

基于 P2P 模式的数字城市服务模式及平台设计

史文勇 李琦 徐建鸿

(北京大学数字地球工作室 北京大学遥感与 GIS 研究所 北京 100871)

摘要 对等网技术(P2P)已经在分布式计算系统中获得了广泛的应用,数字城市中的服务本质上也是对等的。本文提出基于 P2P 模式的数字城市服务管理系统,并对平台设计、实现进行了探讨。

关键词 数字城市, P2P, 服务管理

The P2P Based Service Model and Platform Design for Digital City

SHI Wen-Yong LI Qi XU Jian-Hong

(CyberGIS Studio, PKU Institute of Remote Sensing & GIS, PKU, Beijing 100871)

Abstract The P2P network has been widely used in distributed computing systems, and the information systems in the digital city is essentially peer to peer. In this paper we introduce an P2P based service management system for digital city, and discuss how to design and implement the platform.

Keywords Digital city, P2P, Service management

1 引言

随着 Internet 的发展和计算模式的演化, P2P 模式逐渐为人们所关注。在此背景下, 本文提出了基于 P2P 模式的服务管理系统: P2PSMS。在详细介绍了 P2PSMS 的定义、组成、特征和实现之后, 给出了基于 P2PSMS 的数字城市服务管理平台。该平台的主要目标包括用户的服务请求的分解、应用服务的发现、应用服务的集成与组装、应用服务的调用与执行等。

2 数字城市框架及服务模式

数字城市是物理世界的城市在信息空间中的映射, 是城市的信息化或数字化表现形态。因此, 数字城市实质上是一个基于计算机网络和数字信息资源的开放的、复杂的适应系统, 是城市数字虚拟空间、信息资源配置中心和信息资源服务市场。

按照从信息生产到信息消费的信息价值链, 数字城市的系统体系结构按“数据-信息(知识)-服务-应用”的逻辑层次来组织, 如图 1 所示^[1,2]。

(1) 第一层次是信息体的自治式系统, 提供分布式的数据和应用服务。

(2) 第二层次是数字城市信息资源管理系统, 通过元数据库管理系统和数据仓库管理系统, 对分布式的数据进行集成与管理, 实现数据的共享与互操作, 实现信息融合、数据挖掘与知识发现。

(3) 第三层次是数字城市应用服务平台, 以 Web 服务的框架与机制来整合数据服务和应用服务, 支持服务的发现、配置、交互、装配、运行和管理, 实现数据、业务逻辑与商业流程的共享与互操作。服务管理平台是其重要内容。

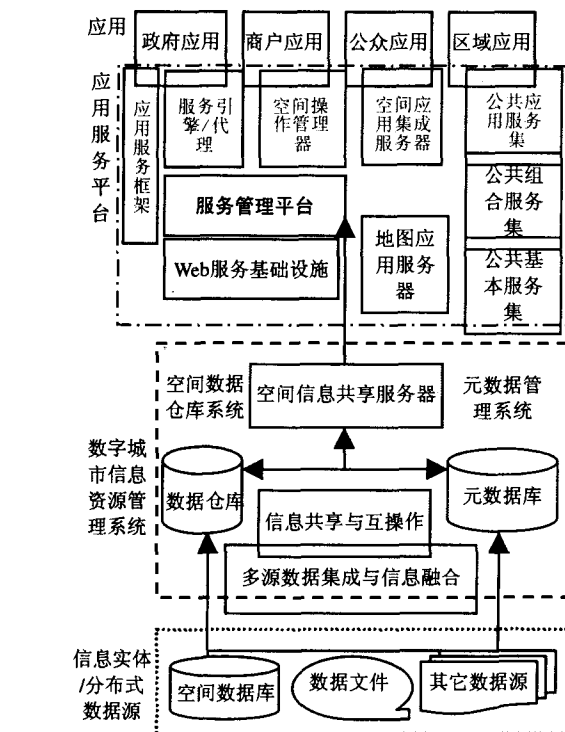


图 1 数字城市的系统体系结构

4) 第四层次是应用系统, 基于统一的信息资源管理平台和应用服务平台来构建和实现各种应用系统, 能够有效避免重复建设和信息孤岛, 实现信息资源的共享与互操作, 使得跨应用系统信息流有序化和信息价值链一体化。

3 P2P

P2P 是 Peer-to-Peer 的缩写, Peer 在英语中有“同事”、“伙伴”等意义, P2P 即指“对等联网”。它是一种网络模型, 在

史文勇 博士研究生, 研究领域: 数字城市服务平台、软件服务总线、数字奥运移动综合信息服务系统等; 李琦 教授、博士生导师, 研究领域: 空间信息科学与技术、数字地球、数字城市等; 徐建鸿 硕士生, 研究领域: 分布式计算、数据挖掘、智能 GIS 等。

