

无线环境下视频服务的若干关键问题^{*}

杨书慧 顾铁成 吴敏强 陈道蓄

(南京大学计算机软件新技术国家重点实验室 南京210093)

(南京大学计算机科学与技术系 南京210093)

Key Issues of Video Services in a Wireless Environment

YANG Shu-Hui GU Tie-Cheng WU Min-Qiang CHEN Dao-Xu

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093)

(Department of Computer Science and Technology, Nanjing University, Nanjing 210093)

Abstract There are two new trends in today's computer technology that are attracting more and more attention. One is multimedia service, especially the video service. The other is mobile wireless computing. The former can be implemented efficiently on the traditional wired network, such as the accession of multimedia information, real-time broadcast. But due to the limitation of the wireless environment, the latter hasn't been able to support the video transport efficiently. Obviously, the combination of these two technologies, video service over wireless network, may be a killer application and in turn advances both of them. This paper summarizes the challenges of this field and introduces some possible solutions.

Keywords Video transport, Wireless environment, Video coding, Multi-resolution video data

1 概述

随着计算机技术的发展,各种各样的视频服务涌现了出来,可视会议(Videoconference)、视频点播(Video-On-Demand)都成为了学术界以及工业界所关注的热点。由于视频服务有着一些特殊的性质,如数据量很大,需要连续传输、实时获取等。学者们将研究重点集中在了提升视频服务有效性的若干个方面。如以获得高吞吐量为目的的磁盘调度,获得低延迟的接入控制与缓存管理,获得高效传输的原码压缩方案等等。另一个新的发展趋势就是无线服务。正如我们所了解的,现在各种移动的用户终端可以用无线的方式任意接入网络,和固定的主机交互,获取信息。这些终端包括手机、PDA(Personal Digital Assistant)等,它们的特点是处理能力有限,电池等能源提供也有限。由于无线环境所固有的缺陷与限制,目前还只有文本和音频信息可以有效地获取到,视频信息还达不到实用的标准^[9]。不过可以肯定的是,如上所述的这类服务一定会有广阔的应用前景,谁不希望只用一个便携的移动终端来进行视频会议,视频点播呢?所以,这一新兴领域值得进一步地探究。

现在无线通讯的能力已经越来越接近视频数据的传输要求,但是实际的视频播放效果仍然是无法接受的。这是因为无线通道不仅是带宽太窄,而且很不稳定,并有较高的出错率。所以,问题的关键就是怎么样在这种情况下保证视频的QoS。在本文中,我们给出了对整个问题一个框架模型,并在此基础上,介绍了两种基于不同层面的解决方案。一个是视频数据在视频服务器上的一种新的布局方法。这种方法可以适应无线信道不断变化的状况,提供各种级别的QoS。另一个是数据在被从服务器上取出以及发送到无线信道中去之前的一些编码,压缩等处理方法。特别地,我们介绍了信源编码与通道编码的“联合编码”理论,这一理论将会影响到今后无线传输中视频编码的具体发展方向。我们还介绍了两种具体的编码与

压缩方法,它们都可以有效地提高此类视频服务的质量。

2 无线网络中视频传送的一个抽象模型

2.1 抽象模型

现有大型主干网络,如因特网,都是有线的,人们所讨论的无线网络基本都是作为有线网络的某种形式的延伸。也就是说,只有最后一段接入网是无线的形式。所以本文中将要讨论的不是要修改甚至替代现有有线网络的构架以及组织结构,而是要给它加进一些内容,让它可以兼容无线环境并且屏蔽掉所有的异构性。在这里,我们给出向无线用户终端传送多媒体信息的系统体系结构(见图1)。如图所示,高速的通讯主干是固定有线的,可以是通常的因特网。移动支撑工作站(MSS—Mobile Support Station)是装有无线接口的固定主机,它在无线与有线媒介间起到类似网关的作用。一个MSS主机负责一个移动区域(Mobile Cell)。MH(Mobile Host)是移动客户端,它通过MSS向多媒体信息服务器(MIS—Multimedia Information Server)获取数据。实际上,这一模型是拓展了现在流行的移动网络的模型,增加了一个无线代理服务器(WPS—Wireless Proxy Server)到每一个MSS中,起到一个中间件的作用。后面我们将详细地介绍它。MSS的其余部分是用来处理其他一些移动接入的常规问题的,如移动接入管理。

