doi :10.3969/j.issn.1671-1122.2009.05.021

计算机数据安全删除和隐私保护

尹燕彬，文伟平

（北京大学软件与微电子学院信息安全系，北京 102600）

摘 要：数据安全是计算机安全问题的核心，政府机关、国防、军队等许多有高度安全保密需求的单位要求，计算机上的机密文件删除时必须要彻底的销毁，不能够被恢复。本文介绍了计算机用户可能导致数据不安全的一些隐患，详细地阐述了数据恢复和数据销毁的原理和方法。另外，计算机用户在使用Windows操作系统的时候，用户的一些操作信息会被存储下来，而用户并不知道这些信息的位置，忽略了对这些信息的及时处理，使得用户的操作信息如上网记录长时间的保存在计算机上，造成隐私泄露的隐患。本文针对这些可能造成隐私泄露的隐患提出了解决的方法。

关键词：数据销毁；数据恢复；隐私保护

中图分类号：TP393 文献标识码：A

Computer Data Secure Delete and Privacy Protect

YIN Yan-bin, WEN Wei-ping

(Department of Information Security, SSM, Peking University, Beijing 102600, *China)*

Abstract: Data security is the core of computer security, the highly confidential files in government, Defense Department and army troops need securely and completely delete, without being recovered. This article introduces some hidden troubles that may lead confidential data to unsafe when the user operates the computer, as well as the principle and method of data recover and destroy. Moreover, the Windows OS sometimes stores the information about the users' operations and the users usually do not know where the information stores. This leads to private information files stay on the computer for a very long time and maybe the leak of private information. This article also presents solutions.

Key words: Data Destroy; Data Recover; Privacy Protect

1 数据销毁的误区

Windows用户在日常的操作中，对于数据的删除操作可能存在认识上的误区。以为使用Delete键就能将文件删除。或者以为格式化就能使数据完全不可能恢复，或者认为数据在硬盘上只有唯一的编码存储。这些对数据销毁认识的偏颇，可能导致数据的泄密。

1.1 常规的Delete方法

“删除(Delete)”是删除数据最便捷的方法，如经常采用的“Delete”的系统删除命令。实际上并没有真正地将数据从硬盘上删除，只是将文件的索引删除而已，让操作系统和使用者认为文件已经删除，又可以腾出空间存储新的数据。

这种方法是最不安全的，数据恢复极易恢复被此类方法删除的数据，而且也有很多专门针对此类删除操作进行数据恢复的软件。

1.2 格式化

“格式化”有许多不同的含义：物理的或低级格式化，操作系统的格式化，快速格式化、分区格式化等等。

大多数情况下，普通用户采用的格式化不会影响到硬盘上的数据。格式化仅仅是为操作系统创建一个全新的空的文件索引，将所有的扇区标记为“未使用”的状态，让操作系统认为硬盘上没有文件。因此，格式化后的硬盘数据也是能够恢复的，也就意味着数据的不安全。

1.3 使用无用文件填充

取一些无用的文件等，复制到磁盘上，直到整个磁盘空间填满为止。如果不能完全填满，则选取更小的文件继续填满，直到磁盘的可用空间为0。这种方法虽然能覆写大部分数据，但对于簇中的剩余空间无能为力，而且费时费力。

1.4 数据随意复制泄密

基于对病毒感染、数据丢失的顾虑，用户经常对重要数据进行备份。有时这种备份操作系统自动进行。整个复制过程很隐蔽，不会留下任何痕迹。如果用户对数据进行擦除时，没有对备份数据进行相应处理，就将导致“副本”数据的残留。这给磁盘信息的保护和保密带来很多难以解决的问题[1]。

1.5 磁盘的内部固有机制导致部分数据无法覆盖

老式磁盘的每一个磁道都有数量相同的扇区，但外圈的磁道比内圈的长，敏感数据有可能隐藏在外圈磁道扇区之间的缝隙中。又如，磁盘的缺陷处理机制。对于磁性存储器来说，通常使用映射的方法来替换受损的磁道或扇区，把坏的和磁介质不稳定的扇区记录下来，做成磁盘缺陷列表，写进磁盘的系统保留区，替换掉原来旧的磁盘缺陷列表，并且通常不再对受损的磁道或扇区进行操作。而且，也有部分软件或病毒程序能将某些扇区故意标记为坏扇区。如果在磁盘的记录间隙、坏的磁道、被故意标记的区域中储存着敏感信息，这些信息仍然可以通过特殊手段被读取。再如纠错机制。许多存储器设备支持不同的纠错方案，以便在设备受到损害时进行数据恢复。因此即使一些数据被可靠地擦除，但通过使用存储器设备内建的纠错能力也可能恢复。支持数据缓冲或高速缓存功能的存储器，操作系统把内存数据写入硬盘前是在缓冲区中收集数据，而写入磁盘上数据的最小单位是一个扇区，因此文件的最后一部分通常不会恰好填满最后一个扇

