

# 基于主题相关的 P2P 网络研究<sup>\*</sup>)

杨 丹 徐传运 文俊浩

(重庆大学软件学院 重庆 400030)

**摘 要** P2P 网络架构是影响 P2P 网络的可扩展性、安全性、资源定位效率、网络利用率的重要因素之一。P2P 网络架构根据是否有中央服务器可分为集中式和非集中式两种,根据节点是否是结构化的组织分为结构化网络和非结构化网络两种,现有的 P2P 网络架构各有优缺点。为了克服现有 P2P 网络架构的缺点,文章提出了一种主题相关的 P2P 网络架构,研讨了主题相关的 P2P 网络架构相关技术问题,对网络的构建、发现算法做了分析。最后举例说明了该网络的应用实例。

**关键词** P2P, 主题相关, 网络架构

## The Peer-to-Peer Network Based on Topic-Relative

YANG Dan XU Chuan-Yun WEN Jun-Hao

(College of Software Engineering, Chongqing University, Chongqing 400030)

**Abstract** The architecture of P2P network is one of the most important factors that affect the expansibility, the security, the efficiency of resource-locations, and the network-using rate of the P2P network. According to whether there is a central server or not, the architecture of P2P network can be classified into the centralized network and the non-centralized network, either of which has its advantages as well as disadvantages. In order to overcome the disadvantages of the architecture of the existing P2P network, the authors brought forward a topic-relative architecture of P2P network, discussed the correlative technical issues about the network architecture of P2P network, then analyzed the algorithms of building and finding the network's resources and finally gave some examples to illustrate this network.

**Keywords** P2P, Topic relative, Network architecture

## 1 引言

当前 P2P 网络架构可以分为以下三类:集中式(Centralized),结构化非集中式(Decentralize but Structured),无结构非集中式(Decentralized but Unstructured)<sup>[1]</sup>。集中式(Centralized)P2P 网络(如 Napster<sup>[2]</sup>)有一个或数个服务器维护所有在线节点的共享数据目录,对等节点(peer to peer)提交的查询首先被发送到服务器,服务器处理后把要查找文件的维护节点的 IP 地址返回给查询节点,最后,查询节点与维护节点建立直接连接并下载查找到的文件,然而,集中式(Centralized)P2P 网络中由存在于中央目录服务器(如 Napster 的中心目录服务器),因而这类系统存在可伸缩性瓶颈和单点出错等局限性。结构化非集中式(Decentralized but Structured)没有中央目录服务器,但 P2P 网络是有一定组织结构,其数据存放是根据一定的算法严格控制,例如:Freenet<sup>[3]</sup>, PAST<sup>[4]</sup>。这类 P2P 网络多数是基于分布式哈希表(DHT: Distributed hash table)<sup>[5]</sup>,根据创建、维护节点路由表的不同,现已提出多种算法,例如:CAN<sup>[5]</sup>, Chord<sup>[6]</sup>, Pastry<sup>[7]</sup>, Tapestry<sup>[8]</sup>等。在结构化的 P2P 网络中,数据在存放时数据根据哈希函数映射到节点上,在搜索时用同样的哈希函数定位数据存放节点,因而这类架构 P2P 系统查询效率比较高。但结构化的 P2P 网络有以下局限性<sup>[9]</sup>:(1)不支持非精确匹配的查询;(2)系统中数据的放置严格控制,这就意味着在动

态的 P2P 环境中维护系统拓扑结构的代价极其昂贵。无结构非集中式(Decentralize but Unstructured)即没有中央目录服务器,也没有严格控制的网拓扑结构和数据存放定位。例如:Gnutella<sup>[10]</sup>。节点独立自主,尽管一个节点可为其他节点提供信息与服务,享用其他节点提供的信息与服务,但每个节点只维护自己关注的文件;在这种网络中查询通常基于“洪泛算法(Flooding Algorithm)”,算法基本思想是:查询一旦提交,提交节点立即搜索本地文件库,并把查询转发给其所有邻居,这些邻居也依次进行类似的操作,直到查询消息的存活期(TTL: time-to-live)减少到 0 为止,所有查询结果(包括击中与失败)沿途返回给提交节点。显然,相对于其他两类数据共享系统,非结构化系统更具有可伸缩性、灵活性和自治性等优良特性。但网络存在问题:如果 TTL 值太小,搜索深度将受到限制,网络资源得不充分利用,这需要 TTL 越大越好;如果 TTL 大了,查询请求将给网络造成极大的负担(现在的 P2P 文件共享软件已占用了 Internet 较大的带宽)。解决该问题的一个思路是:把 P2P 网络根据资源的相关性组织成动态的、逻辑的、有限互联的多个子网,搜索分成发现子网和子网遍历两步。该解决方法能保证在不降低搜索深度的条件下减少网络带宽。

