计算机

病毒清除技术

Design Programs for Removing Viruses

● 刘晓凤 著 ●

●● 陕西科学技术出版社

计算机病毒清除技术

刘晓风　著

陕西科学技术出版社

（陕）新登字第 002 号

计算机病毒清除技术

刘晓风　著

陕西科学技术出版社出版发行

（西安北大街 131 号）

西安理工大学印刷厂印刷

787×1092 毫米　16 开本　16 印张　39 万字

1995 年 6 月第一版　　1995 年 6 月第一次印刷

印数：1—4.000

ISBN 7-5369-2349-X/TP・68

定　价：16.00 元

内 容 简 介

本书详细地介绍了清除病毒程序的设计方法和有关的操作系统知识，同时给出了一些程序实例。这些程序用汇编语言编写，可以自动对全磁盘文件进行检索和处理，同时都写成便于修改的形式。在第 9 章和第 10 章给出的两个清除文件型病毒的程序中，主程序可以完全不动，只需要修改一个文件处理子程序中的几十条指令，就可以改造成清除另一种病毒的程序。第 11，12，13 章中给出的清除三种混合型病毒的程序中只需要修改处理 Boot 区和处理运行文件的两个子程序就可以用来清除另一种混合型病毒。第 15 章给出了一个清除 26 种计算机病毒的程序，只要加入新的子程序，就可以不断地扩充其功能。对于需要编写一些清除病毒程序的读者来说，这本书提供了一个很好的基础。

前　言

计算机病毒作为现代信息社会中的一个公害现象，在我国已有了七年的历史。无论是新病毒不断出现的势头，还是人们对病毒的惧怕现在似乎都渐渐地平息了下来。新病毒的出现速率并没有象人们所担心的那样按指数规律迅速地增长，而由于人们对病毒防范意识的增强，病毒的传播速度和范围也受到了一定的限制。然而随着新的人才不断地加入到计算机技术人员的行列中，其中一定也会有极少数的人加入到病毒制造者的行列中去。但象生物病毒总是首先攻击对其完全不加防范的人们一样，计算机病毒也主要危害那些对其知之甚少的人们，对病毒传播原理稍有了解，受其危害的程度就会大大降低。毕竟在我国目前的计算机应用水平下病毒的传播必须有计算机用户的参与，病毒只能通过软盘的易地使用传播，病毒进入内存也必须要由用户给它这样的机会。对病毒的了解与防毒意识的增强都意味着病毒进入内存机会的减少。

在过去的几年里，有数十种病毒在我国流行过，然后又销声匿迹了，在这个过程中清除病毒程序起了很大的作用。一个病毒制造者用很短的时间编写一段病毒可以传播到很大一个地区，影响很多的用户。而设计一个清除这种病毒的程序也只需要很短的时间，它也可以在计算机用户中传播，消除这种病毒。问题在于病毒样本能否很快地送到分析者手中，以及清除病毒程序能否迅速地到达计算机用户的手中。

在这本书中将主要讨论清除病毒程序的设计，这些程序都是用汇编语言编写的，可以自动地对全磁盘所有目录下的运行文件进行处理。由于对病毒进行分析的人自然会比较熟悉机器指令，同时也了解有关文件处理的一些 DOS 中断调用，这样用汇编语言编写这些程序是很自然的事。如果有了一个这样的程序，那么再编写下一个就很容易。分析病毒与文件的连接方式要困难一些，视病毒程序复杂程度不同及是否有加密反跟踪等措施，可能需要从一两个小时到几天时间。

