

中国网络 视频 史



陆地 靳戈/著

ONLINE VIDEO

中国广播影视出版社

中国网络 视频 史

陆地 靳戈/著

ONLINE
VIDEO

中国广播影视出版社

图书在版编目 (CIP) 数据

中国网络视频史 / 陆地, 靳戈著. — 北京 : 中国广播影视出版社, 2017.8

ISBN 978-7-5043-7937-5

I. ①中… II. ①陆… ②靳… III. ①计算机网络—视频系统—历史—中国 IV. ①TN941.3-092
②TN919.8-092

中国版本图书馆CIP数据核字(2017)第148253号

中国网络视频史

陆地 靳戈 著

责任编辑 余潜飞

封面设计 成晟视觉

责任校对 张 哲

出版发行 中国广播影视出版社

电 话 010-86093580 010-86093583

社 址 北京市西城区真武庙二条9号

邮 编 100045

网 址 www.crtpp.com.cn

电子信箱 crtpp8@sina.com

经 销 全国各地新华书店

印 刷 三河市人民印务有限公司

开 本 710毫米×1000毫米 1/16

字 数 240 (千) 字

印 张 17.5

版 次 2017年8月第1版 2017年8月第1次印刷

书 号 ISBN 978-7-5043-7937-5

定 价 43.00 元

(版权所有 翻印必究 · 印装有误 负责调换)

自序

踏上北京地铁，十有七八的乘客在用手机看视频。如果哪一天忘记带手机，或是前一天晚上忘记下载路上要看的视频，那一定是一场无聊透顶的旅途。

网络视频在国内出现不足十五年，即使算上它的雏形（如电视台网站），这段历史也才刚刚够二十年。但就是这短短的时间，网络视频已经成为都市生活不可缺少的内容、传媒产业重要的组成部分和互联网商业的流量来源。它何以“少年得志”？不仅是作为竞争对手的电视台好奇，就连网络视频从业者自己也好奇。关于它成长的记载散布于各类网站、报道和研究报告中，官方广播电视年鉴也没有详细的记录。加上互联网从业者流动非常频繁，客观上不利于行业历史资料的积累。因而，了解网络视频行业全貌面临资料少、碎的困难。

但是，当下传媒研究绕不开网络视频，它处于媒体融合的交界地带，是互联网与传统媒体的接触面之一。近年来各类学术刊物上也出现了一些研究网络视频的文章，网络视频的发展现状、发展策略等频频成为学术论坛的重要议题。一项研究领域成型的标志是理论、历史和实务三者皆备。目前关于网络视频的研究，尚停留在实务的阶段。关注网络视频生产过程、传播方式、营销策略和盈利模式的研究文章不少，但由于缺少理

论和历史的根基，总显得隔靴搔痒、流于表面。要么是空话、大话连篇，看起来都对、想起来都不对；要么作者在某一篇文章里信誓旦旦提出若干种策略，但很快在另一篇文章中又改变主意拿出新方案。惭愧的是，这种文章笔者也写过。因此才决心痛定思痛，尝试从理论或者历史着手，对网络视频进行一番系统研究。

在笔者看来，理论绝不是空中楼阁，要么是遵循演绎逻辑根据其他理论推导而来，要么是按照归纳逻辑由历史的经验提炼而出。网络视频研究可借鉴的现有理论非常少，自然也谈不上逻辑推导。因此，只能先从历史的经验和研究着手。

严格地说，网络视频史研究属于媒介史研究的子领域。目前，报刊史、图书史、广播史、电视史、互联网史等领域已经有比较多的著作面世，网络视频史的研究尚属空白。个中原因，也许是因为网络视频处于电视与互联网的交叉领域，需要复杂的知识结构应对两方面的史料；也许是因为网络视频始终没有获得官方的认可，就连行业年鉴也提之甚少，收集足够的史料困难较多。当然，也可能是网络视频的历史不长，一切尚处于当下，即所谓当代不修史，活人不立传。

