

数字图书馆的移动服务^{*})

党德鹏 周立柱 邢春晓

(清华大学计算机科学与技术系 北京 100084)

摘要 随着移动计算技术的飞速发展,越来越多的用户将通过无线移动方式访问数字图书馆。但是,目前的数字图书馆系统都假定用户访问信息时是静止的。本文首先给出了移动数字图书馆系统的体系结构,在讨论移动计算环境特点的基础上,阐述了将数字图书馆服务延伸到移动计算环境时面临的关键问题以及研究进展情况,为进一步开展这方面的研究奠定基础。

关键词 数字图书馆,移动服务,查询处理,数据广播

Mobile Services of Digital Library Systems

DANG De-Peng ZHOU Li-Zhu XING Chun-Xiao

(Department of Computer Science & Technology, Tsinghua University, Beijing 100084)

Abstract While mobile computing technology is developing rapidly, more and more users would like to access digital library systems through wireless mobile mode. However, current digital library systems all assume that users are stationary. In this paper, a typical architecture of mobile digital library systems is shown at first. Then, the characteristics of mobile computing environments are discussed. Finally, the future research direction is identified.

Keywords Digital library, Mobile service, Query processing, Data broadcast

1 引言

各种基于海量分布式多媒体数据集成、管理与通讯的数字图书馆系统,在信息社会中将扮演愈来愈重要的角色^[1,2]。目前的数字图书馆系统都假定用户访问信息时是静止的^[2,3]。但是,随着移动计算技术的飞速发展,越来越多的用户将通过无线移动方式访问数字图书馆^[4~6]。通过移动网络提供数字图书馆服务是许多从个人、研究,到特定事务计算等应用的需要。数字图书馆系统提供移动服务,不仅意味着服务方式的扩充,更意味着应用服务领域的进一步扩展。文[1]详细阐明了提供数字图书馆移动服务将给读者带来的诸多好处。当数字图书馆、移动计算这两种正在快速发展的技术结合在一起时,将会出现许多具有挑战性的问题^[1,5,6]。本文首先给出了基于现代无线通讯技术的移动数字图书馆系统的体系结构,进而在讨论移动计算环境特点的基础上,阐述了数字图书馆移动服务研究领域的主要问题以及进展情况,为进一步开展这方面的研究指明方向。

2 移动数字图书馆系统体系结构

图1给出了移动数字图书馆系统的典型体系结构。整个平台类似于分布式数据库系统,但是它为固定和移动客户提供数字图书馆服务。一组通用计算机通过高速有线网络互连。这些计算机分为固定主机(FH)和基站(移动支持基站BS)。移动设备(MU)是由电池供电的便携式计算机,它可在称之为“地理上的移动区域”(G)的范围内自由移动,G是由所有BS覆盖的区域。为了挖掘频率复用、支持MU的移动

性,整个G被分作许多叫单元(Cells)的小区域。每个单元由专门的BS管理,且被分配了一组用于通讯的频率。移动性要求MU必须在G内(单元间)无限制移动,且必须能在任意单元中访问需要的数据。MU穿越单元的界限从一个单元进入另一个单元叫转交(过区切换)。过区切换的处理对MU是透明的。固定主机通过固定网络通信,而移动设备在基站支持下通过无线信道访问数字图书馆。系统中包含有数字图书馆服务器(DLS)。包含DLS丝毫不影响移动网络的基本功能。DLS可安装于BS或FH,即数字图书馆服务器可分布于若干通过高速有线网络连接到整个系统的固定主机。系统服务支持查询和广播两种方式。

从数字图书馆的角度来看,与基于固定网络的传统分布式计算环境相比,移动计算环境的许多显著特征将对数字图书馆系统产生重要影响^[7]。

1) 移动性。在移动计算环境中,一台移动计算机不仅可以在不同的地方连通网络,而且在移动的同时也可以保持网络连接。有的用户需要查询与用户位置相关的信息,由于用户是移动的,在处理查询的过程中,用户的位置可能不断发生变化。

2) 网络条件与移动设备计算能力多样性。移动计算机的移动性使得不同时间可用的网络条件(如网络带宽、通信代价、网络延迟以及服务质量)是变化多端的,甚至相差悬殊。各个用户所用移动计算机的计算能力也千差万别。

3) 频繁断接性。移动计算机在移动过程中,由于受使用方式、电源、无线通信费用、网络条件等因素的限制,一般不采用持续性联网的工作方式,而是主动或被动地间歇性入网、断

^{*} 本文研究得到中国博士后科学基金(No. 2003034146:数字图书馆的移动服务)、国家自然科学基金重大基金(No. 60221120146:数字图书馆的关键理论、方法和技术研究)、国家重点基础研究发展规划(973)(No. G1999032704:数字图书馆系统中的海量信息存储与组织)基金资助。党德鹏

博士后,主要研究方向包括数字图书馆移动服务、移动与嵌入式计算、多媒体通讯。周立柱 教授,博士生及博士后导师,主要研究方向包括数字图书馆、计算机软件、信息系统、数据库系统、GIS研究与应用、Web与海量数据处理技术。邢春晓 博士,副教授,主要研究方向包括数字图书馆等。

接(disconnection)。

4) 网络通信的不对称性。由于物理通信媒介的限制,一般的无线网络通信都是不对称的,表现在固定服务器节点可以拥有强大的发送能力,而移动计算机的发送能力非常有限,下行链路(服务器到移动计算机)的通信带宽与代价和上行链路(移动计算机到服务器)相差很大。此外,许多系统中的信息流动范型也会引起通信的不对称性。

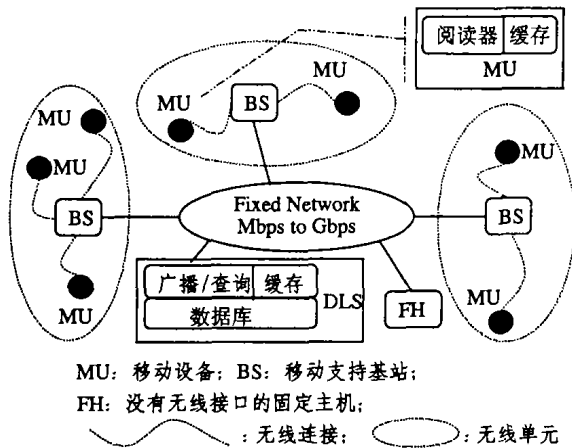


图1 移动数字图书馆系统体系结构

5) 移动计算机的电源能力。移动计算机主要依靠蓄电池供电,而现在电池的容量相当有限,例如一台典型的笔记本电脑在正常的连续使用情况下一般只能维持2~5个小时。不幸的是,电池容量在未来5年内只能提高30%,这远远低于同期CPU速度和存储容量的发展速度。因此,尽管已有不少节能技术应用于移动计算机,但电池容量问题仍将在长时间内存在。

6) 规模。移动环境都要求系统同时支持大量的移动用户并发访问,这就要求移动数字图书馆系统必须具有比传统分布式数字图书馆系统高得多的可扩展性。

这样,通过无线网络提供移动数字图书馆服务,将会面临许多具有挑战性的问题。

3 数字图书馆移动服务涉及的关键问题

如上所述,移动计算环境有许多显著特征,移动数字图书馆系统与传统数字图书馆系统在体系结构上有很大不同。为了提供数字图书馆移动服务,需要对其中涉及的关键问题进行深入研究^[1,5,8,9]。

3.1 数据组织和存储

移动通讯中网络条件呈现多样性,网络带宽、通讯代价、网络延迟、服务质量等变化多端,相差悬殊;移动计算机的计算能力也千差万别。但是,移动数字图书馆系统又需要很高的可扩展性。这使得传统数字图书馆系统中的数据存储方法显然无法适应移动环境,必须研究新的适用于移动应用环境的数据组织和存储策略。由于可以依据可用带宽、存储容量和计算能力等移动主机的物理限制,计算出合适的检索和传输数据量,问题的关键就是用小对象表示大对象,提高信息数据比^[2,5]。目前提出以下三种方案:

· 多版本数据存储 系统中存储一个数据对象从非常详细到非常粗略、具有不同细节的多个版本。检索和传输数据时系统将选择最佳的版本,以便充分利用网络条件并给用户提提供尽可能详细的信息。比如,对于一篇论文,一个版本包含完整全文所有信息;一个版本仅含摘要和各小节子标题;另外一个版本仅含精简的内容摘要。对于一个图像可按照不同颜色

深度、帧尺寸、帧分辨率、编解码模式等存为多个版本。

· 多版本的实时计算 这种模式中,与数据对象一起存有可获得不同对象版本的计算方法。检索和传输数据时系统将依据可用带宽、存储容量和计算能力等移动主机的物理限制和用户需要的响应时间,使用与对象相关的合适的计算方法,从对象大版本计算出对象小版本。数字图书馆数据对象有各种不同的媒体,各种不同媒体有各种不同方法。比如,与一图像对象相关的有损压缩技术可用于压缩图像,不同压缩算法可获得不同压缩比。

· 混合方法 多版本数据存储方法需要占用较多的空间而对对象多版本的实时计算因需要较多计算时间会影响客户响应时间。混合方法仅存储数据对象有代表性的几个版本,检索和传输数据时系统选择已有版本或从最接近需要的版本计算出一个合适的新版本。

3.2 数据发布

移动数字图书馆系统应该有效支持基于内容的移动检索、位置相关的移动查询以及多媒体信息的同步平滑发布^[5],数据发布处理是数字图书馆服务器中最复杂的部分。

3.2.1 查询模式 移动数字图书馆系统查询服务中的关键问题是基于相似性的移动查询和位置相关移动查询。

基于相似性的移动查询:许多多媒体应用中数据库存储的数据只是对原始信息的近似描述,基于内容的多媒体信息查询(比如根据图例或草图检索等)都是基于相似性的^[10,11],基于相似性的移动查询是高度交互的。例如,基于在某本书中什么地方看到过的印象来检索一幅图片。这时,系统的任务就是以反馈机制逐步找出用户确切需要的结果。这可使用标准信息检索技术将所有命中目标按照相似度排序,给用户提供最相似的 n 个命中目标,其中 n 受限于当前可用带宽、移动主机的存储器等物理限制。其它的命中目标可缓存于服务器,分批传输直至用户指明了一个满意的答案。基于相似性的查询是高度交互的,为了有效支持移动环境中基于内容的检索,就需要研究相应的移动检索算法和检索模型等来使得基于相似性查询的处理适应频繁断接、通信不对称、计算能力以及带宽有限的移动计算环境。

位置相关移动查询:由于用户是移动的,在处理查询的过程中,用户的位置可能不断发生变化。许多移动应用中的查询涉及移动用户的位置并具有一定的实时性^[5,7]。需要这类查询服务的典型用户包括旅行者、医疗人员、飞行员、警察等特殊职业的从业者,以及在远程无法直接访问静态基站的人。比如,一个家庭在作长途旅行,通常要知道“附近最近且拥有游泳池的旅店在哪里?去最近的医院怎么走?”。又比如,一个司机需要知道下一站2公里范围内的餐馆信息,并且这个信息只有在他开车穿过下一站之前得到才有效;在这之后,其结果则没有任何意义,而需要另外查找再下一站的情况。这就需要研究相应的位置管理机制、位置相关查询优化技术以及位置相关实时查询处理方法。

3.2.2 广播模式 可以利用移动计算环境中通讯的非对称特性,由数字图书馆服务器把大多数用户频繁访问的数据即热点数据组织起来,通过广播技术周期性地提供给移动客户机访问。广播模式有许多好处,比如具有很好的可扩展性、易于进行断接处理等。最近几年,移动多媒体消费装置的数量与日俱增,从PDA、手机、电子书到PPC,通过这些装置访问数字图书馆系统中的多媒体服务是一种必然需求。为了使现有分布式多媒体系统为移动用户提供高质量的音视频信息服务,MPEG工作组制定了MPEG-4标准。MPEG-4编码/

(下转第15页)

