

一种 P2P 系统索引结构生成算法

姜国华 顾君忠

(华东师范大学计算机科学系, 上海 200062)

E-mail: ghjiang@mmit.stc.sh.cn

摘要 论文分析了 P2P 系统的基本概念, 研究了现有 P2P 系统由于缺乏分布式索引结构而产生不足的现状。从而提出了基于虚拟二叉键树的分布式索引模型及其构建算法。利用虚拟二叉键树可以把信息分散存放在各个节点上, 并通过路由算法迅速有效地找到存放在其它节点上的信息。同时, 建立索引的过程是节点间自发进行的, 不需要中心服务器的支持。

关键词 P2P 分布式索引 虚拟二叉键树

文章编号 1002-8331-(2004)02-0168-02 文献标识码 A 中图分类号 TP393

An Index Construction Algorithm on P2P System

Jiang Guohua Gu Junzhong

(Computer Science Dept., East China Normal University, Shanghai 200062)

Abstract: This paper firstly analyzes the basic concept of P2P system, then studies the actuality of current P2P systems without index structure. Then a model based on virtual binary key tree is put forward and the construction algorithm is given. Using this virtual binary key tree, information can be separated onto individual nodes, and the information on other nodes can be found easily and quickly. Index construction is proceeded spontaneously. No central service is needed.

Keywords: P2P, Distributed index, Virtual binary key tree

1 P2P 概念

P2P 是 Peer-to-Peer 的简写, 译为“对等网络”。P2P 系统中没有客户机和服务器之分, 每个主机既是客户机也是服务器。它使互联网上的个人电脑相互平等地连接, 在进行信息交换时无须通过任何服务器和公司。P2P 使用户感觉到在进行文件交换时象以往登录网络一样方便。这项技术主要应用于企业内部的局域网或局域网之外的任何特定人群间的对等交流。

2 P2P 的典型代表

P2P 的典型代表有两个 Napster 和 Gnutella。Napster 为用户提供了共享 MP3 格式文件的下载, 并在系统的中央服务器上记录所有用户拥有的音乐文件, 用户可以直接连接到歌曲拥有者的电脑, 并从电脑的硬盘中得到这首歌曲。Napster 要求所有的用户登录到中央服务器。Gnutella 是由属于“美国在线”的 Nullsoft 软件公司研制开发的资源完全开放的应用软件, 可以从 <http://gnutella.wego.com> 免费下载不断更新并可以在 Windows、Mac、Linux 上运行。当用户下载 Gnutella 后, 便可以与任何一台同样拥有 Gnutella 的电脑相连。用户要查找信息时, 就像目前我们使用 Windows 的查找文件夹功能一样简单, 只不过搜索范围换成了拥有 Gnutella 的所有电脑。

3 现有 P2P 系统存在的问题

通过 P2P 系统, 使 PC 拥有者之间可以直接通过互联网进行通信而脱离任何服务器的控制。P2P 系统将给我们带来一个资源共享、按需分配、完全互动的新的互联网。但现有的系统不能达到这些目标, 大多数的 P2P 系统要求有中心服务器, 如

Napster, 这使得中心服务器可能成为网络瓶颈, 并且一旦中心服务器瘫痪, 整个系统就无法运作了。而另一些无中心服务器需求的 P2P 系统对网络带宽的要求很高, 如 Gnutella, 在拨号上网的网络中根本无法工作。这种 P2P 系统的查询代价高, 响应速度慢。造成这种现象的原因是它没有一个内在的索引结构, 查询搜索请求是在网络中进行广播。所以问题的症结在于缺乏一种分布式索引结构。

4 探索中的解决方法

在建立分布式索引结构方面已经有了一些研究成果, 主要是在分布式数据库领域, 如搜索结构复制技术、使用变体 B 树技术等。这些方法都假定有一个可靠的执行环境, 需要一定的中心服务支持, 而且只适用于数量较少(几百台)的主机群。P2P 系统的执行环境是不可靠的, 主机数量变化幅度很大, 可以从几十台到几十万台, 而且不需要中心服务器。所以 P2P 系统需要的索引结构要适应这种环境, 同时还要保证较高的搜索效率、较低的网络带宽需求和良好的可伸缩性。

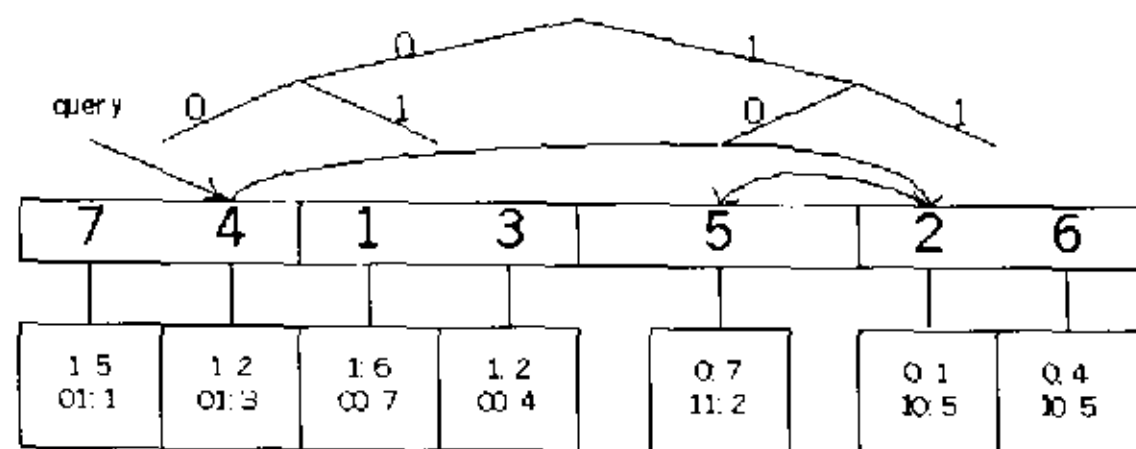
5 解决方法思想

peer 之间随机相遇, 相遇时划分搜索空间, 并保留足够的信息来指向搜索空间的其余部分。peer 之间的相遇是由系统机制保证的, 也可由其它操作连带引起。作者使用虚拟二叉键树为索引结构, 键值 2 个: 0 和 1。每个 peer 的路径就是由根到 peer 所在节点的二进制串值, 记为 $path(peer)=b_1b_2\cdots b_m$ 。路径长度为 m 。每个 peer 所存储的数据 data 对应到一个二进制串值, 记为 $key(data)=b_1b_2\cdots b_n$ 。key 的长度为 n 且长度固定, 路

作者简介: 姜国华(1975-), 男, 硕士生, 研究方向为分布式数据库和多媒体技术。顾君忠, 教授、博导。

168 2004.2 计算机工程与应用

这样每个 peer 就负责了整个搜索空间的一部分, 即以其路径为前缀的数据。同时每个 peer 还要有一个路由表以指向搜索空间的其余部分。一个路径为 $b_1b_2\cdots b_m$ 的 peer, 它的路由表有 m 项。对于路径的每个前缀形式 $b_1b_2\cdots b_k, k=1, 2, \cdots, m$, 在路由表中对应到一条路由记录。记录中指出当要查找的数据的 key 是以 $b_1b_2\cdots b_{k-1}(1-b_k)$ 为前缀时, 把查询请求发送到记录中指定的 peer。



如图 1 所示,7 个 peer 构成了一个二叉键树索引结构。由图中可以看到,每个 peer 负责搜索空间的一部分,如 peer 7 负责 00 开始的 key 值集合,可以有多个 peer 负责同一条路径,以提高可靠性。每个 peer 有路由表以指向搜索空间的其余部分。搜索查询请求可以向任何一个 peer 发出,比如当向 peer 4 发出查询 key 为 1001 的请求,peer 4 根据自己的路由表把请求转发给 peer 2,peer 2 也不负责该键值,根据自己的路由再把请求发给 peer 5,peer 5 负责该键值,它在自己的索引表中查找 key 为 1001 的数据,把结果发给最初发出查询请求的 peer。peer 的路径、路由表和索引表是在二叉键树的构建过程中建立起来的。

