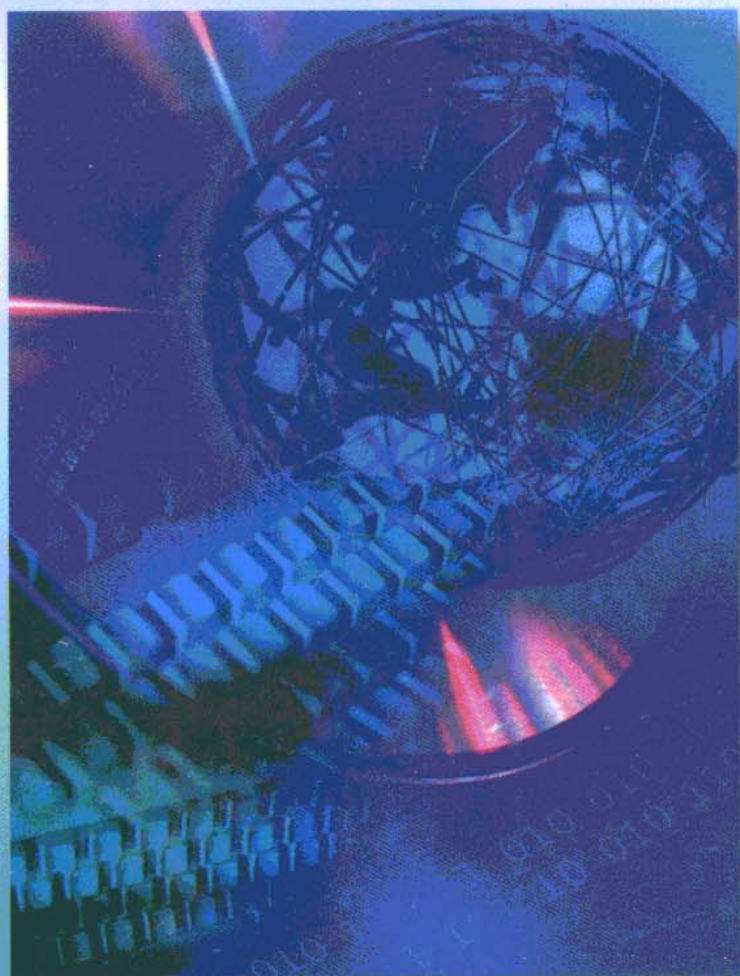


Internet

网络应用教程

裴纯礼 编著



北京师范大学出版社

Internet

网 络 应 用 教 程

裴纯礼 编著

北 京 师 范 大 学 出 版 社

图书在版编目(CIP)数据

Internet 网络应用教程/裴纯礼编著. - 北京:北京师范大学出版社, 1999, 8
ISBN 7-303-05168-6

I . I… II . 裴… III . 因特网 - 教材 IV . TP393.4

中国版本图书馆 CIP 数据核字(1999)第 35272 号

北京师范大学出版社出版发行
(北京新街口外大街 19 号 邮政编码:100875)

出版人:常汝吉

北京师范大学印刷厂印刷 全国新华书店经销
开本:787mm×1 092mm 1/16 印张:23 字数:592 千字
1999 年 8 月第 1 版 1999 年 8 月第 1 次印刷
印数:1~3 000 定价:28.00 元

前 言

一、你有 Internet 的应用意识吗?

在人类历史发展的近 300 年中,伴随着一系列新的技术而发生了一系列革命——18 世纪发明了蒸汽机;19 世纪产生了工业革命;20 世纪最关键的技术涉及信息的收集、处理和传播,其中最突出的成就包括:世界范围内的电话网,无线电、电视的发明与应用,计算机工业的诞生与迅猛发展,通信卫星的发射,Internet 的建立,并由之而引发的信息技术革命。

与汽车、飞机等传统工业相比,计算机工业是最年轻的,但是它的发展速度则是其它任何工业所无法比拟的。在计算机诞生后的前 20 年(我国则要延后 10 年左右),计算机系统是集使用 and 管理的:一个中等规模的大学或研究机构通常拥有一两个集中放于一间或若干房间的大、中型计算机系统。随后的 30 年,计算机和通信技术的结合逐渐将单一的计算机系统用大量分散但相互连接的计算机网络代替。计算机网络中计算机相互连接的主要意义在于相互交换信息,并逐渐消灭在收集、存储、处理和传输信息之间的差别。

本世纪 80 年代,是个人计算机发展的关键 10 年,微机进入家庭,其应用由专业人员迅速地扩展到家庭中的每个成员,于是很自然地产生了“计算机文化”。人们开始逐渐建立起计算机应用意识:就像我们出外旅行很自然要选乘汽车、火车或飞机(这说明我们已经具有了现代化交通工具的使用意识),几乎没有人会选择步行一样,具有了计算机应用意识的非计算机专业的教师们开始用计算机进行科研和辅助教学,作家们开始用计算机代笔进行无纸创作,教学管理人员开始用计算机管理教学,会计人员开始用计算机管理财务,服装设计师开始用计算机设计服装,动画设计师开始用计算机设计动画,孩子们则很容易迷恋上计算机游戏……。

而最早产生于 60 年代末,并最终将全世界各国的计算机主干网连接到一起的 Internet 则把信息技术提高到一个新的高度,向我们展示出更加美妙的未来。

一台孤立的计算机,就像一辆被困在一个大海孤岛中的小卧车。虽然这辆车上具有最先进的照明、空调、冰箱、高级音响、无线电通信设备,甚至配备了舒适的卧床,这些豪华的设施可能是每一个想要购买小卧车的人们的梦想与追求。但是如果没有公路,再高级的小卧车也会寸步难行,失去它的魅力。汽车,作为人类的一种代步工具,它的发明产生了对公路的需求,而有了公路,尤其是有了高速公路才真正使汽车有了用武之地,反过来又促进了汽车工业的发展。高速公路的普及给人们带来了自己开车旅游、购物、探亲访友等各种各样的方便,极大地促进了商品和物资的交流,促进了整个社会的发展。

