刘尊全 著

计算机病毒防范与信息对抗技术

清华大学出版社

计算机病毒防范与信息对抗技术

刘尊全 著

清华大学出版社

内 容 简 介

本书从信息安全角度，阐述了计算机病毒的危害，系统地介绍了防范计算机病毒的方法、技术和理论问题，阐述了计算机网络和大型信息系统防范计算机病毒的有关问题，专门论述了软件保护方法、信息对抗技术和防范计算机病毒的人工智能方法和技术。内容新颖，实用性强。书中给出了防范病毒的程序实例，具有重要的实用价值和理论价值。

本书可作为高等学校计算机专业本科生和研究生的教材，也可作为计算机应用人员、管理人员在信息安全方面的参考读物。

计算机病毒防范与信息对抗技术

刘尊全 著

责任编辑 焦金生

☆

清华大学出版社出版

北京 清华园

北京昌平第一排版厂排版

人民文学印刷厂印装

新华书店总店科技发行所发行

☆

开本：787×1092 1/16 印张：22 1/4 字数：525 千字

1991年5月第1版 1991年5月第1次印刷

印数：00001～10500

塑装ISBN 7-302-00827-2/TP·299 定价：10.00 元

精装ISBN 7-302-00828-0/TP·300 定价：15.00 元

前　言

随着计算机科学技术的迅速发展，特别是微型计算机日益广泛的应用，计算机在现代化建设中的战略地位和作用日益突出。与此同时，近年来计算机犯罪和计算机病毒的蔓延，已构成对国际社会的严重威胁。面临计算机犯罪和计算机病毒的挑战，我国计算机系统（特别是大型信息系统和正在建立过程中的计算机网络）的安全问题日益受到人们的重视和关切。

自1989年春季以来，计算机病毒在我国出现并大量蔓延，已构成了对计算机系统安全的威胁。计算机病毒是计算机犯罪的一种形式，是人为对计算机系统非法入侵和破坏的一种手段。当前，计算机专业人员、广大计算机用户和计算机方面的管理人员需要了解和掌握信息系统的安全技术和防范计算机病毒的知识，并且应有紧迫感。

刘尊全研究员1982年在国外工作期间，就着手信息安全方面的研究工作，起步较早，并多次出席国际会议。他在信息安全方面对应用技术开发和理论研究均取得了重要进展，并引起了国内外有关方面的关注。目前他所负责的科研项目已被列为1990年中国科学院院长基金特别支持项目。

自1988年秋季开始，刘尊全研究员为软件专业人员和研究生讲授《计算机病毒及其防治》课程，并在研究工作的基础上，结合教学过程的实践，编著了《计算机病毒防范与信息对抗技术》一书。书中分析了信息系统的脆弱性，阐述了防范计算机病毒的方法、技术和理论，提出了软件保护和发展信息对抗技术的一系列问题，具有重要的应用价值和学术价值。

为了加速我国计算机安全技术的发展，我向高等学校计算机专业大学生、研究生和从事计算机研制（特别是软件设计）、应用工作的科学技术人员推荐这部著作，并期望它在防范计算机病毒方面发挥积极的作用。

吴几康

1990.6.26于北京

目　录

第一章　信息系统的特征及其脆弱性

计算机是自动化信息加工的工具，它的运行依赖于程序。计算机的应用已从科学计算、过程控制、数据处理等方面，逐步进入到各个领域的知识处理，计算机已经成为知识处理的工具。人工智能技术将计算机高速度运行、大容量存储信息的"量"的方面优势转向知识表示、知识处理、机器证明和创造的"质"的方面，从而使计算机技术的发展进入了一个崭新的阶段。

计算机技术的迅猛发展，特别是微电子学技术的进步，使微型计算机进入社会和家庭，它对计算机的发展日趋产生着巨大的影响。计算机系统和通信系统的结合，使信息传输和加工可以跨越地理位置的障碍，信息加工过程具有动态、随机和瞬时发生的特性。