其实，研究当下既是一种挑战，也是新闻传播学的魅力。历史奔流如江海，谁又能预言哪些信息会沉淀，哪些信息会被抛弃？如果将来发现当时的抛弃和忽视是天大的错误，追悔莫及又有何用？与其将来后悔，不如现在动手记录和研究当下。因此，笔者选择目前还是新闻传播学研究空白的中国网络视频发展史作为研究对象，力图把“流动”的信息用学术研究的方式固定在字里行间，以便为后人研究留下一点基础性的资料。

北京大学视听传播研究中心的研究团队收集整理了现有的文件资料，访谈了部分网络视频从业者和网络视频行政主管部门工作人员，用一手资料补充了一些残缺的记载，采用交叉印证的方法纠正了一些习以为常的谬误记载，基本理清了 1996 年至 2016 年中国网络视频产业发展的线索。关于 1996 年之前的网络视频“史前史”部分，研究团队也收集了一些资料，因为尚嫌零

碎，暂未纳入书中。

此书虽为国内第一本中国网络视频发展史研究专著，且我们也确实为此书面世熬过了无数的不眠之夜，但我们的自我定位也就是抛砖引玉。希望此书能够帮助辛苦而伟大的中国网络视频从业者们理清所走过的路径和得失，为中国网络视频产业的发展和研究略尽点滴心力。

是为序。

陆地 靳戈

2017年6月10日

第一章 雏形：2006年之前 / 1

- 一、网络视频的技术基础 / 1
- 二、中国互联网早期情况 / 9
- 三、视频网站的初级形态 / 17

第二章 萌发：2006年至2008年 / 26

- 一、出现探路者 / 27
- 二、探索经营模式 / 37
- 三、两种发展思路的分歧 / 40
- 四、尚不成熟的产业生态 / 42

第三章 竞争：2009年至2013年 / 47

- 一、“国家队”来了 / 47
- 二、资本市场的青睐 / 58
- 三、版权争夺战 / 64

第四章 进化：2014年至2015年 / 96

- 一、从购买到自制 / 99

- 二、从 UGC 到 PGC / 108
- 三、从互联网到移动互联网 / 111
- 四、从免费到付费 / 122
- 五、网络视频超越电视？ / 128

第五章 未来：2016 年及以后 / 137

- 一、2016 年的遗产与遗憾 / 137
- 二、中国网络视频的未来 / 144

附录 学术团队关于网络视频的研究论文 / 152

后记 / 269

第一章 雏形：2006 年之前

一、网络视频的技术基础

(一) 多媒体技术：从数字到声画

说起互联网的功能，但凡有些互联网使用经验的人大概都会说互联网的主要功能是数据传输。的确，互联网极大地提高了信息传播的效率。书信、电话等传统沟通方式遇上了互联网，在速度上相形见绌是毫无疑问的。在 20 世纪 60 年代，加拿大一位语言学教授麦克卢汉通过考察人类使用媒介的历史，提出了“地球村”的预言：电子信息瞬息万里，使全球生活同步化，全球经济趋同、整合，游戏规则走向同一，网络生活同一，时空差别不复存在，昔日遥不可及的海角天涯刹那可达——谁不说这就是弹丸之地？^①

然而，麦氏所处的传媒生态，仍属于今人所说的“传统媒体”时代，信息传输的速度虽已超越了第一次工业革命，但还不足以支撑麦氏的判断。那个时代传媒生态并不足以支撑麦氏的判断。人们很长时间一直对麦氏这一论断半信半疑，直到互联网带来的信息传输速度革命，为这一预言提供了实现的条件。借助即时通讯软件，“望穿秋水”等待远方来信的场景只存在于诗歌的意象中；电子商务的普及，加快了商品的跨地域甚至跨国流动，在网站上轻轻一点，远方的商品就会很快寄到用户的面前——这已经越来越接近于村口小卖部的体验。