在信息社会,计算机则是人类的另一种代步工具,它也需要自己特殊的“信息公路”——计算机网络。没有网络,目前再高级的计算机能够提供给用户的信息都是有限的,并且很快就会陈旧过时。计算机正是由于有了网络,尤其是 Internet 这一信息高速公路,才真正显示出其巨大的潜力。Internet 是一个“计算机网络的网络”,它已经并将进一步把本世纪最伟大的通信技术与多媒体技术综合到一起。Internet 的发展与普及打破了国界与距离的限制,人们坐在计算机前就能“坐地日行八万里,巡天遥看一千河”,进行虚拟的旅游和参观访问,阅读最新的电子报纸和杂志,查阅与获得最新的软件和各种信息资料,与同行交流,或获得全世界任何一个领域里最优秀专家的帮助,拜访远方的亲人,休闲、娱乐、购物……;而 Internet 的发展则进一步促进了计算机的发展。

目前英特尔(Intel)公司和惠普(HP)公司正合作致力于本世纪结束前推出 64 位的 Merced 芯片。从理论上讲,它可以装备 64 条地址线,其寻址能力可达 $2^{64} = 18.4\text{GG}$ (注:在计算机中 1G 约为 10 亿个字节)。如果能够配以相应的操作系统和内存(18.4GG 字节),以每个中文字符占两个字节计,则该内存容量可以存储约 920 亿亿个中文字符,相当于 3 亿多个图书馆(每个图书馆藏有 3 亿本 100 万字的中文书籍)的全部馆藏总和。就是这样一台具有不可思议的存储功能的计算机,如果不能及时地从 Internet 上获取和交流信息,在信息已成指数发展的今天,其中许多信息也会很快变成历史。

一台计算机的力量与希腊神话里的大力士安泰的力量颇为相似:安泰是无敌的,但他随时需要及时地往地上一躺,以从他的地神母亲盖娅那儿重新获取力量,否则他的力量就会很快耗尽。安泰的敌人正是利用他的这一弱点,而将它高举在空中,在他力量耗尽之时将它活活勒死。Internet 正是每一台计算机的“地神母亲”。

汽车因为高速公路的发展已向人们展示其真正的魅力和价值,计算机也因为信息高速公路的发展而向人们发出难于抵御的吸引与诱惑。今天,如果我们拥有了一台与 Internet 相联的计算机,就可以畅通无阻地到达世界各地。或者说,整个世界就被浓缩在计算机的方寸之中。

90 年代初,当微机在我国已开始普及的时候,我在一本书的前言和若干文章中曾多次向广大读者阐述过在信息时代具有计算机应用意识的必要性;那么,在人类即将迈入 21 世纪的今天,随着我国与 Internet 相联的四大国家主干网不断地“延伸与拓宽路面”,我要向大家提出一个新的问题:你有 Internet 的应用意识吗?你的计算机已经入网“上路”了吗?你会从 Internet 上获取信息和各种服务吗?你会在 Internet 上发布自己的信息吗?

Internet 是一个新鲜的事物,我们可以从它获取得许多东西;但是如果我们只有一台没有与之相联的计算机,就很难对它有深刻的认识。

截至 1998 年 12 月 31 日,我国的上网用户已达 210 万。如果说过去我们因没有“信息公路”而无可奈何的话,那么在今天已经有了“信息公路”,并且这条“公路”还在不断发展的情况下,如果你还没有让你的计算机入网“上路”的话,请不要再等待了!赶快让你的计算机入网“上路”吧,让它不断地从它的“地神母亲”——Internet 上获得新的力量,显示其英雄本色!

二、Internet 的工作特点

谈到 Internet,我们有必要强调它与一个分布式计算机系统之间的异同。Internet 和分布式计算机系统中计算机相互之间均需要联接,就连接的硬件而言两者没有什么差别,它们之间的主要差别在于各自的软件系统。

分布式系统是一个基于计算机网络系统的软件系统,它对整个网络系统自动进行管理,网络中的每台具有自主权的计算机对用户来说都是透明的。当一个用户在该网络的某台计算机上发出命令时,分布式系统将自动为该用户从网络中选择最好的处理器,找到并将用户所需的文件传输给该处理器,最后将结果放到适当的地方。

在 Internet 的应用中,用户通常使用一本地计算机,并通过本地主机上的客户应用程序明确地对 Internet 某个服务器进行访问,即在所访问的网络服务器服务程序与本地主机客户程序之间建立“虚联接”,直到完成文件或信息的相互传输后再断开“虚联接”。

读者了解 Internet 的上述工作特点,对学习 Internet 的应用服务是有帮助的。它向读者表明在学习过程中应特别注意下列两个方面的问题:如何正确表示使用不同协议的服务器的地址,以及如何在访问不同协议的服务时正确选用相应的客户程序。

三、Internet 有关知识的内容组成

我们学习 Internet 及其应用,主要应学习以下六个方面的知识。

①基础知识,包括数据通信、计算机网络和 Internet 的基本知识;

②Internet 提供的各种信息应用服务分类、特点、作用、区别与联系;

③正确获取 Internet 的各种信息服务的方法;

④在 Internet 上发布各类信息(包括各类信息组电子稿件、用 HTML 语言编写的 Web 主页等)的方法;

⑤高级 Web 主页(包括 Java Script 与 Java Applet 程序等)的编写方法;

⑥基于 Sockets API 应用编程界面的 Internet 应用程序的开发。

对于大多数非计算机专业人员来说,只要学习与掌握前四个方面的知识就够用了。

四、编写本教程的背景与目的

综上所述,对于国内大多数非计算机专业学生或读者来讲,迫切需要一本介绍 Internet 及其应用的教材,它应该既有助于学生对数据通信技术、计算机网络和 Internet 的基本知识有一个全面而基本的了解,又能帮助学生全面正确地了解 Internet 上提供的各类服务,并让学生能够熟练掌握获取网上信息的方法,以及初步具有在 Internet 上发布自己信息的能力。

由于 Internet 应用技术本身的迅速发展,许多陈旧的服务已经或正在被淘汰,而保留下来的服务则在不断改进,新的服务还在出现。这些都给编写 Internet 应用教材带来一定的困难。实际上,如果某一本教材是在 1997 年以前出版的,它其中的某些内容可能已经应归类到介绍 Internet 的历史了。

国内现有的一些介绍 Internet 应用的书籍或教材缺乏对 Internet 信息服务的全面介绍和具体的上机应用练习。此外,由于受多种因素的制约,国内广大学生与读者中的绝大多数人很难免费获得服务器帐号与磁盘目录空间,因此也很难有机会全面实践和应用 Internet 的各类服务,结果当然是很难对 Internet 及其应用有一个全面深刻的认识。

作者于 1996 年 5 月~1997 年 11 月期间以高级访问学者身份在美国俄亥俄州立大学进修和从事 Windows Sockets 在 Internet 上的应用开发等,在科研、教学和日常生活中积累了大量应用 Internet 的实践经验;回国后即为首都师范大学非计算机专业本科生和研究生开设了“Internet 网络及其应用”的选课,并利用暑期为新疆地区的教师开办系列讲座。本教程则是结合几轮教学实践经验,并在 1998 年 4 月(初版)和 1998 年 9 月(修改版)讲义的基础上于 1999 年 2 月利用整个寒假的时间认真修改而成。作者把这本教程奉献给自己的学生和广大读者,目的是希望大家都能学习和真正掌握 Internet 的应用,并牢固建立起 Internet 的应用意识,以便适应 21 世纪信息社会的需求。

1999 年初,作者在北师大校园网上建立了教学服务器,并开始通过它进行 Internet 辅助本教程相应课程教学与教学管理的试验。作者向每个注册参加本课程学习的学生都提供该服务器的帐号、磁盘目录空间,以及包括电子信箱、远程登录、FTP 文件服务、电子函递名单、BBS 电子布告板系统、WWW 主页发布等多项服务,其目的是通过言传身教,给未来 21 世纪信息社会的新一代教师们作出全面利用 Internet 服务来辅助课程教学与教学管理的实践示范。“星星之火,可以燎原”,能作为这样的一颗火种,我深深感受到了一名大学教师肩负的重任,并为之自豪和激动不已!

五、本教程及其使用说明

本教程共分十一章:前两章介绍数据通信和计算机网络的基本知识,第三、四章介绍 Internet 的基本知识,第五章介绍 Web 浏览器的基本知识,第六、七、八、九章分类介绍 Internet 网上各种信息服务,第十章介绍 HTML 超文本语言及简单主页的编制和发布,第十一章介绍主页编辑器的使用。

其中第一、二章是为了帮助非计算机专业的学生或读者能比较全面地了解数据通信和计算机网络,以及与 Internet 的关系。教师可根据学生的实际情况选用或跳过不讲,如选用,则授课前应要求学生自学,一般读者则可根据自己的实际情况选学或跳过不看,这部分内容只需初步了解即可。

本教程重点放在 Internet 的具体应用服务上。根据现代 Web 浏览器所具有的特点,全书以第五章介绍的典型 Web 浏览器作为主要工具应用于 Internet 的各类信息服务中,一般读者只要掌握某典型浏览器的应用就可以了。书中从第六章到第九章还介绍了许多专用于某类服务的客户程序,旨在帮助读者熟练地使用 Internet 大多数应用程序。例如,我们将介绍如何找到所需的免费软件,以及如何找到所需的站点及其上面的有关信息。

本教程在第十章详细介绍了 HTML 语言,以及如何使用它来编制 Web 主页的方法,以便在网上发布自己的信息;第十一章在此基础上采用与第十章类似的顺序与结构介绍了主页编辑器的使用。初学者可在学习 HTML 语言的同时,同时学用主页编辑器,并选择某个实例边练边学,有条件的学生或读者还可根据自己的需要进行 Web 主页的编写与发布自己信息的尝试。

限于篇幅,本教程没有介绍 Java Script 与 Java Applet 程序设计语言,以及 Socket API 应用编程界面。有进一步学习要求的读者可参阅其它有关的书籍或 Web 站点。

Internet 的发源地在美国,因此本教程根据目前 Internet 的实际情况,以介绍 Windows 9x 视窗系统的工作环境为主,同时兼顾简单介绍 UNIX 和 Linux 操作系统的工作环境。

非计算机专业的学生或读者在开始接触 Internet 时的一个最大障碍是专业术语多,为此本教程对所用到的专业术语,尤其是英文缩写词均给出英文全称和中文名称,并尽可能进行简单而通俗的解释和说明,但初学者不必把主要精力放在记忆这些名词术语上。

我国于 1997 年 7 月 18 日由全国科学技术名词审定委员会确定:互联网(the internet)用以泛指由多个计算机网络联接而成的一个网络;因特网(Internet)专指由美国阿帕网(ARAPNET)发展而成的全球最大的、开放的、由众多网络相互联接而成的计算机网络;此外还统一规范了许多名词术语,如:万维网 WWW、电子函件 E-Mail、FTP 文件传送、Archie 阿奇工具等,本教程将遵守此规定。

本教程每章后面均附有思考题,以帮助学生或读者复习和归纳总结。从第四章起,每章(或每节)后均附有上机练习题。有条件的学生和读者可按教师要求或根据自己的需要,按照上机练习题进行每一种应用实践,并不断总结经验。例如:当学习到电子函件时,最重要的实践就是发出你的第一个电子函件。

作者在本教程相应课程的教学中将尽可能全面应用 Internet 的信息服务来辅助课程教学与教学管理,作者也鼓励每个学生或读者尽可能利用 Internet 来进行学习,以获得最佳学习效果。