当今信息系统在现代化建设中的重要作用和战略地位，充分显示了它的巨大生命力。同时，它自身的脆弱性和易于攻击的特点，使得信息系统的安全成为发达国家和发展中国家急待解决的重大课题。

1.1　信息系统的构成

从安全角度，我们应该特别注重信息系统内部的物理界面、逻辑界面及环境的分析。

1.1.1　信息系统工作的环境

计算机系统（Computer System）是信息系统（Information System）的核心，它由硬件（Hardware）和软件（Software）组成，完成对信息的自动化加工过程。通信系统（Communication System）包括工作站、计算机通信网络和计算机网络，可以通过线路和计算机之间或通过线路和终端设备之间进行数据传输。信息系统的开放方式和共享资源的环境，使得系统易受攻击和存在潜在威胁（图 1-1）。

信息系统采用开放方式，向多个用户（可以是不同时间、不同地点的用户）提供服务。用户可以根据需要（正常情况下，需经过审核和注册）共享信息系统的资源，其中包括硬件资源和软件资源。

早期的计算机与用户的接口是物理界面，这使得使用者只有在熟悉机器内部构造和细节的情况下，才能有效地发挥系统的作用。软件的发展，使得用户可以通过逻辑界面使用系统设备（图 1-2）。

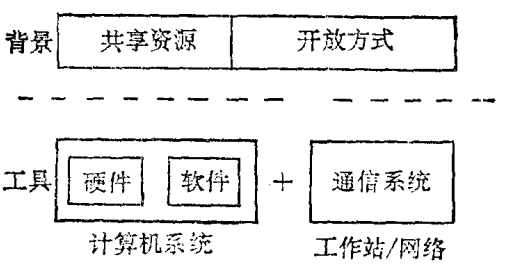


图 1-1　信息系统及其背景

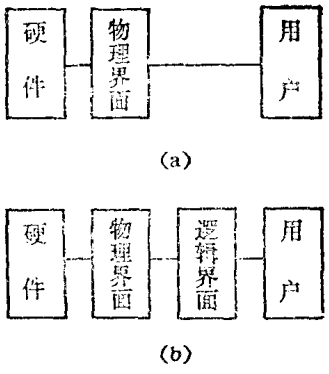


图 1-2 从物理界面到逻辑界面

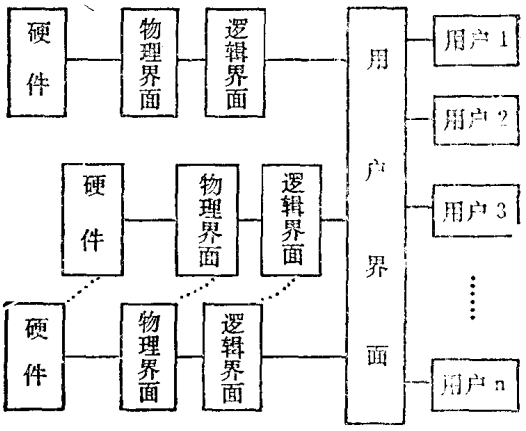


图 1-3 用户—逻辑界面—系统

随着电子技术的发展，硬件性能日趋提高，其结构也日趋复杂，关键部件已由过去的人工生产方式进入到自动化生产方式。系统为用户提供逻辑界面，一般情况下只要阅读手册或说明书，用户就可使用计算机系统。逻辑界面在相当大的程度下提供了信息系统对用户的透明度（图1-3）。系统的透明度在一定程度上反映了系统的脆弱性和易受攻击的程度。

1.1.2 操作系统

操作系统（Operating System）是系统软件中最基本的部分，它用来控制和管理系统资源，将裸机（Bare Machine）转换成虚拟机（Virtual Machine），方便了用户的使用，并有效地将用户的程序运行进行集中管理，是系统的总调度师。

