

第一部分 文件压缩存档

1. 关于所用示例的说明

(1) 用于压缩处理的示范文件

为了说明每一种压缩存档软件的用法和功能,将主要以下面列表中的文件为例叙述。

Directory of C:\FORTRAN			
FOR2	EXE	111776 03-27-85	7:28p
L1	BAT	256 05-12-89	3:57p
FOR1	EXE	127514 04-04-85	4:31p
L2	BAT	256 05-12-89	3:58p
COM	BAT	172 01-04-80	7:14a
P	BAT	128 11-01-89	2:20p
FORTTRAN	LIB	116224 03-27-85	4:50p
MATH	LIB	39936 02-24-85	4:04p
ALTMATH	LIB	39424 02-24-85	4:05p
DECMATH	LIB	27648 02-24-85	4:06p
8087	LIB	28672 02-24-85	4:04p
CEXEC	LIB	12800 03-03-85	8:59a
LINK	EXE	39076 12-30-85	12:00p
13file(s)		543882 bytes	

若想将这些文件保存在一张 360KB 的软盘片上,用一般的拷贝方法显然是放不下的,但使用文件压缩存档软件就可以把这个子目录下 13 个文件共 543882 字节的内容压缩拷贝到一个文件中,并且压缩率至少在 30%~50%。不过,文件压缩程序有很多版本,不同版本软件的文件压缩能力及所选用的压缩方式都有差异。比如有的版本可以用几种不同的方式压缩文件,其效果截然不同,所以你应通过下面章节的内容仔细了解每种文件压缩存档软件的特点和用法,并灵活、择优地选用。

(2) 关于软件应用实例的说明

不同版本的文件压缩存档软件所提供的功能有许多是近似的,为了节省篇幅,不能把每个软件的功能及其操作方法一一列出。笔者将尽量选择软件中比较有特点的操作命令或选项举例,同时又兼顾到每个软件都能有示例说明命令的一般使用格式,所以,文件压缩存档软件均有的基本命令(如:A—把文件加入档案、U—建立或更改档案、D—从档案中删除文件、M—把文件移入档案、V—查看档案、E—释放压缩文件、X—释放压缩文件时连同子目录一起释放)的输出结果等信息,应用实例对每种软件来说都有参考价值。

2. 主要参考资料

由于文件压缩存档软件从一开始就是在网络中提供使用的,例如 LH 系列的软件就是用于日本的一个较大的信息网络 NIFTY-SERVE,资料多采用电子信息方式,发表于正式出版物的很少,所以有必要向读者推荐下列有关的参考文献。

- ① Labo, A. P. , A Hard Disk Cookbook; Shouei Press (1987).
- ② Kurita, T. , Huffman coding ,bit; Jewelry Box of Computing 43, vol. 20, no. 7, pp. 100—101 (1988)
- ③ Miki, K. , Documentation for Larc; LARC. MAN.
- ④ Haruyasu Yoshizaki (Yoshi) , User's Manual for High-Performance File-Compression Program; ICE. MAN
- ⑤ PKWARE Inc. , The file compression system ; MANUAL. DOC
- ⑥ Robert K Jung , User's Manual for the ARJ archiver program ; ARJ. DOC (ARJ version 2.22, ARJ version 2.39d)
- ⑦ ARC File Archive Utility ; ARC520. DOC
- ⑧ Terry A. Welch , A Technique for High Performance Data Compression; IEEE Computer Vol 17 No 6 (June 1984)
- ⑨ Data Compression Algorithms of LARC and LHarc; COMPRESS. TXT
- ⑩ Ziv, J. and Lempel, A. , "Compression of individual sequences via variable-rate coding." IEEE Trans. Inform. Theory IT-24,5 (Sept. 1978)
- ⑪ Haruyasu Yoshizaki (Yoshi), Manual for LHA Version 2.13
- ⑫ Rahul Dhesi, Reference Manual for ZOO; ZOO. MAN
- ⑬ Haruyasu Yoshizaki, Manual for LHA Version 2.13; LHA213. DOC
- ⑭ Jonas I Hammarberg, SQZ—Squeeze v1.08.3 ; SQZ. DOC
- ⑮ NoGate Consulting, PAK——The File Compression Utility Version 2.51; PAK. DOC
- ⑯ PKWARE, Inc. , PKZIP 2 —— The next generation in compression software (SHARE-WARE MANUAL); MANUAL. DOC
- ⑰ Technical background material; APPNOTE. TXT (PKPAK Version 3.6)

第一章 文件压缩存档软件概述

文件压缩存档软件也叫档案实用程序,俗称打包软件。其主要功能是将多个文件压缩拷贝到一个文件中存放,以节省磁盘空间,待使用时再释放还原。

目前国内流行的文件压缩存档程序一般都没有使用说明书,特别是在流传过程中,往往只注重可执行的应用部分,而忽略了软件提供的资料信息文件。为了弥补这一缺陷,笔者将应用中理解的软件使用方法和技巧整理出来,供读者参考。

1.1 用途及要点

1.1.1 文件压缩存档软件的用途

1. 释放占用的硬盘空间

如果你的资料很多,且长时间不使用,占用了大量的磁盘空间,这时就应该将它们打包存放,需要时再释放还原。

2. 节省软盘片

① 用较少的盘片提供软件

随着计算机应用软件的大量涌现,一个微机用户手头可能就拥有几十种,占用了大量的软盘片。如果将其压缩保存,不是可以节省你许多磁盘开销吗?精明的软件开发商已经采用了文件压缩方法来提供他们的产品,以尽可能少的盘片交付用户。例如著名的美国微软公司在交付 MS-DOS 6.2 版软件时,也使用了两个自解压缩档案 1MSDOS62.EXE 和 2MSDOS62.EXE 来存放 73 个 DOS 升级文件。

