证了 100% 的查全率。

B.T.LOO 等^[26]首先发现可以利用两种 P2P 架构的优缺点和共享数据项的流行度来提高 P2P 系统的搜索性能, 提出了混合式 P2P 架构 (HP2P)。其中, 结构化搜索技术用于索引和定位稀有的数据项, 而泛洪查询技术用于定位高度流行的内容。

为了证明这种混合搜索方案, 作者首先对 Gnutella 网络进行测量分析。通过修改 Limewire Gnutella 客户端, 编写了数据抓取的爬虫。在 PlanetLab 的合适位置部署了 30 个爬虫进行数据获取, 所获取数据包括的节点数量超过 100 000, 文

件数量为 2 亿。该 Gnutella 网络是一种层次架构, 其中, 功能相对强大的节点作为超级节点, 其他节点为叶子节点。超级节点之间构成一个无结构 P2P 网络, 叶子节点通过超级节点进行资源查询。数据分析的结果表明, 对流行的数据项来说, Gnutella 网络的搜索机制是非常有效的, 对稀有数据而言, 这种搜索机制的效果很差。Gnutella 网络的查询机制反映出来这样一个事实, 即使所查询的数据项存在, 很大一部分查询都返回很少的结果或者无结果, 查询稀有数据项的响应时间较长。

DHT 的搜索机制同样存在其优缺点。优点为避免把查询消息泛洪到网络中, 保证了完美的查全率; 缺点是构建倒排索引的开销较大, 对大数据集而言, 是不可行的。对于稀有数据, 其倒排表较少, 因此, 构建倒排索引和数据查询的开销相对较小。总之, DHT 适合处理对稀有数据的存储和查询。

根据对两种查询机制的分析, 一种很自然的想法就是能否结合两者的优势构建更好的 P2P 资源搜索系统。泛洪策略用于处理对流行数据项的查询, DHT 用于解析对稀有数据的查询, 本文把这种混合的 P2P 搜索机制称为 HP2P。HP2P 的难点在于如何分布式实施查询策略? 对于一个新的查询, 采用泛洪方式解析还是用 DHT 解析, 这一点是 HP2P 的关键难题, 这个问题又可以归结到如何区分流行数据项和稀有数据项。作者提出了一种简单的解决方案, 利用查询探听机制获取那些获得很少返回结果的查询, 从而确定其要发布的数据项是否为稀有数据项。这里利用了这样的事实, 一个查询的返回结果很少, 表明该查询的对象是稀有的数据项。查询探听机制虽然简单, 但是存在一个问题: 如果一个数据项没有被查询过, 如一些稀有的数据项, 则无法判断其流行度。其他可考虑的方法包括词频统计、邻居节点的数据项采样等。

实验结果表明, HP2P 在查询响应时间和查全率两方面均有较大的提高。就查询响应时间而言, HP2P 在 40 秒内返回第一个查询结果, Gnutella 查询机制在 65 秒内返回第一个查询结果, 减少了 25 秒 (38%) 的查询时延。在查全率方面, 相对 Gnutella 网络的查询机制而言, 减少了 18% 的无结果查询数量。实验的数据分析结果表明, 混合搜索机制虽然简单, 但是其搜索性能更好, 实验中只部署了部分混合超级节点, 因此, 其分析结果并不能完全反映真实混合搜索机制的性能。如果整个系统均采用混合节点, 每个节点各自负责发布其稀有的数据项, 则能够提供更好的查询性能。

通过对混合搜索机制的简单实现和测量分析表明, 混合机制能够有效地提高 P2P 搜索的响应时间和查全率, 是一种值得进一步发展的 P2P 搜索机制。

(2) 基于 Gossip 搜索选择——GAB

在 IPTPS06 会议上, M. Zaharia 等^[27]提出了 GAB 搜索算法来解决 HP2P 网络