较大,因而不适宜将整个文件缓存,一般采用分片缓存的解决办法^[18]。

以上列举的是适应性缓存比较有代表性的几个研究方向,在更进一步的工作中,适应性缓存的研究还可以考虑结合目前流行的 P2P, Grid 等体系结构,构建更高效的适应性缓存系统。

参考文献

- Zhang L, Floyd S, Jacobson V. Adaptive Web Caching. In: Proc. of the NLANR Web Cache Workshop, 1997
- Michel Scoot, Nguyen Khoi, Rosenstein Adam, Zhang Lixia. Adaptive Web Caching: Towards a New Global Caching Architecture. In: Third Intl. Caching Workshop, June 1998
- Kaiser M J, Tsui K C, Liu J. Adaptive distributed caching. In: Proc. of the 2002 Congress on Evolutionary Computation, 2002
- Kaiser M J, Tsui K C, Liu J. Adaptive distributed caching with minimal memory usage. In: Proc. of Simulated Evolution and Automated Learning Conf. 2002
- Kaiser M J, Tsui K C, Liu J. Self-organized Autonomous Web Proxies. AAMAS, 2002
- Mohammad, Salimullah, Raunak. A Survey of Cooperative Caching. [Technical report]. December 1999
- Buchholz S, Schill A. Adaptation-Aware Web Caching: Caching in the Future Pervasive Web. In: Proc. of the GI/ITG Fachtagung Kommunikation in Verteilten Systemen (KiVS 2003)/ Leipzig, Feb 26~28, 2003
- Fox A, Gribble S D, Brewer E A, Amir E. Adaptation to Network and Client Variability via On-Demand Dynamic Distillation. In:

- Sixth Intl. Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS VI), Cambridge, MA, Oct. 1996. ACM
- Gramacy R B, Warmuth M K, Brandt S A, Ari I. Adaptive caching by Refetching. In Advances in Neural Information Processing Systems (NIPS) 2002
- Danzig P B, Hall R S, Worrell K J. A case for caching file objects inside internetwork. In ACM SIGCOMM, Sep. 1993
- Abrams M, Standridge C R, et al. Caching proxies: Limitations and potentials. In: 4th Intl. World Wide Web Conf. Boston, USA, www.w3.org/Conference/www4/Papers/155, 1995
- Bestavros A, Carter R L, Crovella m E. Application-level document caching in the internet. In: Second Intl. Workshop on Services in Distributed and networked Environments, 1995
- Mic B C, Danzig P B, et al. The harvest information discovery and access system. In: Second Intl. World Wide Web Conf. October 1994
- Wang Zheng. Cachesmesh: A distributed cache system for world wide web. In: 2nd NLANR Web Caching Workshop, June 1997
- Calvert B K, Zegura E. Self-Organizing Wide-Area Network Caches. In: Proc. of IEEE Infocorn'98, San Francisco, CA, March 1998
- Kaiser M J, Tsui K C, Liu J. Self-organized autonomous web proxies. In: Proc. of the First Intl. Joint Conference on Autonomous Agents and Multiagent Systems, July 2002. 1397~1404
- Tsui K C, Liu J, Liu H L. Autonomy Oriented Load Balancing in Proxy Cache Servers. Web Intelligence: Research and Development, First Asia-Pacific Conf., WI 2001. 115~124
- Tran D A, Hua K A, Sheu S. A new caching architecture for efficient video services on the internet. In: IEEE Symposium on Applications and the Internet, Orlando, FL, USA, 2003. 29

(上接第5页)

解码和 DAB(Digital Audio Broadcasting)标准相结合,首次实现了快速移动交通工具上无扭曲接收电视节目。但是,无线带宽非常有限(即使是3G),通过无线广播提供多媒体信息仍是一个新课题,需要我们深入研究相应的调度策略,从同步的调度、同步的反馈技术、基于网络的模式、基于缓存的模式四个方面来处理网络抖动、终端系统抖动、时钟漂移、网络条件变化,以及帧内和帧间同步,设计3G无线移动网络以及数字化广播电视网络环境中的同步协议,以充分利用有限的无线带宽,在可接受的访问和调谐时间范围内提供优质的多媒体服务。