然而，如果把视线放得更远一些，超越互联网商业化的历史，触及互联网的起源，就会发现互联网最初的设计动力并非为了加快信息传输速度，而

^① 马歇尔·麦克卢汉著，何道宽译：《理解媒介：论人的延伸》，译林出版社 2011 年版，第 11 页。

是对超级^①计算能力的追求。全球公认美国国防部组建的阿帕网^②(Advanced Research Projects Agency Network, ARPANET)是现代互联网的雏形，而阿帕网实际上是一个提升军方武器设计效率的高性能计算平台。后来，阿帕网的开发者将这一平台向高校和科研单位开放，并最终向企业敞开大门，使商业力量进入互联网领域，互联网“逃离”了单一色调的计算功能，“下凡”到纷繁多彩的世俗世界。

互联网为追求超级计算能力而生，但它借助商业的力量，成为全球通用又惠及全球的信息传播技术。计算与商业的结合，是互联网时代大部分应用形态的中心模式。这对于深入理解之后一系列的互联网服务，如即时通讯、电子商务、网络视频，至关重要。

互联网在商业化伊始即被用于信息的传输。计算机在诞生的早期时候，只能处理由0和1组成的二进制机器语言。在20世纪50年代末60年代初，美国——计算机技术的起源地——制定了二进制数码与英语字母表对应转换的ASCII码^③系统。在ASCII码中，用7位或8位二进制数的组合来表示英语中的字符(含有标点符号和特殊符号)，并使用不同的数字组合区分大小写。诞生于1968年的阿帕网，自然直接使用上了ASCII码，可以直接传输英语文本。后来，计算机之所以能够传输包括汉字在内的多种语言，也是通过ASCII码与0和1的机器语言进行转换的。

在20世纪90年代之前，大部分的计算机还不能显示彩色的图片，无法播放复杂的声音，更不用说连续的动态影像。受制于计算机的处理能力，互联网只能传输文本信息。这是因为彩色图片、复杂声音和动态图像在由计算机处理的过程中必须通过数码来表达，但相对于纯粹的文本，彩色图片、复

^① “超级”并非严肃的学术表达，但的确是计算机领域的常用词汇，如超级计算机、超级互联网等。笔者在此想避开这一词汇，但思量再三，没有找到更合适的替代。

^② 阿帕网：阿帕网(Advanced Research Projects Agency Network, ARPANET)，美国国防部高级研究计划局组建的计算机网，又称ARPA网。现在的Internet是在APRA的基础上才建立起来的。阿帕网于1968年开始组建，1969年第一期工程投入使用。开始时只有4个节点。

^③ ASCII(American Standard Code for Information Interchange, 美国标准信息交换代码)是基于拉丁字母的一套电脑编码系统，主要用于显示现代英语和其他西欧语言。

杂声音和动态图像要使用更多的数码——这需要更大存储空间和更强计算能力的支撑。然而，当时大部分家用计算机的处理能力和互联网的传输带宽都达不到这一要求。

20世纪90年代初，随着计算机性能的不断提升，家用计算机也可以生成彩色图片、复杂声音和动态图像。但此时距离计算机播放视频还差着一步——多媒体。多媒体不是一项单一的技术，而是一系列涉及图形和声音的技术与协议的集合，比如视频格式标准和数据压缩技术。没有声音的配合，动态图像不能称之为视频。没有配套的压缩技术，直接从自然界采样的图片文件和声音文件存在较大的冗余信息，体积比较大，传输效率较低，不利于在民用领域的推广。