索引结构的正式定义如下: 给定有 L 个 peer 的集合 $P = \{p_1, p_2, \dots, p_L\}$, 对于任意一个 $p \in P$, 有 $path(p) = b_1b_2 \dots b_m$, $b_i \in \{0, 1\}$ 。给定一个数据对象集合 D , 对于任意一个 $d \in D$, 有 $key(d) = b_1b_2 \dots b_n$, $b_i \in \{0, 1\}$ 。key 的长度 n 为常数, $m \leq n$ 。peer p 存储的数据对象 d 的索引满足下面的要求: $path(p)$ 是 $key(d)$ 的前缀形式。

下面是图 1 中 peer7 的索引表和路由表例子:

key	location
0001	peer2
0010	peer7
0011	peer5

path	peers
1	peer5
01	peer1

成平衡的键树;MAXR 指出最大的递归次数,该常数影响算法的收敛速度。二叉键树索引结构的描述性构建算法如下,很容易转变成各种具体语言的算法。

对算法中四种情况的说明:最初,每个 peer 负责相同的搜索空间,那个时候,当两个 peer 相遇时,它们把搜索空间一分为二,每个 peer 负责其中一半。它们同时存储了另一个 peer 的地址,以覆盖整个搜索空间。当两个路径相同的 peer 相遇时,处理方法是一样的。这是情况 1。当一个 peer 的路径是另一个的前缀形式的时候,路径较短的一个 peer 将它的路径在相反的方向上扩展一位,这是情况 2 和 3。当两个 peer 的路径不同的时候,可以使用各自在 $lc+1$ 位置的路由信息和对方的 peer 再递归执行该算法,以加快索引结构的构建速度,当 MAXR 设为 0 时,不执行递归。这是情况 4。

万方数据

按照 fcm() 的格式要求输入数据,通过计算得到每一类的数据中心见表 3。

表 3 某一类的数据中心

指标 矿名	塌陷深度 (m)	积水深度 (m)	浅塌 陷率	万吨塌陷率 (亩/万吨)	充填复 垦系数	复垦率
类 1	3.0693	1.5452	0.38465	5.5593	0.16381	0.25348
类 2	6.0959	4.453	0.15531	3.6537	0.2047	0.28323
类 3	11.556	9.7562	0.051435	2.6668	0.071567	0.071898

各类的含义是:类 1 属于复垦条件较好的矿井,类 2 居中,类 3 的复垦难度大。

每个数据相对于每一类的隶属度见表 4。

表 4 相对于某一类的隶属度

指标 矿名	类 1	类 2	类 3
袁庄	0.99118	0.0076391	0.0011832
张庄	0.020647	0.96769	0.011664
朱庄	0.037363	0.93555	0.027088
相城	0.9892	0.0093817	0.0014164
岱河	0.14975	0.82847	0.021787
杨庄	0.014847	0.97948	0.0053727
沈庄	0.99205	0.0068861	0.0010597
芦岭	0.014643	0.032447	0.95291
朔里	0.075073	0.9119	0.013024
石台	0.94958	0.046567	0.0058547
朱仙庄	0.040853	0.12695	0.83219

如果按最大隶属度原则进行划分,可得各矿井的分类统计

结果见表 5。

表 5 分类结果统计表

类别	矿名
I	袁庄,沈庄,相城,石台
II	朱庄,岱河,杨庄,张庄,朔里
III	芦岭,朱仙庄

4 结束语

通过以上两例可以看出 MATLAB 具有强大的计算功能,可以提高土地复垦规划效率。MATLAB 不仅可以帮助我们进行数值分析,而且还可以进行图形图象处理,可视化显示,在矿区土地复垦中有广泛的应用前景。(收稿日期:2003 年 2 月)

参考文献

1.清源计算机工作室.MATLAB 基础及其应用[M].北京:机械工业出版社,2000
2.楼顺天等著.基于 MATLAB 的系统分析与设计-模糊系统[M].西安:西安电子科技大学出版社,2001
3.卞正富等著.矿区土地复垦规划的理论与实践[M].北京:煤炭工业出版社,1996
4.闻新等编著.MATLAB 模糊工具箱的分析与应用[M].北京:科学出版社,2001
5.谢季坚.模糊数学方法及其应用[M].第二版,武汉:华中理工大学出版社,2000

(上接 169 页)

7 讨论

使用虚拟二叉键树作为索引结构,可以获得以下好处。最大的可伸缩性,由于 P2P 系统的天性决定了系统中的节点数量是不受限制的,所有的 P2P 解决方案都要考虑到可伸缩性的要求。算法分析和模拟试验表明该索引结构的建立需要交换的消息总数随着节点数的增加呈现线性增长的趋势,每个节点所需要交换的消息数是个常数,不随着节点总数的增加而增加,因此该结构具有很强的可伸缩性。信息和服务的可用性是该结构的又一个特点,由于有多个节点负责同一个信息和服务空间,提高了信息和服务的可用性。即使负责某个信息和服务空间的若干个节点中的一部分出现故障,不影响该信息的服务空间的运作。灵活性也是一个特点。节点可以很容易地从一个环境切换到另一个环境。比如节点上下线,通信通道的切换,或者从有线切换到无线环境等。另外,管理花费也很少,不需要中心管理服务,节点之间是自发地组织在一起的。

使用该索引结构使得搜索的时间和代价大大下降,搜索可以向任何一个节点发出,该节点根据自己的路由表可以快速地找到存放该信息的节点,不需要象 Gnutella 那样向所有的节点发出广播查询该信息,不但降低了带宽要求,而且使得查询十分快速有效,同时也不象 Napster 那样需要中心服务器的支持,使得管理维护费用几乎为零。路由信息和索引信息是在虚拟二叉键树的自发构建过程中建立起来的。建立该结构每个节点所需交换的消息数很少而且不随着系统节点总数的增加而增加。

综上所述,使用该索引结构使得 P2P 系统的性能得到很大的提高,同时降低了对环境的要求。

8 小结

论文探索了在 P2P 系统中建立分布式索引结构的一种方法并给出了构建算法。对于有 L 个 peer 构成的 P2P 网络,使用该算法产生的索引结构使得搜索时间复杂度降为 $O(\log L)$,而采用中心服务器结构的 P2P 系统如 Napster 和无索引结构的 P2P 系统如 Gnutella 的搜索时间复杂度为 $O(L)$ 。使用该算法构建的 P2P 系统既避免了中心服务器的单点故障问题,同时具有高效的搜索效率,很好地适应了分布式环境的要求。这种索引结构也有缺点,因为没有考虑数据对象分布的不均匀性,在实际使用中会导致 peer 的负担不均匀。解决这个问题的一个方法是使用不平衡的虚拟二叉键树作为索引结构,因为该索引结构是虚拟的,所以不平衡的索引结构实际上不会增加搜索的时间复杂度,这是有待进一步研究的问题。

(收稿日期:2003 年 2 月)

参考文献

1.K Aberer,M Ponceva,M Hauswirth et al.Improving Data Access in P2P Systems[J].IEEE Internet Computing,2002:58~67
2.T Johnson,P Krishna.Lazy Updates for Distributed Search Structure [C].In:SIGMOD Conference 1993:337~346
3.W Litwin,M Neimat,D A Schneider.RP*:A Family of Order Preserving Scalable Distributed Data Structures.VLDB 1994:342~353