本教程适合大学、大专、中专非计算机专业学生。而任何一个读者如果已初步掌握一个简

单的文字处理软件包的使用,他也具有掌握 Internet 应用的能力。

六、鸣谢

感谢美国俄亥俄州立大学电气工程系主任,我的老同学郑元芳教授给我提供的在该系从事 Internet 应用研究的机会,以及从选题和经费等多方面给予的帮助和支持。

感谢美国俄亥俄州立大学电气工程系博士林成江、乔大继、秦林洪等中国留学生们对我留美期间给予的帮助,使我能顺利完成有关课题,并为编写本教程奠定网络理论基础和积累 Internet 应用经验。

感谢北京师范大学教务处领导对建设本教程相应课程的关心、支持和提供的课程启动经费,从而使作者有机会编写讲义,并通过教学实践不断修改而逐步趋于成熟。

感谢北京师范大学计算中心为教师提供备课用计算机,使作者能提前实现建立教学服务器的愿望。

感谢北京师范大学信息学院黄荣怀副院长对作者所建教学服务器所给予的大力支持。

感谢北京师范大学计算中心孙建刚老师对作者建立教学服务器所给予的无私、热心的帮助。可以说,如果没有孙老师的工作,该服务器不可能这么快就向学生提供 Internet 教学服务。

感谢北京师范大学出版社资助本教程的正式出版,感谢责任编辑王安琳副编审在本教程的编辑出版工作中给予的各种帮助。

感谢参加本教程原讲义相应课程学习的北京师范大学 10 多个系数百名研究生和本科生对作者主讲课程所给予的充分肯定与支持;感谢他们从教学内容、教学方式、上机内容、网络辅助教学和管理、发挥学生积极性,以及原讲义内容、编排、错别字等诸多方面提出的许多中肯意见和有益建议,本教程正是根据他们的建议进行了最后的修改和润色。

感谢您选用本教程,衷心希望您在通过本教程学习 Internet 及其应用的过程中,不断利用 Internet 本身学到更多的知识。

由于 Internet 及其有关技术每天都在不断地更新和发展,为了让本教程能尽快到达您的手中,本教程的修改仍十分仓促,来不及更仔细检查和选择更多较好的实例;再加上本人水平所限,书中必定仍存在不少错误与疏漏之处,欢迎您批评指正,在此深表谢意!

裴纯礼

1998 年 4 月第一稿

1998 年 9 月第二稿

1999 年 2 月 25 日定稿

于北京师范大学 计算中心

第一章 数据通信基本知识

数据通信(Data Communication)是计算机网络(Computer Network)和 Internet 的基础,为了帮助非计算机专业的读者对数据通信技术与计算机网络和 Internet 的关系有一个全面的了解,本章将对数据通信知识进行简单介绍。

第一节 传输介质

所有计算机之间的通信都涉及由传输介质传输某种形式的数据编码信号。传输介质在计算机、计算机网络设备间起互连和通信作用,为数据信号提供从一个节点传送到另一个节点的物理通路。

计算机与计算机网络中采用的传输介质可分为有线和无线传输介质两大类。

一、有线传输介质

计算机网络中流行使用的有线传输介质(Wired Transmission Media)为:铜线和玻璃纤维。

1. 铜线

铜线(Copper Wire)由于具有较低的电阻率、价廉和容易安装等优点因而成为最早用于计算机网络中的传输介质,它以介质中传输的电流作为数据信号的载体。为了尽可能减小铜线所传输信号之间的相互干涉(Interference),我们使用两种基本的铜线类型:双绞线和同轴电缆。

(1)双绞线(Twisted Pair)。由两条互相绝缘的铜导线扭绞起来组成,它既可减小流过电流所辐射的能量,也可防止来自其他通信线路上信号的干涉。双绞线分屏蔽和无屏蔽两种,其形状结构如图 1.1 所示。双绞线的线路损耗较大,传输速率低,但价格便宜,容易安装,常用于对通信速率要求不高的网络连接中。

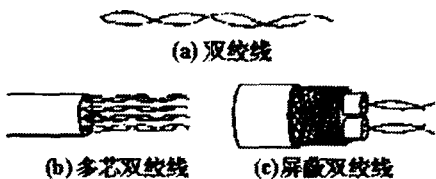


图 1.1 双绞线

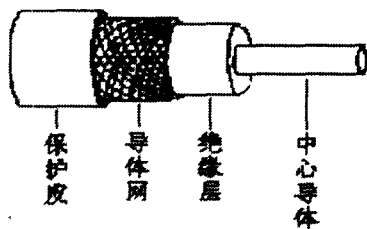


图 1.2 同轴电缆

(2)同轴电缆(Coaxial Cable)。由一对同轴导线组成,其形状结构如图 1.2 所示。同轴电缆频带宽,损耗小,具有比双绞线更强的抗干扰能力和更好的传输性能。按特性阻抗值不同,同轴电缆可分为基带(用于传输单路信号)和宽带(用于同时传输多路信号)两种。同轴电缆是目前 LAN 局域网与有线电视网中普遍采用的比较理想的传输介质。

2. 玻璃纤维

目前,在计算机网络中十分流行使用易弯曲的石英玻璃纤维来作为传输介质,它以介质中传输的光波(光脉冲信号)作为信息载体,因此我们又称之称为光导纤维,简称光纤(Optical Fiber)或光缆(Optical Cable)。

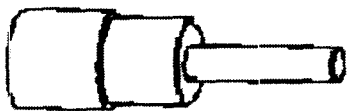


图 1.3 光缆

光缆由能传导光波的石英玻璃纤维(纤芯),外加保护层构成,其形状结构如图 1.3 所示。在光缆一头的发射器使用 LED 光发射二极管(Light Emitting Diode)或激光(Laser)来发射光脉冲,在光缆另一头的接收器使用光敏半导体管探测光脉冲。

与电缆相比,光缆具有体积小、重量轻,频带宽、容量大,传输速率高,不受外界电磁场的影响,抗干扰能力强,安全保密性好,可以单根使用等优点;但它弯成直角时易折断,并难于确定折断位置,安装和连接均需要专用设备。近年来,由于扩展束透镜技术在多节点连接器应用上的突破,再加上诸如光源、检波器和多路转换器等光缆部件的开发和端口化,光缆已成为计算机网络中最理想的传输介质。