这种网络中所有的节点是对等的,各节点具有相同的责任与能力并协同完成任务。IBM 为 P2P 下了如下定义:P2P 系统由若干互联协作的计算机构成,且至少具有如下特征之一:系统依存于边缘化(非中央式服务器)设备的主动协作,每个成员直接从其他成员而不是从服务器的参与中受益;系统中成员同时扮演服务器与客户机的角色;系统应用的用户能够意识到彼此的存在,构成一个虚拟或实际的群体。不难看出,P2P 把网络计算模式从集中式引向分布式。也就是说,网络应用的核心从中央服务器向网络边缘的终端设备扩散:服务器到服务器、服务器到 PC 机、PC 机到 PC 机,PC 机到 WAP 手机,所有网络节点上的设备都可以建立 P2P 对话。

目前 P2P 主要的研究角度有:文件交换、对等计算、协同工作、即时通讯、搜索引擎、网络游戏、基于 P2P 的文件存储系统、基于 P2P 的操作系统等;另外,还有对 P2P 开发平台的研究和 P2P 安全框架的构建等。本文试图从另一个角度进行探讨,即使用 P2P 技术实现数字城市中服务平台,处理数字城市中服务的管理问题。

4 数字城市中的 P2P 服务模式

从上一章对 P2P 的介绍中可以看出,目前 P2P 的主要应用集中在消息交换、文件共享、计算资源共享等方面。实际上,我们可以将它们都视为一种服务,文件传输、CPU 周期,都是服务的一种。因此,我们可以将服务从具体应用中抽象出来,而 P2P 点对点的模式不变,将它们演化成基于 P2P 模式的服务共享系统。

进一步,我们引进专门用于管理在线节点的管理者(Manager),将连入 P2P 网络中的所有节点,无论是提供服务还是请求服务,都视为对等,我们就将这样的系统称之为 P2PSMS,即基于 P2P 模式的服务管理系统(P2P Service Management System)。之所以引进管理者,是为了提高节点相互之间的定位和连接的效率,并对节点间的会话进行管理、调度、集成等工作,而节点之间的地位是完全平等的。

4.1 P2PSMS 特点

P2PSMS 系统实际上是将传统服务系统连接起来,共同提供服务,它是建立在 Internet 体系基础之上的。从用户角度看,它是一个具有无限计算能力和存储能力的、提供全面服务的系统,当然它是一个虚拟系统。它的实际服务容量和能力决定于它的节点而不是 Manager。P2PSMS 并不真正提供具体的应用或服务,而是提供服务的元信息,即对服务的各种描述,以及对服务的管理、调度、集成和互操作。

从系统外部使用者角度看,P2PSMS 将是一个统一服务提供系统,可以满足各个用户的个性化服务需求。

从设计的角度来看,P2P 数字城市服务平台有如下的优点:

(1) 网络带宽的最优化使用。现今 Web 上典型的局部流量集中导致的拥塞不会影响 P2PSMS 网络。

(2) 更低的内容分发成本。P2P 网络能吸纳内容并复制它,使它易于存取。

(3) 来自网络中各个节点的计算能力的均衡。通过异步操作,您可以同时发出许多资源或服务请求,然后让网络为您完成这些工作。

(4) 无限的可伸缩性。P2P 网络可以在不影响可伸缩性限制的情况下横越整个已知的连接着的宇宙;而这在任何集中式模式中是完全不可能的。

(5) 完好的封装性。节点具备的良好封装性,对于管理者和其它节点而言,它们能且仅能看到该节点自愿提供的信息,如功能列表、会话情况等。

(6) 松散耦合。当一个节点的接入端发生变更的时候,只要提供的服务内容不变,调用它的其它节点者是不会感到这一点的。对于调用者来说,只要调用界面不变,服务的实现任何变更对他们来说都是透明的,甚至是当服务的实现平台从 J2EE 迁移到了 .NET 或者是相反的迁移流程,用户也无需知晓。

4.2 P2PSMS 组成

P2PSMS 将由一组基于 P2P 模式的开放协议、一个或一组专门的 P2PSMS 管理者(Manager),以及 P2PSMS 节点(Node)组成。

其中 P2PSMS Manager 主要负责节点管理、在线节点维护、节点会话管理、服务列表、搜索引擎(提供服务检索、节点检索等)、在线即时沟通等方面的内容。它并不像传统的 Server 提供直接的服务,而是提供对服务的管理和用户管理以及相互间的操作。它实际上是把各种服务以及会话(Session)作为对象来进行管理、调配、通信等,并将这些信息提供给用户查询使用,这种查询可以是人工的,也可以是自动的^[11~14]。