操作系统的主要作用是：

（1）管理系统资源，这些资源包括中央处理机、主存储器、输入输出设备（显示器、打印机、绘图仪、磁盘驱动器等）和数据文件等。

（2）对系统资源进行合理调度，使用户能够共享系统资源。

（3）为用户信息的输入输出提供简便的工作方式。

（4）规定用户的接口，发现和处理各种系统发生的可确认的错误。

1.1.2.1 进程

一个进程（Process）是一个程序的执行，它是由一系列可以被解释成机器指令、数据或栈的有格式字节组成的。一个进程的生存周期可以划分为一组状态（State），每个状态都具有一定的描述进程的特点，其基本状态是（图1-4）：

（1）进程正在用户态下运行。

（2）进程正在核心态下运行。

（3）就绪状态，一旦调度程序选中了它，它可以投入运行。调度算法决定哪一个处于就绪状态的进程就成为下一个执行的进程。

（4）睡眠状态，此时的进程不再继续执行。例如等候I/O完成时，进程使自己进入睡眠状态。

需要指出，任何时刻一个处理机只能执行一个进程，所以至多有一个进程可以处在用户运行状态或核心运行状态。

类似进程状态及其转换，计算机系统一旦感染上计算机病毒，当感染病毒的程序体执行时，病毒体本身的逻辑判定就将决定病毒潜伏机制、传染机制或激发机制的选择。

在一个分时方式下，几个进程都可处于运行状态，并且它们有可能都在核心态下运行。**如果对它们在核心态下的运行不加以限制，则有的进程会非授权使用全程核心数据结构。**

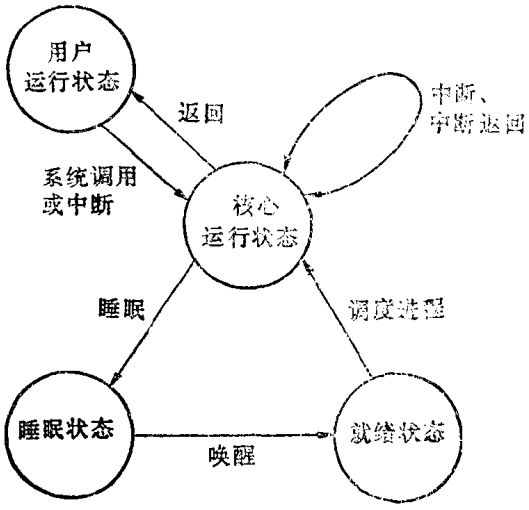


图 1-4 进程状态及转换

1.1.2.2 中断

计算机系统是在操作系统管理和控制下运行的，在一定意义上操作系统是中断驱动，中断对操作系统的重要性如同一台机械装置中的齿轮一样。

中断（Interrupt）是指CPU对系统中发生的无一定时序关系的随机事件的响应。一般情况下，中断大体可以有下面的分类：

（1）机器故障中断：例如电源故障、内存奇偶校验错，机器电路检验错等。

（2）输入输出中断：用以反映输入输出设备和通道的数据传输状态（完成或出错）。

（3）访管中断：用户程序在运行过程中通过访管指令向操作系统请求为其提供某种功能的服务（例如为其分配一块主存，建立进程等）。

（4）外部中断：包括时钟中断，操作员控制台中断，计算机网络中其它机器的通讯要求中断等。

（5）程序中断：程序中的问题引起的中断，如用户错误地使用指令或数据、机器运算过程中溢出等，存储保护、虚拟存储管理中的缺页、缺段等。

计算机系统的中断由硬件和软件（可以是固化的软件）配合起来处理。大型计算机和微型计算机的中断处理大致相同，IBM-PC微型计算机的中断处理过程如(图1-5)：

（1）当处理机接受某中断请求时，先由硬件进行下面的隐操作：

* 将处理机的程序状态字PSW压入堆栈。

* 将程序代码段的相对地址的指令指针IP和基地址寄存器CS的内容压入堆栈，用来保存被中断程序的返回地址。

* 保存被接受中断请求的中断向量地址（包括IP，CS等）。

* 按中断向量地址把中断处理程序的程序状态字存入处理机的程序状态字寄存器中。

（2）转入中断处理程序进行中断处理。

（3）当中断处理完成后，恢复被中断的程序现场，即取出堆栈中的PSW、IP、

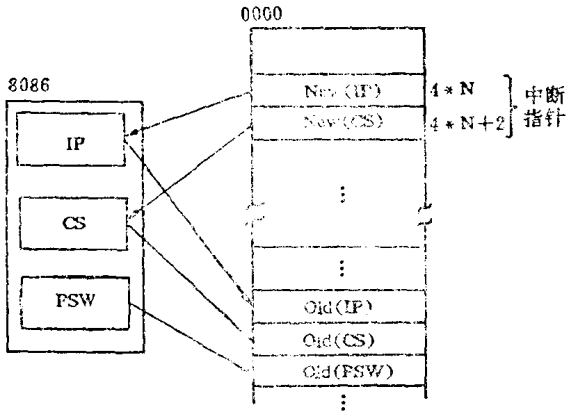


图 1-5 IBM-PC 中断的隐操作

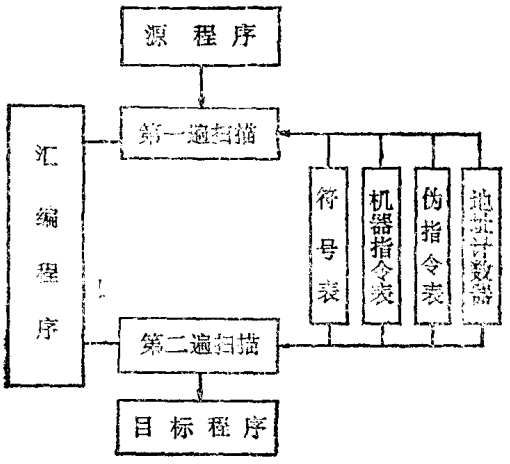


图 1-6 汇编过程

CS 内容，以便返回被中断程序执行。

当前流行的微型计算机病毒，正是利用磁盘文件读写中断将计算机病毒非授权侵入合法用户的程序体中，从而实现计算机病毒的传染机制。

1.1.3 编译系统

编译系统是一个转换器，它将用户书写的算法语言程序转换成相应的机器可直接运行的目标程序，防范计算机病毒要注意程序变换机制及其方向性。

1.1.3.1 汇编程序

汇编程序（Assembler）是将汇编语言写的源程序翻译成机器代码（目标程序）的一种翻译程序（图 1-6）。

通常汇编程序需要对源程序进行两次扫描。在第一次扫描过程中，通过查机器指令表，确定源程序中各指令助记符的字节数，通过处理伪指令，确定各数据区所占单元，并通过地址计数器累计单元数目，以此确定所有标号所对应的地址，并登记在符号表中。第二遍扫描是在第一次扫描的基础上进行翻译，将助记符翻译为机器代码，把符号地址翻译成实际地址（通过查符号表来完成），并给数据区分配单元，这样便完成了汇编的主要工作。

汇编程序可以对源程序中的语法错误给出相应的出错信息，对汇编结果产生列表文件。

浮动汇编程序可以产生浮动目标程序，给目标程序以重定位信息，以便连接装配程序使用。

目前微型计算机的汇编程序都有宏处理功能，因此也可以将汇编程序称为宏汇编程序。

目前相当多的计算机病毒采用汇编语言书写，以便充分利用机器的功能和各种特性。对于已感染病毒的微型计算机，其病毒程序往往是分成几个部分寄生在合法用户的文件中，在程序运行时需要对病毒体程序进行重新定位和安装。

1.1.3.2 解释程序

解释程序（Interpreter）是处理算法语言的另一种形式。解释程序是采用边解释边执行的方法，解释程序的工作结果是源程序的执行结果而不是目标程序状态的运行（图1-7）。