② 用较少的盘片保存大量的数据信息

比如,我们土壤肥料研究所担负着对全国土壤肥力变化的监测任务。分布在各地的监测基地每年都要上报大量的监测数据,包括土壤物理、土壤生化、土壤化学、田间气象及农作物消耗土壤养分的多项测定数据等。采用文件压缩的方法保存数据库早期数据,既可节省磁盘又便于管理。下面列表说明用文件压缩方法保存的多年数据:

1988. ARC	—— 用 PKARC 建立的年度档案文件,必须用 PKXARC 释放、处理
1989. ARC	
1990. ARC	
1991. ZIP	—— 用 PKZIP 建立的年度档案文件,需要用 PKUNZIP 释放、处理
1991. EXE	—— 用 ZIP2EXE 程序(由 PKZIP 软件提供)将 1991. ZIP 转换为自取档案文件 1991. EXE,运行即可释放
1992. LZH	—— 用 LHA 建立的年度档案文件,用其自身的命令就可释放、处理
1992. EXE	—— 用 LHA 的 S 命令建立的自取档案文件,只需运行便可释放

1993. ARJ	——	用 ARJ 建立的年度档案文件,用其自身的命令就可释放、处理
1993. EXE	——	用 ARJ 的 A -JE 命令选项建立的自取档案文件,只需运行便可释放
80-87. ARJ	}	用 ARJ 的 -V 命令选项压缩保存在三张软盘上的多卷档案文件,记录了 1980-1987 年度的所有数据库信息。
80-87. A01		
80-87. A02		

上面列表中是使用不同的文件压缩存档软件生成的档案文件,而在实际应用中一般应选用一种文件压缩软件。

再比如,一个财会人员可以把每年的帐目文件经压缩后保存在一个年度档案文件中,在你的软盘上所能见到的只有一系列年度档案文件,再不像过去那样杂乱无章了。

3. 方便交换数据

在办公自动化、财会电算化和计算机管理信息系统中,数据量越来越大,数据交换也越来越多,上传下递无不需要使用磁盘。进行软件或磁盘文件交换时,采用压缩副本岂不十分方便?

4. 减少病毒感染机会

对计算机病毒,真可谓防不胜防,新型病毒总是走在防病毒软件的前头,所以避免病毒造成损失的最好办法还是及时保存副本,同时,压缩形式的文件不易受到病毒的侵害,它能确保你的数据安全。

另外,文件压缩存档软件中能有效防范病毒的程序,是那些可以运行压缩档案中可执行文件的软件(例如 ARC 和 LOOZ),因为这类软件可以直接从档案中运行程序(它们执行的方法不同,一种是把档案中的文件释放到磁盘上,待运行后再删除;另一种是把档案中的文件直接释放到内存执行),所以你可以把不常用的或者怕被病毒感染的程序用这类软件打包保存,待需要时用命令从档案中直接调用。这样做虽然启动速度会慢一点,但能保证你的程序安全。

5. 用口令为存档文件加密

你还可以使用口令来压缩保存文件,这有一定的加密作用,不知道口令的人是无法正确释放档案文件的。也就是说,要么压缩文件无法释放,要么释放出的文件根本不能使用,不过你自己一定不要忘记口令。

6. 提高通信效率

文件压缩软件最初是为解决网络和通信传输而设计的,目的是要加快信息传递速度,节省通信费用。同时,打包的文件在发送和接收时都更加方便,因为在通信业务中,为防止数据丢失或出错,每份文件通常要发送两到三遍,如果将许多零散的小文件独立发送或接收,那么重复的报文头就要占用相当大的信息量,同时分散的多个小文件也不如打包的大文件容易查找和处理。

随着远程工作(Telework)岗位的不断增长(仅欧美国家的远程工作岗位数目就将从目前的 60 万个,发展到 1997 年的 530 万个,预计到 2000 年将增加到 1170 万个),信息交换量也必定扩大。据伦敦 Ovum 公司调查,在 1993 至 1997 年期间,这一领域的硬件配置需支付 100 亿美元,软件投资需 290 亿美元,数据通信设备需 72 亿美元,而通过 modem 进行的数据交换就需要用户支出 250 亿美元,可见通信费用之高。倘若能充分地利用数据压缩技术,必定提高传输效率,降低数据信息交换的费用,得到非常可观的经济效益和社会效益。

总之,打包软件将成为你用计算机管理、传送数据信息及软件的好帮手。

1.1.2 术语和约定

这里介绍一些在使用文件压缩存档软件时将遇到的术语。

1. 建立压缩档案文件的有关术语

archive (档案)或 arcspec (档案文件标识)

ZIPfile(打包文件, ZIP 是拉拉链的意思, 可以理解为把文件紧密地联结在一起; 也可认为是把文件装入一个“包裹”中, 另外, 这类软件名称中常冠有“PK”字样, PK 来自 PacKing—“打包”)

档案和打包文件都是由文件压缩存档程序生成的文件, 这个文件的特点是:

(1) 装在其中的文件多数都是压缩文件(The compressed files), 压缩率一般在 30%~50%。

(2) 可以把众多的压缩文件装入一个档案, 例如早期的 PKZIP 软件可以在一个档案文件中存放最多 3900 个文件, 而现在只要内存足够, 压缩档案中的文件数目将不受限制, 假如有 500K 以上的内存, 大约可以在一个档案中容纳 8000 个文件。