光缆可分为单模(只提供单路光通道)和多模(提供多路光通道)两种传输方式。

二、无线传输介质

无线传输是指无须架设或铺埋电缆或光缆,把数据信号转换为电磁波后直接通过自由空间进行传送。例如,无线电波、微波、红外线、激光等数据信号载体本身都可以用作无线传输介质(Wireless Transmission Media)。

1. 无线电波

无线电波(Radio)作为传输介质,既可用于无线电和电视广播,也可用于计算机网络与计算机网络之间数据信号的传输。网络通信设备间通过天线来发送和接收无线电波实现数据传输,我们称之为 RF 射频传输(Radio Frequency Transmission)。

射频对应的无线电波不能随地球表面弯曲传输,但可以通过卫星转播传输。通常一颗卫星装有许多对相互独立的发射器和接收器,每一对发射器和接收器使用各自的频道,因此可以同时实现多路通信。

2. 微波

微波(Microwave)使用高于广播与电视所用的电磁波频率,它也可以用作传输介质。微波通信可用于长途电话,也可以用于数据通信传输。

与无线电波向所有方向发射不同,微波只向某个固定方向传输,并可以携带更多的信息。由于微波不能进入金属结构,因次在微波发射器与接收器之间不应有障碍物,微波也可以使用卫星转播传输。

3. 红外线

红外线(Infrared)使用远高于微波而接近于可见光的频率,常用于小范围(例如在一个房间里)的信号传输,它不需要天线,使用时要求发射器直接对准接收器。它可用于笔记本电脑,例如我们可以在一个房间的计算机网络中使用红外线来让笔记本电脑在房间内移动时均保持与网络连接。

4. 激光

激光(Laser)除了可以在光缆中用作传输数据信号的载体外,它也可以直接作为传输介质在空气中传输数据。与微波传输类似,激光只能向一个固定方向传播,它不能穿过金属、植物,甚至雪和雾,因此直接用激光作为传输介质在实用上受到限制。

第二节 模拟数据通信与数字数据通信

一、通信信道与信道容量

通信信道(Communication Channel)是数据传输的通路,在计算机网络中信道分为物理信道和逻辑信道。物理信道指用于传输数据信号的物理通路,它由传输介质与有关通信设备组成;逻辑信道指在物理信道的基础上,发送与接收数据信号的双方通过中间结点所实现的逻辑“联系”,由此为传输数据信号形成的逻辑通路。逻辑信道可以是有连接的,也可以是无连接的。

物理信道还可根据传输介质的不同而分为有线信道和无线信道,也可按传输数据类型的不同分为数字信道和模拟信道。

信道容量(Channel Capacity)指信道传输信息的最大能力:对于数字信道一般用单位时间可以传输的最大二进制位(比特 bit)数来表示,对于模拟信道则由信道的带宽表示。信道容量的大小还受信道质量和可使用时间的影响,当信道质量较差时,实际传输速率将降低。

二、模拟数据通信和数字数据通信

1. 模拟数据与数字数据

我们一般将数据分为模拟数据和数字数据两大类。

模拟数据(Analog Data)是由传感器采集得到的连续变化的值,例如温度、压力,以及目前在电话、无线电和电视广播中的声音和图像。

数字数据(Digital Data)则是模拟数据经量化后得到的离散的值,例如在计算机中用二进制代码表示的字符、图形、音频与视频数据。目前,美国信息交换标准码 ASCII(American Standard Code for Information Interchange)已为 ISO(国际标准化组织)和 CCITT(国际电报电话咨询委员会)所采纳,成为国际通用的信息交换标准代码,使用 7 位二进制数来表示一个英文字母、数字、标点或控制符号;图形、音频与视频数据则可分别采用多种编码格式。

2. 模拟信号与数字信号。

(1)模拟信号与数字信号。不同的数据必须转换为相应的信号才能进行传输:模拟数据一般采用模拟信号(Analog Signal),例如用一系列连续变化的电磁波(如无线电与电视广播中的电磁波),或电压信号(如电话传输中的音频电压信号)来表示;数字数据则采用数字信号(Digital Signal),例如用一系列断续变化的电压脉冲(如我们可用恒定的正电压表示二进制数 1,用恒定的负电压表示二进制数 0),或光脉冲来表示。

当模拟信号采用连续变化的电磁波来表示时,电磁波本身既是信号载体,同时作为传输介质;而当模拟信号采用连续变化的信号电压来表示时,它一般通过传统的模拟信号传输线路(例如电话网、有线电视网)来传输。

当数字信号采用断续变化的电压或光脉冲来表示时,一般则需要用双绞线、电缆或光纤介质将通信双方连接起来,才能将信号从一个节点传到另一个节点。

(2)模拟信号与数字信号之间的相互转换。模拟信号和数字信号之间可以相互转换:模拟信号一般通过 PCM 脉码调制(Pulse Code Modulation)方法量化为数字信号,即让模拟信号的不同幅度分别对应不同的二进制值,例如采用 8 位编码可将模拟信号量化为 $2^8 = 256$ 个量级,实用中常采取 24 位或 30 位编码;数字信号一般通过对载波进行移相的方法转换为模拟信号。

计算机、计算机局域网与城域网中均使用二进制数字信号,目前在计算机广域网中实际传送的则既有二进制数字信号,也有由数字信号转换而得的模拟信号。但是更具应用发展前景的是数字信号。

3. 模拟数据通信与数字数据通信

(1)模拟数据通信(Analog Data Communication)。指用现有的模拟传输线路来传输模拟数据或数字数据对应的模拟信号。例如目前我们广泛使用公用电话线路来传输语音或计算机数字数据对应的模拟信号,我们也可以使用公共有线电视网来传输视频和计算机数字数据对应的模拟信号。

为了用模拟数据通信的方法实现模拟数据和数字数据的远距离传输,我们一般不直接传输模拟信号和数字信号,而是在发送方使用某一频率的电磁波作为载波(Carrier),然后用模拟信号或数字信号对其进行调制(Modulation),调制后的载波信号(为模拟信号)占有以该载波频率为中心的一段频谱,并能在适于该载波频率的介质上传输;而在接收方则通过解调制(De-modulation)还原叠加于载波上的模拟信号或数字信号。我们将可同时完成调制和解调的装置称为调制解调器(MODEM)。