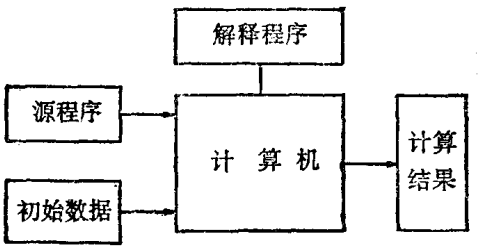


图 1-7 解释方式

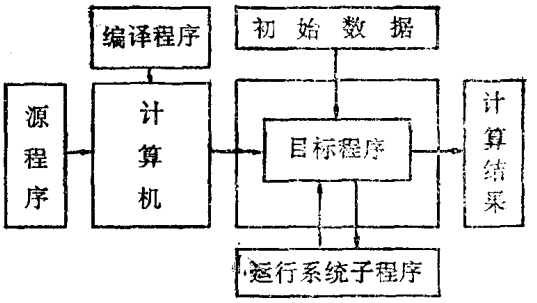


图 1-8 编译方式

从本质上讲，解释程序是典型的造表、查表技术。

对于数据文件中的计算机病毒，往往采用解释器进行病毒体的再生或激发。

1.1.3.3 *编译程序*

编译程序（Compiler）将用户用算法语言写成的源程序转换成面向机器的代码，这种代码也可以由汇编程序或装配程序作进一步加工，得出机器可直接运行的目标程序（图1-8）。

编译程序和汇编程序的主要区别是，汇编语言是机器语言（指令）的符号化，汇编语言和指令往往是一对一的；对于编译程序，源语言的每个语句可以等于多条指令，相应的算法语言更易于为用户掌握和使用。

一般情况下，编译程序要完成以下工作：

（1）词法分析：扫描源程序的ASCII码序列，拼出每一个单词并检查词法错误。将单词的ASCII码序列替换为机内表示。

（2）语法分析：检查源程序的形式语法错误，发现语法错误将输出有关的信息。

（3）中间代码生成：按照语义规则将语法分析过的语法单位翻译成中间代码形式，例如四元式

算 符	左操作数	右操作数	结 果

它的意义是，对左、右操作数进行算符指定的操作，将运算结果保留下来。中间代码接近或易于转换成计算机的机器指令。

（4）中间代码优化。

（5）目标代码生成。

编译过程中源程序的各种信息被保留在各种不同的表格中，编译的各个阶段都涉及到与构造、查找、更新有关的表格，是一个表格处理过程（图1-9）。

实践表明，计算机病毒可以侵入用算法语言书写的源程序。特别是计算机病毒能够存在于大多数编译器中，并且可以隐藏在各个层次当中，这样就造成每次调用编译程序

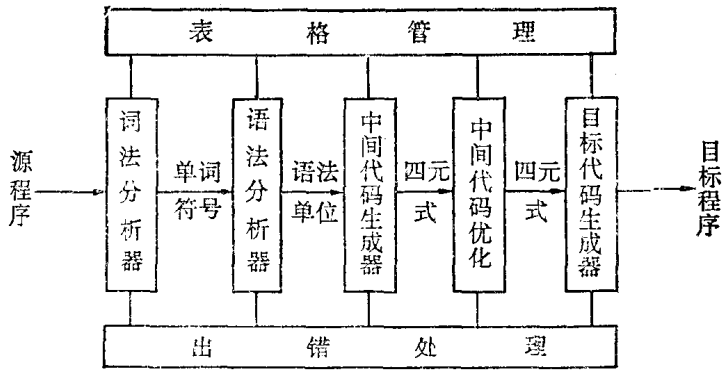


图 1-9 源程序至目标程序的数据转换

就是一次潜在的计算机病毒攻击或侵入。

1.1.4 数据库管理系统

数据库管理系统（DBMS，即 Data Base Management System）提供了用户使用逻辑手段共享信息系统资源的工具，是用户与数据库之间的接口，提供用户对数据库使用和加工的命令，包括对数据库的建立、修改、检索、计算、删除和统计等。当前数据库管理系统防范计算机病毒，已经成为国际上急待解决的问题。

1.1.4.1 结构和数据独立性

为了提高系统效率，减少冗余，便于增加新的数据，方便用户共享数据资源，目前数据库采取三级结构和两级数据独立，即在用户数据逻辑结构与数据的物理存储结构之间通过数据的整体逻辑结构来沟通，使得数据的物理存储结构的变化尽量不影响数据的整体逻辑结构或用户的应用程序，进一步数据整体逻辑结构的改变也尽量不影响用户的应用程序（图1-10）。