正因为档案文件的最大特点就是能保存大量的压缩文件, 所以许多人也常称之为“压缩文件”, 比如从档案文件中取出文件就被称为“释放压缩文件”。

(3) 档案文件有其特殊的格式, 所以必须使用软件自身提供的命令去操作。比如你用 PKZIP 软件建立的档案文件, 还要用它提供的命令或与之配套的 PKUNZIP 程序中的命令来处理, 就连自取档案文件也是如此(除了能自己释放外)。当你使用的 PKUNZIP 与建立档案文件所用的 PKZIP 不是同一版本时, 操作都无法进行, 程序将给出信息“I don't know how to handle”, 告诉你不知如何处理。也有个别的软件可以将其它类型的档案文件转换成自己能处理的格式。

不同版本的文件压缩存档程序所建立的档案文件通常都采用不同的文件扩展名, 常见的有如下几种:

.ZOO .ARC .PAK .ZIP .LZH .ARJ .SQZ

见到带有这类扩展名的文件, 你就可以断定它们是压缩档案文件了。

(4) 有一种档案文件的扩展名是 EXE, 这类压缩档案被称为“自取档案文件”和“自释放档案文件”, 有关的细节请参见 1.1.3 节。

(5) 一般都是使用文件压缩存档软件的 A 命令来建立档案文件, 这一过程被称为“压缩文件”、“打包”或“建立档案文件”。

下面是不同打包软件显示压缩文件过程时使用的术语:

deflating	(“缩小”, 排除空隙使文件尽量缩小, 即压缩文件)
crunching	(“碾压”, 指压缩文件)
squashing	(“压扁”, 指压缩文件)
shrinking	(“收缩”, 指压缩文件)
imploding	(“内部破碎”, 亦指压缩文件)
freezing	(“冻结”, 指压缩文件)
packing	(“打包”, 即压缩文件)
squeezing	(“压榨”, 即压缩文件)
distill	(“提取精华”, 指压缩文件)
storing	(“储存”, 即不压缩, 仅保存文件)

2. 释放压缩档案文件的有关术语

Extract (取出、提取)

严格地说, Extract 操作就是把存放在档案中的压缩文件取出来, 可是, 因为在取出文件的同时, 软件需要把压缩的文件展开, 即从压缩态还原, 且从不能使用到可以使用, 按照汉语的意思——把禁锢的东西放出来, 完全可以称其为“释放”; 而“解包”是相对于“打包”而言的民间说法。不过无论怎样称呼, 只要理解其所表示的操作过程就可以了。

下面是不同打包软件显示文件释放过程时使用的术语:

extracting (从档案中取出压缩文件, 即释放压缩文件)
 uncompressing (将压缩文件展开, 即“压缩”的逆操作)
 melting (“融化”, “冻结”的逆操作, 指释放压缩文件)
 expanding (胀大、展开, 指释放文件)
 unshrinking (shrinking 的逆操作, 即展开压缩文件)
 invoking (引用档案中的文件, 出现于执行档案中的可执行文件时)
 inflating (膨胀, 使文件胀大, 用于展开压缩文件)

3. “命令”、“命令选项”(“命令开关”)及“参数”

Command (命令)
 Options (选项)
 Switches (开关)
 Parameter (参数)

以上列出的英文名称均是在使用文件压缩存档程序要遇到和使用的术语。为了说明这些术语在命令行中的位置及作用, 让我们先看一下典型的命令行格式:

```
PKARC [-oct, -nct] options[g<passwd>] archive [filename... ]
PKXARC [options] archive [d:path\] [file... ]
PKZIP [-b[path]] [options] zipfile [@list] [files... ]
PKUNZIP [options] zipfile [d:path\] [@list] [file... ]
LHA <command> [/option] <archive[.LZH]> [DIR\] [@list|filenames]
ARJ <command> [-<-sw> [-<sw>...]] <archive_name> [<list|file_names>...]
```

(1) 关于一级命令的约定

当文件压缩存档软件只提供一级命令时, 并且允许它们组合使用(在组合使用时这些命令的排放位置没有严格限制), 所以认定它们都是跟在程序名后面的选项(实际上是把程序名作为命令), 而把执行程序时所要求的具体操作选择定义为“options”。

例如:

```
PKZIP -A -XCOM.BAT FB *.BAT
PKZIP -XCOM.BAT -A FB *.BAT
```

这两条命令的结果是一样的, 都是把当前目录下除 COM.BAT 文件外的所有批处理文件. BAT 压缩保存到档案文件 FB.ZIP 中。

(2) 关于两级命令的约定

那些提供了二级命令的软件, 就严格要求二级命令必须跟在一级命令的后面, 并且二级命令是不可单独执行的。也就是说, 在一个命令行中可以没有二级命令, 但不能没有一级命令。为使它们

的身份有明显的区别,所以才用“命令”和“选项”来确定其权限,“选项”总是为命令服务的。在命令行中,“命令”的前面没有符号,而“选项”的前面必须有减号“-”或斜杠“/”。LHA 和 ARJ 都是提供了二级命令的文件压缩存档软件,它们决不会接受下面这样的命令:

LHA -A COM *.COM (如果 COM.LZH 存在,则仅显示其内容,无论直接使用哪个选项均如此)

或 ARJ -N FB (显示命令求助信息,无论直接使用哪个选项均如此)

所以,一定要在“命令”的后面使用“选项”。正确地输入命令行应该是:

LHA A -A COM *.COM (使用 LHA 的 A 命令及 -A 选项,不顾文件属性的限制,把所有的 .COM 文件,包括系统和隐含文件都压缩存入 COM.LZH 中)

或 ARJ A -N FB (使用 ARJ 的 A 命令及 -N 选项,把当前目录下的新文件,即档案中没有的文件,加入 FB.ARJ 文件中)