P2PSMS 节点(Node)实际上由三部分组成:代理(Agent)、驱动(Driver)、接入端,如图 2 所示。

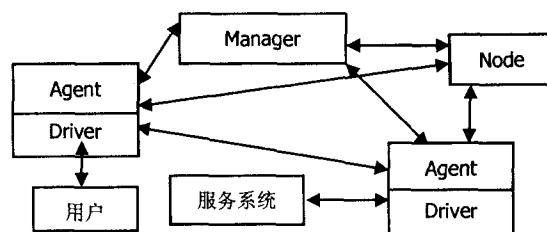


图 2 P2PSMS 节点构成

Agent 连入 P2PSMS 中,并负责与管理者、其它节点进行交互。实际上,根据接入端的不同,从功能上可以分为两种 Agent,即服务请求代理(Request Agent)和服务提高代理(Service Agent)。作为前者,它主要是向用户提供标准界面,将用户需求转化为标准的 P2PSMS 请求,并通过和 P2PSMS Manager 进行交互(浏览、查询等),得到各种服务的列表及相关信息(服务的元数据、服务接口等),然后去访问相应的其他节点,以获得服务;作为后者,它定期向管理者发送自己的状态描述(可提供的服务、相应的服务接口等元数据,以及目前的会话情况、系统状况等等)。但这两者中,Agent 只是功能上不同,它们所遵守的协议是相同的,互相之间的地位也是平等的,因此 P2PSMS 节点间可以进行互操作,提供相互的对等服务^[3~5,10]。

P2PSMS 节点通过 Driver 来连接接入端,接入端可以是一个浏览器或是传统意义上的客户端,也可以提供具体服务的系统,即传统意义上的服务器。不同的接入端分别有不同的 Driver,但所有的 Driver 均遵守相同的协议、规范。

实际上,Node 通过不同的 Driver 屏蔽接入者的不同,通过 Agent 保证相互的平等性。这正是基于 P2P 模式的 P2PSMS 最重要的特征。

P2PSMS 协议组是 P2PSMS 系统的关键和基础,它包含 P2PSMS Manager 和 P2PSMS Node 进行交互的各种协议;它

包含 P2PSMS Node 之间的互操作协议以及 Agent 和 Driver 的实现规范;它还包含对服务的描述协议,既包含各种服务共性的描述,也包含具体服务特性的描述,以及如何获得服务(如数据格式、参数、所需环境等元数据)。此外,它还应包含相应的安全机制、认证机制、授权机制等等。

P2PSMS Manager 本身不提供具体的应用服务,P2PSMS 所能提供的服务是所有 P2PSMS 节点所能提供的服务的总和,其服务能力与在线节点的数目有直接关系,理论上,它的可扩展性和服务能力是无限的。

表 1 P2PSMS 协议组

协议名称	描述
管道 (Pipe) 协议	管道是 P2PSMS 中各节点以及管理者相互之间用来通信的虚拟通道,本协议是最基本的连接协议。包括管道的申请、建立、绑定、取消、等待等过程。
注册协议	用于节点向管理者注册、变更、注销节点或服务的相关信息。
查询协议	用与节点查询节点、服务、会话、节点状态等信息。
授权与认证协议	节点间以及节点加入系统时所需的协议,用于身份认证以及相应的权限的申请、设定、修改等。
服务协议	用于获取或提供服务的协议。包括服务的基本属性及扩展属性(用于描述具体服务)的设置和服务申请、使用等等。
消息队列协议	用于系统节点间同步或异步方式的消息传输,通过队列方式来实现消息的存储转发,以保证信息的可靠传输。
代理协议	规定了实现一个 P2PSMS Node 的 Agent 所应符合的规范
驱动协议	规定了实现一个 P2PSMS Node 的 Driver 所应符合的规范
节点缓存协议	实现对已访问节点相关信息的缓存。

5 平台设计

5.1 JXTA

JXTA 技术是由 SUN 公司在 2001 年 2 月提出的一项新技术,主要用于提供 P2P 程序所需的基础服务。该技术致力于创建一个通用的平台,以简单而有效的方式构建特定的对等式和分布式服务与应用,使得开发者不需要过多考虑如何解决对等计算的技术问题,而可以专注于如何实现与完善可扩展、互操作性强且具有高可用性的高层应用。可以说,JXTA 是 P2P 发展史上一个里程碑式的标志。JXTA 具有互操作性、跨平台性、通用性等特点,这正是数字城市服务管理这个复杂系统所需要的,因此我们选择了 JXTA 为技术实现的平台。