有人会问，计算机的显示器看起来和电视机差不多，为什么计算机多媒体技术会比电视机“彩色”技术出现得晚许多？需要指出的是，计算机作为一种集信号制作、传输和接收于一体的设备，与电视机这种单纯的信号接收设备在技术方案上差异很大。电视机不需要考虑编码和发射的问题，只是作为一个信息的接收终端。计算机作为一个综合体，既要解决编码问题，又要兼顾传输和接收的需要，这就要求计算机必须有强大的多任务处理能力，还要把不同编码格式的文件（如文字、图像和声音）按照统一的格式进行处理。更重要的是，人与电视机的互动形式非常有限，屈指可数；而计算机与人的互动形式极其丰富，有无限种可能。因此，在计算机上实现多媒体功能，在20世纪90年代之前一直未能有突破。

有需求就有研发的动力。一系列的视频协议、音频协议和压缩协议在20世纪90年代被提出，声画协同技术和文件压缩技术得以低成本地实现，并在民用领域快速普及。家用计算机也能进行音视频采集和输出工作。计算机由此进入多媒体时代。

多媒体技术是计算机实现娱乐功能的基础性条件。在多媒体技术出现之前，缺少能够把图像和声音进行统一编码的技术和标准。多媒体技术使文字、声音、图像可以同时糅合在一种文件格式下，三者通过一定的逻辑组合，转

变成具有意义的动态图像——这就是视频的雏形。多媒体技术极大地拓展了计算机所能处理的文件类型，使计算机与人的交互方式在文字、图像、声音之外又增加了视频。这一新形式远比文字更具体、比声音和图像更生动，更“贴合”人的自然交互习惯。

视频的出现，使计算机与人的交互更加友善，用户可以用更加“自然”的方式向计算机发布命令和接收计算机的反馈。但是，直到 20 世纪末人们仍无法在互联网上流畅地点播视频。因为，多媒体技术仅解决了将声音和图像“生成”视频这一问题。如何将视频搬到互联网上并提供流畅的观看体验，还需要合适的传输技术。这是下一节的主题。

(二) 流媒体技术：从等待到即时

读罢上一节，读者可能会问，计算机不是已经可以制作视频、实现了多媒体功能吗？借助当时的互联网技术将视频传输到世界各地，不就是如今的网络视频吗？

实现在线观看视频的功能并非像想象中的那样简单——仅仅是把视频上传到互联网上。若是这般，网络视频就与“视频下载”没有区别。在 20 世纪 90 年代末和 21 世纪初，国内许多地区尚未普及宽带互联网，一些用户仍然在使用带宽峰值只有 56k 的电话拨号方式上网。在这种硬件条件下，即使使用多线程^① 下载软件（如网络蚂蚁、网际快车等），下载一部电影仍需要数十个小时甚至几天的时间。

21 世纪初，ADSL^② 开始在全国普及，居民上网速度有了明显提升。由于不同环境下 ADSL 的传输差异较大，因此很难说 ADSL 比普通的电话拨号具体能快多少，但可以确定的是 ADSL 的传输速度已经和电话拨号上网不在

① 多线程下载：线程可以理解为下载的通道，一个线程就是一个文件的下载通道，多线程也就是同时开启好几个下载通道。当服务器提供下载服务时，使用下载者是共享带宽的，在优先级相同的情况下，总服务器会对总下载线程进行平均分配。线程越多，下载就越快。

② ADSL 技术采用频分复用技术把普通的电话线分成了电话、上行和下行三个相对独立的信道，从而避免了相互之间的干扰。用户可以边打电话边上网，不用担心上网速率和通话质量下降的情况。理论上，ADSL 可在 5 km 的范围内，在一对铜缆双绞线上提供最高 1 Mbps 的上行速率和最高 8Mbps 的下行速率，能同时提供话音和数据业务。

一个数量级。然而，即使 ADSL 和更快的光纤宽带^① 普及之后，下载一部电影的时间也至少要以分钟计算。不过，本书这里要讨论的视频下载，重点并不是它的下载时间，而是这种模式与用户的互动机制——用户下载视频必须要付出时间成本（下载时间）和空间成本（磁盘空间）。这些成本，尤其是时间成本，使用户与视频之间的互动并不是直接的——就像你去百货商场买了一件东西，却被告知只能带回家后才能拆开包装看。不妨把网络视频看作是淋浴，把视频下载看作是浴缸——相比较来说，淋浴比浴缸省空间，也省去了预先灌满水的时间。

网络视频作为一种新型互联网应用，其革命性的意义在于能够提供实时的观看体验，不需要用户提前支付时间成本，对空间成本的占用也比较少。达到这种实时观看体验的基本思路是内容随下（载）随播（放）、随播（放）随删（除）。从内容与介质（计算机硬盘）的物理关系上看，它是一种临时的存在——所占用的时间与空间都是转瞬即逝的；从这种关系的社会属性上看，它又是永久的——用户可以通过任何一台接入互联网的设备观看网络视频。临时存在，却时刻触手可及；触手可及，却不属于你（网民）。这种极有张力的关系，是网络视频区别于视频下载的关键特征。

实现这种具有张力的关系，需要与之配套的网络传输技术 / 协议的支持。这种传输技术 / 协议必须要实现以下功能：将完整的视频文件在发送端拆解成若干个小文件，传输端要保证这些拆解后的小文件按顺序准确无误地发送到客户端，客户端把这些破碎的小文件按时间顺序无缝播放。在这种一边拆、一边传、一边播的机制中，上文提到的临时又永久的关系就建立起来——用户可以实时在线观看网络视频。这种技术具有若干种实现路径，被统称为“流媒体”（Stream Media）技术。流媒体，恰如其名地概括了上文提到

^① 光纤宽带：光纤宽带就是把要传送的数据由电信号转换为光信号进行通讯。在光纤的两端分别都装有“光猫”进行信号转换。光纤是宽带网络中多种传输媒介中最理想的一种，它的特点是传输容量大，传输质量好，损耗小，中继距离长等。光纤传输使用的是波分复用，即是把小区里的多个用户的数据利用 PON 技术汇接成为高速信号，然后调制到不同波长的光信号在一根光纤里传输。光纤宽带和 ADSL 接入方式的区别就是：ADSL 是电信号传播，光纤宽带是光信号传播。

的一边拆、一边传、一边播的模式——如同溪流一般，低头看到的流水是小溪的一部分，抬头看到的却是整个蜿蜒的溪流。好似永远流动，却又总是一体。由于这种多媒体文件能够在网络上像水一样不断流动，用户可以在接收新文件的同时观看已经收到的部分，而不需要等待整个文件全部下载完成。因此，流媒体传输具有“缩短等待时间、节省存储空间和能够实时播放”等特点。

最早将流媒体技术商业化的是美国 RealNetworks 公司，大名鼎鼎的流媒体播放软件 Realplayer 就是这家公司的产品。Realplayer 的技术方案即公司名——RealNetworks。之后，美国微软公司以 Windows Media 方案、苹果公司以 Quick Time 方案也加入了流媒体阵营。这是目前流媒体技术的三大主流方案。在这三大主流方案中，还可以细分为两种传输方式：顺序流 (Progressive Streaming) 和实时流 (Real Time Streaming)。所谓顺序流，是指在传输端被拆分的若干小文件必须按照时间顺序进行传输，客户端也要按照时间顺序进行接收和播出。用户在客户端不能进行除了开始和终止以外的其他操作，只能按照线性播出顺序观看视频，与电视节目的播放方式十分类似。这种传输方式禁止来自用户的快进、快退指令，降低了信息传输的压力，适合对传输时效要求较高的网络直播。实时流默认采用的是顺序流的传输逻辑，但允许用户的前进和后退的操作。发送端在接到用户的前进或后退的指令后，会实时调取指定的内容传输到客户端。这种模式对传输带宽的要求较高，会产生一定的延迟，但允许用户快进、快退等简单命令，适合网络视频点播。

流媒体技术可以在网络带宽提升有限的情况下，通过优化传输算法，实现高码率视频的实时传送。多媒体文件这艘巨轮，借流媒体技术对内容进行打碎和重组，化整为零在互联网弯弯曲曲的小溪中畅行无阻。