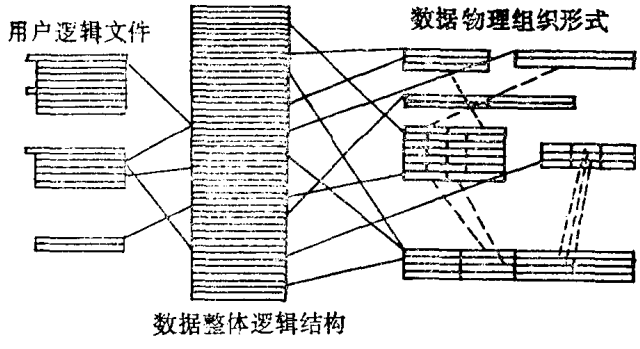


图 1-10 数据库系统

1.1.4.2 数据模型

目前数据库按其所使用的数据模型或数据结构，可分为下面三种：层次数据库，网络数据库，关系数据库。

（1）层次数据模型

如图1-11在层次数据模型中，数据的记录按层次或树形结构排列，记录之间存在有上下级关系，在层次数据模型中一个记录只可以从属于一个其它记录。

层次数据结构是一种有根的定向有序数。

（2）网状数据模型

如图1-12，在网络数据模型中，数据的记录之间可以任意发生联系，一个记录可以从属于网络中的几个其它记录。

层次数据模型可以看成是网络数据模型的一个子模型。

（3）关系数据模型

在关系数据模型中，一个记录不从属于另一个记录，使用一些特殊的关系数据模型的命令可以协调记录或选择记录。关系数据结构把数据的逻辑结构归结为满足一定条件的二维表形式，对数据进行逻辑运算或关系运算（图1-13）。

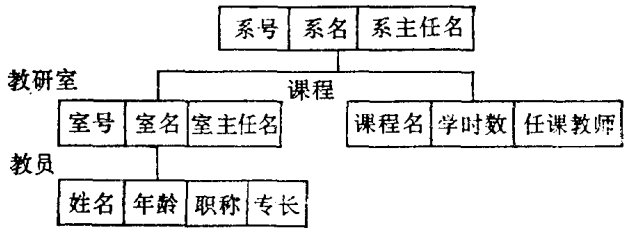


图 1-11 层次模型

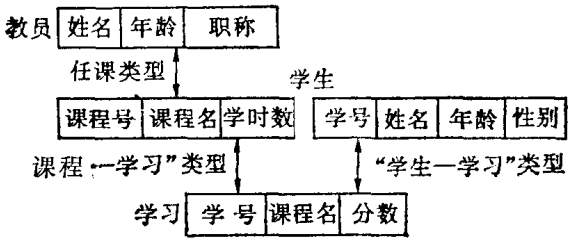
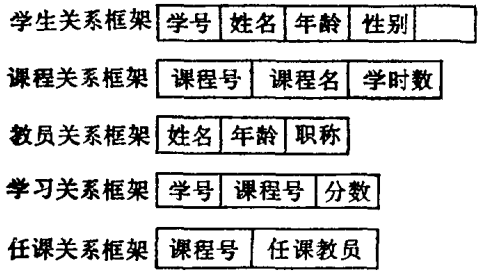


图 1-12 网络模型

（4）数据模型的属性与转换

每一种模型都可用来描述数据库中数据的实体、属性和关系。

层次数据模型是网络数据模型的特殊情况，对网络数据模型加以限制就能变成层次模型。把网络数据模型中的每一个记录类型和络类型变成关系框架，网络数据模型就转换成关系数据模型；把描述对象和描述联系的



课程关系

课程号	课程名	学时数
CS1	高等数学	300
CS2	数据结构	120
⋮	⋮	⋮
CSn	信息安全	80

图 1-13 关系模型

关系框架都变成记录类型和络类型，则关系数据模型就变成网络数据模型。

注意，不同数据模型的数据存储和访问方法是不同的。

这三类数据模型之间的主要区别在于信息表示的方式。在表示与用户有关的信息时，关系数据模型只用了数据记录的内容，而层次数据模型和网络数据模型用到数据记录间的联系以及它们在存储结构中的布局。关系数据模型不问存取形态，不要求用户了解数据记录的联系和顺序，它向使用者提供了一个简单、中性的应用形态。