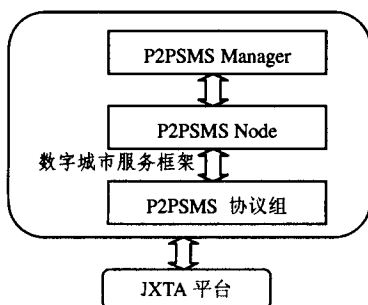


图 3 框架系统

框架系统可使用 JXTA 平台的 JAVA 参考来实现。主要包括 3 部分(图 3):

- (1) P2PSMS 协议组;
- (2) P2PSMS Node;
- (3) P2PSMS Manager.

5.2 P2PSMS 协议组

P2PSMS 协议组是 P2PSMS 系统实现的关键和基础,它包含表 1 所示的核心协议。

从协议的层次上看,P2PSMS 的协议组属于应用层的协议,是建立在 HTTP、WAP 等协议基础上的一种协议,不涉及到网络中传输细节,大部分可以基于 XML 来定义,以保证系统的开放性和良好的可读性。例如,设计成数据报的消息,包含一个信封和带主体的协议头堆栈,以及使用称为 UUID 的 128 位数据的标识来引用诸如节点、代理或服务等实体。

另外,现有的很多标准和协议可以根据实际需要直接使用或作为参照,例如 SOAP、WSDL、UDDI 等等。

而其他如安全性、加密、服务质量(QOS)等方面的内容也可以作为 P2PSMS 协议组的附加协议逐步加入。

5.3 P2PSMS Node

对于 P2PSMS 节点,其代理、驱动、接入端 3 部分是相对独立的。原则上,我们对接入端不做改动,主要功能都通过代理和驱动间接完成。而驱动要根据不同类型的接入端进行专门的开发。

Agent 的实现主要是基于 P2PSMS 协议组的各种操作,如管道连接、查询、数据交换、服务获取等。原则上采用 JAVA 语言,以保证其平台无关性。通常,我们认为 Agent 具有自主性、反应性、交互性、主动性等特点。

我们可以知道,不同的接入端会使节点有完全不同的表现(通过不同的 Agent 和 Driver 来体现)。例如,可以在传统的非开放、非网络化的服务系统中安装,那么它就相当于该服务系统的虚拟用户,负责将整个 P2PSMS 系统的用户的需求反馈到该服务系统中,并将结果返回给用户,完成对用户的透明转化,而且可以支持与其他服务系统的互操作。也可以装在 WebGIS 服务器上,相当于一个 Web 三层架构中的中间件,是对 WebGIS 用于服务器的扩展,协助实现 WebGIS 的动态负载均衡、数据的缓存、WebGIS Server 与其它系统之间的动态数据交换等功能。还可以将它进一步简化,装到嵌入式系统中(如手持设备等),作为用户的智能助理,实现导航、定位、监控、数据传输等功能。

5.4 P2PSMS Manager

P2PSMS 的管理者将不作为整个系统中的重心和主要的服务,而更多的是对整个虚拟系统进行管理,以及节点之间的会话控制、调度管理等。因此,它基本上可由一个中小型关系数据库、搜索引擎和一些常用组件构成。

其中数据库用来保存所有节点的信息,包括在线和曾经在线节点的历史信息,所有系统可用/曾经可用服务信息,所有服务会话信息。另外,数据库还用来实现消息队列的存储转发功能。

相应地,在主存中维护如下内容:在线节点表 ANT (Active Node Table)、当前服务表 CST (Current Service Table)、活动会话表 AST (Active Session Table)、消息队列 MQ (Message Queue)、系统缓存 SC (System Cache)、其它;同时,运行如下 Daemon:节点、服务注册/注销进程、服务搜索引擎、

(下转第 222 页)



图3 一个访谈对话场景的镜头关键帧序列。镜头158被选为最终的场景关键帧

表1 图3中部分关键帧的FF-ISF值

	#157	#158	#159	#160	#161
FF	0.125	0.875	0.125	0.875	0.75
ISF	0.005	0.083	0.001	0.072	0.008
FF-ISF	6.25e-4	7.26e-2	1.25e-4	6.3e-2	6.0e-3

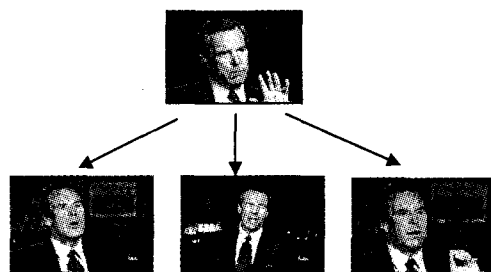


图4

图3是一段主持人访谈的新闻场景,其中镜头157、159

等为主持人镜头,镜头158、160等为被采访对象镜头。对于这一段新闻故事,我们更希望用含被采访对象的帧作为关键帧,因为它给我们带来更多的信息。而主持人经常出现在新闻中,只含有主持人的帧能给我们的信息就相对地少一些。我们的算法最终选择镜头158的关键帧作为整个场景的关键帧。从表1可以看出,镜头158的关键帧有相对比较高的FF和ISF值。所有的主持人镜头帧都因为经常出现在其他场景中,所以ISF值比较低,从而没有被选中。图4给出了一些此场景以外镜头关键帧的例子,这些帧在算法中被认为与镜头159的关键帧相似。

结论 本文提出了一种基于FF-ISF的关键帧提取方法。它的优点在于同时利用了镜头内和镜头间的视觉信息来选择关键帧。利用镜头间的信息来选择关键帧可以使关键帧更加个性化。通过实验分析,我们的算法在用于场景关键帧提取时更加有效和省时。实验结果说明了我们的方法能够有效地提取有价值的键关键帧。

参考文献

- 1 Pentland, et al. Video and image semantics, advanced tools for telecommunications. IEEE Multimedia, summer 1994
- 2 Ngo C W. Analysis of spatio-temporal slice for video content representation; [PH. D. thesis]. 2000
- 3 Wolf W. Key frame selection by motion analysis. In: Proc. of IEEE Intl. Conf. on Acoust, Speech and Signal Processing, 1996
- 4 Hanjalic A, Zhang H J. An integrated scheme for automated video abstraction based on unsupervised cluster-validity analysis. IEEE Trans on CSVT, 1999,9
- 5 Zhang Yueting, et al. Adaptive key frame extraction using unsupervised clustering. In: Proc. of IEEE Intl. Conf. on Image Processing, 1998
- 6 Xiong W, et al. A novel technique for automatic key frame computing. In: Proc. of SPIE, 1997
- 7 Rasheed Z, et al. Scene detection in Hollywood movies and TV shows. In: Proc. of Intl. Conf. on Multimedia and Expro, 2003

(上接第103页)

会话管理进程、节点调度进程、消息进程、缓存同步进程。

小结 在数字城市框架上,对数字城市的服务进行归纳和抽象,提出了数字城市中基本服务的概念,并在此基础上设计了基于P2P模式的数字城市服务管理平台P2PSMS,专门用来管理数字城市这个复杂的服务系统,实现通用化的服务管理。它具有高可扩展性、高度动态性、良好的封装性以及松散耦合的特点,都是符合数字城市服务管理平台的需要的。

参考文献

- 1 林绍福,李琦. 数字城市:创建21世纪的智能服务平台. 计算机科学,2002(7)
- 2 林绍福,李琦,董宝青. 数字城市应用服务平台体系结构研究. 计算机科学,2002(12)
- 3 沈达阳,林作铨. Internet上的软件Agent. 计算机科学,1997(4)
- 4 史忠植,王文杰,田启家. 智能主体研究现状与发展趋势. 计算机世界,1998(1)
- 5 丛升日,罗英伟,等. Geo-Agent的体系结构. 中国图象图形学报,1999(4)

- 6 Shuey R L, Spooner D L, Frieder O. The architecture of distributed computer systems: a data engineering perspective on information systems. Addison Wesley Pub Co, 1997
- 7 Szyperski C. Component Software: Beyond Object-Oriented Programming. Addison-Wesley, 1997
- 8 Bradley N. The XML Companion. Addison-Wesley, 1998
- 9 Yang Chaowei, Li Qi. Web publishing of Spatial Information. IS-DE'99, 1999
- 10 Maes P. Artificial life meets entertainment; life like autonomous agents, Communications of the ACM 1995,138(11)
- 11 Shuey R L, Spooner D L, Ophir Frieder O. The architecture of distributed computer systems: a data engineering perspective on information systems. Addison-Wesley Pub Co, 1997
- 12 Szyperski C. Component Software: Beyond Object-Oriented Programming. Addison-Wesley, 1997
- 13 Goodchild M F. Interoperating Geographic Information Systems. Boston: Kluwer Academic Publishers, 1999
- 14 Yang Chaowei, Li Qi. Web publishing of Spatial Information. IS-DE'99, 1999