

基于用户需求的内容分发点对点网络系统研究*

张玉洁^{1,2}, 何明^{1,2}, 孟祥武^{1,2}

¹(智能通信软件与多媒体北京市重点实验室(北京邮电大学), 北京 100876)

²(北京邮电大学 计算机学院, 北京 100876)

通讯作者: 张玉洁, E-mail: zhangyj@bupt.edu.cn

摘要: 点对点内容分发网络技术已成为近年来研究热点领域之一。为用户推荐有价值的资源,提高用户资源定位的准确率和分发效率,是 CDN-P2P 技术面临的巨大挑战。从用户需求的角度出发,综述了近年来 CDN-P2P 领域的研究状况;讨论了用户需求并介绍了需求获取方法;对基于用户需求的 CDN-P2P 系统模型、节点特征及用户相似性、用户之间的相互关系、节点安全性和搜索机制等进行了对比讨论和概括总结;对基于用户需求 CDN-P2P 领域存在的难点和热点问题作了深入的剖析;最后给出了基于用户需求的 CDN-P2P 系统的发展趋势及其展望。

关键词: 用户需求;CDN-P2P;需求获取;用户相似性;用户信任度

中图法分类号: TP393 **文献标识码:** A

中文引用格式: 张玉洁,何明,孟祥武.基于用户需求的内容分发点对点网络系统研究.软件学报,2014,25(1):98-117.
<http://www.jos.org.cn/1000-9825/4454.htm>

英文引用格式: Zhang YJ, He M, Meng XW. Research on CDN-P2P system over user requirements. Ruan Jian Xue Bao/Journal of Software, 2014, 25(1): 98-117 (in Chinese). <http://www.jos.org.cn/1000-9825/4454.htm>

Research on CDN-P2P System over User Requirements

ZHANG Yu-Jie^{1,2}, HE Ming^{1,2}, MENG Xiang-Wu^{1,2}

¹(Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia (Beijing University of Posts and Telecommunications), Beijing 100876, China)

²(School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Corresponding author: ZHANG Yu-Jie, E-mail: zhangyj@bupt.edu.cn

Abstract: CDN-P2P technology has recently become one of the hottest research fields. Recommending the valuable resources and improving accuracy of resource location and efficiency of content distribution for users, are CDN-P2P's challenges. This paper presents an overview of the field of CDN-P2P research in recent years from user requirements' perspective. It discusses user requirements and introduces ways of requirements acquirement. Based on CDN-P2P system over user requirements, this paper provides a comprehensive discussion on node characteristics and user similarity, relationships between nodes, nodes security and search mechanism. The difficult and hot issues about CDN-P2P system over user requirements are analyzed in depth. The future development trend and prospects for CDN-P2P system over user requirements are also discussed.

Key words: user requirement; CDN-P2P; requirement acquirement; user similarity; user trust

网络资源的分发模式已成为近年来研究的热点。目前,Internet 上的资源分发模式大多采用中央集中式,这种分发模式存在诸多的弊端^[1]:中央节点(服务器)对整个网络服务具有决定性的作用,如果服务器失效,则整个服务终止;大部分的分发机制都是在服务器上完成的,这样会导致服务器计算上的瓶颈,而且用户机端也会造成

* 基金项目: 国家自然科学基金(60872051, 61001118); 北京市教育委员会共建项目专项资助

收稿时间: 2012-09-28; 定稿时间: 2013-06-26; jos 在线出版时间: 2013-07-25

CNKI 网络优先出版: 2013-07-25 14:07, <http://www.cnki.net/kcms/detail/11.2560.TP.20130725.1407.005.html>

极大的资源浪费.基于用户需求的点对点内容分发网络能很好地改善这一问题,它是将点对点技术(peer to peer,简称 P2P)与内容分发技术(content distributed network,简称 CDN)相结合,并针对用户需求提出来的一种网络资源分发模式.该模式既拥有 P2P 非集中的分布式计算性能,又兼具 CDN 的内容分发功能^[2].该模式下的每台计算机都具有同样的地位,既可以请求服务,也可以提供服务.即,P2P 信息的分发与共享无需通过 Web 服务器,也不受任何文档格式和宿主设备的制约,完全按照用户需求,去中心化思想,这将是传统的 C/S 系统架构所不能比拟的.

在 Web 2.0 时代,网络用户同时也可以成为内容提供者,这种“可读写”的网络表现出一种双通道的交流模式.作为其核心技术,P2P 代表真正的互联网精神,平等、自由、开放^[3],最能体现 Web 2.0 的去中心化、用户共享与互助、用户资源整合与聚类思想.基于 P2P 共享理念,使用户在 Internet 上的共享行为提升到一个更高的层次,使人们更主动地参与到网络中.从技术上说^[4],不同用户可以看成是信息网络中多个相对独立的节点,这些节点可以随时加入和离开网络,而不用关注其他节点的行为状态,具有对等网络系统管理的特点.

基于用户需求的 CDN 技术也是一种成熟的网络构建方式,它是把资源部署在网络边缘,采用重定向技术将用户的请求转向离用户最近的边缘服务器,在传统网络中,形成了特别优化的并能发布丰富媒体资源的网络架构.根据所采用网络拓扑结构和面向对象的不同,CDN 技术可以分为 4 类^[5]:中心化服务内容分发、面向用户需求 CDN 分发、P2P 网络内容分发、混合内容分发.基于用户需求的 CDN 网络^[6]是一个从用户角度出发,分工详细的系统:包括用户资源的分布式存储、资源传输过程中的负载均衡、网络资源请求的重定向和用户所需资源的内容管理,而资源内容管理和全局的网络流量管理是用户需求 CDN 的核心所在.

本文第 1 节讨论用户需求的概念和需求获取方法.第 2 节~第 5 节介绍基于用户需求的 CDN-P2P 的理论方法与研究现状,其中,第 2 节阐述基于用户需求的 CDN-P2P 系统优势;第 3 节、第 4 节分别引入基于用户需求的 CDN-P2P 体系结构模型和基于用户需求的 CDN-P2P 结合策略;第 5 节对用户需求 CDN-P2P 环境下用户节点的相似性计算方法进行归类 and 对比分析.第 6 节~第 8 节分析并总结用户需求 CDN-P2P 环境下用户节点之间的关系、节点资源、CDN-P2P 应用等.第 9 节对基于用户需求 CDN-P2P 难点、热点和发展趋势进行总结和展望.最后是结束语.

1 用户需求

用户对于网络各项性能指标的偏好是不一样的,因此,网络运营商应该依据用户需求来为用户提供服务.如有的用户希望所选用的节点计算能力强,而有的用户希望节点存储量大、信誉度高.此外,网络中备选节点自身的性能也不尽相同,在选取节点时,会根据对性能指标的喜好选出符合自身需求的节点.

1.1 用户需求

本文针对的用户需求是用户需求,包括用户资源搜索与定位、用户资源分布式存储、用户资源内容管理、资源传送过程中的动态调整,以有效提高资源分发效率为目的.文献[7]提出了用户如何根据对资源的需求,适应性地选取 CDN-P2P 超级节点的机制.文献[8]将这种需求自适应的超级节点选取机制进一步加以改进,通过负载均衡策略,改善了用户查询、搜索、资源分发的效果.

1.2 用户需求与拓扑之间的关系

在用户需求和拓扑之间还有更深层次的关系没有得到重视,即拓扑与用户资源之间的相关性.在结构化 CDN-P2P 系统(通常是 distributed hash table,简称 DHT)^[9,10]中,拓扑不仅仅是用户之间的位置关系,更重要的是用户资源的分布和定位所形成的用户需求的架构^[11].因此,在结构化 CDN-P2P 中,用户资源可以被精确定位,其查询算法也是基于单播的;而目前所有的非结构化 CDN-P2P 系统都是资源和网络拓扑相互独立的,其查询、分发机制多是基于泛洪搜索机制.

1.3 用户需求获取

为了准确、有效地获取 CDN-P2P 系统中潜在的用户需求,文献[12]提出了一种基于可拓数据挖掘的关联规

则挖掘方法,通过将用户需求分类,构成具有不同功能属性的用户资源需求集;利用粒子群算法挖掘出用户需求集中的关联规则^[12,13],结合可拓变换原理,推理出新的关联规则.用户需求信息中的关联规则挖掘意味着生成一系列有价值的 if-then 规则^[13],将用户需求信息数据分为若干属性的集合,确定条件属性集和结论属性集,每个规则由来自用户需求信息数据库的条件属性和结论属性组成.

条件属性集为 $C=(c_1, c_2, \dots, c_m)$, 每个 c_i 表示一个条件属性;结论属性集 $D=(d_1, d_2, \dots, d_m)$, 每个 d_i 表示一个结论属性,则基于用户需求信息的一条关联规则可以表示为公式(1):

$$\begin{aligned} \text{IF } (c_1 = c_1^3 \ \& \ c_2 = c_2^1 \ \& \ \dots \ \& \ c_m = c_m^2), \\ \text{THEN } (d_1 = d_1^2 \ \& \ d_2 = d_2^1 \ \& \ \dots \ \& \ d_n = d_n^4) \end{aligned} \quad (1)$$

支持度与置信度是描述用户需求关联规则的两个主要指标,支持度描述了用户需求关联规则的普遍性,置信度用以衡量用户需求关联规则的准确性.对于上述的一条关联规则,描述为 if E then F , 其中, E 包括了若干条件属性, F 包含了若干结论属性.由此支持度 sp 与置信度 cn 的计算公式(2)和公式(3)如下所示:

$$sp(E \Rightarrow F) = \frac{T(E \ \& \ F)}{|T|} \quad (2)$$

$$cn(E \Rightarrow F) = \frac{T(E \ \& \ F)}{T(E)} \quad (3)$$

其中, $|T|$ 为用户需求信息的数量, $T(E \ \& \ F)$ 为同时出现 E, F 的样本数目, $T(E)$ 为出现 E 的样本数目.文献[14-17]通过 QFD(quality function deployment) 技术进行需求分析,为研究用户资源需求提供了通用的框架.文献[18]采用 Apriori 算法挖掘用户资源需求信息,实现了对关联规则的有效提取.文献[19]运用规则挖掘和模糊聚类的方法,完成了用户需求与专家知识的相互转换.

2 基于用户需求的 CDN-P2P 系统的优势

作为对传统 CDN-P2P 网络的补充,对于用户需求的 CDN-P2P 可以将用户需求的内容和服务推送到网络边缘,使得用户在满足自己需求的情况下从最近的地方获取服务,既保证了自治域内节点组织间的独立性,又能充分发挥节点间资源共享的特性.