区[2]，操作系统就会随机提取缓冲区中称为内存渣滓的数据来填充空余区域，而这些内存渣滓数据在进行删除操作时可能滞留在缓冲区中没有真正得到执行。最后一个簇中没有用到的扇区就原封不动保留原来称为磁盘渣滓的数据。这些被称为渣滓的地方可能包含大量的敏感信息不能被彻底销毁。

1.6 剩磁效应

所有磁介质都存在剩磁效应问题，磁介质会不同程度地永久性磁化，所以磁介质上记载的信息在一定程度上抹除不净。由于每次写入数据时磁场强度并不完全一致，这种不一致性导致新旧数据之间产生“层次”差。剩余磁化及“层次”差都可能通过高灵敏的显微镜探测方法探测到，经过分析与计算，对原始数据进行“深层信号还原”可以恢复以前的影子数据[1]。

2 数据恢复的方法

由于数据恢复技术的存在，使得数据残留的安全性面临挑战。当今数据恢复路径主要有两种方式：软件方式和实验室方式。

2.1 软件方式

目前用于数据恢复的软件较多，如硬盘诊断软件、反删除、反格式化软件以及数据恢复软件都可以在数据擦除不完全的情况下进行数据恢复。恢复的效果既取决于软件所采用的算法，更取决于数据被重写的程度。

2.2 实验室方式

主要是借助精密设备的协助进行的数据恢复提取。尽管该方法非常复杂，但却是残留数据的最大威胁。对于磁性介质来说，最常用的方法是磁力显微镜法 MFM 。它可以通过最小的样品数量，高分辨率地显示磁性介质磁化时的图像。从而通过分析与计算来恢复出原始数据。

3 数据销毁的方法

3.1 覆写

数据覆写是将非保密数据写入以前存有敏感数据的存储位置的过程。硬盘上的数据都是以二进制的“1”和“0”形式存储的。使用预先定义的无意义、无规律的信息覆盖硬盘上原先存储的数据， 完全覆写后就无法知道原先的数据是“1”还是“0”，也就达到了清除数据的目的。

根据数据覆写时的具体顺序， 软件覆写分为逐位覆写、跳位覆写、随机覆写等模式。根据时间、密级的不同要求， 可组合使用上述模式。美国国防部 Network & Computer Security 的 DOD 5220122_M 标准和北约 NATO 的多次覆写标准规定了覆写数据的次数， 覆写数据的形式。美国国防部订立的磁盘清洗规范， 要求数据必须对所要清除的数据区进行三次覆盖：第一次用一个 8 位的字符覆盖，第二次用该字符的( ) 补码 0 和1 全反转的字符覆盖， 最后再用随机字符覆盖。如先用 0011 0101 覆盖，接着用1100 1010，然后用1001 0111。覆写必须完成的次数与存储介质有关， 有时与其敏感性有关， 有时因国防部门的需求有所不同。在不了解存储器实际编码方式的情况下， 为了尽量增强数据覆写的有效性， 正确确定覆写的次数与覆写数据的样式非常重要。

对要删除的数据的存储位置进行多次覆写的方法， 是数据销毁的有效途径。处理后的硬盘可以循环使用， 适应于密级要求不是很高的场合。特别是需要对某一具体文件进行删除而其它文件不能破坏时， 这种方法更为可取。到目前为止，数据覆写是最安全、最经济的销毁数据的方法之一。

覆写软件必须能确保对介质上所有的可寻址部分执行连续写入。如果在覆写期间发生了错误或坏扇区不能被覆写； 软件本身遭到非授权修改时， 处理后的硬盘仍有恢复数据的可能。 因此该方法不适用于包含高度机密信息的介质。

3.1.1 覆写原理

覆写是指使用预先定义的格式——无意义、无规律的信息来覆盖硬盘上原先存储的数据。这是销毁数据的既有效又可操作的方法。如果数据被“成功”地完全覆写，即使只覆写一次，也可以认为数据是不可恢复的。硬盘上的数据都是以二进制的“1”和“0”形式存储的。完全覆写后就无法知道原先的数据是“1”还是“0”，就达到了清除数据的目的[3]。