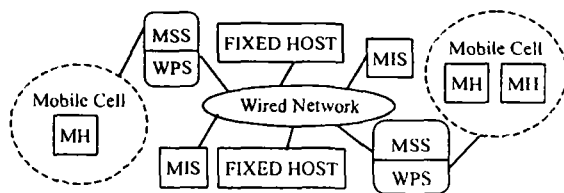


Figure 1 System Model

^{*} 本文的工作得到了国家863项目“应用服务器的运行、管理及调度技术”(项目编号:2001AA113050)及江苏省高技术产业化推进项目“基于多服务器群的大规模多媒体信息服务平台的研究和开发”的资助。

2.2 无线代理服务器 WPS

无线代理服务器(WPS)位于有线网络的“边界”区域,负责将多媒体数据透明地传送到移动客户端。对于 MIS 来说,无论向固定主机(Fixed Host)或是 MH 发送数据都是一样的。WPS 的具体功能如下:

•QoS 管理 无论在有线或是无线网络中,QoS 都是非常重要的因素。考虑到无线环境的低带宽、不稳定、以及 MH 的低处理能力等问题,WPS 必须能够有效地支持各种级别的多媒体流。所以,它应可以感受到网络环境的变化,并且是自适应的。例如,当遇到通讯失败时,WPS 可以根据当时可获得的资源调整发送内容与发送策略,以保持向用户提供连续传送,并能够做到在资源缺乏时平滑降低视频质量。

•协议转换以及视频编码转换 在有线网络中,我们使用 TCP/IP 协议来传送数据。但对于无线网络,我们必须选择其它合适的协议来适应无线网络的特殊需要。例如,我们可以使用 H. 223 推荐协议,因为它可以提供一种 QoS 参数,并且它对于短数据包的开销是最小的,这非常适用于低数据率的视频传送^[4]。WPS 从 MIS 上下载所需的视频数据,这些数据通常是以 MPEG 的格式存放的,并不适合于无线传输。所以 WPS 将根据当时的网络状况、MH 终端的处理能力、容量大小、分辨率等因素选择一种合适的方案对视频序列进行重新编码(Transcoding)。并且可以在编码中加入错误检测、错误恢复功能。

•高速缓存,视频流与重传 因为无线网络存在着不稳定等问题,数据经常会出错以致需要重传。WPS 对此进行管理。WPS 上设有高速缓存区,确保需要实时传输的视频数据可以停留在 Caching 里。当错误发生时,仅需要从与 MH 在同一个移动区域的 WPS 上重传而无需再向 MIS 申请,从而可以节约时间降低延迟。另外,如果 MIS 不支持流式传送的话,可以让 WPS 来支持这一功能。

•保存 MH 的能力参数信息 最后,WPS 将对它所在移动域中 MH 的参数信息予以存储。这些信息包括显示分辨率、帧速率、存储容量、计算速度等。如上所述,当 WPS 决定它的传送策略时将用到这些信息。

3 多分辨率视频数据布局方案

在传统的多媒体信息服务器设计中,主要的瓶颈之一是磁盘的 I/O 带宽。因为磁盘系统的带宽小于有线网络的带宽,所以为了有效地利用网络带宽,视频数据都被分割到了多个磁盘上,以并行输出的方式来保证实时传送。但是无线网络的情况不是这样,它的带宽(通常在1Mbps 左右)比磁盘 I/O 带宽还要小。所以,原有的数据分割方法将不再适用。针对这一新情况,多分辨率视频数据布局方案被提了出来。当我们将此方案应用于高带宽的磁盘阵列时,就可以获得适应网络变化的低延迟的视频传送。实际上,在传统的多媒体信息服务器中,有时候会保存同一视频文件的不同分辨率的版本,在传送时,就可以根据网络状态以及用户端情况选择适合的版本。这里,这一思想被进一步深化了。即使用一定的数据放置以及分割方法使在同一视频文件中可以抽取不同分辨率版本的传送实例。

基于这种多分辨率视频数据的概念,一种可行的视频数据在服务器上的分布方案在文[1]中被提了出来,这种方案可以有效地解决无线环境视频传送的若干问题。在方案^[1]中,K 层的多分辨率数据流是指包含一帧或几帧的一个视频片