(三) 网络宽带技术：从小溪到江海

尽管流媒体技术可以在带宽有限的条件下提升网络视频的画质和观看流畅度，但实际上提升带宽对提升用户观看体验更加直接有效。如果把流媒

体技术带来的观看效果提升比作“盖上锅盖”，那么提升带宽就好比“釜底添薪”。带宽，通俗意义上讲是指网络传输信道的容量，一般以比特率（bps）作为计量单位，指的是每秒传输的数据大小。自“计算机之父”冯·诺依曼于20世纪30年代提出在计算机设计中使用二进制替代十进制以来，虽然计算机的硬件和软件更迭了多代，但如今的计算机依然是冯·诺依曼理论的产物，仍然采用二进制。二进制，是通过0和1表示信息的方法。在电路设计中，关闭可以代表0，开启可以代表1，因此二进制是电路设计中最容易实现的计数方法。0和1通过ASCII码表转换为字母，进而再通过字母表达自然语言。相应地，也存在一套转换机制将0和1转化为汉字、图像、声音和视频。一般来说，记录一个汉字所需的磁盘空间要比记录一个字母大（一个英文字母占一个字节的空间，一个中文汉字占两个字节的空间），声音、图像和视频则需要更多的字节来表示。因此，互联网在传输声音、图像和视频时，要比传输单纯的文本占用更多的带宽。这也就是为什么文字传输最先在互联网上实现，而实时点播的网络视频要在互联网出现近五十年才开始普及。

相较于传输电子邮件等简单的文本或图片信息，互联网在传输视频时需要占用更多的带宽。在2004年，笔者第一次使用56k的电话拨号上网，使用多线程下载软件网络蚂蚁获取一首MP3格式的音乐需要近十分钟。在2007年，笔者第一次使用ADSL时，下载一首同大小的MP3音乐只需要不到两分钟。今天，笔者已经使用上了光纤专线，现在下载一首MP3音乐基本可以瞬时完成。足够的网络带宽能够明显提升数据传输的速度，接收数据的体验也会更好。而且，数据量越大，带宽增加带来的体验提升就越明显。

任何一门产业的发展都需要基础设施的配套支持，网络视频也是一样——网络带宽就是重要的基础条件。根据中国互联网络信息中心（CNNIC）的数据，自1997年以来，我国国际出口带宽每年均有较大的增幅（如图1-1所示）。

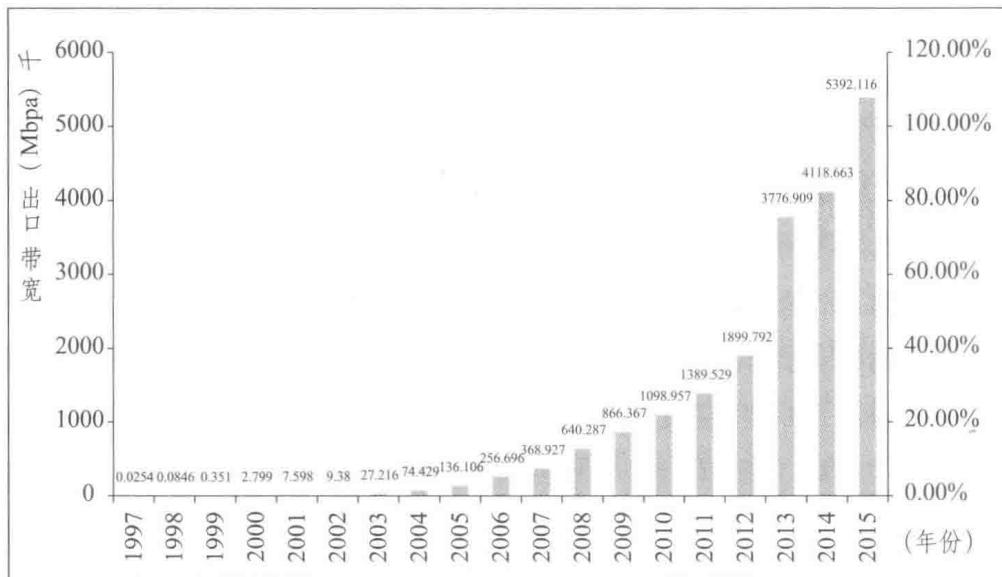


图 1-1 1997 年至 2015 年中国国际出口带宽变化情况

在 2005 年之前，中国政府推出了一系列加速互联网带宽提升、带动互联网产业发展的政策。在 1997 年的全国信息化工作会议上，互联网被列入国家信息基础设施建设目录。1998 年，中国公用计算机互联网二期工程启动，将主干网带宽扩充至 155 兆比特每秒。2000 年，北京国家级互联网交换中心开通，使中国主要互联网网间互通带宽由原来的每秒不足 10 兆比特提高到每秒 100 兆比特。2002 年，作为国家“十五”规划的信息化子规划，《国民经济和社会发展第十个五年计划信息化重点专项规划》明确提出大力发展高速信息网，提高信息网络传输能力，满足社会日益增长的带宽需求，大力发展战略性新兴产业。在一系列政策的支持下，中国互联网的带宽发展得到了快速提升（见图 1-1），虽然与发达国家相比仍有差距，但进步依然可观。

本节分析了网络视频传输的三大基础性技术：多媒体、流媒体和宽带。这是网络视频得以出现的基本技术条件。下一章节将分析网络视频在中国诞生的社会性条件，其中最主要的因素是互联网在全社会的普及——尤其是在人们意识中的启蒙。

二、中国互联网早期情况

互联网在中国的发展道路与美国很相似——首先服务于科研领域，之后向商用领域开放，最终实现全面普及。20世纪80年代，中国科学院高能物理研究所参与了欧洲核子中心高能电子对撞机计划国际合作项目，该研究所的工作人员于1986年8月25日从北京通过国际互联网向瑞士日内瓦欧洲核子中心的科学家斯坦伯格发了一封“跨越长城，走向世界”^①的电子邮件。这是中国互联网史上第一封发自我国境内的国际电子邮件。

从1989年开始，国家着手建设四张重点骨干互联网，包括中关村教育与科研示范网络（中国科技网的前身）、中国公用计算机互联网（China Net）、中国教育和科研计算机网（CERNET）、中国金桥信息网（China GBN）。其中，中国科技网（CSTNET）先后经历了“中关村教育与科研示范网络”、“中国科学院院网”和“中国科技网”三个阶段，是最早与美国国家科学基金网（NSFNET）直接互联的中国骨干网，将中国互联网带入国际互联网的“大世界”。

中国公用计算机互联网是由中国邮电电信总局^②负责建设的公共计算机互联网，承担面向全国提供普遍服务的义务。该网络与其他三个骨干网相连，与国际互联网相连，主要向国内民用市场提供互联网服务。

中国教育和科研计算机网是由国家投资建设、教育部负责管理，华北、西北、华南、西南、华东北、华东南、华中、东北八个地区科研实力较强的高等学校^③承担建设和管理运行的全国性学术计算机互联网络。该网络分四级管理，分别是全国网络中心、地区网络中心和地区主结点、省教育科研网、大学校园网。全国网络中心设在清华大学，负责全国主干网运行管理，且设有专门的国际出口。

^① 原文为：“Across the Great Wall we can reach every corner in the world”。

^② 中国邮电电信总局前身是1994年从国家邮电部分离出来的企业局——“电信总局”，后以“中国邮电电信总局”的名义进行企业法人登记。

^③ 八所院校分别为：清华大学、西安交通大学、华南理工大学、电子科技大学、东南大学、上海交通大学、华中科技大学、东北大学。