3.1.2 覆写的方法

根据覆写时的具体顺序，软件覆写分为逐位覆写、跳位覆写、随机覆写等模式。根据时间、密级的不同要求，可组合使用上述模式。可靠的专业清除软件应同时支持多种模式。

3.1.3 覆写的局限性

既然一次完全的覆写就可以彻底清除数据，那标准为什么还要规定须多次覆写呢[4]？因为磁信号泄露了数据的历史痕迹，可以通过特殊的专业设备来识别痕迹进而恢复被覆盖的数据。这也是相关标准的制定原因。由于磁盘可以重复使用，前面的数据被后面的数据覆写后，前面的数据被还原的可能性就大大降低了，随着被覆写次数的增多，能够被还原的可能性就趋 0，但相应的时间支出也就越多。密级要求的高低对应着不同的标准，低密级要求的就是一次性将磁盘全部覆写，高密级要求则须进行多次多规则覆写。

如果磁头定位系统不是足够精确，那么在进行覆写时，原先的数据不会被完全精确的覆盖。由于磁道的偏移，可以将原先的数据从当前磁道的痕迹辨别出来。但这是需要专门的特殊设备的，而且这种设备也是受控的、限制销售的，通常只有特别的部门才可能拥有[5]；因此，被覆盖的数据使用现有的技术是不可恢复的，到目前为止，数据清除是最安全的、最经济的销毁数据的方法之一。

3.2 物理销毁

3.2.1 消磁（Degauss）

“消磁”是指使用专门的消磁设备来磁化磁盘表面的磁介质。它产生电磁场来磁化磁盘。销毁前， 硬盘盘面上的磁性颗粒沿磁道方向排列，不同的 N/ S 极连接方向分别代表数据“0”或“1”， 对存储介质施加瞬间强磁场， 磁性颗粒就会改变

沿场强方向顺序排列，使介质表面的磁性颗粒极性方向发生改变，失去表示数据的意义。

具体的消磁办法和技术有很多种，但实质上可分为直流消磁法和交流消磁法两种。直流消磁法是使用直流磁头将磁盘上原先记录信息的剩余磁通，全部以一种形式的恒定值所代替。交流消磁法是使用交流磁头将磁盘上原先所记录信息的剩余磁通变得极小。这种方法的消磁效果比直流消磁法要好得多，消磁后磁盘上的残留信息强度可比消磁前下降 90 db，即消磁后能将测试信号减小到初始强度的十亿分之一，满足 NSA/ CSS L14 - 4 - A 规范对信号强度 90db 消减的需求。

消磁时，介质放在强磁场中，为了使消磁更为有效，一般至少要使用相当于磁性介质矫顽磁性 5 倍的磁力，确保信息真正被消磁。如果使用手持式磁铁，除了中间隔一层防止划伤磁盘的保护片，磁铁必须几乎直接接触磁盘。消磁是净化多数磁存储介质的最佳方法，但也有风险。在消磁周期完成之前，介质被拿走、消磁器出现故障或随着使用年限的增长性能降低都会影响消磁效果。如果整个磁性介质上的数据不加选择的被全部销毁，那么消磁是一种有效的方法。对一些曾记载过较高密级信息的磁盘，必须使用消磁技术进行处理。

消磁最突出的特点就是高效，后果是磁盘再也不能使用。如果希望能够循环使用硬盘，就不能够采用这种方法。

3.2.2 物理破坏

常见的是盘片划损、硬盘回炉、外力破损等方法，这些方法的共同点是费时、费力、效果差，基本未被广泛采用。

3.3 化学腐蚀

据 2005 年的一则报道，美国一家科技公司宣称，开发出了一款专门针对军方、银行、高度商业机密用途的（DoD）自毁式硬盘。一旦电脑遭到偷窃、系统遇到不法的入侵，抑或硬盘被拆卸，硬盘内部会自动释放出化学喷雾剂，将所有硬盘盘片磁介质层完全破坏，从而彻底销毁数据。不仅如此，还采用了独特的触发机制，除常规的联网触发、手动触发等触发模式，还可通过内置的 GPS 全球定位系统来确定硬盘许可范围，一旦越界就会触发[5]。