断(segment)被划分成了 k 个部分(component)。这样当提供第 j 等级视频质量时,就从最底层开始送出去共 j 个部分即可。所有部分全部送出时就可以获得原始分辨率。相应的,一共有 k 个等级的质量标准。QoS 参数也就可以根据质量级别来设定了。文[1]中具体描述了这一方案。用 i, j 来表示第 i 个片断的第 j 个部分,并且用 d 来表示磁盘数目。以下具体方案将视频流分割放置到了磁盘上,分割的单位是部分。如图2所示(k=d=5,i=2),连续的部分被以 Z 字形放置在磁盘阵列上。这样,就在满足连续实时传送的同时实现了对多分辨率数据输出的支持。例如,当需要以第二质量级别发送第一,二个片断时,1. 1,1. 2,2. 1,2. 2 被从磁盘1,磁盘2,磁盘4,磁盘5 同时取出。由1. 1,1. 2 组成片断1,2. 1,2. 2 组成片断2。当然这只是最简单的一种情况。有的时候 k 和 d 并不相等。如果 k 小于 d,那么一个片断的 k 个部分将放在 k 个磁盘上,而另外(d-k)个磁盘的同一位置将空闲,并且各个片断对应的空闲磁盘是不同的。当 k 比 d 大时,更加复杂的放置策略会降低它的有效性,所以应该尽量避免这种情况。

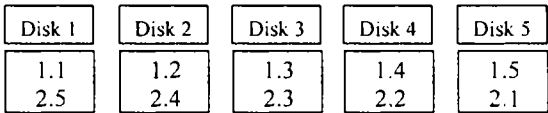


Figure 2 Multi-resolution video layout

虽然这种视频数据布局方案较为简单,却十分有效,可以满足对 QoS 的管理。很明显,应用这种方案可以简单地对同一数据流实施不同级别的 QoS,计算 QoS 参数,实时地改变传送质量且不需要额外的开销。同时还可以提供一些 MH 所需要的交互^[1]。如快进操作。见图2,如果原来的传送流是第四等级的,那么在不增加带宽要求的情况下可以将等级降低到原来的一半,以获得原来两倍的速度。其他一些策略,如 skip-ping 方案也可以和这里的方案联合使用,以进一步提高性能。

4 视频编码解决方案

由于多媒体信息的数据量之大以及现有无线移动网络的低带宽和不稳定性,传送数据的编码方法成为了对于传送的实现以及传送质量都至关重要的一件事。我们知道,现在动态图像的最重要的标准就是 MPEG (Moving Pictures Expert Group)^[8,10,11]。MPEG-1 可以以 1.5Mbps 到 4Mbps 的速率传送数字视频信息,这样的速度可以和 CD-ROM 以及 DAT (Digital Audio Tape)相匹配。MPEG-2 是在 4Mbps 到 20Mbps 的范围内为 HDTV 提供高质量的视频信息的,也可以达到 80Mbps^[2]。MPEG-4 是最适合移动计算的一个版本,不仅是因为它是可视电话、可视游戏等低传送速率的应用为目标的,而且它强调对多媒体数据进行基于内容的处理^[8]。另一个针对可视电话和可视会议等应用而开发的标准是 ITU (International Telecommunication Union) 提出的 H. 263。与 MPEG 相比,它需要较少的 CPU 资源来进行实时解码,这对于移动计算来说是一个很有利的条件,因为绝大多数移动终端都是缺乏计算资源的。这些标准都有适合于无线传送的特性,但它们都不是专为无线传送设计的,所以不可避免地会有不足之处。以下将介绍针对这一问题的一些专门的技术。

4.1 联合编码理论

现在的多媒体信息传送,尤其是有线传送,信源编码器和通道编码器是分开设计的。这种做法的依据是 Shannon 的著

名分离法则(separation principle)^[12]。这一法则从理论上证明了将信源编码和通道编码分开进行可以达到网络带宽利用率的最优化。我们可以简单地来理解它^[3]。例如当使用一条不稳定的通道进行点到点传输时,若通道容量为C比特/秒,那么将通道编码的结果设为R,则 $R \leq C$ 即可,然后信源编码的任务就是尽一切可能地压缩多媒体数据,使其适合现有的通道性能,也就是R。很明显,没有其他方案可以比这种做法更优了。这样就可以保证正确有效的多媒体数据传输。但是在无线环境中,信道是很不稳定的,并且极易出错,现在最流行的可变长信源编码方案(Variable Length Sources Coding Schemes)对于错误十分敏感。所以哪怕是很少几个比特位的错误,都会使整帧图像被认为无效而要求重传,从而降低了实际的通道利用率。联合编码的另一个原因是为了通过统计复用来最好地利用通道资源。

据统计,无线环境中的传输出错率平均可以达到 10^{-2} 左右^[7]。所以错误检测与错误恢复都至关重要。传统的回旋编码(convolution codes)和块编码(block codes)方案以及自动重复请求(ARQ)形式的重发可以在一定程度上解决这一问题。当然这一类的开销不可能降低到零,但是我们可以寻找另一种解决方法,那就是让信源编码部分也来做一些这一类工作^[3]。

4.2 基于 H. 263 的视频编码器

ITU 推荐标准 H. 263 是专为窄带信道中的低速率传送的视频通讯而设计的一种基于离散余弦变换的编码方法^[7]。与 H. 261 相比,它有更低的输出速率以及更高的图像质量。但是,H. 263 设计之初是为了在 POTS 以及 PSTN (Public Switch Telephone Network) 上进行可视通讯的,这些网络环境的特点是低带宽、稳定性较好、出错率小于 10^{-6} 。在这样的环境下,H. 263 可以很好地工作。但对于出错率较高的无线网络来说,如果不加改进而直接使用它是不合适的。

在文[7]中提出的基于 H. 263 的视频编码器是一种较好的解决方案。其基本思路就是在压缩编码前加入一项预处理,即将图像分割成不同重要性等级的区域,并对它们分别用 H. 263 进行编码。根据文[7]中的试验,比特位错误对整个视频质量的影响大小在于此错误在整个图像流中所处的空间位置以及时序位置。并且由于带宽等原因,我们不可能对整幅图像进行百分之百的防错处理。所以我们可以得出结论,最好的方法就是对其中最关键的部分进行有效的防错保护。在压缩编码之前,可以将一帧图像分解成N子帧。可以用一些专门的算法来决定子帧的划分,如形态过滤(morphological filtering),小波算法,金字塔分解算法(pyramidal decomposition)等。在一些特殊应用如可视会议中,大部分图像相对来说比较“静态”,只有头部和肩部变化较大,所以利用一些简单的启发式方法就可以做好分割的工作了。然后每一子帧可以分别地用各种方案进行容错,交叉等处理,具体要视它的重要级数而定。若用 S_{ij} 来表示图像i的第j个子帧,则在MH端,整幅图

像可以用公式 $S_i = \sum_{j=1}^n S_{ij}$ ^[7]来获得。当然在传送过程中会产生错误位,但是错误一定会被限制在一个子帧中。如果某一子帧被发现包含错误,可以简单地将其丢弃,并且使用MH中的帧存储器中相应子帧的内容代替它。帧存储器中存有每个子帧的前一个正确传送的内容。这样如果前一子帧和当前子帧的差别不大,仍然可以获得高质量的图像,如果差别很大以致不能保证图像允许的最低质量,那么就要请求WPS重发这

一子帧。并且在下一次传送时,这一子任务将被赋给较其它传送子帧高的优先级。并且这种设计方案对MH中帧存储器的大小要求并不大。

实际上,这种方案也可以利用其它具体的压缩方法进行子帧压缩。而且,这一压缩预处理可以和压缩过程联合完成以获得更高的效率。这一方案的优点是显而易见的。因为子帧被单独编码传送,所以某一子帧中的错误不会传播到其它子帧,而且即使需要重传,也只是要重传一个子帧而非原先的一整帧。当网络带宽有限时,我们可以只传送重要级数高的若干个子帧,其它部分用帧存储器中的内容。这样仍然可以保持MH端有可以接受的质量的连续播放视频。根据文[7]中的性能测试,这一方法在减缓延迟,减少出错等方面都有良好性能。

4.3 3D-SLCCA 视频编码技术

虽然 H. 263 是适用于低速率传送的,并且可以提供可以接受的视频质量,它的计算复杂性以及错误敏感性都降低了它的实用性。所以开发一种新的压缩算法可能是一种更好的解决办法。这里,我们将对3D-SLCCA,一种适合低速率传送的高度可扩展小波视频编码方法,进行讨论,它可能成为无线传输压缩编码的又一热点。

由2D到3D的拓展是加入了时序一维,成为了(时序*空间*空间)的三维的模式。如果设计合理的话,这样的压缩编码算法将具有可扩展性,低计算复杂度,以及有效抑制错误传播的特性。3D-SLCCA 是利用了 SLCCA 方法的三维压缩方法。其主要处理流程是这样的,前后两个连续帧将作为一个处理单元,让它们的和以及差分别进入时序低通滤波器和时序高通滤波器进行处理,然后分别进行二阶空间小波分解。再用 SLCCA (Significance-Linked Connected Component Analysis) 方法来组织与表示有效小波系数(significant wavelet coefficients)。更多的技术细节可以参阅文[5,6]。

3D-SLCCA 不加修改地利用了 SLCCA 的有效的数据组织与表示策略。它将二阶的空间小波分解方法应用到了低通滤波器以及高通滤波器中的子带上,这不仅可以提高最终的质量,还可以降低计算复杂性。根据可获得的带宽和计算能力,编码器以及解码器可以自动地调整空间分辨率以及时序分辨率。文[6]中的性能评估指出,这种新的编码方法可以获得比 H. 263 更好的性能。

结论及今后工作 在本篇论文中,我们讨论了一种新的服务,即在无线网络中的视频服务。对这一领域的背景知识以及困难所在进行了分析与归纳。我们提出了一个抽象模型以作为解决问题的基础,并且这一模型并没有影响到现有的有线传输系统的构架。出于对 QoS 的考虑,我们引入了一种多分辨率视频数据的布局方案,它可以满足无线视频传输的多项要求。然后我们着重讨论了视频编码的问题,并引入了信源与信道的“联合编码”原则。

本文中介绍的这些解决方案都可以一定程度上地提高无线网络中视频传送的性能。但是这些对于实际需求来说还远远不够。以后的工作可以是对压缩算法,容错编码方法等的进一步研究。我们现在的讨论还基本只局限在数据编码这一方面,多分辨率视频分布方案也只是稍微涉及到了点多媒体服务器一端的设计问题。我们认为今后这一领域的研究可以拓展到多媒体服务器的设计层面,即视频信息在 MIS 上的组织,存取,甚至磁盘调度,接入控制等方面都可以根据无线传送的特点来进行设计改进,从根本上来兼容无线网络。

参考文献

- 1 Cho J, Shin H. Design issues for multimedia information servers in mobile computing environment. in: Proc of the 4th Intl. Workshop on Mobile Multimedia Communications, Sep. 1997. 336~339
- 2 Alonso R, et al. Managing video data in a mobile environment. SIGMOD Record, 1995, 24(4): 28~33
- 3 Bahl P. Supporting digital video in a managed wireless network. IEEE Communications Magazine, June 1998. 94~102
- 4 Vass J, et al. Mobile video communications in wireless environment. In: Proc. of IEEE Workshop on Multimedia Signal Processing, Copenhagen, Denmark, Sept. 1999. 45~50
- 5 Vass J, et al. Error resilient video communications in wireless ATM. IEEE Journal on Selected Areas in Communications, June 2000
- 6 Vass J, et al. 3DSLCCA-A high scalable very low bit rate soft-

ware-only wavelet video codec. In: Proc. of IEEE Workshop on Multimedia Signal Processing, Dec. 1998

- 7 Bahl P, Chlamtac I. H. 263 based video codec for real-time visual communications over wireless radio network. In: Proc. IEEE Int. Conf. on Universal Personal Communications, San Diego, CA, Oct. 1997
- 8 Gall D L. MPEG: A video compression standard for multimedia applications. Communications of ACM, 1991, 34(4): 46~58
- 9 Imielinski T, Badrinath B R. Mobile wireless computing: Challenges in data management. Communications of ACM, 1994, 37(4): 38~47
- 10 Filippini L. MPEG-FAQ. <http://www.crs4.it/HTML/LUIGI/MPEG/mpegfaq.html>, July, 1995
- 11 Ozer J. Video compression for multimedia. Academic Press, 1995
- 12 Shannon C E. Communications in the presence of noise. In: Proc. IRE, vol. 37, Jan, 1949

(上接第65页)

最好的数据作为解决方案的提供者,然后跳到算法的第9步。

7 请求需要数据笛卡儿乘积操作才可能获得解决,则根据 D 中 Item 对应的元组 H,找到另外一个数据 D'的元组 H',使得 $H \times H' = G$ 。

8 递归执行该算法以获得 D'。如果仍然需要 D'',则继续递归,直到获得能够满足请求的数据集合,如果搜索了所有的相关的数据仍然不能满足需要,则请求无解,算法结束。在选择 D', D''等的时候,要遵循尽量在一个物理元数据库的原则。

9 找到了满足请求的数据集合 (D, D', D'', ...), 如果请求本身含有交操作,则试图通过数据的 and 操作来完成;同理还有或、差、选择、衍生等操作作用物理级的 or, not, where 等代替。

10 在物理层实现每个 Data 的演算,在同一个物理数据源的笛卡儿乘积等操作则在物理层通过表格的连接操作(join)实现。最后把 XML 格式的 Data 集合提交到虚拟层。

11 在虚拟层完成底层无法完成的一些请求演算。

12 最后将规范的结果集合(XML 格式)提交到应用层。

13 将问题的解决方案写入虚拟数据库日志,以便于数据模型的自身优化。

4.4 算法分析

·算法的动态有穷性 本算法是机械执行的,并且是在一个有限的数据集合上进行请求解决方案的搜索,所以总是能够终止的。

·算法的有效性 算法可以回答虚拟数据库可以回答的所有问题,因为依靠维度、粒度、度量、主题等概念,算法从最可能的、效果最好的数据开始搜索一直到数据集穷举完毕。并且,虽然算法不能保证总给出最优解,但算法总是能给出一个相对稳定的较好的解。

·算法的优越性 基于查询优化的虚拟数据库算法比传统的虚拟数据库算法而言的优势在于其底层的数据操作更加灵活地根据需求的变化而变化。虚拟数据库淡化了数据仓库集成的特点,而本模型进一步淡化了数据仓库不可修改的特点,目的都是为了能够更灵活更方便地为 OLAP/DSS 提供服务。

5 相关工作的比较以及继续研究的问题

1. 关于数据库的多维数据模型

在数据仓库的构建过程中,数据模型的设计至关重要,因而也成为了数据仓库研究的核心问题之一。文[1]以“偏序”和“映射”为基础,给出了一种多维数据模型以及表达数据仓库的复杂数据结构和语义,并提供了一个以 OLAP 操作为核心的操作代数。

2. 关于虚拟数据库

需求驱动的数据仓库是作为供给驱动型的物理数据库的替代物而产生的。在虚拟数据库中,其自身并不实际集中存储大量的数据,而仅仅是由一层(或多层)用于将多种不同的数据系统连接在一起的中间件所构成的,这些中间件还负责将数据呈现给用户,通常是通过 Web 界面。这样,无论最终实现的是集中式的数据仓库,数据集市或者其他,虚拟数据库都能提供访问各个存储地点所存放的数据的方法。对用户来讲,并不需要去关心物理数据的实际存放位置,因而是“虚拟”了的。文[2]给出了虚拟数据库的特点,并解释了虚拟数据库与物理数据库的关系。

3. 关于元数据模型

在物理上分布的数据的集成过程中,元数据模型发挥着至关重要的作用,管理数据的关键主要来自于对元数据的管理。文[3]提供了一种用于分布式信息集成的模型驱动的体系结构,这种体系结构主要基于元数据模型。

目前,我们正在本所提出的基于查询优化的虚拟数据库模型的基础上实现面向 Web 的数据查询优化和处理方法。

参考文献

- 1 李建中,高宏.一种数据库的多维数据模型.软件学报,2000,11(7):908~917
- 2 Bischoff J. Data Warehouse - Practical Advice From The Experts. 1998. 196~202
- 3 A model-driven architecture for Distributed Information Integration. <http://www.omg.org>
- 4 Inmon W H. Building the Data Warehouse Second Edition. 2000. 116~143
- 5 Zhang Yulong, Li Shanping. Web-based XML Document Management. In: Proc. of APWEB' 2000. Xi'an, China, 2000, 10
- 6 Lou Agosta. 数据库技术指南. 人民邮电出版社, 2001