2.1 CDN与P2P结合的优势

P2P 技术有很好的扩展性和容错性,能够提高系统的效率和稳定性;而 CDN 技术能够在不改变现有服务模式 and 客户端的情况下优化网络结构,提供更好、更有效的服务.将 P2P 和 CDN 技术全面结合用于用户资源分发平台,将体现如下优势^[20-22]:

- (1) 大大减少中心节点数据源的压力.传统 CDN 中,每个服务节点若要获得中心节点数据源数据,均需要直接访问数据源服务器,这将造成数据源压力大,消耗带宽多.而通过对 CDN 服务节点采用 P2P 方式进行组织,节点之间可以通过 P2P 方式共享、缓存数据,极大地降低了中心数据源的压力;
- (2) 中心数据源数据多点备份,提高系统服务能力和可靠性.不同服务节点之间对中心数据多点备份,这一策略使得 CDN 系统整体冗余能力、服务的自我恢复能力得以提高,进一步增强了系统的稳定性;
- (3) 增加可提供服务节点的数量,提升服务灵活性.通过 P2P 方式共享、缓存数据,使得可以提供服务的节点数量大为增加,同时使得服务节点可以提供更加灵活而智能的服务;
- (4) 提高网络的可管理性,避免流量无序.系统将 P2P 的范围严格限制在某一边缘服务节点的服务区域内,避免了传统 P2P 技术的过多跨区域、跨 ISP 而造成骨干网拥塞、流量无序等问题,增强了网络可管理性和服务可靠性;
- (5) 增加了系统的可扩展性,提高了服务效率.下层内容分发采用 P2P 技术实现,使得系统的可扩展能力大大提高.整体系统具有良好弹性,可随时应对用户访问突发性和随意性的情况,保持良好的服务效率.

如表 1 所示^[2],CDN 和 P2P 技术各有所长,具有一定的互补性,将这两种技术有效结合,将是一种高效、实用

的用户资源分发方法.

Table 1 Advantages compared of CDN and P2P

表 1 CDN 与 P2P 优劣势对比

各项性能指标	CDN	P2P
可扩展性	扩展成本高	扩展成本低
服务器带宽	占用大	占用小
建造成本	高	低
并发用户数	有限	无限
内容版权	可监管	不可监管
QoS 服务	可保障服务	无法保障
数量有序性	流量区域控制	无序
技术优势	基于 C/S 模式的 Web 加速业务	大规模流媒体分发和负载分担

2.2 用户需求与CDN-P2P系统结合优势

CDN-P2P 是目前 Internet 中提供资源服务的主流技术,将用户需求与 CDN-P2P 技术相结合,能够将 CDN 的高可靠性服务、可管理性和 P2P 非中心化、可扩展性相互融合,能够为用户推荐和分发最贴近用户需求的网络资源,在不增加成本的情况下,满足用户激增的要求.文献[21]提出了基于用户需求 CDN-P2P 网络架构下系统稳定状态的流媒体调度策略:首先使用可靠性算法^[23]为用户提供可靠性服务,避免由于可靠性不高的数据块在传输过程中中断服务现象的发生,同时可预防由于大量用户请求失败导致的重复请求对服务器和带宽资源造成的压力而引发的服务性能不佳.

此外,文献[24]将用户需求与 CDN-P2P 系统结合,通过计算用户需求的相似度,将用户需求相似的资源进行分类和聚类,有效减少了系统所在边缘服务器的数量,增大了系统资源容量,从而降低了总体成本,形成以 CDN 为可靠的内容核心、以 P2P 为服务边缘的架构.通过引入用户的需求,在不增加 CDN 成本的同时大幅度提升 CDN 服务能力,有效避开了 P2P 应用的诸多弊端,使用户访问媒体的质量得到极大的提高^[24].

基于用户需求的 CDN-P2P 系统能够根据用户偏好,智能运用代理模式和信息推荐模式满足用户.建立用户需求模型跟踪用户行为、学习且记忆用户兴趣,并通过描述用户兴趣构建用户需求信息测试集,利用机器学习思想,将个性化信息从全局信息空间中分离出来.相比其他 CDN-P2P 系统,基于用户需求的 CDN-P2P 系统在实现个性化资源推荐服务以及推荐资源的准确性方面有明显优势.

3 基于用户需求的 CDN-P2P 系统模型

从用户需求的角度出发,基于用户资源需求相似性的 CDN-P2P 系统架构,起着内容服务载体的作用,目的是有效捕捉用户需求规律,建立合理的用户需求模型.

3.1 基于用户需求的CDN-P2P结构模型

基于用户需求的 CDN-P2P 结构模型如图 1 所示^[25],当用户访问资源内容时,中心服务器首先通过请求路由系统确定最接近用户的最佳代理服务器,并将该用户请求转向该节点.当用户的请求到达指定节点时,如果缓存空间已经存储了所请求的内容,则将内容资源提供给用户;否则,从其他代理服务器获取内容并转发给用户,同时将资源内容缓存下来.这种结构以 CDN 为基础,在骨干网层次保留了原有 CDN 系统的架构和功能,边缘节点引入了带有用户资源需求流的 P2P 技术来进行用户所需资源的分发和共享.由单个或多个边缘服务器及其覆盖区域的接入用户共同构成一个 P2P 自治域,在域内利用 P2P 技术实现了用户需求与内容分发目标的融合.该方法有效减少了边缘服务器的数量,增大了系统的存储空间,降低了总体成本.该结构采用随机拓扑的查询算法,每个用户节点自身的资源整体类型倾向决定了用户自身的兴趣,并采用树形结构来表达用户兴趣,通过“用户兴趣”^[26,27]这个量化值,建立起资源与网络拓扑间的相关性.

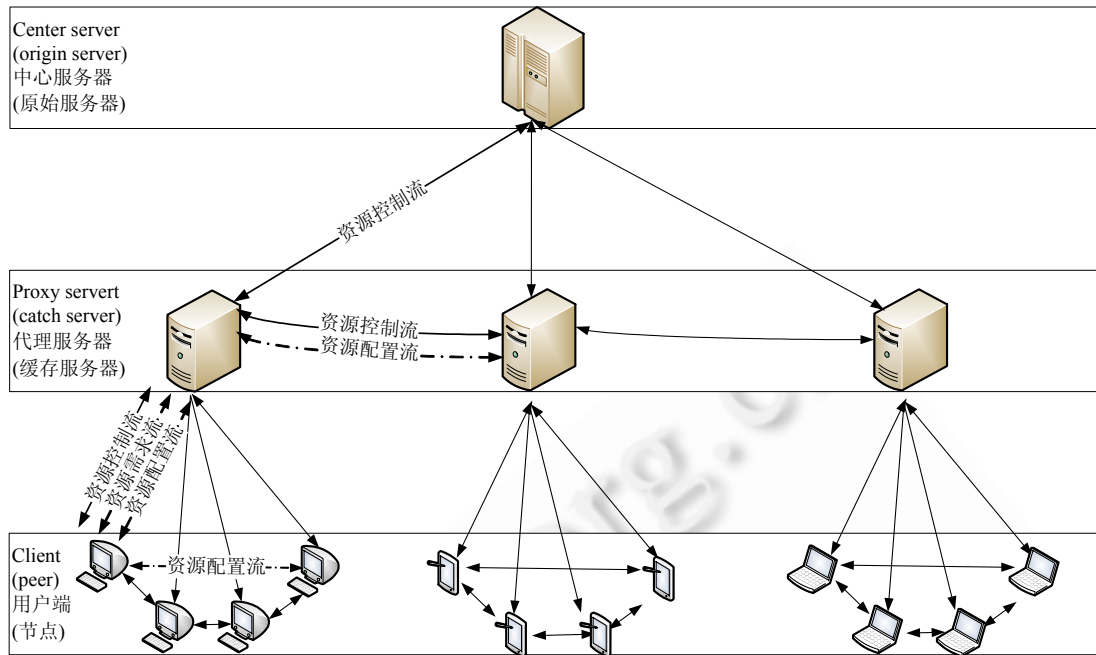


Fig.1 Model structure of CDN-P2P system based on user need

图 1 基于用户需求的 CDN-P2P 结构模型

3.2 不同内容分发方法理论模型的有效性对比

内容分发的实现,在技术上要求是比较高的,它所采用的关键技术策略包括:动态内容路由策略、高速缓存机制、动态内容分发策略与复制策略等.目前,基于用户需求的内容分发理论模型能够根据用户的相似性结果进行自组织^[28],可以随着用户相似性的不断变化,自行进行网络拓扑和内容资源分布上的调整,使用户节点能够较快地查询和下载到所需要的资源.文献[29]提出了内容分发用户自组织策略模型,该模型可以有效地提高自治域内的分发效率,减少转发次数.与传统意义上的内容分发系统相比,在提高用户服务质量的同时,有效降低了服务器的负载,并改善了用户对所需资源的查全率和查准率.Afergan^[30]申请公开了一项专利,它采用一个具有代表性的内容分发理论模型,将用户的请求引导到相应合理的 CDN 服务器上,使多个用户节点与 CDN 服务器产生关联,其自身包含的映射系统可以使一个用户节点在网络中定位另一个用户节点或是 CDN 服务器.这样使得内容分发在结构上实现了多点复制,即使某一节点由于某种原因发生故障,对网站的访问也能够被自动地导向其他健康节点进行响应.该模型的布设,尤其是对于偏远地区的用户,可能会遇到更少的网络抖动、网络峰谷和浪涌,并改善数据流的质量.

由于网络中资源存放的方式不同,内容分发策略模型分为单源模型和多源模型.此外,根据分发过程中数据的传输路径的不同,又分为单路径模型和多路径模型;根据资源种类的不同,分发策略模型总体上分为 4 类:单源单路径模型、单源多路径模型、单源分代模型、多源多路径模型^[31].

单源单路径模型又分为链式模型和树型模型,最具代表性的是 DirectStream^[32].单源多路径模型是为了解决单源单路径负载不均衡问题而提出来的,该模型具有较强的容错性.单源分代模型采用的是混合式拓扑结构,采用该结构的内容分发模型主要有 P2PVoD^[33]等.多源多路径模型拥有去中心化的内容分发架构,经过一定时间的资源浏览或下载,资源已经分布在不同的节点上,因此,可以利用多个视频源来分发资源.采用多源多路径策略的模型有 Gnustream^[34]等.