基于以上思想,本文提出了一种新的基于主题相关的 P2P 网络。基于主题相关的 P2P 网络有可伸缩性好、资源发现效率高、容错性好等优点。

<sup>\*</sup>)基金项目:重庆市信息产业发展基金(200501012)。杨 丹 教授,博士生导师,主要研究方向为:软件工程;徐传运 硕士研究生,主要研究方向:软件工程;文俊浩 副教授,博士研究生,主要研究方向为:软件工程、数据挖掘。

## 2 主题相关网络

**定义 1(主题 Topic)** 信息资源(包括:文件、网页、Web Service、计算能力等)所表达的主要信息,例如:一个关于程序设计的网站可能有编程语言、算法、测试等主题。需要指出,主题不一定是一个词,可以是一些关键词的集合,也可以是一个句子,甚至一段话,关键在于它传递了一个主要的信息,这个主要信息就是资源表达的主题。

**定义 2(主题相关网络, topic relative network)** 具有相同主题的节点建立链接,由这些节点组成网络。主题相关网络如图 1 所示,主题相关网络根据主题形成多个主题子网。主题相关网络中建立链接的依据是有相同的主题,两个节点间可能有多个链接(如果有多个相同的主题)。基于控制主题网络对建立成本的有要求,节点间通常不是全连接。连接密度与资源发现成本成反比,连接密度越高,发现代价越小。连接密度的高低需要根据具体应用环境而定,也就是说应用是更关注建立成本还是更关注发现成本。一般来说,一个相对稳定的网络,可以提高连接密度,达到提高系统总体性能的目的。

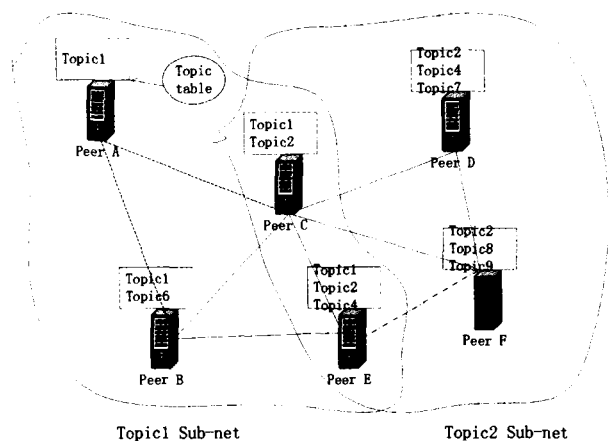


图 1 主题相关网络

**定义 3(对等主题点, topic peer)** 组成主题相关网络各 P2P 对等点,对等主题点主要由以下几部分组成:资源、主题表、远程相同主题对等点表、本地相同主题对等点表。

**定义 4(主题表)** 是本地资源中主题的集合(如图 1 所示)。

**定义 5(远程相同主题对等点表)** 记录了与本节点有相同主题的其他节点的地址。

**定义 6(本地相同主题对等点表)** 记录了本地与其它节点建立了链接的节点的地址。

**定义 7(主题子网, Topic subnet)** 由拥有相同主题的对等主题点(topic peer)组成的网络,主题子网由远程相同主题对等点表中的节点地址链接而成。其中的节点拥有相同主题的资源。在主题子网中搜索能够搜索到更多的目标资源。

## 3 主题网络建立算法分析

主题网络构建的基本思想是:节点  $p_i$  主题网络中定位拥有相关主题的节点  $p_j$ , 并把节点  $p_i$  的地址存入到节点  $p_j$  的远程相同主题对等点表中,建立链接,在建立主题网络时应考虑节点性能不同、负载均衡等问题。新增节点新增主题  $T_i$  时的算法讨论。

(1)定位主题网络中的任一个拥有  $T_i$  的节点  $p_j$ ;

(2)如果  $p_j$  的远程相同主题对等点表未满载(各节点远程相同主题对等点表的大小是不相同的,根据节点的性能而定),则  $T_i$  和  $p_i$  存入,把  $p_j$  返回给  $p_i$ ,存入到  $p_i$  的本地相同主题对等点表,并把主题允许最大链接数(节点主题与其它节点建立链接是有限制的)减 1,已满执行下一步;

(3)如果允许最大链接数为 0,则完成  $T_i$  的链接建立;

(4)搜索  $p_j$  的远程相同主题对等点表。随机选择一个与  $p_i$  根据  $T_i$  建立链接的节点  $p_h$ ,如果  $p_h$  已与  $p_i$  建立链接,则重新选择一个节点,做同样的判断。如果没有找到  $p_h$ ,则结束  $T_i$  的链接建立,否则转到步骤(2)。