载波信号可以表示为 $u(t) = A(t)\sin(\omega t + \varphi)$,其中:振幅 A 、角频率 ω 和相位 φ 是载波的调制控制参数,我们可以通过改变这三个参量实现对载波的不同调制。

模拟信号对载波的调制一般采用调幅(AM)或调频(FM)的方法;数字信号对载波的调制可以采用调幅(AM,即分别用幅度为 1 和幅度为 0 的载波信号代表二进制数字 1 和 0)、调频(FM,即分别用两种不同频率的载波信号代表二进制数字 1 和 0)和调相(PM,即分别用两种不同相位的载波信号代表二进制数字 1 和 0,称为两相调制,其调制一次载波信号只能传输 1 位数字信号)的方法。

在数字信号调制中,由于幅度调制抗干扰能力差,而频率调制的频带利用率低,因而很少使用;相位调制占用频带窄,抗干扰性能好,在实际中应用较多。

调相又称为移相(Phase Shift),为了每调制一次载波相位能传输多位数字信号,对于数字信号一般采用多相调制的方法。例如:将数字信号按两个比特一组进行编码(每组编码调制一次载波相位),两位二进制数可以有 $2^2 = 4$ 种不同的数(00,01,10,11),每种数对应一种调制相位,因此称为四相调制,其调制一次载波相位则传输二位数字信号;依次类推,将数字信号按三个比特一组进行编码,三位二进制数可以有 $2^3 = 8$ 种不同的数(000,001,010,011,100,101,110,111),每种数对应一种调制相位,因此称为八相调制,其调制一次载波相位则传输三位数字信号……。

显然,调制本身就可以将数字信号转换为模拟信号。为了使数字数据能在普通的音频电话线上传输,目前微机上使用的调制解调器一般选用音频作为载波频率,它可以实现数字信号与音频模拟信号之间的调制与解调制转换。

(2)数字数据通信(Digital Data Communication)。指直接利用数字传输技术在数字设备之间传输数字数据,或模拟数据对应的数字信号。由于计算机使用二进制数字信号,因而计算机与其外部设备之间,以及计算机局域网、城域网大多直接采用数字数据通信。此外,目前北美采用的 24 路 PCM 脉码调制(速率为 1.544Mbps),以及欧洲和我国采用的 30 路 PCM 脉码调制(速率为 2.048Mbps)电话系统均是数字数据通信系统。

由于数字数据通信传送的是离散的数字信号,即逐位传送二进制数字代码,因此要求系统应能确知传输线上正在传送的数位是 0 还是 1。

(3)数字数据通信的优点。与模拟数据通信相比较,数字数据通信具有下列优点:

①数字技术比模拟技术发展更快,数字设备很容易通过集成电路来实现,并与计算机相结合,而由于超大规模集成电路技术的迅速发展,数字设备的体积与成本的下降速度大大超过模拟设备,性能/价格比高;

②来自声音、视频和其它数据源的各类数据均可统一为数字信号的形式,并通过数字通信系统传输;

③在长距离数字通信中可通过中继器放大和整形来保证数字信号的完整及不累积噪音;

④以数据帧为单位传输数据,并通过检错编码和重发数据帧来发现与纠正通信错误,从而有效保证通信的可靠性;

⑤使用加密技术可有效增强通信的安全性;

⑥多路光纤技术的发展大大提高了数字通信的效率。

需要指出,鉴于传统公用电话网已在世界范围普及,目前家庭个人计算机用户大都通过电话线路与计算机网络相连;此外,随着卫星通信的发展,高容量、高宽带的多路复用传输也大大提高了模拟通信的传输效率。但是,如果在两台计算机的通信线路之间,只有部分电路采用数字通信,则数字通信的优点并不能充分地得到发挥。因此,为了提高通信效率,有条件的用户应安装数字数据通信专线;此外,应大力发展陆上和海底的洲际光缆。

近 20 年来,数字数据通信技术已开始发展并得到广泛应用。目前,数字通信已开始 in 长距离话音和数字数据领域逐渐替代传统的模拟通信。计算机网络技术的应用发展,则大大推动了数字通信技术的迅速发展。可以预言,数字数据通信最终将取代模拟数据通信。

第三节 数据通信的主要技术指标

在数据通信中,我们一般使用带宽、比特率、波特率和误码率来分别描述信道传输能力,数据信号传输速率的大小和传输质量的好坏等。

1. 带宽

在模拟信道中,我们常用带宽表示信道传输信息的能力,带宽即传输信号的最高频率与最低频率之差。理论分析表明,模拟信道的带宽或信噪比越大,信道的极限传输速率也越高。这也是为什么我们总是努力提高通信信道带宽的原因。

2. 比特率

在数字信道中,比特率是数字信号的传输速率,它用单位时间内传输的二进制代码的有效位(bit)数来表示,其单位为每秒比特数 bit/s (bps)、每秒千比特数 (Kbps) 或每秒兆比特数 (Mbps) 来表示(此处 K 和 M 分别为 1 000 和 1 000 000,而不是涉及计算机存储器容量时的 1 024 和 1 048 576)。

3. 波特率

波特率指数据信号对载波的调制速率,它用单位时间内载波调制状态改变次数来表示,其单位为波特(Baud)。

波特率与比特率的关系为:比特率 = 波特率 × 单个调制状态对应的二进制位数。

显然,两相调制(单个调制状态对应 1 个二进制位)的比特率等于波特率;四相调制(单个调制状态对应 2 个二进制位)的比特率为波特率的两倍;八相调制(单个调制状态对应 3 个二进制位)的比特率为波特率的三倍;依次类推。

4. 误码率

误码率指在数据传输中的错误率。在计算机网络中一般要求数字信号误码率低于 10^{-6} 。

第四节 数据传输方式

一、基带信号与宽带信号以及它们的传输

1. 基带信号与基带传输

基带信号(Baseband Signal)直接用两种不同的电压来表示数字信号 1 和 0, 因此我们将对应矩形电脉冲信号的固有频率称为“基带”, 相应的信号称为基带信号。

基带传输(Baseband Transmission)指通过有线信道直接传输基带信号, 一般用于传输距离较近的数字通信系统, 如基带局域网系统。

2. 宽带信号

宽带信号(Wideband Signal)用多组基带信号 1 和 0 分别调制不同频率的载波, 并由这些分别占用不同频段的调制载波组成。

3. 多路复用

为了充分利用通信干线的通信能力, 人们广泛使用多路复用(Multiplex)技术, 即让多路通信信道同时共用一条线路。多路复用可分为频分多路复用和时分多路复用。

· 频分多路复用

当我们采用宽带信号时, 由于同一线路上不同频率的各路信道互不干扰地同时传输各自的信号, 我们称之为频分多路复用(Frequency-Division Multiplexing)。频分多路复用常用于宽带网络中。

· 时分多路复用

当我们采用基带信号时, 如让各路通信按时间顺序瞬时地分别占有线路的整个频带, 并周期性地重复此过程, 该线路就按时间分隔成了多个逻辑信道, 我们称之为时分多路复用(Time Multiplexing)。其中, 同步分时多路通信可以确定每个信道何时使用线路; 反之则称为异步分时多路通信。时分多路复用常用于基带网络中。

二、并行与串行方式

根据一次传输数位的多少可将基带传输分为并行(Parallel)方式和串行(Serial)方式, 前者是通过一组传输线多位同时传输数字数据, 后者是通过一对传输线逐位传输数字代码。通常, 计算机内部以及计算机与并行打印机之间采用并行方式, 而传输距离较远的数字通信系统多采用串行方式。

并行传输方式要求并行的各条线路同步, 因此需要传输定时和控制信号, 而并行的各路信号在经过转发与放大处理时, 将引起不同的延迟与畸变, 故较难实现并行同步。若采用更复杂的技术、设备与线路, 其成本会显著上升。故在远距离数字通信中一般不使用并行方式。

串行通信双方常以数据帧为单位传输信息, 但由于串行方式只能逐位传输数据, 因此, 在发送方需要进行信号的并/串转换, 而接收方则需要进行信号的串/并转换。

三、单工、半双工和全双工方式

根据通信双方的分工和信号传输方向可将通信分为三种方式: 单工、半双工与全双工, 在

计算机网络中主要采用双工方式,其中:局域网采用半双工方式,城域网和广域网采用全双工方式。

①单工(Simplex)方式:通信双方设备中发送器与接收器分工明确,只能在由发送器向接收器的单一固定方向上传送数据。采用单工通信的典型发送设备如早期计算机的读卡器,典型的接收设备如打印机。

②半双工(Half Duplex)方式:通信双方设备既是发送器,也是接收器,两台设备可以相互传送数据,但某一时刻则只能向一个方向传送数据。例如,步话机是半双工设备,因为在一个时刻只能有一方说话。

③全双工(Full Duplex)方式:通信双方设备既是发送器,也是接收器,两台设备可以同时两个方向上传送数据。例如,电话是全双工设备,因为双方可同时说话。

四、异步传输与同步传输

1. 同步问题的重要性

在数字通信中,同步(Synchronous)是十分重要的。当发送器通过传输介质向接收器传输数据信息时,如每次发出一个字符(或一个数据帧)的数据信号,接收器必须识别出该字符(或该帧)数据信号的开始位和结束位,以便在适当的时刻正确地读取该字符(或该帧)数据信号的每一位信息,这就是接收器与发送器之间的基本同步问题。

当以数据帧传输数据信号时,为了保证传输信号的完整性和准确性,除了要求接收器应能识别每个数据帧对应信号的起止,以保证在正确的时刻开始和结束读取信号,也即保持传输信号的完整性外;还要求使其时钟与发送器保持相同的频率,以保证单位时间读取的信号单元数相同,也即保证传输信号的准确性。

因此当以数据帧传输数据信号时,要求发送器应对所发送的信号采取以下两个措施:①在每帧数据对应信号的前面和后面分别添加有别于数据信号的开始信号和停止信号;②在每帧数据信号的前面添加时钟同步信号,以控制接收器的时钟同步。

2. 异步传输与同步传输

异步传输与同步传输均存在上述基本同步问题:一般采用字符同步或帧同步信号来识别传输字符信号或数据帧信号的开始和结束。两者之间的主要区别在于发送器或接收器之一是否向对方发送时钟同步信号。

异步传输(Asynchronous Transmission)以字符为单位传输数据,采用位形式的字符同步信号,发送器和接收器具有相互独立的时钟(频率相差不能太多),并且两者中任一方都不向对方提供时钟同步信号。异步传输的发送器与接收器双方在数据传送之前不需要协调,即发送器可以在任何时刻发送数据,而接收器必须随时都处于准备接收数据的状态。计算机主机与输入、输出设备之间一般采用异步传输方式,如键盘、典型的 RS-232 串口(用于计算机与调制解调器或 ASCII 码终端设备之间)。发送方可以在任何时刻发送一个字符(由一个开始位引导,然后连续发完该字符的各位,后跟一个位长以上的哑位)。

同步传输(Synchronous Transmission)以数据帧为单位传输数据,可采用字符形式或位组合形式的帧同步信号(后者的传输效率和可靠性高),由发送器或接收器提供专用于同步的时钟信号。在短距离的高速传输中,该时钟信号可由专门的时钟线路传输;计算机网络采用同步传输方式时,常将时钟同步信号植入数据信号帧中,以实现接收器与发送器的时钟同步。

第五节 错误检测与修正

在数字数据通信中,由发送器发送的数据信号帧(Frame)在经由网络传到接收器后,由于多种原因可能导致错误位(bit errors)的出现,因此必须由接收器采取一定的措施探测出所有的错误位,并进而采取一定的措施予以修正。