3.3 用户界面

目前,传统数字图书馆系统通常都采用 B/S 结构,用户通过 Web 浏览器访问数字图书馆资源。但是,对手持移动计算机来说,未来发展趋势是更小、更易于携带和使用,屏幕尺寸也会进一步减小,使用传统 Web 浏览器作界面显然不合适。目前,手持移动计算机的输入方式主要有笔输入、语音输入、特制键盘输入等三种。随着移动设备输入技术的不断发展,针对主流输入方式,在非常有限的显示区域内设计类似于 Web 浏览器、非常友好的用户界面是我们面临的又一个新挑战。可在现有语义 Web 和个性化服务领域研究成果的基础上,进一步研究动态自适应、设备无关及个性化的 Web 处理机制。首先提出一种普遍适应各种移动设备的自适应处理框架;在 CC/PP 的基础上研究设备 Profile 在无线移动网络上的传输机制;基于语义 Web 方法,研究动态自适应的 Web 处理算法,并通过建立服务质量模型,进一步提高自适应处理的有效性。

3.4 缓存

传统数字图书馆系统中常使用缓存来改进对静态用户的响应时间^[9]。与传统分布式计算环境不同,移动计算机频繁处于断接状态,这使得当移动用户访问数字图书馆时数据缓存显得更重要,需要用缓存来支持断接时的数据连续使用。断接时间经常是可预知的,这样就可以将断接期间需要显示的数据量计算出来,并预取到缓存。比如在视频点播技术支持下,

用户欣赏数字图书馆中的视频数据。当静态服务器给用户发送视频文件块时,可以一边传输视频数据,一边将收到的数据展示给用户。若在断接前已经预取了足够的数据并存储于缓存中,那么断接期间的用户就可以连续欣赏而感觉不到断接的存在^[5]。这就需要我们深入研究存储容量非常有限的移动计算机适合于多媒体数据传输和播放的缓存机制。

小结 随着移动通讯和移动计算机技术的飞速发展,移动计算已经成为现实。另一方面,各种基于海量分布式多媒体数据集成、管理与通讯的数字图书馆系统,在未来信息社会中将扮演愈来愈重要的角色。尽管数字图书馆的固有目标就是支持用户随时随地访问大量在线信息,数字图书馆的研究已十余年,但是数字图书馆移动服务系统的研究仍是国内外信息科学研究领域的全新课题。本文在讨论移动计算环境特点的基础上,给出了移动数字图书馆系统的典型体系结构,阐述了将数字图书馆服务延伸到移动计算环境时面临的关键问题以及研究进展情况,为进一步开展这方面的研究指明了方向。

参考文献

- Marshall C, Golovchinsky G, Price M N. Digital libraries and mobility. Communications of the ACM, May 2001
- Madria S K, Mohania M, Bhowmick S S, Bhargava B. Mobile data and transaction management. Information Sciences, 2002, 141: 279~309
- Imielinski T, Badrinath B R. Mobile wireless computing: challenges in data management. Communications of ACM, 1994, 37(10): 18~28
- Imielinski T, Viswanathan S, Badrinath B R. Data on air: organization and access. IEEE Transactions on Knowledge and Data Engineering, 1997, 9(3): 353~372
- Bhargava B, Annamalai M, Pitoura E. Digital library services in mobile computing. ACM SIGMOD Record, 1999, 24(4): 34~39
- Barbara D. Mobile computing and databases-A survey. IEEE Transactions on Knowledge and Data Engineering, 1999, 11(1): 108~117
- 贾焰,李霖,等.分布式数据库技术.国防工业出版社,2000
- Li Chao, Xing Chunxiao, Zhou Lizhu. Using a Light-weighted Cache Pool to Leverage the Performance of Web Picture Databases, 2003
- Pitoura E, Samaras G. Locating objects in mobile computing. IEEE Transactions on Knowledge and Data Engineering, 2001, 13(4): 571~592
- 马洪超,李德仁.一种定义影像相似性的新方法及其在影像检索中的应用.武汉大学学报,2001,26(5):405~411
- 段文娟,高文,林守勋,马继伟.图像检索中的动态相似性度量方法.计算机学报,2001,24(11):1156~1162