需要指出“主题允许最大链接数”是根据节点拥有资源数量而定,拥有较多资源的节点的“主题允许最大链接数”可以适当增大。“主题允许最大链接数”的大小决定着节点被发现的概率。“主题允许最大链接数”大的节点可以适当增大远程相同主题对等点表,这样可以提高网络的整体性能。

如果节点主题  $T_i$  删除,与之建立的链接也要同时删除。删除链接时先遍历本地相同主题对等点表,向已基于  $T_i$  建立链接的节点发出链接删除请求。

在节点异常时永久退出网络后,网络将留下无效链接,为了保持网络可用性,每个节点应定期检查远程相同主题对等点表,即向建立链接的节点发出状态检测请求,如果在多个状态检测请求没有响应的情况下,则删除与之建立的所有链接。状态检测请求应根据一定的周期(例如:一天,一周,周期短可以提高网络的可用性,但会加重网络的负担)发出。如果节点在异常退出并且离开时间大于网络允许离开时,再次加入网络,应把该节点当作一个新节点加入,执行新节点加入的相关处理。

## 4 资源发现方法分析

在主题相关网络(topic relative network)中发现资源分为两步:一是定位主题子网(Topic subnet),二是遍历主题子网。

定位主题子网是指在网络中定位一个或多个拥有目标主题的节点,定位到一个拥有目标主题的节点,就说明已进入一个目标主题子网,就执行遍历子网的操作。需要指出的是,在前面的主题网络建立算法中,没有保证所有拥有相同主题的子网是连通的,因此在网络中发一个拥有主题的节点,并不一定能遍历所有拥有这个主题的节点。当然可以改进前面的主题网络建立算法,保证所有拥有相同主题的子网连通,无疑要保证这一点,必须以增加网络建立成本为代价。是否需要连通要根据的应用环境而定。在不同应用环境中,有不同的定位主题子网的方法。下面给出几种参考方法:

(1)运用“洪泛(flooding)算法”发现一个拥用目标主题的节点作为子网入口。这种方法适用范围广,但是效率较低。

(2)建立一个主题目录服务器,服务器为每个主题维护一个或多个拥有这个主题的节点。发现时查询主题服务器,得到一到多个节点,作为进入子网的入口。这种方法有单点失败的缺点。

(3)运用“哈希算法”把主题映射到一个或多个节点上,这些节点作为子网入口。这种方法的效率很高,但也存在单点失败的缺点。

(4)发现请求发起节点维护主题目录。这种方法适用的

(下转第 91 页)

基于秘密共享方案的可生存存储系统可用性的详细评价。

**结论** 本文在借鉴基于纠删码的可生存存储系统可用性评价方法的基础上,面向基于秘密共享的可生存存储系统的设计需求,提出了一种简单可用性评价方法(SEMA),并基于马尔可夫链模型提出了一种新型的可用性评价方法(MEMA),建立起了基于秘密共享方案的可生存存储系统可用性的科学评价标准,最后给出了基于两种评价方法的评价结果,并进行了科学分析。

本文所提出的两种评价方法为基于秘密共享可生存存储系统可用性的评价提供了一种解决方案,具有较好的实用性。

## 参考文献

1 Wylie J J, Bigrigg M W, Strunk J D, et al. Survivable Informa-

tion Storage Systems[J]. Computer, 2000,33(8):60~68

- 2 Weatherspoon H, Kubiatowicz J. Erasure coding vs. replication: A quantitative comparison. In: Proceedings of the First international Workshop on Peer-to-Peer Systems (IPTPS 2002)
- 3 Lin W K, Chiu D M, Lee Y B. Erasure Code Replication Revisited. In: Fourth International Conference on Peer-to-Peer Computing (P2P'04), (Zurich, Switzerland), August 2004. 90~97
- 4 Shamir A. How to Share a Secret[J]. Communication of the ACM, 1979, 22(11):612~613
- 5 林晓东, 刘心松. 高可用性系统结构的研究[J]. 计算机应用, 1997, 17(4):15~18
- 6 毛用才, 胡奇英编著. 随机过程. 西安电子科技大学出版社, 1998
- 7 茆诗松, 王玲玲编著. 可靠性统计. 华东师范大学出版社, 1984

(上接第 63 页)

环境有很多限制,例如:要求主题相对稳定,有一部分静态节点(节点 ID 不变,拥有的主题不变)。

遍历主题子网有“深度优先(Depth-First)”和“广度优先(Breadth-First)”两种方法。在实际应用环境中,常常不需要遍历整个子网,只要满足一定的条件就可以结束,例如:发现第一个目标资源结束,遍历到一定深度结束,发现一定数量的目标资源结束。需要指出,在实际应用中,通常从主题子网中的几个不同的节点开始遍历,因此遍历算法需要保证不会出现重复遍历。