化学腐蚀的方法可以使用浓缩氢碘酸 浓度 55% 到 58% 溶解磁盘表面的三氧化二铁微粒。也可以使用酸活化剂 Dubais Race A 8010 181 7171 和剥离剂 Dubais Race B 8010 181 7170 处理磁鼓记录表面，然后使用工业丙酮 6810 184 4796 清除磁鼓表面的残余物。化学腐蚀的方法只能由得到批准的专业人员在通风良好的环境中进行。

4 数据销毁软件

现在一些数据销毁软件，例如 BCWipe，Eraser 等软件，主要提供的数据销毁的功能就是针对文件，剩余空间和物理磁盘的销毁。所采用的方式都是通过软件对磁盘进行相应的覆写。

4.1 文件文件夹销毁

目前主流的数据销毁软件都是采用覆写的方式来销毁文件。覆写是一种简单、经济的方式。但是，有些机构声称他们可以恢复被重写 21 次的文件。文件的所有者应当权衡文件的价值大小和恢复文件的资源时间开销。文件被覆写一次，软件的方式就不能恢复该文件。当然，文件覆写的次数越多，文件被恢复的可能性就越小。下面是几种覆写的方式：

一次覆写，数据被一次覆写，使用 0 或 1，或者是随机数据；DOD，DOD 是美国国防部数据销毁的一个标准，数据要重写七次，每次先用 0 或 1 来重写，然后再用随机数据进行重写，最后用 0 来重写；GUTMN，数据被覆写 35 次，利用伪随机数字去重写磁盘，并且考虑到不同磁盘的编码方法。

文件的销毁不但包括内容销毁，还包括文件名的销毁。当使用覆写的方式销毁文件内容后，还要将文件名进行随机化处理，从而彻底销毁文件的所有信息。

4.2 剩余空间销毁

在磁盘的剩余空间中，存在着大量的文件信息，包括被用户使用 delete 命令删除掉的信息。如果不对剩余空间进行销毁，则大量数据面临被恢复的危险[6]。

要注意到的问题是，Windows 的磁盘碎片整理程序会在剩余空间中产生大量的可恢复的文件信息。在碎片整理的过程中，一些文件的位置移动了，但原来的位置并没有被其它文件所覆盖，因此这些文件就有可能被恢复，造成信息泄露。因此在碎片整理后，还要进行一次剩余空间销毁。

4.3 物理磁盘销毁

物理磁盘的销毁非常简单，从第一个扇区开始覆写，直到写完最后一个扇区为止。当然，如果一个扇区一个扇区的去覆写，会很慢。所以可以采用一个块作为覆写的一个单位。最后一部分如果不足一个块的大小，则要进行扇区的覆写[7]。

5 Windows 系统使用痕迹的清理

Windows 操作系统为了提高系统运行效率，保证被操作数据的完整性和可靠性，采用了很多机制，如虚拟内存，缓存文件，临时文件等。这些机制在保证 Windows 系统高效运行的同时，也留下了隐私泄露的隐患。因此，我们应该了解 Windows 的所作所为，并及时地清楚一些隐患数据。

5.1 虚拟内存交换页面

内存在计算机中的作用很大，电脑中所有运行的程序都需要经过内存来执行，如果执行的程序很大或很多，就会导致内存消耗殆尽。为了解决这个问题，Windows 中运用了虚拟内存技术，即拿出一部分硬盘空间来充当内存使用，当内存占用完时，电脑就会自动调用硬盘来充当内存，以缓解内存的紧张。举一个例子来说，如果电脑只有 128MB 物理内存的话，当读取一个容量为 200MB 的文件时，就必须要用到比较大的虚拟内存，文件被内存读取之后就会先储存到虚拟内存，等待内存把文件全部储存到虚拟内存之后，跟着就会把虚拟内里储存的文件释放到原来的安装目录里了。

Windows 的所有会话，包括应用程序的剩余数据，word

程序的数据，上网浏览的数据，都被包含在Windows的页面交换文件中，存在在虚拟内存上[6] [7]。

由于虚拟内存的文件是动态创建的，每次Windows启动都会创建一个新的交换文件，这些文件可能会被存储在磁盘上的不同位置，这些数据就有可能造成保密或隐私数据的泄漏[2]。

用户可以进行设置，使操作系统在关机时对虚拟内存文件进行清理。在控制面板—〉管理工具—〉本地安全设置—〉本地策略—〉安全选项，双击“关机:清理虚拟内存页面文件”，启用。