在命令行中,“命令”只能有一个,而“选项”可以有多个。

(3) 关于命令选项和命令开关

“命令选项”和“命令开关”是称呼不同、意义相同的计算机术语,它们几乎是与计算机操作命令同时诞生的,只是那时也被称为 qualifiers——“命令限定”,不过含义一样。计算机软件(系统软件或应用软件)的一个命令常常可以提供多项服务或输出多个不同的结果,就拿最简单的 DOS 命令 DIR 来说,也可以用七种方式来显示文件目录清单(这里以 MS-DOS 5 为例):

/P 显示时满屏暂停

/W 使用宽的格式来显示信息

/A 显示指定属性的文件

/O 用排序列表方式显示

/S 显示所有子目录的文件列表

/B 显示简单列表

/L 用小写字母显示文件列表

可见,命令选项是扩展命令功能的最好方法。人们也常用 switches(开关)来表示命令的可选项,这是因为 switch 在专业术语中表示计算机程序中的参数,它用来控制分支且在到达分支点之前必须置值,而一个带有多个命令或可选项的软件即是一个具有多个控制分支的程序,而操作者所输入的选择即是为之置值。所以使用“开关”来表示可选操作功能,更符合专业习惯。

(4) 关于参数

下面有两个带参数选择的命令选项示例:

-p[a,b,c][1,2,3]

-m[0,1,2,3,4]

方括号中的字符就是可选参数。参数有几个特点:

① 命令或命令选项都可以有可选参数。虽然在上面示例中 -p 的前面也标记有“-”号(PKUNZIP 约定的命令格式),但它仍是一个命令,第二条是 ARJ 的 -m 命令选项。

② 直接跟在命令或命令选项后面,比如上面第一条, -p 是 PKUNZIP 的一个命令,可以把文件释放到打印机,每个参数的意义是:

a ASCII 方式(该操作每打印一个文件后都走纸,并从下一页开始新的文件);

b 二进制方式;

c 指定串口打印机(缺省为并行口)

1,2,3 通讯口号(缺省为1)

如果选择参数为:ac1,那么-p命令的正确格式是:-pac1。若写成-p -a -cl就错了,PKUNZIP肯定不认识它。

③ 参数后面仍允许有可选参数。第一个示例就有两组参数,第二组参数是为第一组参数中的c准备的,当选择用串行口打印机输出时,还可以进一步指定用哪个串口。

④ 参数可以缺省。在软件编制时,一般都命令准备了缺省参数,这些参数多为常用输出形式或最佳效果形式,因而可以说,参数往往是为一些特殊需要准备的。比如第二个示例的-m选项是为ARJ程序的文件压缩命令(包括:A、U、F等)指定压缩方式的:

-m[0,1,2,3,4]: 选择用方式0,1,2,3,4压缩

m0: 仅保存不压缩

m1: 最大限度地压缩(是缺省方式)

m2: 压缩时占用较少的内存(压缩比同m1,但压缩速度较慢)

m3: 快速压缩(压缩量较少)

m4: 最快速压缩(压缩量最少)

从上面每个参数的意义不难看出,缺省方式——最大限度地压缩,显然是人们最常用的方式了。然而,当你在操作时,磁盘空间绰绰有余,但时间紧迫,不妨采用参数4,用最快的速度完成文件压缩工作,不是更好吗?

也正因为参数有缺省值,所以在带可选参数的命令或选项时,应首先查看有关参数的缺省值,以便确定适合自己需要的选择。

4. 关于“预选文件列表”

Pre-selected file list (预选文件列表)

预选文件列表的用途是用来指定要处理(或不被处理)的一批文件。

(1) 什么是预选文件列表

“预选文件列表”实际上是一个含有多个文件名的文本文件,可以用COPY CON filename直接由控制台建立或者使用字处理软件建立。例如有个名为LXP.LST的预选文件列表,它的内容如下:

```
FORTRAN.LIB
*.EXE
USER.*
INC???.FOR
; This is a comment
emp\census.txt ; This is also a valid comment
E:\DATA\FER.WK1
E:\PRO\QC\*.*
D:\QWK\*.REP
@C:\PK\PKB.LST
```

这个文件的用途是,在使用文件压缩存档程序时想要处理多个文件,但又不便于用文件通配符(即万用符*或?)在一个命令行中表示出来,即使在命令行可以输入也十分麻烦,特别是当需要使用多次命令重复操作这些文件时,用压缩软件规定的符号调用这个文件将十分方便快捷。在上面的

例子中,有两处由分号“;”引出的注释说明,一个是独立的注释行,另一个是在文件说明后面加注的说明。

(2) 关于预选文件列表的一般约定

① 必须是 ASCII 格式文件。

② 预选文件列表中的文件名最好按列输入(如上面例子的格式),这是因为,多数文件压缩软件命令执行时都不接受按行顺序排列的文件名(本书提及的软件中只有 LHA 可以接受横排格式)。

③ 可以为不同目的的需要建立多个预选文件列表。

④ 预选文件列表中仍可以使用文件通配符“*”或“?”。

⑤ 预选文件列表中的文件名可以使用 DOS 支持的所有形式,如包括驱动器名和目录名的文件名。

⑥ 在预选文件列表中还可以调用其它的预选列表文件。

⑦ 有的软件允许在预选文件列表加注释说明。

⑧ 预选文件列表的文件名可以不要扩展名,但通常为了容易辨认,最好加上文件名后缀。

LST。

(3) 怎样调用预选文件列表

① 一般的调用格式

多数文件压缩存档软件都使用符号@或!来调用预选文件列表,有的软件为避免与文件名首字母冲突,可以由用户根据自己的需要在配置文件中定义。下面是不同打包软件调用预选文件列表的命令格式:

PKZIP -C FOR @FOR.LST	给 FOR.ZIP 中由 FOR.LST 预选列表文件所定义的文件加入注释)
PKUNZIP -E FOR @FOR.LST	(从 FOR.ZIP 释放文件,只释放 FOR.LST 中包含的文件)
LHA V FOR @FOR.LST	(显示 FOR.LZH 的内容,只显示 FOR.LST 中所包含的文件信息)
ARJ D FOR ! FOR.LST	(从 FOR.ARJ 中删除 FOR.LST 指定的文件)
SQZ A FOR @FOR.LST	(把 FOR.LST 中指定的文件压缩后存入档案文件 FOR.SQZ)

预选文件列表简化了命令行的输入。

② 所允许的多样化格式

下面是 PKPAK/PKUNPAK 软件使用预选文件列表的几种命令格式:

● 在命令行中既有用文件标识符指定的文件,又有预选文件列表:

PKPAK F FOR *.BAT @LXP.LST

这条命令将更新档案文件 FOR.ARC 中所有的.BAT 文件和由预选文件列表 LXP.LST 所指定的所有文件。

● 在命令行中使用多个预选文件列表:

PKPAK A FOR @LXP.LST @DATA.LST

这条命令中出现了两个预选文件列表,PKPAK 软件将把它们所指定的文件全部加入档案 FOR.ARC 中。

5. 关于口令的说明

口令也被称为“关键字”，是用来为档案或档案中的文件加密的。档案中的所有文件可以使用相同的口令，也可以为某一个文件或某一些文件定义专门的口令。下面以 PKZIP 软件为例来说明与口令有关的几个问题。

(1) 口令的几种形式

① 用一句话作为口令(长口令形式)

下面是一个例子：

```
C:\>PKZIP TEST.ZIP -S"MARY HAD A LITTLE LAMB" *.DBF
```

其中英文“MARY HAD A LITTLE LAMB”(玛丽有只小羊羔)被作为所有加入档案 TEST.ZIP 的数据库文件的口令。

② 用几个字符作为口令(短口令形式)

```
C:\>PKZIP TEST.ZIP -S"LXP" SOIL.DBF
```

在这条命令中口令只有简单的三个字符，“LXP”作为文件 SOIL.DBF 的口令被加入档案文件 TEST.ZIP 中。

(2) 输入口令的方法

① 直接在命令行中输入

上面“口令的几种形式”中所给出的例子使用的都是命令行直接输入方式，这种方式由于受到 DOS 命令行长度的限制，一般口令都不可能太长。

② 经压缩软件提示后输入

```
C:\>PKZIP TEST.ZIP -S SOIL.DBF
```

由于在命令行中没有给出具体的口令，所以 PKZIP 软件会提示你输入口令：

Password ?

这时便可以输入口令，由于不再受 DOS 命令行长度的限制，通常可以输入较长的口令，PKZIP 限制不超过 64 个字符。不过从以后使用文件的角度考虑，太长的口令肯定不方便。

本书 2.1.1 节、2.1.2 节、5.4.2 节介绍了用口令加密和解密的应用实例。

(3) 不同软件在用口令加密时使用的术语

encrypting (“加密”，用口令给文件加密)

scramble (“搅乱”，用口令使被压缩的文件“杂乱无章”，起到加密作用)

garble (“歪曲，篡改”，用口令使文件不能正常操作)

6. 命令行中的有关符号

— 和 / (引出命令或选项的标识符)

[] (可缺省参数或选项符号，表示其中的内容可缺省)

< > (必选项符号，表示其中指定的参数必须输入)

{ } (必选项符号，表示其中指定的参数必须输入)

| (相当于“或”，即该符号两侧的参数可从中选一)

@ 和 ! (预选文件列表调用符号，它的后面跟预选列表文件名)

+ 和 - (某些命令开关选项的开启和关闭)

7. 命令行中字母的大小写限制

在多数情况下打包软件对命令行中的字母不限制大小写，也就是说，它认为大小写都是一样的，但特殊情况例外，那就是当打包软件定义某个字母的大小写分别表示不同的命令或选项，例如：

PKZIP 软件在处理日期限制时将字母 T 定义为指定日期前;而该字母的小写 t 表示指定日期后。这类命令格式的说明中常用符号< | >来表示,比如<T|t>、<P|p>、<J|j>……,对这类命令在输入时应该严格遵守软件的约定,决不能忽略字母的大小写。这个限制仅对“命令”本身,而命令行中的其它内容可以忽略大小写。

8. 关于压缩率和压缩比

不同文件压缩存档软件表示压缩量的方法不同,有的用与被压缩掉的量之比来表示,而有的则用与压缩后剩余的部分之比来表示,为了区别它们,我们用压缩率和压缩比这两个不同的词来定义。下面是两个不同软件压缩同一文件的压缩量比较:

软件名称	压缩前大小	压缩后大小	压缩量及定义	计 算 公 式
PKARC	734	323	56%(压缩率)	$734 \times (1 - 0.56) = 323$
ARJ	734	154	0.21(压缩比)	$734 \times 0.21 = 154$

从表中可以看出, PKARC 软件把文件压缩掉了 56%,而 ARJ 软件是用压缩后与压缩前的文件大小之比来表示压缩量的。

1.1.3 关于自取档案文件

1. 什么是自取档案文件

Self Extracting archive (自取档案文件)

Self-unpacking archive (自解包档案文件)

“自取档案文件”也叫“自解包档案文件”,或称“自释放压缩文件”,这种文件有如下特点:

(1) 它的文件扩展名是 .EXE 而不是通常的 .ZIP、.ARJ 等,只要将其运行就可把压缩保存在其中的文件全部释放出来。

(2) 它不需要文件压缩存档软件的支持便可从档案中独立取出(释放)压缩文件,故被称为“自取档案文件”。

(3) 仍可使用文件压缩存档软件提供的多数命令或选项来处理“自取档案文件”,像处理其它档案文件一样,例如:

ARJ V FOR (用 ARJ 的 V 命令显示档案文件 FOR. ARJ 的内容)

ARJ V FOR. EXE (用 ARJ 的 V 命令显示自取档案文件 FOR. EXE 的内容)

必须注意的一点是,在命令行中自取档案文件的扩展名 .EXE 不能省略,否则文件压缩软件会以为你是要处理一个标准档案文件,如上例中的 .ARJ 类型文件。

(4) 自取档案文件总是比装有同样内容的标准档案文件大,如上例中两种不同方式的档案文件的列表:

Directory of C:\FORTRAN

FOR ARJ 265601 07-26-93 9:49a

FOR EXE 280459 07-26-93 9:53a

2. 自取档案文件的两种模式

针对自取档案文件的不同用途,软件开发商提供了建立两种模式的自取档案:小的(small)自

取档案文件和大的 (large) 自取档案。这两种模式也可称为“简功能”和“全功能”自取档案文件。本人在计算机世界报 1993 年 44 期 171 版中曾经介绍了 ARJ 软件可以建立两种方式的自释放档案文件, 其中 SFXJR 就是小的自取档案, 而 ARJSFX 就是大的自取档案。有的读者来信询问其它压缩软件是否也有这种能力, 笔者将有关问题作一简要说明。

(1) 能建立两种模式自取档案的文件压缩存档软件

下表给出不同文件压缩存档软件建立的两种模式自取档案的效果比较(注: 这些软件均采用缺省压缩方式操作), 它们都是将共有 137365 字节的若干个文件压缩成不同模式的自取档案, 表中的数字表示文件的字节数。

文件压缩软件	小模式自取档案	大模式自取档案	大一 - 小 =
LHA	69165	69474	309
LHARC	71843	72418	575
LHICE	71843	72418	575
ARJ	72314	81843	9529
SQZ	80051	87201	7150
PKZIP	79739	89589	9850

由于 LH 系列压缩软件建立的大模式自取档案提供的命令也很少, 所以两种模式文件的大小差异不大。

(2) 两种模式自取档案文件的特点

① 小模式自取档案的主要特点

a. 生成的自取档案文件比较小, 通常比大模式的自取档案文件小几十到几千个字节, 见上表。这种模式特别适合软件开发商, 或用于智能化电子产品中固化软件。他们不需要自取档案提供更多的其它功能, 只要能适时释放压缩文件。为了节省记录媒体和存储芯片, 降低产品成本, 所以希望自取压缩文件越小越好。

b. 由于小型自取档案文件一般不处理口令及覆盖写操作等询问, 所以释放文件时速度要快一些。

c. 还有资料介绍, 小的自取档案的设计是使之适合工作于内存, 而大的则依赖于磁盘可用空间(见 ICE User's Manual)。

② 大模式自取档案的特点

a. 它能够提供较多的功能, 如处理口令、处理子目录等。例如, LH 系列的大模式自取档案可以把文件释放到指定目录中, 如果该目录不存在, 它会自动建立目录, 而 LH 系列小模式自取档案只能把文件释放到当前位置。

b. 除 LH 系列的压缩软件外, 多数大模式自取档案都提供若干个操作命令, 如列表显示档案中的文件情况、把文件释放到标准输出设备、档案完整性测试等, 使用户在释放文件时能根据实际需要作出选择, 有较大的灵活性, 使用起来比较方便。所以这种模式更适合于一般计算机用户。

(3) 自取档案文件模式类型的表示方法

当你拿到一个自取档案文件时, 怎样才能识别它的模式类型呢?

当你自己建立了自取档案文件后, 怎样才能确认它是大模式的还是小模式的呢?

每个自取档案文件在运行时,都会显示其软件版本及档案模式类型信息,它们通常都显示在屏幕信息的第一行。但是不同压缩软件表示自取档案模式类型的标记方法不同,笔者收集到的文件压缩存档软件中主要有如下三种表示方法:

① 用 SFX 和 SFXJR 表示

SFX —— 表示大模式自取档案文件

SFXJR —— 表示小模式自取档案文件,字母 JR 取自英文“JunioR”,意思是等级较低的。

例如 ARJ 和 SQZ 软件都使用上述表示方法,当运行自取档案文件时,输出屏幕显示的信息中会给出有关自取档案模式的信息。

a. ARJ 软件表示自取档案模式的信息

● 大模式自取档案文件:

ARJSFX 2.30 Copyright (c) 1990,91 Robert K Jung. All rights reserved.

ARJSFX -? for help.

Usage: ARJSFX [-command] [-switch(s)] [directory\] [file(s)]

Commands:

e: Extract files	v: Verbosely list contents
l: List contents	x: eXtract files with pathname (default)
t: Test contents	

Switches:

a: show ANSI comments	n: only New files (not exist)
c: skip time stamp Check	p: match with Pathname
f: Freshen existing files	u: Update files (new + newer)
g: unGarble with password	y: assume Yes on queries

这些信息中除了在开始处的“ARJSFX”字样表示该自取档案是大模式类型外,还提供了软件版本、软件作者和自取档案的有关命令格式及开关选项说明。

● 小模式自取档案文件:

ARJSFXJR 2.30 Copyright (c) 1990,91 Robert K Jung. All Rights Reserved.