本书第 1 章介绍 DOS 的磁盘数据组织。第 2 章介绍有关的操作系统中断调用。第 3 章介绍系统引导型病毒的识别方法并给出一个通用的系统引导型病毒的清除程序。第 4 章介绍了硬盘分区表的结构与修复的方法，分区表被破坏的故障是比较少见的，但如不能恢复则意味着硬盘全部数据的丢失。第 5 章讨论运行文件的结构与文件型病毒的连接方式，对清除病毒程序的设计者来说，这是必须了解的。第 6 章介绍了 debug 和 Codeview 这两种调试器的使用方法，以及用 debug 清除文件型病毒的方法。使用调试器对程序设计的重要性怎样强调都是不过分的，否则将面对程序中的错误一筹莫展，寸步难行。第 7 章讨论清除病毒程序设计与调试中的一些共同的问题。第 8 章介绍一个全磁盘运行文件的检索程序，在本书所有清除文件型病毒的程序中都使用了这一程序。第 9，第 10 两章是两种文件型病毒的清除程序，第 11，12，13 章是三种混合型病毒的清除程序，这些程序都写成很容易修改的形式，主程序可以完全不动，只需修改一个或两个子程序中的几十行指令就可用来清除另一种病毒。第 14 章是 Dir II 病毒的清除程序。在最后一章里我们介绍了一个多种病毒的清除程序，它可以清除 26

种病毒，这样的程序更适合向一个较大的地区的用户提供。如果仔细地阅读这个程序就会发现，程序功能的扩充，即增加能清除的病毒种类，是很容易的。

本书作者近年来编写过一些称之为 Virus Eraser（VE）的清除病毒程序，这种程序并不很难写。现在通过这本书将设计方法介绍给大家，希望它能对读者编写新的清除病毒程序有所助益。我们在保持理想的性能的前提下对程序尽可能地进行了简化，最后的程序是相当简单的。

作　者

1995 年 5 月

第1章　DOS的磁盘数据组织

1.1　逻辑驱动器

逻辑驱动器是由引导扇区，文件分配表，根目录和数据区组成的一个文件存储系统，一个逻辑驱动器可以是一个软盘(A盘或B盘)，也可以是硬盘或硬盘上的一部分(C，D，E，...)。一个逻辑驱动器也称为一个逻辑卷，它可以有一个卷标，也可以没有卷标。如果有卷标，则卷标填写在根目录里。一个逻辑驱动器是由磁盘上若干个地址相连的扇区组成的，这些扇区可以排成编号由0开始的一个逻辑扇区序列。下面以360KB软盘为例看一下逻辑扇区号的编排。360KB软盘有两个面，由两个磁头来读写，两个磁头的编号是0和1，每一面有40个磁道，编号从0到39，每个磁道划分为9个扇区，编号从1到9。磁盘的读写是以扇区为单位的，每个扇区为512字节，这些扇区按下面的次序排成一个序列：

0道0头1～9扇区	0道1头1～9扇区	1道0头1～9扇区	1道1头1～9扇区	…

这样一个逻辑扇区序列编号从0开始到719，逻辑0扇区就是0道0头1扇区。下面再看看1.2MB的软盘。1.2MB软盘有两个面，每面有80道，每道有15个扇区。1.44MB软盘每面有80道，每道有18个扇区。分别按相同的规律组成一个逻辑扇区序列。

下面我们讨论硬盘的情况。以最简单的情况为例，假设硬盘有4个磁头，每个头读写612个柱面，每个柱面有17个扇区。那么磁头的编号将是从0到3，柱面编号从0到611，然而扇区编号却是从1到17。这些扇区可以按下面的规律排成一列：

0柱0头1～17扇区	0柱1头1～17扇区	0柱2头1～17扇区	0柱3头1～17扇区	
1柱0头1～17扇区	1柱1头1～17扇区	1柱2头1～17扇区	1柱3头1～17扇区	…

各种不同规格的硬盘的磁头数和柱面数以及每柱扇区数都可能有不同的数值，但组成一个扇区序列所遵循的规律是相同的。

如果硬盘是用DOS 2.X版分区和格式化的，那么逻辑卷在硬盘中的位置如下：

0柱0头1扇区	逻辑卷从0柱0头2扇区开始

这时硬盘最前面的0柱0头1扇区是一个隐藏扇区，称为主引导扇区(Master Boot)或分区表(Partition Table)。逻辑卷从0柱0头2扇区开始，逻辑0扇区就是0柱0头2扇区。

如果硬盘是以DOS3.0以上版的系统分区的，而且只有一个分区，那么逻辑卷在硬盘中的位置如下：

隐藏扇区(0柱0头1～17扇区)	逻辑卷从0柱1头1扇区开始

其中隐藏扇区包括了0柱0头的全部扇区，0柱0头1扇区用来存放主引导块与分区表，其余的扇区闲置。逻辑卷从0柱1头1扇区开始，0柱1头1扇区就是逻辑0扇区。

如果硬盘用DOS的FDISK程序划分为两个分区，那么第一个分区称为是基本的DOS分区，同时也是一个逻辑驱动器(C盘)。第二个分区称为是扩展的DOS分区，可以是一个逻辑驱动器(D盘)，也可以划分为若干个逻辑驱动器(D,E,F盘等)。它们在硬盘中的位置如下：

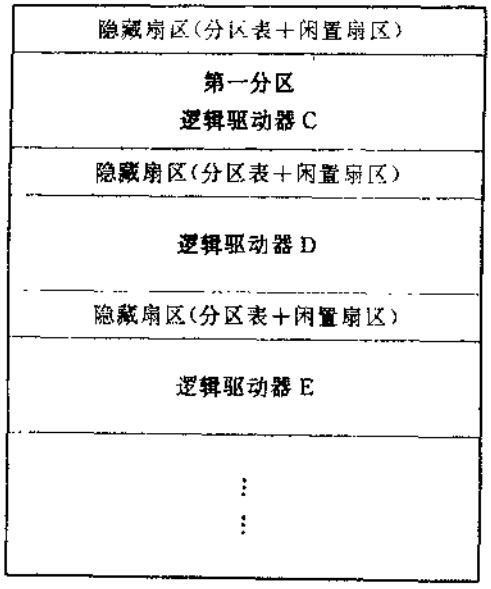

图1.1 硬盘的分区

如果硬盘划分为多个逻辑驱动器，则每个逻辑驱动器都是从某一个完整的柱面开始的。该柱面0头的全部扇区都将是隐藏扇区，第一扇区里有一个分区表，其它的扇区闲置。逻辑驱动器的实际内容从同一柱面的1头1扇区开始。

一个逻辑卷，无论是软盘还是硬盘或是硬盘的一部分，经过格式化后都由下面的一些部分组成：

引导扇区	文件分配表1	文件分配表2	根目录	数据区

引导扇区只有一个扇区，其内容是在格式化时写入的。文件分配表与根目录的划分都是在格式化时完成的。

1.2 硬盘分区表

硬盘的第一个扇区即0柱0头1扇区称为主引导扇区(Master Boot)，也称为分区表(Partition Table)。这个扇区有以下两个内容，前面的200多个字节是一段程序，称为主引导程序，这段程序的作用是从本分区逻辑0扇区中装入DOS引导程序(DOS Boot)，然后

转到DOS引导程序执行。后面从偏移为01BEH开始的64个字节是分区表。如下图所示：

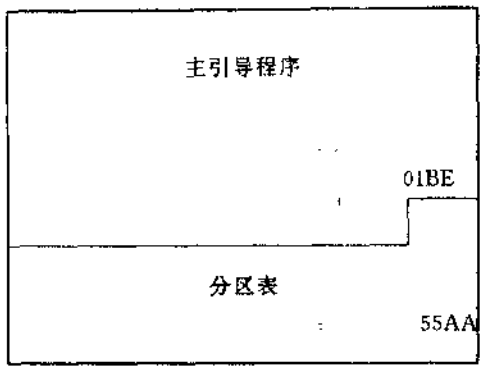

图 1.2　硬盘主引导扇区

分区表从偏移为01BEH处开始，共64字节，分为四栏，每栏16字节，用来描述一个分区。各字节的含义如下表所示：

表 1.1　分区表的一栏

00H	标志字节 活动DOS分区为80 其它为00
01H	本分区逻辑0扇区(DOS Boot)的磁头号
02H	逻辑0扇区在所在柱面中的扇区号
03H	逻辑0扇区所在柱面号
04H	分区类型标志
05H	本分区最后一个扇区的磁头号
06H	本分区最后一个扇区的扇区号
07H	本分区最后一个柱面的柱面号
08H	硬盘上在本分区之前的扇区总数 用双字表示
0CH	本分区的扇区总数 从逻辑0扇区计起，不含隐藏扇区， 用双字表示

在上面的表中给出的柱面号与扇区号虽各占一个字节，但实际上扇区号用6位表示，柱号用10位表示，扇区号所在字节的最高二位实际上是柱号的最高二位。这与INT 13H中这两个参数的表示方法一致。

在上面的表中偏移为04处的分区类型字节的含意如下：

01　　活动的DOS分区，小于12MB，12位文件分配表；

04　活动的 DOS 分区，小于 32MB，16 位文件分配表；

06　活动的 DOS 分区，大于 32MB，16 位文件分配表；

05　非活动的 DOS 分区；

02　XENIX 分区。

分区表虽然有四栏，但用 DOS 的 FDISK 程序分区后，最多只用两栏，第一栏描述基本的 DOS 分区，第二栏描述扩展的 DOS 分区。

如果硬盘是用 DM(Disk Manager)分区的，则分区表可能有 1 至 16 栏，描述所有可能的分区，详见第 4 章第 3 节。

用 debug 可以很容易地将硬盘的主引导块与分区表读出来检查。由于隐藏扇区不属于逻辑扇区，所以需编写一小段程序运行以后才可读出。操作步骤如下：

```
c>debug
-a100
14DE:0100  mov ax,0201
14DE:0103  mov bx,0200
14DE:0106  mov cx,0001
14DE:0109  mov dx,0080
14DE:010B  int 13
14DE:010D  int 3
14DE:010E  回车
-g=100
```

运行从 100 开始的这六条指令的结果是将硬盘 0 柱 0 头 1 扇区的内容读到内存中地址为 200H 处，打入下面的命令就可以看到该扇区的全部内容，

```
-d 200 3ff
14DE:0200  FA 33 C0 8E D0 BC 00 7C-8B F4 50 07 50 1F FB FC   .3.....|..P.P...
14DE:0210  BF 00 06 B9 00 01 F2 A5-EA 1D 06 00 00 BE BE 07   ................
14DE:0220  B3 04 80 3C 80 74 0E 80-3C 00 75 1C 83 C6 10 FE   ...<.t..<.u.....
14DE:0230  CB 75 EF CD 18 8B 14 8B-4C 02 8B EE 83 C6 10 FE   .u......L.......
14DE:0240  CB 74 1A 80 3C 00 74 F4-BE 8B 06 AC 3C 00 74 0B   .t..<.t. ...<.t.
14DE:0250  56 BB 07 00 B4 0E CD 10-5E EB F0 EB FE BF 05 00   V.......^.......
14DE:0260  BB 00 7C B8 01 02 57 CD-13 5F 73 0C 33 C0 CD 13   ..|...W.._s.3...
14DE:0270  4F 75 ED BE A3 06 EB D3-BE C2 06 BF FE 7D 81 3D   Ou...........}.=
14DE:0280  55 AA 75 C7 8B F5 EA 00-7C 00 00 49 6E 76 61 6C   U.u.....|..Inval
14DE:0290  69 64 20 70 61 72 74 69-74 69 6F 6E 20 74 61 62   id partition tab
14DE:02A0  6C 65 00 45 72 72 6F 72-20 6C 6F 61 64 69 6E 67   le.Error loading
14DE:02B0  20 6F 70 65 72 61 74 69-6E 67 20 73 79 73 74 65    operating syste
14DE:02C0  6D 00 4D 69 73 73 69 6E-67 20 6F 70 65 72 61 74   m.Missing operat
14DE:02D0  69 6E 67 20 73 79 73 74-65 6D 00 00 00 00 00 00   ing system......
14DE:02E0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
```

```
14DE:02F0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ..... .....
14DE:0300  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ..... .....
14DE:0310  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:0320  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ..... .....
14DE:0330  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ..... .....
14DE:0340  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ..... .....
14DE:0350  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:0360  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ..... .....
14DE:0370  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:0380  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:0390  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:03A0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:03B0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 80 01   ................
14DE:03C0  01 00 04 04 D1 02 11 00-00 00 EE FF 00 00 00 00   ................
14DE:03D0  C1 03 05 04 D1 CF FF FF-00 00 11 44 00 00 00 00   ..........D....
14DE:03E0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 00   ................
14DE:03F0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 55 AA   ..............U.
```

这里为了方便，将该扇区的内容全部列了出来，实际上在25行显示方式下一部分内容将移出屏幕上方。可以看出前218字节是主引导程序，后面有64字节的分区表，最后两个字节是分区表的标志。中间有228个字节全为0，当该扇区感染病毒时，除分区表外的446个字节大部分都会被病毒代码占据，这时就看不到那么多的0了。不过也有例外的情况，如在感染Flip病毒时病毒只修改主引导块最前面的66个字节，后面内容不变，这显然是为了迷惑人。要判断是否感染病毒，主要还是要检查反汇编后的代码。

如果分区表内容被破坏，那么硬盘将不能启动，用软盘启动后系统也不能识别硬盘的各逻辑驱动器。这时如果不能恢复丢失的分区表，硬盘数据将全部丢失。我们将在第4章中详细地讨论硬盘分区的划分与分区表的修复方法。

1.3 引导扇区

引导扇区(DOS Boot)也就是逻辑0扇区，它与前面讲过的硬盘主引导扇区是两种不同的数据。引导扇区最前面的3个字节一般总是EB XX 90，EB XX是一条短跳转指令，它跳过接着的几十个字节的BPB参数块，转到后面的引导程序去，90H是一条空指令。在这3个字节后面是8个字节的OEM识别码，用来写入各种系统的标志。这以后从偏移0BH开始是几十个字节的磁盘参数表，也称为BPB参数块。参数表后面是引导程序，其功能是检查磁盘上有无IO.SYS和MSDOS.SYS(或IBMBIO.COM和IBMDOS.COM)这两个系统文件，如果有则将IO.SYS(或IBMBIO.COM)装入内存，两个系统文件必须登录在根目录的第一和第二项，并且必须在数据区连续地存放。如果系统文件不存在则显示一些提示信息。引导扇区的内容如下：

表 1.2　引导扇区

00H	EB XX 90
03H	OEM 识别码(8 字节)
0BH	每一扇区的字节数(字)
0DH	每一分配单元(簇)的扇区数(字节)
0EH	保留扇区数(字)
10H	FAT 的个数(字节)
11H	根目录的目录项数(字)
13H	逻辑卷的扇区总数(字),若该字为 0, 则扇区总数是偏移为 20H 处的双字
15H	介质描述符(字节)
16H	每 FAT 的扇区数(字节)
18H	每道扇区数(字)
1AH	磁头数(字)
1CH	本分区逻辑 0 扇区前隐藏扇区数(字)
	引导程序 55AA

用不同版本的系统格式化的磁盘的参数块长度略有不同,用 debug 或 PCTools 可以很容易地查看引导扇区的内容。下面用 debug 查看一个 360KB 软盘的引导扇区内容,软盘插在 A 驱动器中:

```
c>debug
-L 100 0 0 1
-D 100 2ff
14DE:0100  EB 34 90 49 42 4D 20 20-33 2E 33 00 02 02 01 00   .4.IBM  3.3.....
14DE:0110  02 70 00 D0 02 FD 02 00-09 00 02 00 00 00 00 00   .p..............
14DE:0120  00 00 00 00 00 00 00 00-00 00 00 00 00 00 00 12   ................
14DE:0130  00 00 00 00 01 00 FA 33-C0 8E D0 BC 00 7C 16 07   .......3.....|..
14DE:0140  BB 78 00 36 C5 37 1E 56-16 53 BF 2B 7C B9 0B 00   .x.6.7.V.S.+|...
14DE:0150  FC AC 26 80 3D 00 74 03-26 8A 05 AA 8A C4 E2 F1   ..&.=.t &.......
14DE:0160  06 1F 89 47 02 C7 07 2B-7C FB CD 13 72 67 A0 10   ...G...+|...rg..
14DF:0170  7C 98 F7 26 16 7C 03 06-1C 7C 03 06 0E 7C A3 3F   |..&.|...|...|.?
14DE:0180  7C A3 37 7C B8 20 00 F7-26 11 7C 8B 1E 0B 7C 03   |.7|. ..&.|...|.
14DE:0190  C3 48 F7 F3 01 06 37 7C-BB 00 05 A1 3F 7C E8 9F   .H....7|....?|..
```

```
14DE:01A0  00 B8 01 02 E8 B3 00 72-19 8B FB B9 0B 00 BE D6   .......r... ....
14DE:01B0  7D F3 A6 75 0D 8D 7F 20-BE E1 7D B9 0B 00 F3 A6   }..u... ..} ....
14DE:01C0  74 18 BE 77 7D E8 6A 00-32 E4 CD 16 5E 1F 8F 04   t..w}.j.2...^...
14DE:01D0  8F 44 02 CD 19 BE C0 7D-EB EB A1 1C 05 33 D2 F7   .D.....}.....3..
14DE:01E0  36 0B 7C FE C0 A2 3C 7C-A1 37 7C A3 3D 7C BB 00   6.|...<|.7| =|..
14DE:01F0  07 A1 37 7C E8 49 00 A1-18 7C 2A 06 3B 7C 40 38   ..7|.I...|* ;|@8
14DE:0200  06 3C 7C 73 03 A0 3C 7C-50 E8 4E 00 58 72 C6 28   .<|s..<|P.N Xr.(
14DE:0210  06 3C 7C 74 0C 01 06 37-7C F7 26 0B 7C 03 D8 EB   .<|t...7|.& |...
14DE:0220  D0 8A 2E 15 7C 8A 16 FD-7D 8B 1E 3D 7C EA 00 00   ....|...}..=|...
14DE:0230  70 00 AC 0A C0 74 22 B4-0E BB 07 00 CD 10 EB F2   p....t"....  ....
14DE:0240  33 D2 F7 36 18 7C FE C2-88 16 3B 7C 33 D2 F7 36   3..6.|....;|3..6
14DE:0250  1A 7C 88 16 2A 7C A3 39-7C C3 B4 02 8B 16 39 7C   .|..*|.9|.. ..9|
14DE:0260  B1 06 D2 E6 0A 36 3B 7C-8B CA 86 E9 8A 16 FD 7D   .....6;|... ...}
14DE:0270  8A 36 2A 7C CD 13 C3 0D-0A 4E 6F 6E 2D 53 79 73   .6*|.....Non-Sys
14DE:0280  74 65 6D 20 64 69 73 6B-20 6F 72 20 64 69 73 6B   tem disk or disk
14DE:0290  20 65 72 72 6F 72 0D 0A-52 65 70 6C 61 63 65 20    error..Replace
14DE:02A0  61 6E 64 20 73 74 72 69-6B 65 20 61 6E 79 20 6B   and strike any k
14DE:02B0  65 79 20 77 68 65 6E 20-72 65 61 64 79 0D 0A 00   ey when ready...
14DE:02C0  0D 0A 44 69 73 6B 20 42-6F 6F 74 20 66 61 69 6C   ..Disk Boot fail
14DE:02D0  75 72 65 0D 0A 00 49 42-4D 42 49 4F 20 20 43 4F   ure...IBMBIO  CO
14DE:02E0  4D 49 42 4D 44 4F 53 20-20 43 4F 4D 00 00 00 00   MIBMDOS  COM....
14DE:02F0  00 00 00 00 00 00 00 00-00 00 00 00 00 00 55 AA   ..........  ..U.
```

从读出的内容中可看到前面的参数块，在该扇区的后部有引导失败时显示的信息，以及两个系统文件的文件名。

1.4 文件分配表

文件分配表 FAT(File Allocation Table)用来控制磁盘存储空间的分配，在格式化时生成，有两个完全相同的拷贝和固定的扇区数目，例如 360KB 软盘每个 FAT 有两个扇区，1.2MB 软盘有 7 个扇区，硬盘分区的 FAT 则随分区大小及系统版本而不同。磁盘数据区的存储空间在分配给文件时，是以一个或几个扇区为一组分配的，这样的一组扇区称为是一个分配单元或一簇(Cluster)，每簇的扇区个数是固定的，而且总是 2 的幂。例如 360KB 软盘的每簇扇区数是 2，1.2MB 及 1.44MB 软盘的每簇扇区数是 1，20MB 硬盘用 DOS2.0 格式化时每簇扇区数是 16，而用 DOS3.3 格式化时每簇扇区数是 4。这样磁盘数据区可划分为许多个连续的簇，文件分配表则划分为许多个域，每个域与一个簇相对应，记录着这个簇的分配情况。文件分配表有两种类型，一种由 12 位的域组成，另一种由 16 位的域组成。所有的软盘都是 12 位的文件分配表，硬盘分区的总簇数小于 4087 时也为 12 位，4087 簇以上则为 16 位的文件分配表。

文件分配表的第一个字节是介质描述符，它的值与引导扇区 BPB 参数块中偏移为 15H 处的介质描述符一样，硬盘为 0F8H，软盘为 0FDH(360KB)，0F9H(720KB 和 1.2MB)和 0F0H(1.44MB)，如果是 12 位的文件分配表，则在介质描述符后面有两个字节为 0FFH，16 位则后面有 3 个字节为 0FFH，以后就是与磁盘的簇对应的域了。

磁盘上可供分配的簇的编号从 2 开始，每一簇对应着文件分配表的一个域，对于 12 位的文件分配表，这种对应关系如下：

簇号乘以 1.5，以乘积的整数部分作为偏移量在 FAT 中取一个字，如果乘积的小数部分为零，则该字的低 12 位是对应的域，若乘积的小数部分不为零，则该字的高 12 位是对应的域。

16 位 FAT 的情况比较简单，只要将簇号乘以 2 作为偏移在 FAT 中取一个字就是对应的域。

FAT 的每一个域中填入的值有以下几种情况：

(0)000H　　可以使用的簇；

(F)FF7H　　坏簇；

(F)FFFH　　文件的最后一簇；

(X)XXXH　　文件下一簇的簇号。

当一个文件存入磁盘的时候，在目录中给出了这个文件首簇的簇号。在文件分配表中与首簇对应的域给出了下一簇的簇号，在下一簇对应的域中又给出再下一簇的簇号，直到遇到(F)FFFH 时为止，这样组成一个链表。用 PCTools 跟踪检查一下一个文件的链就很清楚了。

1.5 目录

目录有两种类型，根目录与子目录。根目录是在格式化时生成的，有固定的大小。而子目录是可以生长的，是一种特殊的文件。无论根目录还是子目录都是由若干个连续的扇区组成的。一个目录划分为许多栏，每一栏有 32 个字节，用来登录一个文件，其内容如下：

表 1.3　　目录的一栏

00H	文件名(8 字节)
08H	扩展名(3 字节)
0BH	文件属性(1 字节)
0CH	保留区(10 字节)
16H	生成或最后修改的时间(2 字节)
18H	生成或最后修改的日期(2 字节)
1AH	首簇簇号(2 字节)
1CH	文件长度(4 字节)

文件名第一个字节若是 00H，则说明该栏从未使用过，如第一个字节为 0E5H，则说

明该文件已被删除。

属性字节的各位含义如下：

0　只读，不允许写入或删除；

1　隐藏，普通的文件搜索找不到；

2　系统，普通的文件搜索找不到；

3　卷标，只在根目录出现；

4　子目录；

5　修改标志，文件一经修改即置 1；

6　保留；

7　保留。

子目录正常属性应是 10H，但如果改为 11H，13H 或 17H 也是可以的，当属性第 1 位置 1 时 DOS 的 Dir 命令将不显示该子目录。这样修改子目录属性并无实际意义。但在编写磁盘检索程序时需考虑到这一点，才可将所有的子目录都检索出来。用来表示时间的字的各位的含义如下：

00H－04H　实际秒数的二分之一

05H－0AH　分钟数

0BH－0FH　小时数

用来表示日期的字的各位的含义如下：

00H－04H　日子

05H－08H　月份

0BH－0FH　年份(相对 1980)

用来表示文件长度的是一个双字，前面是低字，后面是高字，如果该目录项表示一个子目录，则这个双字的值为 0。

一个磁盘的根目录是在格式化时生成的，它有固定的大小，根目录后面就是数据区。而子目录是一种特殊的文件，子目录在生成时是磁盘上的一簇，随着目录项的增加，子目录也可以生长。在 DOS 的树形目录结构下，在每一子目录下还可以生成新的子目录。与根目录不同的是，子目录生成时前两栏就填入了固定的内容，第一栏的文件名是'.'，代表当前目录，其首簇号指向当前目录的第一簇。第二栏文件名是'..'，代表当前目录的上一层目录，其首簇号指向上一层目录的首簇。从第三栏开始是子目录内的文件或子目录。

第 2 章　操作系统中断调用

在清除病毒程序中需要用到许多操作系统功能，我们所用到的大多数功能都是 DOS 中断 INT 21H 实现的，另外也有少数其它 DOS 中断和 BIOS 中断。在本书中不可能全面地讨论中断功能，但为了便于读者阅读本书中的程序以及编写新的程序，我们将对需要用到的功能作比较详细的介绍。

2.1　磁盘输入输出

在清除病毒时可能用到两类磁盘输入输出功能，一类是 BIOS 中的 INT 13H，另一类是 DOS 中的 INT 25H 和 INT 26H。这些在清除文件型病毒时是不需要的，只有在清除系统引导型病毒或是处理 Dir II 这样比较特殊的病毒时才需要用到。

INT 13H 功能 00H 复位软盘系统

调用参数：AH＝0

返　　回：无

这一功能用来使软盘控制器复位，但不需要在对每一个软盘进行读写前都调用此功能。

INT 13H 功能 02H 读磁盘扇区

调用参数：AH＝2

AL＝要读的扇区个数

CH＝磁道号(柱面号)

CL＝起始扇区号

DH＝磁头号

DL＝驱动器号

ES:BX＝读出数据在内存中的存放地址

这里 CH 对软盘是磁道号，对硬盘则是柱面号。实际上柱面号是用 10 位表示的，其最高两位放在 CL 的最高两位里，扇区号只用 CL 的低 6 位表示。驱动器号对 A 盘是 0，B 盘是 1，第一硬盘为 80H，第二硬盘为 81H，这里第一硬盘和第二硬盘均指物理上的硬盘，而不是指分区或逻辑驱动器。

返　　回：成功时　CF＝0

AH＝0

AL＝读出的扇区个数

失败时　CF＝1

AH＝错误代码

01H 非法驱动器命令
02H 找不到地址标记
03H 软盘写保护
04H 扇区未找到
08H DMA 出错
09H DMA 越界
0BH 坏道
10H 数据错误(CRC)
20H 磁盘控制器出错
40H 寻道失败
80H 操作超时

在用 INT 13H 读一个新插入的软盘时，即使软盘一切正常，并用 INT 13 的 0 号功能使软盘控制器复位后，也有可能读盘失败。所以在编程时首次读软盘如失败应重复试读三次，而在第一次读盘成功后，以后再对同一软盘进行读或写时都只需要试一次即可。对硬盘则无论读或写都只需试一次。

INT 13H **功能** 03H **写磁盘扇区**

调用参数：AH＝3
AL＝要写的扇区个数
CH＝磁道号(柱面号)
CL＝扇区号
DH＝磁头号
DL＝驱动器号
ES:BX＝要写入的数据在内存中的存放地址

这里柱面号与扇区的格式以及驱动器号均与上面 02H 功能的规定一致。

返　　回：成功时　CF＝0
AH＝0
AL＝实际写入的扇区个数
失败时　CF＝1
AH＝错误代码，其含义同功能 2。

在清除病毒程序中在向磁盘写入数据前一般都已有过成功的读盘操作，在写盘时只需调用一次。

INT 25H **绝对磁盘读**

这一功能是按磁盘的逻辑扇区号来读磁盘的，其入口参数有两种格式，第一种格式是 2.0 以上所有 DOS 版本都支持的，只能读软盘和小于 32MB 的硬盘分区，入口参数如下：

AL＝驱动器号(A＝0,B＝1,C＝2,...)
CX＝要读的扇区个数
DX＝起始扇区的逻辑扇区号
DS:BX＝读入数据的存放地址