计算机病毒可以隐藏在数据文件中，资源共享为计算机病毒的扩散提供了必要的条件和环境。

1.1.5 计算机网络

计算机网络是计算机技术和通信技术相结合的产物，它使得信息传输和加工可以跨

越地理位置的障碍，实现网络中的资源共享。当前，计算机病毒已经构成对计算机网络安全的严重威胁。

1.1.5.1 计算机通信网络

计算机通信网络是将不同地理位置的具有独立功能的多台计算机、终端及附属设备，用通信链路连接起来，采用网络软件实现通信过程中的资源共享（图1-14）。

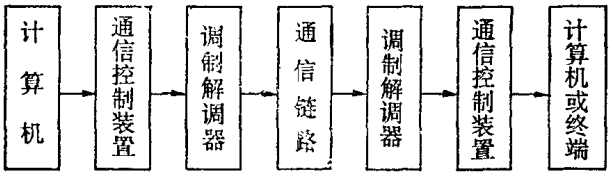


图 1-14 计算机通信网络

为了实现计算机通信，需要有网络协议、通信软件、网络操作系统的支持。通信控制装置（包括接口处理机、前端处理机、集中器和通信控制器等）、调制解调器、通信子网提供的通信链路是计算机通信网络中硬件的基本部分。

计算机通信网络按拓扑结构，可以有四种基本形式：星形、树形、环形和网格形，如图1-15。

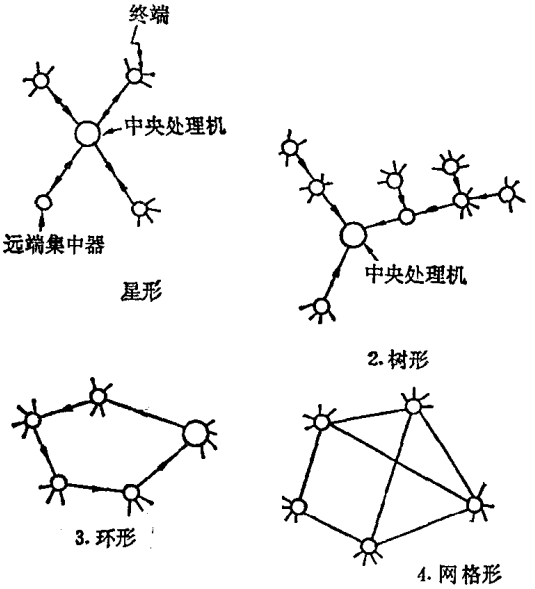


图 1-15 通信网络的拓扑结构

在计算机通信网络中，用户通过网络接口与每一个子系统建立物理通路，用户与每个子系统有一条逻辑通道，从而构成了若干个具有不同功能的计算机子系统的松散耦合（图1-16）。

在计算机通信网络中可以实现资源共享：

（1）硬件共享

如图1-17，本地用户A共享远地用户B的硬件E，当本地用户把软件C和数据D同时送到远地B，利用硬件E进行信息加工，最后将结果送回A。

（2）软件共享

如图1-18，本地用户A共享远地用户B的软件资源D，可以采用两种方式：将数据C送至B，然后B用软件D处理数据C并将结果送回A；也可以指定远地B将软件D送至A，用软件D处理数据C并得到结果。

（3）数据共享和控制信息的传送

如图1-19，本地用户A把控制信息C和软件D送到远地B，利用远地B的数据E进行信息加工，最后将结果送回本地用户A。

1.1.5.2 计算机网络

用户使用计算机网络，不需要先熟悉所需资源和文件所处的位置才能进行访问调用，而是通过网络操作系统(NOS,Network Operating System)来进行网络中文件或资源的访问和调用（图1-20）。