在一些应用环境(例如:P2P Web 内容搜索)中,可能出现跨子网遍历的情况,即在遍历资源的过程中,根据一定的条件进入另一个主题网络。

## 5 应用示例

基于主题相关的 P2P 网络可以应用于各种 P2P 资源共享系统,例如:Web 缓存共享<sup>[10]</sup>、Web Service 发现<sup>[11]</sup>,基于 P2P 的商品网络、文件共享、Web 内容搜索等。在这些应用中,都可以建立相关的主题,把拥有相同主题的资源节点互联起来,形成动态的、松散的主题子网。如果把这些应用建立在基于主题相关的 P2P 网络上,可以大大的改善系统的质量。

下面以 Web Service 发现为例,说明基于主题相关的 P2P 网络的应用。Web 服务技术的兴起为 Web 的使用提出了新的计模型,构造基于面向服务的构架(SOA)的应用是未来分布式计算的发展方向。在 SOA 所要解决的技术问题中,服务的自动发现是难点和关键。文<sup>[11]</sup>提供的解决方案是把具有语义相关度的节点组成组(Group),由组服务器维护各节点的目录,负责组间路由,这是一种半结构化的 P2P 网络。这种解方案存在如下问题:(1)存在单结点失败问题,组服务器是一个分散的中间服务器,组服务器的异常将给网络造成较大影响;(2)一个节点常常提供多类 Web 服务,有多种相关性较小的语义,把这些拥有不同语义的 Web 服务的节点强制归入一个组是不合理的,不利于服务的发现,例如一个软件企业的 Web 服务可能提供产品询价服务,产品自动升级服务,Bug 报告服务,要把这些语义相关不大的服务归入一个组是比较困难的,即使勉强归入一个组,对服务发现也是极为不利的。

基于主题相关的 P2P 网络的解决方案是:把各节点的 Web 服务根据语义分析,把语义相关较紧密的 Web 服务归入一个主题。一个节点可以拥有多个主题,一个 Web 服务也可

以属于多个主题,这样基于主题的归类更合乎实际,也更有助于提高 Web 服务发现的查准率。拥有相同主题节点组成主题子网,服务发现时只需要遍历整个主题子网就能发现所有 Web 服务,这样能大大提高查全率,提高网络的利用率,缩短发现时延(latency)。这种解决方案也能消除单点失败的缺点,提高系统的容错性,即使网络中多个节点异常,剩下的节点也能正常提供服务。

**结束语与展望** 上述在分析现有 P2P 网络优缺点的基础上提出了基于主题相关的 P2P 网络,讨论了相关概念及网络建立算法,并用示例说明了可能的应用。后续研究工作主要集中在两个方面:一是建立最优性能主题网络的数学模型,为实际应用提供理论指导,二是将主题应用于内容搜索,建立应用案例。

## 参考文献

- 1 Lu Qin, Cao Pei, Cohen Edith, Li Kai, Scott Shenker Search and Replication in Unstructured Peer-to-Peer Networks. ACM, 2002. 84~95
- 2 Napster Home Page. <http://www.napster.com/>
- 3 Freenet website <http://freenet.sourceforge.net>.
- 4 Druschel P, Rowstron A. PAST: A large-scale persistent peer-to-peer storage utility. In: Elphinstone K, ed. Proc. of the HotOS VIII. Schloss Elmau: IEEE Press, 2001. 65~70
- 5 Ratnasamy S, Francis P, Handley M, et al. A scalable content-addressable network. In: Proc. ACM SIGCOMM, San Diego, CA, August 2001. 161~172
- 6 Stoica I, Morris R, Karger D, et al. Chord: A scalable peer-to-peer lookup service for internet applications. In: Proceedings of SIGCOMM, Aug. 2001
- 7 Rowstron A, Druschel P. Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In: Proceedings of the 18th IFIP/ACM International Conference on Distributed Systems Platforms (Middleware 2001), Nov. 2001
- 8 Hildrum K, Kubatowicz J D, et al. Distributed Object Location in a Dynamic Network. In: Proc. 14th ACM Symp. on Parallel Algorithms and Architectures (Aug. 2002)
- 9 Balakrishnan H, Kaashoek M F, Karger D, Morris R. LOOKING UP DATA IN P2P SYSTEMS
- 10 凌波, 陆志国, 黄维雄, 钱卫宁, 周傲英. PeerIS: 基于 Peer-to-Peer 的信息检索系统. 软件学报, 15(9):1375~1384
- 11 LING Bo, WANG Xiao-Yu, ZHOU Ao-Ying, Ng Wee-Siong. A Collaborative Web Caching System Based on Peer-to-Peer Architecture. Chinese Journal of Computers, 2005, 28(2):170~178
- 12 Cheng De-Wei, Xu Bin, Cai Yue-Ru, et al. A P2P Based Web Service Discovery Mechanism with Bounding Deployment and Publication. Chinese Journal of Computers, 2005, 28(4):615~626