5.2 临时文件

许多应用程序为了提高效率和数据的可靠性，都会在程序运行的过程中对程序的数据创建一个副本，也就是一个临时文件。应用程序在关闭的时候，才会去删除这个临时文件。但是，这个临时文件是存储在磁盘上的，即使创建他的应用程序删除了它，磁盘上的文件信息并没有被擦除掉，除非存储临时文件的磁盘位置被别的文件覆盖。因此，即使用户安全删除了某个原文件，如果没考虑到临时文件，也有可能造成数据泄露[2]。

删除临时文件的方法，开始—〉程序—〉附件—〉系统工具—〉磁盘清理。

5.3 上网记录

用户在使用Windows IE浏览器上网的时候，为了加速浏览，会创建一些临时文件夹。这些文件夹包含了被浏览网页的信息，这些信息可能泄露我们访问的网站和我们的上网习惯，我们称这些信息为上网记录[8]。

上网记录存在于以下目录中：

%userprofile%\Local Settings\Temporary Internet Files\

%userprofile%\Local Settings\History\

%userprofile%\Cookies

同时，Windows还创建了index。dat文件，用来创建对上网记录临时文件的索引。这个文件包含了用户浏览网页的地址等信息。

如果对这些文件进行手动删除的话，需要关闭所有的应用程序，因为某些临时文件可能正在被应用程序使用。删除index。dat的时候还需要关闭explorer.exe进程。

另外，Windows还提供了WinInet函数调用。我们可以使用FindFirstUrlCacheEntry，FindNextUrlCacheEntry，和DeleteUrlCacheEntry这三个函数来完成对上网记录临时文件和对应index。dat文件中索引的删除。

5.4 删除注册项

在Windows NT系统中，注册表示一个数据库，当用户删除注册表中的项时，被删除的项实际上并没有从注册表数据库中移出。彻底删除注册表项还需要重建注册表数据库[6]。

5.5 缩略图缓存文件

在Windows XP/Server 2003操作系统中，当我们在资源管理器中以缩略图或幻灯方式浏览BMP、JPEG、GIF、GIF等格式的图片时，甚至在浏览PDF、HTM格式的文档时，默认设置下会自动生成相关的缩略图缓存，这就是删除之后又重新出现的Thumms.db文件，这些文件总是在图片文件夹下反复出现。

当文件被安全删除后，文件的相关信息还有可能在这些缩略图缓存文件里找到。因此对这些文件的清除也很重要。

如果希望在Windows XP下取消缓存缩略图的功能，可以打开“文件夹选项→查看”对话框，在这里勾选“不缓存缩略图”复选框即可。

6 数据销毁的结果测试

当前很多软件只提供了数据销毁的接口，并未对数据销毁的效果进行评估和检测。用户虽然使用了数据销毁和清理软件，但向软件使用者展示数据销毁的结果也是很必要的。在目前数据销毁软件没有集成这个功能的时候，我们可以使用其他软件来进行测试。例如使用WinHex软件察看是否存在可恢复的文件，如果使用0或1进行覆盖之后，磁盘的剩余空间是否全为0或1。

图1 数据销毁结果测试 （责编 程斌）

参考资料：

[1] 徐菁，朱有佃，赖凡．论磁性存储介质的数据销毁技术．

[2] John R. Mallery Secure File Deletion: Fact or Fiction? 7/16/01, Updated 6/12/06.

[3]“DoD 5220. 22-M National Industry Security Program Operating Manual（NISPOM）”Chapter 8，January 1995.

[4] Peter Gutmann,“Secure Deletion of Data from Magnetic and Solid-State Memory”，1996，July: 22-55. http://www.cs.auckland.ac.nz/~pgut001/pubs/secure_del.html

[5] 王建锋 数据销毁：数据安全领域的重要分支北京瑞星公司

[6] Dean，Sarah.“Disk and File Shredders: A Comparison.” http://www.sdean12.org/Comparison_Shredders.htm （5/27/06）

[7] New Technologies，Inc.“Windows Swap File Defined.” URL:http://www.forensics-intl.com/def7.html ，5/27/06.

[8] Marcel Lambert CodeProject A Cleanup API for Windows_ Free source code and programming help.

作者简介：尹燕彬（1984－），男，北京大学软件与微电子学院，硕士研究生，主要研究方向：网络安全；文伟平（1976－），男，北京大学，副教授，主要研究方向：网络攻击与防范、恶意代码研究、信息系统逆向工程和可信计算技术等。