文献[31]对这 4 种内容分发模型进行了比较.单源多路径模型和多源多路径模型在负载均衡、资源平衡、

容错能力、异构性调整和稳定性等方面具有较好的性能,但需要较高的管理开销;单源单路径模型和单源分代模型的管理开销要低于前两种模型,但在异构性调整方面较差。

3.3 CDN-P2P系统模型概况

目前,多媒体技术领域中最关键的问题是如何确立一个新一代的内容分发方案,支持海量节点的集群访问,同时具有监控机制完善、性价比高、网络友好性强、地域可控性好、安全、稳定性高等特性。文献[21]提出的模型是在一般 P2P 流媒体系统的基础上选取高性能、高带宽、在线时间稳定的用户节点作为内容分发网络的边缘服务器,同时,为了更好地服务于非 P2P 终端用户节点,选择最优的边缘服务器,为其提供流媒体服务。文献[33]提出 CDN-P2P 融合架构的 IPTV 系统模型,该模型针对 IPTV 的网络特点以及可运营和可管理的要求,将 P2P 融合到运营商自营的 IPTV 系统中,通过与 CDN 系统的融合,以高性价比提供一种具备 TV 级媒体质量、电信级服务质量、可靠的交互式 IPTV 业务,能够适应未来大规模用户并发的需要,在大量用户并发时,用户可以自服务。而在诸如文件共享无中心的 CDN-P2P 网络中,对等节点具有匿名性和高度自治的特点,由于缺乏对与之交互节点可信程度的知识,节点需要应对交互过程中可能出现的威胁。文献[34]提出了一种基于节点评分行为相似度加权推荐的 CDN-P2P 环境下的全局信任模型,该模型可用于量化和评估节点的可信程度,在遏制恶意节点攻击、改善迭代收敛速度和网络中的成功下载性能指标上有较大提高。文献[35]使用博弈论模型,理论分析表明,引入依据节点“可信程度”的高低对节点区分服务这种机制,是解决音频污染和 Free Riders 两个问题的有效办法。

4 用户需求和 CDN-P2P 的结合策略

用户需求与 CDN-P2P 结合策略,并不是将 CDN-P2P 技术与用户需求信息简单的叠加,而是一方面在 CDN 网络核心层(即从中心到边缘服务器层)采用 P2P 技术进行优化,利用用户需求信息对子网进行重新组织,对用户需要的资源节点进行动态调整,这是原有 C/S 架构模式所不能比拟的;另一方面,在 P2P 网络中也融入了 CDN 技术的一些优点,使其能在可控、可管理性上得以提高。在基于用户需求信息的基础上使得两种技术充分融合、渗透,形成一个真正优化的 CDN 与 P2P 用户资源分发平台。

4.1 用户需求的CDN-P2P结合方式

目前,基于用户需求的 P2P 和 CDN 结合方式有两种^[21]:一种是 P2P over CDN based on user requirements,另一种是 CDN over P2P based on user requirements。前者将用户需求资源以 P2P 方式组织,存储到 CDN Cache 设备中,利用 P2P 的目录服务和多点传输能力,实现 CDN Cache 设备之间的用户资源交换,提升 CDN 的内容分发能力;后者将用户需求信息利用 CDN 的可管理机制和高服务能力引入 P2P 网络,形成以 CDN 为可靠内容核心、P2P 为服务边缘的架构,在这种架构下,用户通过 P2P 的客户端获取服务。

比较而言,前者是针对用户需求的管理和控制平台功能上有所扩展的服务,后者可以大幅度地提高用户需求的网路资源分发能力。

4.2 流媒体领域用户需求的CDN-P2P融合

为了缓解流媒体带宽占用率高、持续时间长对主干网络造成的压力,将用户需求与 CDN,P2P 相结合是目前流媒体主流的服务技术。CDN 是把流媒体服务器布置在网络的边缘,采用重定向策略,将用户需求转向距用户最近的边缘服务器。P2P 技术充分利用了用户端的资源,能够在不增加成本的情况下满足用户激增的需要。

目前,基于用户需求的 CDN,P2P 在流媒体领域的融合主要表现为^[21]:以 P2P 方式组织 CDN 节点设备,利用 P2P 目录服务和多点传输能力,依据用户需求,实现 CDN 设备之间的内容交换和备份,由此增强了 CDN 网络中心节点到边缘节点的内容传送效率,从而提升流媒体的内容分发能力。这种融合的优势在于:一方面,在 CDN 网络核心层采用 P2P 技术进行优化,重新组织原有 C/S 架构的各服务节点;另一方面,在 P2P 网络中融入 CDN 技术的优势,使其在可控、可管理性上得到大幅度提升。在两种技术充分融合、渗透、改造后,又结合用户需求,将三者合为一体,形成一个真正优化、完善的 CDN 与 P2P 流媒体内容分发平台。

5 用户需求 CDN-P2P 环境下用户节点的相似性

本节从 4 个方面对 CDN-P2P 环境下的用户节点相似性进行总结,分别是邻居节点的相似性、用户节点内容的相似性、用户群体结构的相似性、用户行为的相似性.其中,用户节点内容的相似性是核心问题,如何计算不同用户之间的相似度以及计算用户相似度所采用的方法是研究重点.

5.1 邻居节点的相似性

通过对众多基于邻居节点相似性算法的研究,Liben-Nowell 和 Kleinberg^[36]发现,Common Neighbors(CN)算法^[36]和 Adamic Adar(AA)^[37]算法比其他已知算法有更好的实验效果.周涛等人^[38]提出的基于邻居节点相似性的 Resource Allocation Index(RA)算法,其在预测准确性方面要高于 CN 和 AA 算法.RA 算法针对网络中互不直接相连的两个节点 x 和 y 的情况作了进一步的分析,当资源从节点 x 传递到 y 时,它们的共同邻居节点 z 成为传递资源的媒介.节点 z 接收到的资源总数反映了节点 x 和 y 的相似程度.文献^[39]将邻居节点相似性度量分为两个阶段:首先提出了一种图结构相似度量机制,来判断两个邻居节点建立连接的方式;然后利用邻居节点含有的用户信息,提出一种基于用户配置文件的邻居节点相似度量策略.该策略有效降低了计算相似性时系统的复杂度.文献^[40]结合已有的基于节点相似性算法,提出了一种新的预测算法——节点引力指数算法.该算法在考虑被预测两个节点共同邻居度数的基础上,同时考虑被预测节点与共同邻居组成的小型网络对预测结果的影响.

网络中节点之间的关系与现实世界中的社会具有一定的相似性.具有相同或相似兴趣偏好的一组节点可以看作一个节点群,节点间的相似可以互相满足节点部分的需求,但不能满足所有节点的所有需求.当一个节点群的所有节点都不能满足查询需求,进行跨节点群查询内容时,如何较快地找到所需内容,是当前基于用户需求的 CDN-P2P 所面临的问题.针对这些问题,文献^[41]提出了用户节点群相似性的概念和基于关键词的节点群相似性度量模型,计算结果不受关键词的顺序和节点排列顺序的影响,且与实际情况更接近.

向量空间模型 VSM(vector space model)^[42]算法是计算邻居节点相似度的常用算法,通过将用户查询向量化、文档向量化、计算两向量夹角的 cosine 值等步骤来计算用户查询与文档的相似度(见公式(4)).进而构造计算模型,计算系统相似度如公式(5)所示.

$$Similarity(x, y) = \frac{\sum_{i=1}^N x_i y_i}{\sqrt{\sum_{i=1}^N x_i^2} \sqrt{\sum_{i=1}^N y_i^2}} \quad (4)$$

$$Q(A, B) = \frac{\sum_{i=1}^p \mu_i x_i^2}{\sqrt{\sum_{i=1}^m x_i^2} \sqrt{\sum_{i=1}^p \mu_i^2 x_i^2 + \sum_{i=p+1}^n y_i^2}} \quad (5)$$

皮尔森聚类度量公式⁽⁶⁾^[43]可以度量出两个用户动态变化的相似度,公式(6)在如下假设中成立:两个用户动态变量必须呈线性关系,它们也必须是动态变化、正态分布、相互独立的.

$$sim(v, u) = \frac{\sum_{h=1}^n (r_{v,i_h} - \bar{r}_v)(r_{u,i_h} - \bar{r}_u)}{\sqrt{\sum_{h=1}^n (r_{v,i_h} - \bar{r}_v)^2} \sqrt{\sum_{h=1}^n (r_{u,i_h} - \bar{r}_u)^2}} \quad (6)$$

5.2 用户节点内容的相似性

在 CDN-P2P 中,用户兴趣范围是相对集中且稳定的,他们所关注的主题相对固定.因此,信息交流与传递关系更多发生在主题内容相似的用户对之间,这些内容相似的主题体现了用户共同关注的话题.根据用户节点内容相似性来确定用户之间的信息传递关系,能够及时发现整个网络社会的用户特性.用户节点内容的相似性包括:用户资源列表的相似性、用户发表资源的相似性、用户资源数据清洗和过滤的相似性、对各个用户建立资源向量空间模型的相似性、用户两两之间文本相似度计算、用户交流关系网络的建立以及利用网络软件对网

络分析^[41]等。

文献[44]提出了一种基于论坛用户发帖内容相似度的用户社会网络的构建与挖掘方法,借助文本挖掘技术与社会网络分析技术分析了网络用户的基本特性,从用户发送的文本内容角度来构建用户社会网络.其中,用户内容相似算法各项执行的时间见表 2^[44].

Table 2 Time of items on user node content similarity algorithm

表 2 节点内容相似度算法各项执行时间

算法各项执行时间	时间消耗(m)
内容抓取执行时间	60
文本分词执行时间	1.52
生成 VSM 模型和相似度矩阵时间	12.45

定义两个用户节点内容是否相似,既要判断内容的共性,还要考虑内容的差异性.文献[45]定义了一个用户节点内容的共性函数 $c(v,u)$ 和一个用户节点异性函数 $d(v,u)$,然后通过公式(7)计算用户 v,u 内容相似度.接着,利用指数函数将用户内容相似度进行扩展,相似度扩展公式如公式(8)所示, α,β 是平滑因子.

$$sim(v,u) = \frac{c(v,u)}{c(v,u) + d(v,u)} \quad (7)$$

$$sim(v,u) = \frac{e^{\alpha c(v,u)} - 1}{e^{\alpha c(v,u)} + e^{\beta d(v,u)} - 2} \quad (8)$$

5.3 用户群体结构的相似性

基于用户需求的 CDN-P2P 网络系统能够利用用户群体结构的相似性发挥强大的智慧效用,这是一种自上而下的数学结构所导致的必然结果.用户在某个特定的社会网络中,以自己的心理取向和志趣进行着特定的网络行为,一旦选择定向,就带来了网络的频繁互动;这种互动一旦形成习惯,就具有了某种群体交流的结构,这样就产生了以用户倾向为基础的群体思维,用户就会聚合形成或发现自己所属的群体^[46].

用户群体结构的相似性还体现在社会网络中,社会网络中的用户在彼此之间交往的过程中,会发现对方没有与自己相同之处,比如性格、爱好、对某些事物的观点等.如果对方与自己本身的性质有相同之处,就会保持联系,继续交往下去.用户往往喜欢与自己各方面都相似的用户交往,这是一种相互同化的效应.因此,网络用户由于自身的需求表现出群体性;其次,用户需求在网络中寻求认同感,找到与自己兴趣相投的用户,在相互交流和分享中就形成一定的群体.具有共同需求或兴趣的用户,通过各种网络行为联系在一起,以获取信息和相互沟通.

Pecora 和 Carroll^[47]首次提出网络混沌同步以后,人们开始逐步地研究混沌网络下用户群体结构的相似性.一个重要的问题是:如何利用用户群体结构的相似性来解决复杂网络中网络资源搜索、存储和计算同步的问题?大量节点在已有的现实网络混沌同步系统^[46]中所具有的特征是同结构性,仅在系统的初始条件上存在差异.用户需要在网络中寻求存在感和认同感,同时,用户行为不受时间和空间的限制.

5.4 用户行为的相似性

Web 2.0 环境下,用户在网络上发布和分享信息知识,同时更期望与具有相同兴趣的其他用户进行关联、互动和协作,表现出强烈的群体性,并影响周围其他用户的网络意识和行为.用户在使用网络的过程中,由于具有各自的兴趣爱好和访问习惯各异,其关注的内容、使用网络的时间和服务也就不尽相同;并且,不同的网络用户群体不可避免地带有各自的群体行为特征.网络环境下的群体性趋势和群体效应不但能够聚合志趣相投的用户,而且对用户的行为具有很大的引导作用.由于用户行为存在从众心理,而从众行为是在群体的压力下个体寻求解决自身与群体之间冲突、增强安全感的手段,用户在不知不觉中就会倾向或从属于某一群体;当个体数量达到一定程度时,群体的行为就表现得有章可循,就会在混沌中表现出秩序性和稳定性,用户的行为就会出现特定的群体性,这就是用户群体行为的相似性^[48].