“ARJSFXJR”字样表示该自取档案文件是小模式类型的。

b. SQZ 软件表示自取档案模式的信息

● 大模式自取档案文件:

SQZ —— SFX, v1.02.3, Sep 11 1992, Copyright J I Hammarberg (/? for help)

Syntax: sqz command /! -[options] [destination directory\|/] [file ...]

x: eXtract files from archive.	e: Extract files w/o path.
t: Test files in archive.	p: Display files on standard output.
v: Verbose listing of contents.	l: List contents of archive.

Options: MPASHR; Method/Path/Archive/System/Hidden/Readonly

d[a|b]ymmddhhmmss Date after|before.

Set afterwards|Let it be|Reset.

o[0|1|2] Overwrite if newer|always overwrite|ask first.

● 小模式自取档案文件:

SQZ —— SFXjr, v1.02.3, Sep 11 1992, Copyright J I Hammarberg

② 用 L(Large)和 S(Small)表示

L —— 表示大模式自取档案文件

S —— 表示小模式自取档案文件

LH 系列的文件压缩存档实用程序(LHARC、ICE 及 LHA)用这一方法表示,它们会在屏幕信息中的软件版本号后面用“L”或“S”字样标记自取档案的模式类型。

a. LHARC 软件表示自取档案模式的信息

● 大模式自取档案文件:

LHarc's SFX 1.13L (c) Yoshi, 1989

请注意“SFX 1.13L”中的“L”字样,它表示该自取档案文件是大模式类型的。信息行中 Yoshi 是软件作者 Haruyasu Yoshizaki 的缩写形式。

● 小模式自取档案文件:

LHarc's SFX 1.13S (c) Yoshi, 1989

信息行中“SFX 1.13S”表示该自取档案是小模式类型。

b. ICE 软件表示自取档案模式的信息

● 大模式自取档案文件:

LHice's SFX 1.14L (c) Yoshi, 1989

● 小模式自取档案文件:

LHice's SFX 1.14S (c) Yoshi, 1989

c. LHA 软件表示自取档案模式的信息

● 大模式自取档案文件:

LHA's SFX 2.13L (c) Yoshi, 1991

● 小模式自取档案文件:

LHA's SFX 2.13S (c) Yoshi, 1991

大家熟悉的汉字系统软件 2.13H,其安装时所运行的自取压缩文件模式是 LHice's SFX 1.14S,而 2.13I 所用的则是 LHA's SFX 2.12S,都是缩小模式的自取档案文件,一个是用 ICE 软件建立的,另一个是用 LHA 软件建立的。

③ 用 Mini 表示小模式自取档案

—— 表示大模式自取档案文件

Mini —— 表示小模式自取档案

下面是 PKZIP/PKUNZIP/PKSFX 软件表示自取档案模式的方法:

● 大模式自取档案文件:

PKSFX (R) FAST! Self Extract Utility Version 1.1 03--15--90

Copr. 1989--1990 PKWARE Inc. All Rights Reserved. PKSFX/h for help

PKSFX Reg. U.S. Pat. and Tm. Off.

大模式自取档案可以用/H 命令获得帮助信息。

● 小模式自取档案文件:

PKSFX (R) FAST!! Mini Self Extract Utility Version 1.1 03--15--90

Copr. 1989--1990 PKWARE Inc. All Rights Reserved.

PKSFX Reg. U.S. Pat. and Tm. Off.

它用“Mini”表示是小模式的自取档案,同时没有帮助命令。

3. 建立自取档案的几种典型方法

不同软件建立自取档案文件的方法不同,一般有四种类型:

(1) 用专门的程序转换

先用文件压缩存档程序建立标准档案文件,而后再用专门建立自取档案文件的程序将标准档案文件转换成自取档案文件。例如 PKZIP 软件就是如此,必须先建立一个 .ZIP 文件,再用 ZIP2EXE 程序将 .ZIP 文件转换为自取档案文件 .EXE。

(2) 用专门的命令转换

这一种与第一种方法很相似,所不同的是,它使用文件压缩软件自身的命令就可完成从标准档案文件到自取档案文件的转换,例如 LHA 软件就是如此,下面是它建立自取档案的操作示例:

LHA A FOR (用 LHA 的 A 命令建立标准档案文件 FOR.LZH,如果你想省略这一步,LHA 将显示信息:Archive not found: 'FOR.LZH')

LHA S FOR (用 S 命令将 FOR.LZH 转换成可自取的档案文件 FOR.EXE)

(3) 将标准档案中加入自解头

早期的 PKWARE 系列软件(例如 PKPAK、PKARC 和 PKZIP 老版本)多使用这种方法,它需要如下操作过程:

① 建立自解头文件

用 MAKESFX 程序建立自解头文件 PKSFX.PGM,在这一过程中必须用专门的程序支持,例如 PKPAK3.61 版需要 PK361.EXE 支持;PKARC3.5 版需要 PKX35A35.EXE 支持。如果运行 MAKESFX 程序时支持模块不存在,屏幕信息中会显示出它所需要的支持文件名。所以在得到这类软件的时候,应该查看软件的文件清单,看是否有 PKSFX.PGM 文件,或者是 MAKESFX.COM 和相应的支持模块。

② 把自解头加入标准档案

使用如下 DOS 命令完成这一操作:

COPY/B PKSFX.PGM+标准档案[自取档案文件名]

假如有一个标准档案文件名为 LXP.ARC,命令行应该如下:

COPY/B PKSFX.PGM+LXP

由于在命令行中没有给出自取档案文件名,所以新建立的自取档案被命名为 LXP.EXE。COPY 命令的开关/B 表示是进行二进制文件拷贝。

(4) 用命令直接建立

这是最简便的一种方法,它可以直接用文件压缩软件提供的命令或选项生成自取档案文件。例如 ARJ 就是这样,下面是它所用的命令行:

ARJ A -JE FOR (命令行中选项-JE 的功能就是直接建立自取档案文件 FOR.EXE)