一、错误检测的基本原理

发送器向所发送的数据信号帧添加错误检验码(Check Bits),并取该错误检测码作为该被传输数据信号的函数;接收器根据该函数的定义进行同样的计算,然后将两个结果进行比较:如果结果相同,则认为无错误位;否则认为该数据帧存在有错误位。

一般说来,错误检测可能出现三种结果:

- ①在所传输的数据帧中未探测到,也不存在错误位;
- ②所传输的数据帧中有一个或多个被探测到的错误位,但不存在未探测到的错误位;
- ③被传输的数据帧中有一个或多个没有被探测到的错误位。

显然我们希望尽可能好地选择该检测函数,使检测结果可靠,即:所有的错误最好都能被检测出来;如检测出现无错结果,则应不再存在任何未被检测出来的错误。

实际采用的错误检测方法主要有两类:奇偶校验(Parity)和CRC循环冗余校验(Cyclic Redundancy Check)。

二、奇偶校验

1. 单向奇偶校验

单向奇偶校验(Row Parity)由于一次只采用单个校验值,因此又称为单个位奇偶校验(Single Bit Parity)。发送器在数据帧每个字符的信号位后添一个奇偶校验位,接收器对该奇偶校验位进行检查。典型的例子是面向ASCII码的数据信号帧的传输,由于ASCII码是七位码,因此用第八个位码作为奇偶校验位。

单向奇偶校验又分为奇校验(Odd Parity)和偶校验(Even Parity),发送器通过校验位对所传输信号值的校验方法如下:奇校验保证所传输每个字符的8个位中1的总数为奇数;偶校验则保证每个字符的8个位中1的总数为偶数。

显然,如果被传输字符的7个信号位中同时有奇数个(例如1、3、5、7)位出现错误,均可以被检测出来;但如果同时有偶数个(例如2、4、6)位出现错误,单向奇偶校验是检查不出来的。

一般在同步传输方式中常采用奇校验,而在异步传输方式中常采用偶校验。

2. 双向奇偶校验

为了提高奇偶校验的检错能力,可采用双向奇偶校验(Row and Column Parity),也称为双向冗余校验(Vertical and Longitudinal Redundancy Checks)。

图1.4给出了由5个ASCII码字符数据信号及其双向偶校验位组成的典型传输矩阵。其中:前五行各字符的偶校验位组成的校验位列(最右边一列),最下面一行由各列信号位的偶校验位组成,该矩阵右下角的校验位则可按行或按列取校验值。

传输方向	校验位列							
←	1	0	1	1	0	1	1	1
←	1	1	0	1	0	1	1	1
←	0	0	1	1	1	0	1	0
←	1	0	0	0	1	0	1	1
←	0	1	0	1	1	1	1	1
←	1	0	0	0	1	1	1	0
								校验位行

图 1.4 典型双向偶校验传输矩阵

显然,如果被传输的字符信号矩阵中同时在偶数个行中的偶数个相同的列中出现错误,双向奇偶校验也是检查不出来的。

三、CRC 循环冗余校验

1. CRC 循环冗余校验的基本原理

发送器和接收器约定选择同一个由 $n+1$ 个位组成的二进制位列 P 作为校验列,发送器在数据帧的 K 个位信号后添加 n 个位($n < K$)组成的 FCS 帧检验列(Frame Check Sequence),以保证新组成的全部信号列值可以被预定的校验二进制位列 P 的值对二取模整除;接收器检验所接收到数据信号列值(含有数据信号帧和 FCS 帧检验列)是否能被校验列 P 对二取模整除,如果不能,则存在传输错误位。 P 被称为 CRC 循环冗余校验列,正确选择 P 可以提高 CRC 冗余校验的能力。(注:对二取模的四则运算指参与运算的两个二进制数各位之间凡涉及加减运算时均进行 XOR 异或运算,即:1 XOR 1 = 0, 0 XOR 0 = 0, 1 XOR 0 = 1)

可以证明,只要数据帧信号列 M 和校验列 P 是确定的,则可以唯一确定 FCS 帧检验列(也称为 CRC 冗余检验值)的各个位。

FCS 帧检验列可由下列方法求得:在 M 后添加 n 个零后对二取模整除以 P 所得的余数。

例如:如要传输的 $M = 7$ 位列为 1011101,选定的 P 校验二进制位列为 10101(共有 $n+1=5$ 位),对应的 FCS 帧校验列即为用 10111010000(共有 $M+n=7+4=11$ 位)。对二取模整除以 10101 后的余数 0111(共有 $n=4$ 位)。因此,发送方应发送的全部数据列为 10111010111。接收方将收到的 11 位数据对二取模整除以 P 校验二进制位列 10101,如余数非 0,则认为有传输错误位。

2. CRC 循环冗余校验标准多项式 $P(X)$

为了表示方便,实用时发送器和接收器共同约定选择的校验二进制位列 P 常被表示为具有二进制系数(1 或 0)的 CRC 标准校验多项式 $P(X)$ 。

(1)CRC 循环冗余校验常用的标准多项式 $P(X)$ 。常用的 CRC 循环冗余校验标准多项式如下:

$$\text{CRC}(16 \text{ 位}) = X^{16} + X^{15} + X^2 + 1$$

$$\text{CRC}(\text{CCITT}) = X^{16} + X^{12} + X^5 + 1$$

$$\text{CRC}(32 \text{ 位}) = X^{32} + X^{26} + X^{23} + X^{16} + X^{12} + X^{11} + X^{10} + X^8 + X^7 + X^5 + X^4 + X^2 + X + 1$$

以 CRC(16 位)多项式为例,其对应校验二进制位列为 1 1000 0000 0000 0101。

注意:这儿列出的标准校验多项式 $P(X)$ 都含有 $(X+1)$ 的多项式因子;各多项式的系数均为二进制数,所涉及的四则运算仍遵循对二取模的运算规则。

(2)CRC 循环冗余校验标准多项式 $P(X)$ 的检错能力。CRC 循环冗余校验具有比奇偶校验