由于位置服务网络的流行,近些年来,对位置服务网络用户行为的分析成为了研究重点之一.文献[49]的研

究表明,用户的签到次数和访问的位置服从正态分布;同时给出结论:用户的签到次数不随用户好友数的增加而增加.文献[50]主要分析位置服务网络用户的使用习惯,例如用户使用时间、主要访问地区等.进而也有文献利用大量的地理位置数据,研究用户行为轨迹的相似性.为了给用户推荐合适的好友和感兴趣的地点,仅仅通过语义相似是不够的,在用户活动轨迹记录并不连续的情况下,利用用户访问的实际地理位置坐标计算用户的相似性是较为合理的方法.文献[51]基于用户历史签到的地理位置研究用户行为的相似性,提出基于划层次密度的聚类方法.该方法首先对用户签到的地理位置进行聚类操作,得到用户访问的位置区域,并通过改变聚类的邻居半径,在不同空间位置比例下观察用户访问各个位置区域的情况,进而利用向量空间模型,计算用户在不同空间比例下的相似性,并最终得到用户行为的轨迹相似性.使得用户推荐的访问轨迹具有相似的地点和好友.通过分析用户行为轨迹,可以得出用户在真实世界的地理位置上的行为方式.在判断用户地理位置的行为相似性时,可参考如下经验:

- (1) 用户所访问的地理位置越接近,用户行为轨迹越相似;
- (2) 用户访问相近的地理位置次数越多越相似.

通过研究用户行为的相似性,建立基于 CDN-P2P 系统的用户细分体系;用改进的 k 均值(k -means)算法和分类决策树算法对实际的用户行为数据进行分析,文献[51]表明,这样可极大地改善系统为用户提供服务的体验.

6 用户需求 CDN-P2P 环境下用户节点之间的关系

基于用户需求的 CDN-P2P 网络结构不同于客户端/服务器(C/S)结构,P2P 中所有节点都是平等的,也不存在权威中心对这些节点进行管理.因此,如何建立互信机制使用户彼此协作,以及如何防止用户之间的恶意行为和黑客攻击造成用户对系统安全的担忧,是近年来基于用户需求的 CDN-P2P 系统面临的主要难题.

6.1 用户之间的相互协作

基于用户需求的 CDN-P2P 系统是用用户相互协作进行资源分发的系统,在该系统中,用户之间提供彼此需求的资源,在一个用户下载资源的同时,拥有该资源的其他用户充当服务器角色,为该用户提供资源下载.此类用户越多,下载速度越快.用户也可以将感兴趣的资源推荐发送给其他用户,还可以将资源搜索记录发送给其他用户作为参考;另一方面,基于用户需求的 CDN-P2P 系统还可以对用户之间相互推荐的搜索结果进行排序,并将排序结果推荐给用户.

在私有数据库广泛应用的同时,对信息共享的要求也越来越高,保护参与各方的私有信息成为了热点问题.这要求在用户相互协作查询结果时,其他参与方的私有信息对该用户透明.目前,私有信息检索(private information retrieval,简称 PIR)被广泛应用于各种领域.2003 年,Agrawal 等人^[52]提出了一个有效的方案,该方案在不借助第三方的情况下,凭借双方对各自明文加密的结果再次对对方的密文进行二次加密,得到两组二次加密的结果,再进行查询操作,在不外泄用户私密信息的前提下,完成了查询协议.该方案极大地保护了双方私有信息的安全,但也使得参与方陷入复杂而繁重的计算中.2007 年,Malaga 等人^[53]提出了一个借助第三方服务器的解决方案,该方案使得查询过程在线完成,极大地降低了通信复杂度.

6.2 用户之间的信任度

在线信誉系统为建立在线信任关系提供了新的有效保障,在线信誉系统中,信任模型处于核心地位,是度量用户信任度、惩戒欺诈交易行为的关键.文献[54]从理论上验证了当前 CDN-P2P 模型存在的缺陷及危害,并提出了评分用户可信度信任模型.在社会网络中,个体间的信任度取决于其他个体的推荐,推荐者的可信度决定了其推荐个体的可信度.基于该思想,Zachara 等人^[55]对在线信誉系统中的信任模型进行了改善,所提出的 Sporas 模型直接将评分用户的信任度作为信任计算的因子,但它仅是一个概念模型,并未给出理论证明,其有效性和正确性有待验证.文献[26]提出了一个基于评分可信度的信任模型,并对评分用户可信度的选取给出了理论证明.模型利用评分用户提供的有关潜在交易伙伴的信誉反馈评分信息,通过考察评分用户信任度,计算信任度模型.Zucker^[56]将信任的产生机制分为 3 种:

- (1) 声誉产生的信任:通过其过去的行为以及声誉的了解给予一定的信任,可以理解为他人的熟悉性产生的信任;
- (2) 由社会相似性产生的信任:根据其家庭背景、种族、价值观念的相似程度给予一定的信任;
- (3) 由法制产生的信任.

在分布式 P2P 网络计算环境中,网络安全管理是一个难题.文献[57]针对现有信任模型存在的问题提出了一种用户行为信任模型 NATM,以评估用户的信誉和名声,并通过评估结果进行用户权威度的计算,用以区分用户的可信度,从而根据用户行为对用户进行有效分类.

文献[58]提出了一种同时基于用户聚类 and 资源聚类协同完成推荐和内容推送的方法,并证明其具有较高的可信度,有效地解决了数据可靠性问题.

6.3 用户之间的安全机制

AT&T 实验室首先提出信任管理的概念,这类管理系统通常被称为基于能力的授权系统.这种系统通常需要服务方预先为请求方颁发制定操作权限的信任证,但无法与陌生方建立动态的信任关系.依赖主体属性授权,是为陌生方之间建立信任关系的一种有效方法.

文献[59]针对 CDN-P2P 网络的分布式计算特点以及现有 CDN-P2P 网络安全管理机制存在的弊端,提出了基于信任机制的 CDN-P2P 网络安全管理模型,引入信任机制对用户行为进行评估,并将用户的信誉度与用户权限进行动态绑定,设计实现了基于信任的动态访问控制关键算法,及时阻止用户的恶意行为.

在基于用户需求的 CDN-P2P 软件系统中,系统中的安全需求及用户安全机制分成 3 个要素层^[60]:

- (1) 功能对象层:根据 CDN-P2P 的基本架构,形成完整的功能树.其中的每个叶节点就是一个功能对象,根据实际的功能,赋予相应的访问属性,如只读、读写、隐藏、有效期等.在 CDN-P2P 系统中,由于使用的对象主要是独立的用户,涉及的功能也有所不同,所以对对象层选择功能树中的叶子节点作为最底层的安全要素;
- (2) 身份对象层:由于每个用户当前处理的问题不同,身份也有所不同.为了方便用户,系统中引入了身份层.每个用户有唯一的账户/密码,但可以有多重身份.同时,在基于用户需求的 CDN-P2P 系统中可以短暂地将自己的部分权限授予临时用户,因此引入授权机制.而授权也是身份的一种.如果选择了授权身份,则需要进一步判定授权人.换句话说,一个用户可以将自己的某种身份授权给多个其他用户代理,而一个用户也可以同时代理多个其他用户的多个身份;
- (3) 用户对象层:用户对象层处于用户安全机制的最上层,每个用户利用自己的账户/密码登录,选定身份后进行操作.

7 基于用户需求 CDN-P2P 环境下用户节点资源

Internet 上充斥着海量的信息,并且还在不断地迅速增长.人们依靠传统搜索引擎寻找信息和资源,其效果已经不尽如人意.CDN-P2P 技术带来了一种全新的检索理念和方式,必将成为未来搜索引擎的发展方向之一.本节针对基于用户需求的 CDN-P2P 资源搜索和资源定位,进行了简单的总结和比较,并指出了基于用户需求的 CDN-P2P 检索技术未来的发展方向.

7.1 用户节点的搜索机制

搜索是 CDN-P2P 信息共享系统中的研究重点,它直接影响信息共享服务的质量和系统性能.在 CDN-P2P 搜索中,发起查询的节点发送查询数据包给所选的部分或全部邻居节点,如果收到查询的节点含有要查询的内容,则应答查询节点,否则转发该查询包.这个简单的搜索过程体现了 CDN-P2P 搜索的鲁棒性:动态变化对搜索几乎不产生影响,搜索可在任何网络拓扑上顺利进行.

7.1.1 基于用户需求的 CDN-P2P 搜索机制

无论是 Google、百度还是其他搜索引擎,其搜索任务都要依靠服务器来完成,信息检索资源存放于固定的

数据库中,它们找到的只是静态网页的索引,研究表明,它们仅能覆盖网络中 28%~50%的资源^[61].而基于用户需求的 CDN-P2P 技术无需通过 Web 服务器就能使用户深度搜索文档,用户节点可以自由加入或离开网络,检索到动态变化的分布式信息资源,并且对宿主设备和信息文档的格式无任何要求,远远超过传统目录式搜索引擎的搜索深度.

7.1.2 基于用户需求的 CDN-P2P 搜索机制分类

根据所利用的结构模式的不同,可以把基于用户需求的 CDN-P2P 搜索技术分成两大类^[62],如表 3 所示:一类是基于结构化系统,主要是采用分布式哈希表(distributed Hash table,简称 DHT)技术来组织网络中的节点;另一类是基于非结构化系统,根据其资源定位的方法不同,又可以分为集中式、全分布式和混合式.

Table 3 Type of CDN-P2P search mechanism

表 3 CDN-P2P 搜索机制分类

CDN-P2P 类型	搜索机制分类
有结构	全分布式
无结构	集中式
	全分布式
	混合式

7.2 用户节点资源定位

在 CDN-P2P 网络中,如何快速、准确地对资源定位,是当前研究的热点问题,也是 CDN-P2P 应用面临的核心问题之一.基于用户需求的 CDN-P2P 系统是为用户资源查询和定位而设计的,从诞生之日起,就以其优越性征服了使用者.随着 CDN-P2P 网络规模的扩大以及用户节点数量的增多,网络节点之间的最大路径也在不断增大.在这种情况下,泛洪算法^[63]的广度优先搜索和随机查找法深度搜索的弊端愈发凸显出来.泛洪算法中的海量搜索严重增加了网络的负载,随机查找法需要查找更大的深度,从而延长了查找时间,不能及时反馈查询结果.

7.2.1 基于用户需求的泛洪算法(flooding)

文献[61]将泛洪算法总结为启发式泛洪搜索和盲目式泛洪搜索两种类型,在无线传感器网络平台的基础上分析经典的泛洪路由算法,提出了改进型的泛洪路由算法.文献[62]提出了基于用户需求的泛洪算法改进策略,包括基于重叠改进策略和基于局部知识的改进策略.在此基础上,文献[64]提出了基于用户需求的局部观测的概率论转播泛洪算法,认为每个用户节点的转播概率都是时间和局部观测的一个函数,算法实验结果表明,在开销和延迟上,相比其他算法有明显的优势.

7.2.2 基于用户需求的随机漫步法(random walk)

在 CDN-P2P 领域中,另一种用作资源定位的算法是 RW(random walk)算法.文献[65]提出了一种用户需求式的随机漫步动态进化算法,并且分析了该算法的时间覆盖机制.该算法能够减轻网络风暴的发生,并能降低随机转发的能源损失.文献[66]提出了一种用户需求下的软件结构测试产生的随机漫步算法,该算法可以提供带有随机漫步参数的最佳搜索策略框架,提高了搜索效率,实验结果表明,在查全率和查准率方面有明显的改善.

7.2.3 基于用户需求的 CDN-P2P 资源定位策略

CDN-P2P 系统的查找定位行为对存储空间和计算能力的需求分布在不同节点上,文献[60]提出的查找策略是通过每个用户节点对系统资源位置的掌握和处理,减小了查找的深度和广度,提高了查找效率和命中率.与传统查找策略相比,对单个用户节点存储能力和计算能力要求更高,单个用户节点的相对处理时间较长.此外,在模拟实验中,大量模拟节点的行为在模拟环境中发生,对模拟环境所处的服务器内存和计算能力要求较高.模拟结果显示:在 P2P 网络中,若能发挥用户节点的能动性,则文件的查找定位效率会随着文件查找次数的增加而大幅提升,并逐渐趋于稳定.

8 基于用户需求环境下 CDN-P2P 的应用

近年来,CDN-P2P 技术在文件共享、流媒体方面得到了广泛关注.CDN-P2P 流媒体即基于 CDN-P2P 技术实现的网络流媒体,具有用户越多,播放越流畅、越稳定的特性,能够支持大量用户同时在线的大规模访问^[67]. CDN-P2P 流媒体提高了资源共享的利用率,充分利用空闲资源分发数据,避免拥塞,提供具备高实时性和容错性的流服务,具有分散化、可扩展性、健壮性、低成本、高性能等特点^[68].

8.1 CDN-P2P在流媒体上的应用

CDN-P2P 流媒体技术已经成为全球互联网上增长最快、带宽占用最大的业务流量,并继续快速增长.预计到 2015 年左右^[46],整个 P2P 流媒体市场将趋于成熟.目前,各类 P2P 流量占全部互联网流量的 70%~85%,其中, P2P 流媒体流量已占全部 P2P 流量的 60%,占全部互联网流量的 36%.当前的 P2P 流媒体业务运营商主要有 PPLive,PPStream,QQLive,Anysee,悠视等,其中,PPLive 是当前用户数量最多的 P2P 流媒体软件,同时在线最高人数为 1 000 万,平时在线人数为 500 多万^[69].

CDN-P2P 系统有效缓解了 Internet 网络拥塞状况,提高用户访问网站的响应速度,从技术上全面解决了网络带宽小、用户访问量、网点分布不均而造成的响应速度慢的问题^[70].CDN 是目前采用比较普遍、技术成熟度比较高的一种平台.2007 年,全球范围内 CDN 市场发展迅速^[71],预计未来 CDN 市场将持续快速发展,动态网页、视频点播、在线交易应用等 CDN 加速类业务将迅速成长.

互动电视 VOD^[72]作为一种新型的媒体服务,具有巨大的市场潜力,但目前的服务器无法承受大规模的并行请求.当发生热点事件及出现浪涌时,会使应用服务器因过载而退出服务.为解决这一问题,文献[72]提出了一种融合 P2P 和 CDN 技术的基于用户需求的有线电视网络模型,用户可以访问所需内容,系统能够实现全局负载均衡.由二八定理可知^[73],用户基本能够在边缘节点找到所需内容.可见:CDN 的加入,解决了并发数据流访问源服务器而导致的服务器性能瓶颈问题.

8.2 CDN-P2P在IPTV中的应用

目前,基于用户需求的 CDN-P2P 技术是 IPTV 领域的一个研究热点^[74].由于 CDN 网络是一种分域网络,在一个特定域内的边缘服务器并不是很多,在这个域内的边缘服务器之间实现 P2P 并不能给整个网络的容量带来质的飞跃.IPTV 作为三网融合的必由之路^[75],已得到全球电信运营商的广泛青睐.文献[76]提出了 P2P 与 CDN 结合实现 IPTV 业务,将 P2P 技术引入到 IPTV 系统,有效减少了边缘服务器的数量,增大了系统容量,从而降低了总体成本.将 CDN 的管理和内容分发机制引入到 IPTV 系统中,形成以 CDN 为可靠的内容核心、以 P2P 为服务边缘的架构、以用户需求为最终导向的模式.由此,在不增加 CDN 成本的同时有效提升 CDN 服务能力,使用户访问媒体的质量得到极大的提高.

8.3 CDN-P2P在移动计算领域中的应用

随着移动计算的快速发展,用户在移动平台访问互联网资源的场景和需求越来越多.为了提升消费者的用户体验,引入 CDN 和 P2P 技术将显得越来越重要.同时,GPRS,UMTS,B3G 等蜂窝移动网络和 Wi-Fi,WiMAX,UWB 等宽带无线接入网络,为移动用户提供了随时随地访问无线网的可能性.文献[77]提出了一种全 IP 宽带移动 CDN-P2P 网络分发模型,该模型通过在数据网关上增加分布式用户服务器,实现异构网络的互联,形成虚拟二维覆盖网络;采用 IMS 为核心控制网络,基于此,结构可开展各种移动 CDN-P2P 业务.针对移动网络中节点移动性强、网络拓扑结构变化剧烈而导致的系统低效和不稳定的问题,文献[78]提出了一种在移动 CDN-P2P 网络中的超级节点选择算法,该算法能够选择性能好、在线时间长的节点作为内容预存的超级节点,每当进行内容分发时,这些超级节点采用候补技术,有效改善了移动网络的分发效率.Wichtlhuber 等人^[79]在此基础上将内容分发和资源定位与移动平台结合起来,提出了一种移动 CDN-P2P 网络资源定位调度算法.该算法能够将移动终端资源信息依据资源类型进行归纳,以集中管理,并将移动 P2P 网络中的节点在共享过程中快速定位到其他节点,实现了就近连接与分发机制,以消除移动终端与资源类型的差异性所造成的移动 CDN-P2P 网络节点无法访

问资源信息并影响到共享分发效率等问题。

虽然移动终端已经有了长足的发展,但与互联网中的主机相比,移动终端仍有其自身的局限性:

- (1) 有限的处理和存储能力,很难满足云计算的需求;
- (2) 有限的无线资源.在移动环境下,无线资源非常宝贵,而传统的 CDN-P2P 架构需要消耗大量的网络资源,这需要在 P2P 业务和无线资源之间取得一种平衡.

此外,移动终端操作系统的统一也给 P2P 应用程序或用户在终端上的部署带来了难度.

9 基于用户需求环境下 CDN-P2P 的难点和热点

目前,基于用户需求的 CDN-P2P 主要应用于互联网,其本身并不考虑控制、管理等问题.但这些问题在应用中是不可回避的,因此在内容分发系统中如果引入 P2P 技术,就必须考虑和解决这些问题,如业务时延问题、网络不可控问题、ADSL 节点上下行带宽不对称问题、NAT/FW 穿越的问题、安全性问题等.

基于用户需求环境下 CDN-P2P 技术难点和热点可归纳为以下几个方面:

(1) 用户需求获取的局限性

用户需求获取,是基于用户需求 CDN-P2P 系统进行网络分发的前提.由于用户偏好的多样性和差异性,使得用户需求获取变得复杂.随着时间的推移,用户偏好不是一成不变的^[80](例如,用户原本喜欢看动作类电影,后来喜欢看喜剧类电影),因此,这需要更为合理的需求获取检测机制.在不影响时间和空间复杂度的情况下,如何更合理地优化用户需求获取算法,是值得关注的研究方向.

(2) 基于用户流媒体的对等网络传输模型和算法

流媒体作为一种重要的网络应用,其用户规模迅速扩大,所需的服务器带宽也急剧增加.基于用户需求的 CDN-P2P 技术可以在一定程度上加速流媒体分发,实现下载、直播和点播等业务,但是从体系结构的角度来看,基于用户需求的 CDN-P2P 系统的流媒体分发过程仍然是集中式的.随着用户规模的迅速增加,用户在获取资源时,用户体验得不到保证;当大规模的视频流媒体业务运营时,需要扩建大量的 CDN 设备,投资巨大,同时开发维护成本也大为增加^[79].

(3) 用户查询模式中的语义映射方法

当前,基于用户需求的 CDN-P2P 查询模式研究的主要内容是:如何在节点间建立语义映射,为节点间语义相互操作提供基础.未来还需要一种全新的语义映射方法来解决查询过程中语义带来的冲突问题;同时,也需要能够优化大量数据源的查询语义映射方法^[81,82];并且随着网络的进化,查询过程会产生大量的语义映射,因此需要研究管理这些语义映射的技术,包括:找出应该建立而尚未建立语义映射的节点集、发现 P2P 局部交互产生的语义矛盾映射、解决语义矛盾映射方法;寻求处理不同质量的语义映射途径,逐步改善语义映射,从而改进用户节点从其他节点获得数据的能力.

(4) 提高用户即兴查询处理能力、优化用户即兴查询算法

在用户即兴查询方面,基于用户需求的 CDN-P2P 系统的关键字搜索仍然是悬而未决的问题;分级关键字搜索需要对用户资源文档进行全局统计(如文档频率),目前还没有收集和维持统计信息的健壮、高效、分布式用户即兴查询算法^[83].另外,在数据库查询领域,现有的范围查询(除 GridVine 外)、连接查询、聚集查询算法外,均没有与用户需求建立必要的联系,大多是利用高可扩展实例级索引(绝大多数利用重叠网络)进行分布式查询处理.然而,基于用户需求的 CDN-P2P 网络的异构性是不可避免的.因此,今后需要研究利用高可扩展实例级索引提供上述查询途径.

(5) 高可扩展性用户资源索引构造算法的缺失

现阶段,只有基于用户需求的 CDN-P2P 索引机制考虑用户资源索引模式异构问题,虽然当前此类索引的可扩展性有限,但却提供用户资源语义支持^[84],其余各类索引均假定系统所有节点都采用相同模式,仅提供数据级索引.如何将用户资源异构支持和高可扩展性相融合、构造高可扩展的用户资源索引,是当今基于用户需求的 CDN-P2P 系统所面临的难题.为了给用户精确和详细的资源定位,扩展性高的用户资源索引构造算法需要

进一步加以优化。

(6) 用户资源定位 Flooding 泛洪算法产生“网络风暴”^[62]和“信息重叠”^[64]

在非结构化 CDN-P2P 网络中,如何快速、准确地对用户资源定位,是当前研究的热点问题。而作为基于用户需求环境下的 CDN-P2P 系统的主流用户资源定位算法——泛洪算法,存在“网络风暴”和“信息重叠”等问题。这种算法不要求维护网络的拓扑结构,接收到用户查询消息的节点以广播的形式转发分组,这会造成严重的网络带宽负担以及巨大的索引维护开销。如何在基于用户需求的环境下产生节约带宽、容易维护、查询命中率高的泛洪算法,是下一步研究的方向。

(7) 用户资源定位 Random Walk 算法搜索命中率低、网络资源消耗过大

基于用户需求环境下的 CDN-P2P 技术的另一种用户资源定位算法——随机漫步算法,存在过于依赖历史搜索记录而导致动态网络环境下搜索命中率低、网络开销过高和稀有资源的搜索成功率提高不明显等问题,这些问题在动态的网络环境中降低了搜索成功率和资源定位的准确性^[65,66]。由于 P2P 网络本身具有强扰动性,邻居节点、资源状态不断发生变化,过多地依赖历史记录进行搜索,可靠性不高。因此,在请求搜索资源的同时,如何摒弃依赖节点历史记录、利用随机漫步的搜索特性、增强算法的抗扰动能力,是下一步研究的方向。

(8) 网络动态性及复杂性导致用户流量检测和识别上的困难

基于用户需求的 CDN-P2P 网络技术是网络流量工程的研究热点。近年来,随着 CDN-P2P 技术的不断发展,P2P 用户端流量已经占用整个 CDN-P2P 网络整体流量的 70%^[85],一方面,给网络带宽造成严重的负担,以近乎对称的流量模式加剧了网络拥堵状况;另一方面,基于用户端的恶意流量也频繁出现在互联网上,大量的非法连接加快了带宽的消耗,甚至导致网络瘫痪。因此,如何识别和控制用户端流量,已经成为网络运营商和管理者面临的难题,其主要困难有^[86]:P2P 流量的动态性使得某些状态下,流量的类别及流量的识别面临一定的挑战:一是负载内容加密协议文档不公开;二是下载软件支持多种协议下载方式,网络行为模式不显著,这样加剧了网络拥塞状态。

(9) 基于用户需求的 CDN-P2P 系统实时性差,存在时延

实时性不好是基于用户需求的 CDN-P2P 的一个问题,这使得来自远程服务器的网络内容与副本服务器或缓存器中的网页不能保持同步,造成用户访问网络资源时出现信息不同步现象,影响用户的访问体验^[87]。另外,由于网络行为的复杂性,使得利用机器学习和数据挖掘技术设计出来的基于用户需求的 CDN-P2P 系统极易出现主题更新不同步现象^[87],大大降低了其搜索引擎的查全率和查准率。如何避免访问信息不同步、主题更新不同步现象的发生,可以作为进一步的研究方向。

(10) 用户资源存储机制不完善,节点选择不合理

基于用户需求的 CDN-P2P 存储机制一直是学术界所关注的热点问题,也被认为是 P2P 最具前途的应用之一^[87]。用户资源数据的持久存储,是制约 CDN-P2P 存储系统发展的难题。它带给数据备份业务的代价是巨大的,最直接的问题是:系统如何选择恰当的节点来存放用户所需的资源数据。因此,从微观的角度看,选择一个合适的节点对于基于用户需求的 CDN-P2P 系统是非常重要的,只有选择合适,才能保证用户资源存储速度快、可靠性高等特点。

(11) 用户的隐私和安全性问题

用户的隐私和安全性问题制约着 CDN-P2P 系统地发展。基于用户需求的 CDN-P2P 系统为了给用户提供所需资源,必须记录并分析用户的信息、行为、位置等;但是,出于隐私与信息安全的考虑,用户不愿意提供完整和准确的信息。通过一定时期的用户需求获取,利用数据挖掘和机器学习经典算法,可以挖掘出用户偏好轨迹,而用户偏好轨迹属于一种特殊的个人隐私,它本身包括敏感信息,如用户访问 ID、用户访问地址、用户访问资源、用户生活习惯等。因此,如何保护用户偏好也是需要进一步研究的方向^[60]。

(12) 用户在移动计算平台下的信任机制问题

在信任管理方面,移动 CDN-P2P 系统因为缺少可信任的中央控制节点,节点间的信任和授权不能采用有线网络中基于中央控制节点授权的安全机制。对于任何节点都必须采用有效的方法来判定可信节点和非可信节

点,因此必须针对移动 CDN-P2P 网络分散化的特点来设计有效的分布式认证协议^[78]。目前,移动平台下的 CDN-P2P 信任机制的研究还处于起步阶段,研究规模和成果不多,可以考虑借鉴和改进其他系统中的安全机制、信任模型、匿名算法和认证鉴权流程,或是根据移动 CDN-P2P 自身的特点提出全新的解决方案。

10 总结及展望

作为一种新型的网络技术,基于用户需求的 CDN-P2P 技术正在不断地发展和完善,是未来流媒体传输和信息检索、网络资源定位的关键技术。为了满足用户对 CDN-P2P 系统各项指标以及在数据库中可扩展性的需求,基于用户需求的 CDN-P2P 应用软件正在不断地更新换代。新的应用软件的涌现、系统成员的增加,使得该系统具有丰富性、多样性、健壮性、可扩展性,具有更强的信息资源查询处理能力和存储能力。此外,通过许多 Peer 间路由请求和复制内容,系统可以隐藏数据提供者和服务者的身份,使得用户的隐私得到保护,它在搜索引擎、数据流管理、语义网、协作信息过滤等领域具有广阔的应用前景。目前,国内外对基于用户需求的 CDN-P2P 研究领域的综述性文献较少,本文将该领域的研究进展和趋势进行归纳总结,希望促进该领域的研究工作。

References:

- [1] Liu J, Zhang L, Fan SH, Guo C, He S, Chang GK. A novel architecture for peer-to-peer interconnect in millimeter-wave radio-over-fiber access network. *Progress in Electromagnetics Research*, 2012,126(4):139–148.
- [2] Pakkala D, Latvakoski J. Towards a peer-to-peer extended content delivery network. *VTT Electronic*, 2012,38(7):145–156.
- [3] Parameswaran M, Susaria A, Whinston AB. P2P networking: An information-sharing alternative. *IEEE Computer Society*, 2001, 34(7):31–38. [doi: 10.1109/2.933501]
- [4] Stavrou A, Rubenstein D, Sahu S. A lightweight, robust P2P system to handle flash crowds. *IEEE Journal on Selected Areas in Communications*, 2012,22(1):6–17. [doi: 10.1109/JSAC.2003.818778]
- [5] Stamos K, Pallis G, Vakali A, Dikaiakos MD. Evaluating the utility of content delivery networks. In: *Proc. of the 4th Edition of the UPGRADE-CN Workshop on Use of P2P, GRID and Agents for the Development of Content Netw, UPGRADE-CN, Colocated Int'l Symp. High Perform. Distrib. Comput. Conf. (HPDC 2009)*. New York: ACM Press, 2009. 734–745. [doi: 10.1145/1552486.1552509]
- [6] Hawa M, Rahhal JS, Abu-Al-Nadi DI. File size models for shared content over the BitTorrent peer-to-peer network. *Peer-to-Peer Networking and Applications*, 2012,5(3):279–291. [doi: 10.1007/s12083-011-0122-6]
- [7] Wang Z, Jiang H, Sun Y, Li J, Liu J, Dutkiewicz E. A k -coordinated decentralized replica placement algorithm for the ring-based CDN-P2P architecture. In: *Proc. of the IEEE Symp. on Computers and Communications (ISCC)*. IEEE Computer Society, 2010. 811–816. [doi: 10.1109/ISCC.2010.5546711]
- [8] Gopalakrishnan V, Silaghi B, Bhattacharjee B, Keleher P. Adaptive replication in peer-to-peer systems. In: *Proc. of the Int'l Conf. on Distributed Computing Systems*. Tokyo: IEEE Computer Society, 2004. 360–369. [doi: 10.1109/ICDCS.2004.1281601]
- [9] Manku GS. Balanced binary trees for ID management and load balance in distributed Hash tables. In: *Proc. of the 23rd Annual ACM Symp. on Principles of Distributed Computing*. New York: ACM Press, 2004. 197–205. [doi: 10.1145/1011767.1011797]
- [10] Zhu YW, Hu YM. Efficient, proximity-aware load balancing for DHT-based P2P systems. *IEEE Trans. on Parallel and Distributed Systems*, 2005,16(4):349–361. [doi: 10.1109/TPDS.2005.46]
- [11] Jin H, Guo XZ, Zhang C, Yuan Q. A new layered P2P architecture with efficient resource location strategy. In: *Proc. of the IEEE Int'l Conf. on E-Commerce Technology for Dynamic E-Business (CEC-East 2004)*. Beijing: IEEE Computer Society, 2004. 298–301. [doi: 10.1109/CEC-EAST.2004.5]
- [12] Shen YG, Liu J, Shen J. The further development of weka base on positive and negative association rules. In: *Proc. of the Intelligent Computation Technology and Automation (ICICTA 2010)*. IEEE Computer Society, 2010. 811–814. [doi: 10.1109/ICICTA.2010.676]
- [13] Sulthana AR, Murugeswari B. ARIPSO: Association rule interactive postmining using schemas and ontologies. In: *Proc. of the 2011 Int'l Conf. on Emerging Trends in Electrical and Computer Technology (ICETECT 2011)*. IEEE Computer Society, 2011. 941–946. [doi: 10.1109/ICETECT.2011.5760255]

- [14] Meng QL, Jiang XJ. A method for rating customer requirements' final importance in QFD based on quantitative KANO model. In: Proc. of the 8th Int'l Conf. on Service Systems and Service Management (ICSSSM 2011). Tianjin: IEEE Computer Society, 2011. 1-6. [doi: 10.1109/ICSSSM.2011.5959396]
- [15] Hong L, Wang W, Zhao HQ. A service quality management approach based on QFD. In: Proc. of the Service Systems and Service Management (ICSSSM). West Sussex: John Wiley and Sons Ltd., 2012. 11-14. [doi: 10.1109/ICSSSM.2012.6252179]
- [16] Yuan F. Performance evaluation of technological innovation for aeronautical enterprise based on QFD. In: Proc. of the 2011 4th Int'l Conf. on Information Management, Innovation Management and Industrial Engineering (ICII 2011). Shenzhen: IEEE Computer Society, 2011. 31-34. [doi: 10.1109/ICII.2011.15]
- [17] Chen CL, Shi WH, Yan HG. Equipment quality monitoring analysis based on QFD. In: Proc. of the Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQ2MSE). Beijing: IEEE Computer Society, 2012. 1329-1332. [doi: 10.1109/ICQ2MSE.2012.6246466]
- [18] Liu YX. Study on application of apriori algorithm in data mining. In: Proc. of the 2010 2nd Int'l Conf. on Computer Modeling and Simulation. Taiyuan: IEEE Computer Society, 2010. 111-114. [doi: 10.1109/ICCMS.2010.398]
- [19] Liu D, Lung CH. P2P traffic identification and optimization using fuzzy *c*-means clustering. In: Proc. of the 2011 IEEE Int'l Conf. on Fuzzy Systems. Institute of Electrical and Electronics Engineers Inc., 2011. 2245-2252. [doi: 10.1109/FUZZY.2011.6007613]
- [20] Hu JL, Wu QY, Zhou B, Liu JH. Robust feedback credibility-based distributed P2P trust model. Ruan Jian Xue Bao/Journal of Software, 2009,20(10):2885-2898 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3554.htm> [doi: 10.3724/SP.J.1001.2009.03554]
- [21] Xu G, Tang RC, Song YJ. A stability based streaming media scheme over CDN and P2P hybrid network. Journal of Ocean University of China, 2012,42(1-2):163-168 (in Chinese with English abstract).
- [22] Pakkaia D, Koivukoski A, Paaso T, Latvakoski J. P2P middleware for extending the reach, scale and functionality of content delivery networks. In: Proc. of the 2nd Int'l Conf. on Internet and Web Applications and Services (ICIW 2007). IEEE Computer Society, 2007. 1-8. [doi: 10.1109/ICIW.2007.46]
- [23] Ma L, Zhang WM. Synthesize trust degree evaluating motel for an information grid environment. Journal of Tsinghua University (Sci. & Tech.), 2009,49(4):599-603 (in Chinese with English abstract). <http://qhxbw.chinajournal.net.cn>
- [24] Yin H, Liu XN, Zhan TY, Sekar V, Qiu F, Lin C, Zhang H, Li B. Design and deployment of a hybrid CDN-P2P system for live video streaming: Experiences with LiveSky. In: Proc. of the 2009 ACM Multimedia Conf., with Co-located Workshops and Symp. Beijing: ACM Press, 2009. 25-34. [doi: 10.1145/1631272.1631279]
- [25] Tong JJ, Xu K, Pi RJ. A new Web service structure of combining P2P and CDN technologies. In: Proc. of the 2010 IEEE 2nd Symp. on Web Society (SWS 2010). New York: ACM Press, 2010. 475-479. [doi: 10.1109/SWS.2010.5607404]
- [26] Li JT. The research on trust model and replication scheme in P2P environments [Ph.D. Thesis]. Shanghai: Fudan University, 2006 (in Chinese with English abstract).
- [27] Wu YY, Lei Q, Chen DS, YokoTa H. A method of discovering relation information from XML data. Ruan Jian Xue Bao/Journal of Software, 2008,19(6):1422-1427 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/19/1422.htm> [doi: 10.3724/SP.J.1001.2008.01422]
- [28] Shi PC, Wang HM, Yin H, Ding B, Wang TZ, Wang M. CDN scalability improvement using a moderate peer-assisted method. Korean Society for Internet Information, 2012,3(6):954-972. [doi: 10.3837/tiis.2012.03.011]
- [29] Kamiyama N, Mori T, Kawahara R, Hasegawa H. Optimally designing ISP-operated CDN. IEICE Trans. on Communications, 2013, 3(E96-B):790-801. [doi: 10.1587/transcom.E96.B.790]
- [30] Afergan MM, Leighton FT, Parikn JG. Hybrid content delivery network (CDN) and peer-to-peer (P2P) network. In: Proc. of the United States Patent Application. Assignee: Akamai, 2008. <http://www.freepatentsonline.com/y2008/0155061.html>
- [31] Zheng CY, Wang X, Zhao J, Xue XY. P2P video-on-demand content distribution schemes. Ruan Jian Xue Bao/Journal of Software, 2007,18(11):2942-2954 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/18/2942.htm> [doi: 10.1360/jos182942]
- [32] Guo Y, Sun K, Kurose J, Towsley D. A peer-to-peer on-demand streaming service and its performance evaluation. In: Proc. of the IEEE ICME 2003. IEEE Computer Society, 2003. 649-652. [doi: 10.1109/ICME.2003.1221700]

- [33] Xu LJ, Gao GH, Li XY, Lv JN. A content distribution mechanism based on P2P for video-on-demand service in IPTV. In: Proc. of the 2010 2nd Int'l Conf. on Communication Systems, Networks and Applications (ICCSNA 2010). Hong Kong: IEEE Computer Society, 2010. 228–231. [doi: 10.1109/ICCSNA.2010.5588876]
- [34] Jiang XX, Dong Y, Xu DY, Bhargava B. GnuStream: A P2P media streaming system prototype. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo, Vol.2. Los Alamitos: IEEE Computer Society, 2003. 325–328. [doi: 10.1109/ICME.2003.1221619]
- [35] Luo XZ, Qian PD, Zhu YQ, Liu JW. Secure computation against convert adversaries based on game theory. Nanjing Hangkong Hangtian Daxue Xuebao/Journal of Nanjing University of Aeronautics and Astronautics, 2012,44(1):70–74 (in Chinese with English abstract).
- [36] Kleinberg J, Liben-Nowell D. The syntenic diameter of the space of N -chromosome genomes. In: Sankoff D, ed. Proc. of the Comparative Genomics: Empirical and Analytical Approaches to Gene Order Dynamic, Map Alignment, and the Evolution of Gene Families. New York: Kluwer Academic Press, 2000. 185–197. [doi: 10.1007/978-94-011-4309-7_17]
- [37] De A, Desarkar MS, Ganguly N, Mitra P. Local learning of item dissimilarity using content and link structure. In: Proc. of the 2010 5th Int'l Conf. on Ubiquitous Information Technologies and Applications (CUTE 2010). Dublin: ACM Press, 2010. 833–845. [doi: 10.1145/2365952.2365999]
- [38] Zhou T, Lü LY, Zhang YC. Predicting missing links via local information. The European Physical Journal, 2009,71(4):63–69. [doi: 10.1140/epjb/e2009-00335-8]
- [39] Akcora CG, Carminati B. Network and profile based measures for user similarities on social networks. In: Proc. of the 4th ACM Int'l Conf. on Web Search and Data Mining. Las Vegas: IEEE Computer Society, 2011. 292–298. [doi: 10.1109/IRI.2011.6009562]
- [40] Krishnamuthy B, Wills CE. On the leakage of personally identifiable information via online social networks. ACM SIGCOMM, 2010,40(1):112–117.
- [41] Lindamood J, Heatherly R, Kantarcioglu M, Thuraisingham B. Inferring private information using social network data. In: Proc. of the 18th WWW. New York: ACM Press, 2009. 1145–1146. [doi: 10.1145/1526709.1526899]
- [42] Tasi CS, Huang YM, Liu CH, Huang YM. Applying VSM and LCS to develop an integrated text retrieval mechanism. Expert Systems with Application, 2012,39(4):3974–3982. [doi: 10.1016/j.eswa.2011.09.039]
- [43] Kim TK, Lee TH. Reliability-Based design optimization using enhanced Pearson system. Trans. of the Korean Society of Mechanical Engineers, 2011,35(2):125–130. [doi: 10.3795/KSME-A.2011.35.2.125]
- [44] Li K, Wang Y. Similarity based ranging method in wireless sensor networks. Ruan Jian Xue Bao/Journal of Software, 2009,20: 95–103 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/09012.htm>
- [45] Shen L, Zhou YM. A new user similarity measure for collaborative filtering algorithm. In: Proc. of the 2010 Int'l Conf. on Computer Modeling and Simulation (ICCMS 2010). Sanya: IEEE Computer Society, 2010. 375–379. [doi: 10.1109/ICCMS.2010.67]
- [46] Huang YS, Meng XW, Zhang YJ. Strategy of content location of P2P based on the social network. Ruan Jian Xue Bao/Journal of Software, 2010,21(10):2622–2630 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3647.htm> [doi: 10.3724/SP.J.1001.2010.03647]
- [47] Louis MP, Thomas LC. Synchronization in chaotic systems. Physical Review Letters, 1990,64(8):821–825.
- [48] Anand D, Bharadwaj KK. Adaptive user similarity measures for recommender systems: A genetic programming approach. In: Proc. of the 2010 3rd IEEE Int'l Conf. on Computer Science and Information Technology (ICCSIT 2010). Chengdu: IEEE Computer Society, 2010. 121–125. [doi: 10.1109/ICCSIT.2010.5563737]
- [49] Ming L, Zhao G, Xie GH, Wang CL. Research on smart space oriented location awareness method. Ruan Jian Xue Bao/Journal of Software, 2009,20(3):671–681 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3228.htm> [doi: 10.3724/SP.J.1001.2009.03228]
- [50] Yuan SH, Chen WB, Fu SK. User behavior similarity analysis of location based social network. Journal of Computer Applications, 2012,32(2):322–325 (in Chinese with English abstract).
- [51] Wang Y, Shu H, Fang DY. A geographic surface routing algorithm in 3D ad hoc networks. Ruan Jian Xue Bao/Journal of Software, 2010,21:318–329 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/10033.htm>

- [52] Rakesh A, Evfimievski A, Srikant R. Information sharing across private databases. In: Proc. of the ACM SIGMOD Int'l Conf. on Management of Data. San Diego: ACM SIGMOD, 2003. 86–97. [doi: 10.1145/872757.872771]
- [53] Malaga RA. Information System Technology. Beijing: Tsinghua University Press, 2006.
- [54] Chen YF, He YX, Cao JN, Wu J. A greedy algorithm for capacity-constrained surrogate placement in CDNs. Ruan Jian Xue Bao/Journal of Software, 2007,18(1):146–156 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/18/146.htm> [doi: 10.1360/jos180146]
- [55] Zachara J, Lesiuk M, Balawender R. Higher order alchemical derivatives from coupled perturbed self-consistent field theory. Journal of Chemical Physics, 2012,136(3):478–496. [doi: 10.1063/1.3674163]
- [56] Zucker LG. Production of trust: Institutional sources of economic structure. In: Cummings LL, ed. Proc. of the Research in Organizational Behavior. Greenwich CT: JAI Press, 1986. 53–111. [doi: 10.1504/IJBT.2008.018356]
- [57] Liu W, Duan HX, Zhang H, Wu JP. Trust based user behavior analysis and security management. Periodical of Ocean University of China, 2008,29(11):12–16 (in Chinese with English abstract).
- [58] Xu X, Tan J. Efficient secure message routing for structured peer to peer systems. In: Proc. of the Int'l Conf. on Networks Security, Wireless Communications and Trusted Computing (NSWCTC 2009). IEEE Computer Society, 2009. 354–357. [doi: 10.1109/NSWCTC.2009.124]
- [59] Rafaeli S, Hutchison D. A survey of key management for secure group communication. ACM Computing Surveys, 2003,35(3): 309–329. [doi: 10.1145/937503.937506]
- [60] Zhang PW, Jiao XJ, Zhou RJ. Network security technology in P2P. In: Proc. of the 2011 Communications in Computer and Information Science. Springer-Verlag, 2011. 145–149. [doi: 10.1007/978-3-642-21411-0_23]
- [61] Li ZT, Yu WN, Liu G. DPFSL: Decreasing probability flooding algorithm in P2P network. Tongxin Xuebao/Journal on Communications, 2006,27(11A):246–250 (in Chinese with English abstract). [doi: CNKI:CDMD:2.2009.035436]
- [62] Liu YH, Zhu LQ, Gao JJ, Li YL. A segment strategy based on flooding search in unstructured P2P network. In: Proc. of the 2008 2nd Int'l Conf. on Future Generation Communication and Networking (FGCN 2008). IEEE Computer Society, 2008. 252–255. [doi: 10.1109/FGCN.2008.201]
- [63] Jung J, Cho Y, Kim Y, Kim I. An adaptive, efficient, and reliable flooding mechanism for MANET. In: Proc. of the 2010 ACM Symp. on Applied Computing. Sierre: Association for Computing Machinery, 2010. 731–732. [doi: 10.1145/1774088.1774239]
- [64] Barjini H, Othman M, Ibrahim H, Udzir NI. Shortcoming, problems and analytical comparison for flooding-based search techniques in unstructured P2P networks. Peer-to-Peer Networking and Applications, 2012,5(1):1–13. [doi: 10.1007/s12083-011-0101-y]
- [65] Ma WM, Meng XW, Zhang YJ. Bidirectional random walk search mechanism for unstructured P2P network. Ruan Jian Xue Bao/Journal of Software, 2012,23(4):894–911 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4086.htm> [doi: 10.3724/SP.J.1001.2012.04086]
- [66] Cheng MM, Zhang GX. Connectedness of random walk segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2011,33(1):200–202. [doi: 10.1109/TPAMI.2010.138]
- [67] Yu I, Newman R. A topology-aware random walk. IEICE Trans. on Communications, 2012,E95-B(3):995–998. [doi: 10.1587/transcom.E95.B.995]
- [68] Zhang CH, Qiu XF, Ji Y. Comprehensive Analysis of P2P Technology. Beijing: Posts & Telecom Press, 2010 (in Chinese).
- [69] Chen JZ, Xu TY, Li WZ, Lu SL, Chen DX, Chan E. P2P interactive streaming system based on derivative tree. Ruan Jian Xue Bao/Journal of Software, 2010,21(5):1138–1152 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3599.htm> [doi: 10.3724/SP.J.1001.2010.03599]
- [70] Zhang ZM, Zhou J, Chen Z, Li J. Modeling and algorithm for network coding based P2P streaming. Ruan Jian Xue Bao/Journal of Software, 2010,23(3):648–661 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3991.htm> [doi: 10.3724/SP.J.1001.2012.03991]
- [71] Liu YJ, Dou WH. A video-on-demand streaming service architecture in P2P environment. Ruan Jian Xue Bao/Journal of Software, 2006,17(4):876–884 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/876.htm> [doi: 10.1360/jos170876]

- [72] Huang LC, Nie JF. Using Pareto principle to improve efficiency for selection of QoS Web services. In: Proc. of the 2010 7th IEEE Consumer Communications and Networking Conf. (CCNC 2010). IEEE Computer Society, 2010. 546–553. [doi: 10.1109/CCNC.2010.5421583]
- [73] Park DG, Park YS, Kim YI, Cheol W, Kyu W. Implementation of IPTV set-top for distance education services using grid-CDN. In: Proc. of the 2010 5th Int'l Conf. on Ubiquitous Information Technologies and Applications (CUTE 2010). IEEE Computer Society, 2010. 13–24. [doi: 10.1109/ICUT.2010.5677638]
- [74] Wählisch M, Schmidt TC, Witteburg G. On predictable large-scale data delivery in prefix-based virtualized content networks. *Computer Networks*, 2011,55(18):4086–4100. [doi: 10.1016/j.comnet.2011.07.020]
- [75] Lu ZH, Wu J, Chen LJ, Huang SJ, Huang Y. CPH-VoD: A novel CDN-P2P-hybrid architecture based VoD scheme. In: Proc. of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Heidelberg: Springer-Verlag, 2011. 578–586. [doi: 10.1007/978-3-642-17616-6_50]
- [76] Garmeni M, Analoui M. An economical mechanism for multicasting of content among servers of hybrid CDN-P2P networks. In: Proc. of the 2011 Int'l Conf. for Internet Technology and Secured Transactions (ICITST 2011). IEEE Computer Society, 2011. 566–571. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6148400>
- [77] Sheu ST, Huang CH. Mixed P2P-CDN system for media steaming in mobile environment. In: Proc. of the 2011 7th Int'l Wireless Communications and Mobile Computing Conf. (IWCMC 2011). IEEE Computer Society, 2011. 657–660. [doi: 10.1109/IWCMC.2011.5982624]
- [78] Wang HT, Song LH. Applications and challenges of mobile P2P systems in ad hoc network. In: Proc. of the Advanced Materials Research 2011. Clausthal-Zellerfeld: Trans. on Tech. Publications, 2011. 575–579. [doi: 10.4028/www.scientific.net/AMR.171-172.575]
- [79] Wichtlhuber M, Rüchert J, Stingl D, Schulz M, Hausheer D. Energy-Efficient mobile P2P video streaming. In: Proc. of the 2012 IEEE 12th Int'l Conf. on Peer-to-Peer Computing (P2P 2012). Washington, 2012. 63–64. [doi: 10.1109/P2P.2012.6335812]
- [80] Xu LJ, Gao GH, Li XY, Lü JN. A content distribution mechanism based on P2P for video-on-demand service in IPTV. In: Proc. of the 2010 2nd Int'l Conf. on Communication Systems, Networks and Applications (ICCSNA 2010). IEEE Computer Society, 2010. 228–231. [doi: 10.1109/ICCSNA.2010.5588876]
- [81] Yang HG, Hsieh MY, Yu HF, Tseng LM. A replication-aware CDN-P2P architecture based on two-step server selection and network coding. In: Proc. of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Heidelberg: Springer-Verlag, 2008. 738–747. [doi: 10.1007/978-3-540-89796-5_76]
- [82] Hsieh MY, Yang HC, Tseng LM. Finding nearest neighbors in replication-aware CDN-P2P architecture. *Journal of Internet Technology*, 2005,6(2):133–140. <http://jit.ndhu.edu.tw/jitcontent.php?getserial=223>
- [83] Lu ZH, Wang Y, Yang YR. An analysis and comparison of CDN-P2P hybrid content delivery system and model. *Journal of Communications*, 2012,7(3):232–245. [doi: 10.4304/jcm.7.3.232-245]
- [84] Kim TN, Jeon S, Kim Y. A CDN-P2P hybrid architecture with content/location awareness for live streaming service networks. In: Proc. of the Int'l Symp. on Consumer Electronics (ISCE). Institute of Electrical and Electronics Engineers Inc, 2011. 438–441. [doi: 10.1109/ISCE.2011.5973865]
- [85] Yin H, Liu XN, Zhan TY, Sekar V, Qiu F, Lin C, Zhang H, Li B. LiveSky: Enhancing CDN with P2P. In: Proc. of the ACM Trans. on Multimedia Computing, Communications and Applications. New York: ACM Press, 2010. [doi: 10.1145/1823746.1823750]
- [86] Korn R, Kuntze N, Repp J. Performance evaluation in trust enhanced decentralised content distribution networks. In: Proc. of the Expert Systems with Applications. Florida: IEEE Computer Society, 2011. 368–375. [doi: 10.1109/CQR.2011.5996084]
- [87] Preston JA, Prasad SK. P2P document tree management in a real-time collaborative editing system. In: Proc. of the Lecture Notes in Computer Science. Springer-Verlag, 2007. 418–431. [doi: 10.1007/978-3-540-77220-0_39]

附中文参考文献:

- [20] 胡建理,吴泉源,周斌,刘家红.一种基于反馈可信度的分布式 P2P 信任模型.软件学报,2009,20(10):2885–2898. <http://www.jos.org.cn/1000-9825/3554.htm> [doi: 10.3724/SP.J.1001.2009.03554]

- [21] 徐广,唐瑞春,宋佑举.CDN 与 P2P 混合网络架构下的一种基于系统稳定状态的流媒体调度策略.中国海洋大学学报,2012,42(1-2):163-168.
- [23] 马礼,郑纬民.信息网络环境下的综合信任度评价模型.清华大学学报(自然科学版),2009,49(4):599-603. <http://qhxw.chinajournal.net.cn>
- [26] 李景涛.P2P 环境下的信任模型与副本方案研究[博士学位论文].上海:复旦大学,2006.
- [27] 吴扬扬,雷庆,陈敏生,YokoTa H.一种从 XML 数据中发现关系信息的方法.软件学报,2008,19(6):1422-1427. <http://www.jos.org.cn/1000-9825/19/1422.htm> [doi: 10.3724/SP.J.1001.2008.01422]
- [31] 郑常熠,王新,赵进,薛向阳.P2P 视频点播内容分发策略.软件学报,2007,18(11):2942-2954. <http://www.jos.org.cn/1000-9825/18/2942.htm> [doi: 10.1360/jos182942]
- [35] 罗喜召,钱培德,朱艳琴,刘建伟.防范秘密攻击的安全计算的博弈论实现.南京航空航天大学学报,2012,44(1):70-74.
- [44] 李凯,汪芸.传感器网络节点间距离估计方法.软件学报,2009,20:95-103. <http://www.jos.org.cn/1000-9825/09012.htm>
- [46] 黄永生,孟祥武,张玉洁.基于社会化网络特征的 P2P 内容定位策略.软件学报,2010,21(10):2622-2630. <http://www.jos.org.cn/1000-9825/3647.htm> [doi: 10.3724/SP.J.1001.2010.03647]
- [49] 明亮,赵刚,谢桂海,王春雷.面向智能空间的位置感知方法研究.软件学报,2009,20(3):671-681. <http://www.jos.org.cn/1000-9825/3228.htm> [doi: 10.3724/SP.J.1001.2009.03228]
- [50] 袁书寒,陈维斌,傅顺开.位置服务社交网络用户行为相似性分析.计算机应用,2012,32(2):322-325.
- [51] 汪芸,苏瀚,房鼎益.三维 AD-HOC 网络地理信息路由算法.软件学报,2010,21:318-329. <http://www.jos.org.cn/1000-9825/10033.htm>
- [54] 陈益峰,何炎祥,曹建农.内容传递网络处理能力受限代理放置贪婪算法.软件学报,2007,18(1):146-156. <http://www.jos.org.cn/1000-9825/18/146.htm> [doi: 10.1360/jos180146]
- [57] 刘武,段海新,张洪,吴建平.基于信任的 P2P 网络安全管理体系结构研究.中国海洋大学学报,2008,29(11):12-16 [doi: 10.1109/FGCN.2008.201]
- [61] 李之棠,余万能,刘刚.DPFSL:P2P 网络的递减概率洪泛算法.通信学报,2006,27(11A):246-250. [doi: CNKI:CDMD:2.2009.035436]
- [65] 马文明,孟祥武,张玉洁.面向非结构化 P2P 网络的双向随机漫步搜索机制.软件学报,2012,23(4):894-911. <http://www.jos.org.cn/1000-9825/4086.htm> [doi: 10.3724/SP.J.1001.2012.04086]
- [68] 张春红,裘晓峰,弭伟,纪阳.P2P 技术全面解析.北京:人民邮电出版社,2010.
- [69] 陈建忠,徐天音,李文中,陆桑璐,陈道蓄,CHAN E.一种基于衍生树的交互式 P2P 流媒体系统.软件学报,2010,21(5): 1138-1152. <http://www.jos.org.cn/1000-9825/3599.htm> [doi: 10.3724/SP.J.1001.2010.03599]
- [70] 张志明,周晋,陈震,李军.基于网络编码的对等网流媒体传输模型和算法.软件学报,2012,23(3):648-661 <http://www.jos.org.cn/1000-9825/3991.htm> [doi: 10.3724/SP.J.1001.2012.03991]
- [71] 刘亚杰,窦文华.一种 P2P 环境下的 VoD 流媒体服务体系.软件学报,2006,17(4):876-884. <http://www.jos.org.cn/1000-9825/17/876.htm> [doi: 10.1360/jos170876]



张玉洁(1969-),女,天津人,讲师,主要研究领域为通信软件,网络服务,CDN-P2P 技术.

E-mail: zhangyj@bupt.edu.cn



孟祥武(1966-),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为网络服务,通信软件,人工智能.

E-mail: mengxw@bupt.edu.cn



何明(1982-),男,博士生,主要研究领域为 CDN-P2P 技术,网络服务,通信软件.

E-mail: heming0405@163.com