下面给出不同软件建立两种模式自取档案的操作命令速查表:

打包软件名称	使用命令格式	
	建立小模式自取档案	建立大模式自取档案
ARJ	ARJ <A U F> -JE1 档案文件名	ARJ <A U F> -JE 档案文件名

SQZ	SQZ S /Z2 档案文件名(用 SFXJR. EXE 支持)	SQZ S 档案文件名(用 SFX. EXE 支持)
LH 系列: 如 LHA	LHA S 档案文件名	LHA S /X 档案文件名
PKZIP 1.10	ZIP2EXE -J 档案文件名	ZIP2EXE 档案文件名

注:其中<A|U|F>表示,可以使用 A 命令,或 U 命令,或 F 命令。

4. 自取档案的释放确认提示

由于自取档案是可执行文件,只要一运行便会释放档案中的文件,如果操作者事先并不知道自己执行的是一个自解档案,往往会造成一些麻烦。为了避免出现被动,有的压缩软件允许在自解档案释放时给出确认提示信息,常见的有两种类型。

(1) 自动给出确认提示

ARJ 软件建立的自取档案,在运行时会自动显示类似如下的提示信息:

Processing archive: C:\FORTRAN\BAT1.EXE (正在处理的档案文件标识)

Archive date : 1991-07-19 16:22:06 (档案文件生成日期)

Continue extraction? (是否继续进行释放操作,回答 N,将放弃释放操作)

其中“Continue extraction?”就是确认询问提示,操作者可以根据屏幕显示的档案文件标识符和档案生成日期等信息,来确定是否要继续进行操作。

ARJ 软件只为大模式的自取档案(使用 -JE 选项建立的)提供确认信息。

(2) 由用户控制确认提示功能

PKZIP 和 LHA 软件的自取档案释放确认功能,必须由用户采用特定的方法设置才能实现。例如 PKZIP 软件需要在档案中用 CTRL+S 作为档案注释内容,而 LHA 软件则需要在档案中加入专门的文件。笔者在本书 2.2.4 节(第 73 页)和 3.3.2 节(第 108 页)将分别对这两个软件的确认提示控制问题作专门叙述。

1.1.4 文件压缩存档操作要点

1. 文件打包前的准备

(1) 确定哪些文件需要打包

以便选择用什么方法指定被压缩文件:

① 如果只有一、两个文件,可以在命令行中指定要处理的文件名。

② 如果有多个相同文件名或文件扩展名的文件要被处理,可以在命令行中用通配符 * 或 ? 来指定。

③ 如果要处理的多个文件不宜用通配符表示,可以建立一个预选列表文件。

(2) 确定是否将子目录一同打包

由于有的文件压缩程序处理带子目录的文件和不带子目录的文件时,所使用的命令不同,因此,在操作之前应作出选择。

(3) 确定是否要保留原来磁盘上的文件

文件压缩程序提供了移入文件命令 M,如果你不需要保留原来磁盘上的文件,应该选用移入文件操作命令。

(4) 确定是否需要多卷操作

多数文件压缩程序没有多卷处理功能,如果你的文件很大、很多,且希望跨卷存放,以便节省盘片,这时就应该选用具有处理多卷能力的压缩软件 ARJ。

(5) 选择压缩方法

有的文件压缩软件能够提供不同的压缩方式,通常考虑到几个方面:

- ① 使文件压缩得尽可能小;
- ② 较快的压缩速度;
- ③ 压缩操作时占用比较小的内存。

你可根据实际需要选择压缩方法。

(6) 根据以上考虑,选用文件压缩程序版本

不同版本的压缩软件所能提供的服务功能是不同的,有的仅能满足一般需要,而有的软件则能提供较多的服务项目(读者可以通过本章 1.2 节中对每个压缩软件功能的介绍了解到)。如果你有特殊需要,就应选用具有特殊处理功能的软件。

2. 建立压缩档案

(1) 主要命令

A(Add) 命令——把指定的文件加入档案,这一方法对于具有备份能力的压缩存档软件可以在档案中为老的同名文件保留备份。

M(Move) 命令——把指定的文件移入档案。

U(Update) 命令——用指定的文件说明更改档案文件,如果指定的文件不存在则把它们加入档案。

F(Freshen) 命令——按指定的文件说明刷新档案,它与 U 命令的区别是,只处理档案中存在的文件,绝不加入新文件。

(2) 处理目录问题

① 跨目录操作

跨目录操作主要有两种情况:一是,处理当前目录时包括其中所有的子目录;二是,在所有的目录下搜索指定的文件。许多软件都用 -R(recurse subdir)选项开关来实现,也有的软件用其它的字母表示,如 PAK 软件用 -I(Include subdir)、SQZ 软件用 -S,而 ARC 软件则用 -Z 选项开关。

② 是否保存目录名

在档案中保存目录名势必要占用一定的空间,但是有时又必需保存目录名,例如把不同目录下的同名文件压缩存档,如果没有目录名的话连压缩操作都无法完成,因为压缩软件无法将它们在档案中区分,所以应该根据实际需要来确定是否保存目录名。

有的软件在使用跨目录操作开关时就自动将目录名一同加入档案(如不想保存目录名还需用其它的选项开关限定,如 ZOO 软件用冒号“:”来限定),而有的软件必须用特定的选项开关来确定保存目录名。

指定保存目录名通常使用 -P、-X 选项开关,有的软件还有其它参数,例如 SQZ 软件要使用 -P2。有关具体细节,请参见软件命令及选项开关的说明。

(3) 处理文件属性问题

文件压缩存档软件处理文件属性(这里主要是指文件的“隐含”和“系统”属性)有三种情况: