

计算机取证关键技术分析

金波, 陶明明

公安部第三研究所 上海 200035

上海金诺网络安全技术发展股份有限公司 上海 200122

摘要:

电子证据是一种新的证据类型, 为提高电子证据的证据力, 本文分析了电子取证的取证程序与取证的关键技术, 提出了取证的一般性原则、数据采集方法、取证的设备和装置要求。

关键词:

计算机取证, 电子证据,

Analysis for Key Technology of Computer Forensic

Jin Bo, Tao Mingming

The Third Research Institute of Ministry of Public Security, 200035 Shanghai

Kingnet Security Inc., 200122 Shanghai

Abstract:

Electronic evidence is a sort of new style evidence. To improve the probative value of electronic evidence, the paper analysis computer forensic process and key technology, provided the rule of computer forensic, data acquire method and the requirement of forensic device.

Keywords:

Computer Forensic, Electronic Evidence

1 概述

随着计算机和互联网络技术的迅速发展，电子商务、网络教育、各类网络服务和电子政务在经济社会的人际交往、经营活动中被大量应用。随之，各类经济纠纷、民事纠纷和刑事案件也会时有出现。判定或处置各类纠纷和刑事案件过程中，电子文档已经成为重要证据之一。

许多计算机化的产品中都会存有电子证据，例如：移动电话、PDA、路由器等，也有许多形式的存储介质，包括：硬盘、光盘、U 盘等。另外，网线、电缆甚至空气也能携带数字信息，通过适当的设备，就能将这些数字信息提取出来，以备使用。本文以计算机证据的重要载体—硬盘为例，研究分析计算机取证中的关键技术要求，包括：取证的一般性原则、数据采集方法、取证的设备和装置要求。

2 取证程序

电子证据处理总共分 3 个阶段：证据获取、证据分析和证据表现[1]。

证据获取阶段的工作是固定证据。电子证据容易修改，一旦决定需要获取电子证据，应该首先进行证据固定，防止有用证据的丢失。在本阶段要求将电子证据的状态固定起来，使之在后续的分析、陈述过程中不会改变。并能够在法庭展示证据固定的有效性，比如展示原始证据和固定后证据的 Hash 校验值。

证据分析阶段的工作是分析证据与案件的关联性。电子证据包含的数据量往往很大，而且数据类型往往杂乱无章，收集的所有证据需要进行提取、整理和筛选后才能被使用。在本阶段要求能够对证据进行全面分析，并在全面分析的基础上能够进行数据挖掘和整合，使之清晰呈现案情相关信息。

证据表现阶段要就电子证据与案件的关联性进行陈述。在此阶段要求能够证实电子证据取得的途径、分析过程，并合理引用电子证据分析结果对案情进行陈述。

3 证据获取

当采集电子证据时，应将注意力放在计算机内容而不是硬件上。当从计算机中采集数据时有两种选择，一种是采集所需要的数据，另一种是采集所有的数据。采集所需要的数据有遗失线索和损害证据的风险，因此一般情况下，取证人员将从涉案的计算机硬盘中完整采集

出所有数据。通过硬盘克隆机或者数据获取软件是两种常用的方法。

3.1 应用硬盘克隆机获取证据

从硬盘中采集数据时最直接的方法是在记录了硬盘和主板的连接方式后,将硬盘从计算机上拆卸下来,然后用取证专用的硬盘克隆机制作原硬盘的克隆品[2]。硬盘数据物理复制采集的原理框图如图 1 所示。

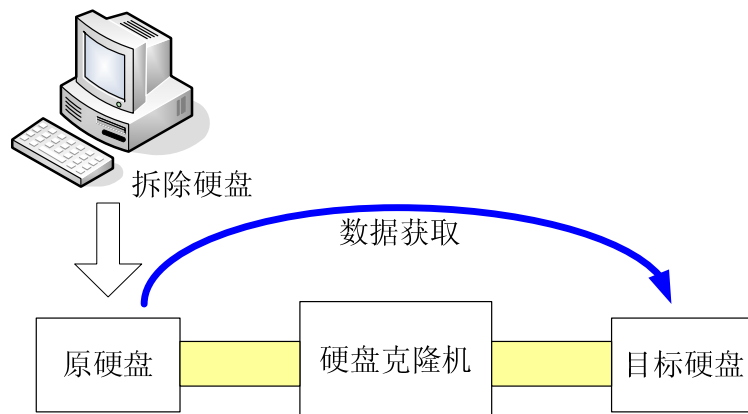


图 1、硬盘数据采集原理框图

硬盘数据物理复制采集可按以下步骤操作:

- (1) 拆除将接受数据取证的硬盘;
- (2) 检测接受数据取证硬盘未被物理故障造成硬盘数据无法读取;
- (3) 打开预备的目标硬盘,对目标硬盘进行格式化处理,清除目标硬盘内所有内容;
- (4) 复制原硬盘数据到目标硬盘。

硬盘复制机必须是专用、特制,具有获取完整数据的复制机。复制机在工作状态时必须对被复制硬盘数据写保护,获取的所有数据应在复制前、后保持一致。复制机应通过物理级复制技术获取文件系统的完整数据,包括文件 Slack 区和未使用的空间,并能提供和原硬盘数据完全一致的副本。

经复制机复制在目标硬盘上的数据应以位 (Bit) 的形式存在,复制机必须将原硬盘的数据全部复制到目标硬盘或镜像文件中。复制范围从硬盘的逻辑第一扇区开始,一直到硬盘逻辑最末扇区结束。复制机应具有复制传输单方向功能;即原硬盘数据向目标硬盘传输,不可逆向。复制机应具备数据校验功能;检验目标硬盘和原硬盘数据完全一致。复制机应具有硬盘擦除和格式化功能;可擦除目标硬盘不正确的数据。或对目标硬盘进行格式化处理。

复制机应遵循严格的工作流程进行操作,确保数据获取的精确性和原始数据的完整性。下图是硬盘复制机工作流程:

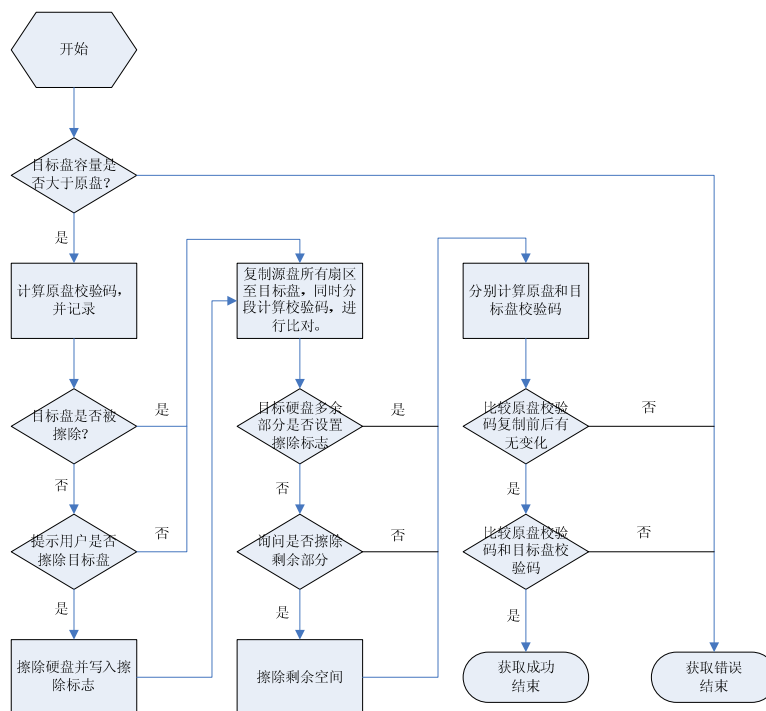


图 2、硬盘复制机工作流程

3.2 数据获取软件获取证据

数据获取软件采集数据是另外一种方法。目前 Windows 和 Linux 下都有相应的数据获取软件。由于 Windows 是会自动在检测到的硬盘上写入数据，因此 Windows 的数据获取软件在使用时必须结合硬盘写保护器进行。硬盘写保护器通过 USB 或 1394 接口和取证计算机连接。

3.2.1 Windows 数据获取软件和硬盘写保护器

其原理如图 3 所示。

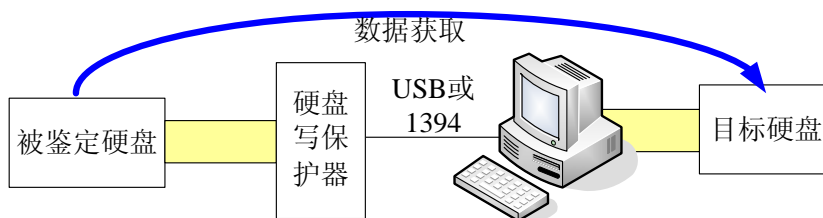


图 3、硬盘数据采集原理图

硬盘数据镜像采集可按以下步骤操作：

- (1) 拆除将接受数据取证的硬盘；

(2) 通过硬盘写保护器将接受数据取证的硬盘连接到另外一台计算机上。利用计算机应用软件复制硬盘原数据到目标硬盘。

硬盘写保护器是保护原硬盘数据不更改的设备，写保护器应在主机发送对取证硬盘复制指令后，不传输任何修改指令给取证硬盘。写保护器与取证硬盘的接口及应用程序应具有单向功能，单向的方向必须是取证硬盘向复制机的目标硬盘，不可逆向。写保护器在收到一个来自主机的读指令操作类的操作后，应通过读操作返回请求的数据。读指令操作可包括：从某个存储介质的特定位置请求数据并把请求的数据返回给主机的操作。一个读操作从存储设备的介质里请求一个或多个数据块，每个数据块都有关于存储位置和长度的说明。写保护器在收到一个来自主机的信息指令操作类的操作后，应返回主机一个包含不修改任何重要访问信息的回复。写保护器应将受保护硬盘的任何错误情况立即报告给主机。其他非修改的指令操作应包括：任何不属于其他的请求存储设备执行一个非破坏性动作的操作类的操作。

3.2.2 Linux 数据获取启动光盘

Linux 数据获取软件一般是以可启动光盘形式出现[3]。取证人员可以利用可启动的、不会改写硬盘数据的光盘启动计算机，将数据采集出来。其原理如图 4 所示。

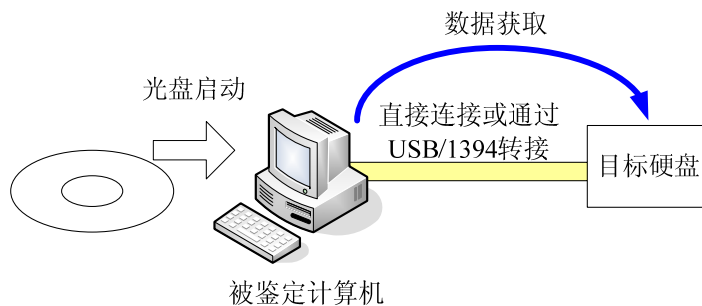


图 4、Linux 启动光盘采集原理图

此方法的数据采集可按以下步骤操作：

- (1) 将接受数据取证的硬盘连接到预置的专用计算机上。
- (2) 通过光盘启动计算机，直接将原硬盘的数据采集到目标硬盘上。

4 证据分析

主要的证据分析技术有：删除数据的恢复、加密数据破解和海量数据的线索分析。

4.1 删除数据的恢复

被删除数据的恢复主要从两个层次进行：文件系统级和应用级的数据恢复。

4.1.1 文件系统级恢复

文件系统用于存储数据，供计算机系统访问。文件系统中的数据通常以文件目录和文件的形式存放于树状结构中。文件系统通常以元数据描述每个文件的信息，包括文件名称、属性、时间戳。一般元数据都位于目录入口，但另外有些元数据位于特定的文件（如 NTFS 的 \$MFT）或者其他位置（如 Unix 的 i-node）[4]。

但一个文件或者目录被删除，通常相关的元数据被设置标志为不激活的。然而在绝大多数文件系统中，元数据和文件真实内容都未被真正删除。因此，被删除的元数据仍然能够被访问到。然而，并不是所有文件内容都能够被访问到，这依赖于元数据的结构和原始格式。比如一个文件目录是分散在硬盘不连续区域的，如果被删除，目录的第一个数据块通常都能恢复，但是其余的数据块就无法恢复[5]。

过去的存储设备都比较昂贵和缓慢（和 RAM 相比），在设计文件系统时均允许操作系统以高效和快速的方式访问二级存储。虽然不同的文件系统执行方式不同，但总体上，都具备恢复删除文件的能力。两个最关键点是：持续写入和文件系统活动方式。文件系统在一般情况下都尽量使用持续写入：绝大多数操作系统以连续数据块的方式往驱动器上写数据。一个给定的数据文件，在写入磁盘后未作任何修改，这个文件数据将在连续的扇区上。这将使得读写速度加快，因为磁头无需转移到另外的区域去读写数据。这样，文件即使被删除，存在于连续区域的几率也比较大，根据此原理，可以找回文件数据。[6]

为了尽快和高效的文件操作，文件系统的许多活动都将变化控制在最小。比如文件删除，绝大多数情况下，只进行逻辑删除，这就意味着真正的数据并未删除，而只是对信息进行索引的元数据发生了变化，标记或者删除。用这项技术，文件内容不管多大，删除时都只是简单修改或者删除索引结构。最简单的例子就是 Windows FAT32 文件系统删除文件。它只是找到被删除文件在目录中的入口，并将首字节改为"0xE5"，并将文件分配表清空。绝大多数元数据和文件内容都保留。

绝大多数情况下，这些通常的属性都能帮助进行数据恢复，不管数据位于何种文件系统中。许多工具都可以定位潜在的文件系统对象，找到残余的元数据，并恢复最大数量的连续文件。

4.1.2 应用级恢复

除了文件系统级的数据恢复，还可以在应用级进行数据恢复。通过识别应用文档特征，在整个硬盘中搜索符合此类特征的数据，达到恢复数据的目的。以 JPG 文件恢复为例，JPEG 文件头都有特征，如果根据类特征在硬盘上寻找所有的 JPG 文件，将比只从文件系统中恢

复带有 JPG 扩展名的文件更加彻底。

4.2 加密文档的解密

加密给证据分析提出了难题，像 Office 这样的办公软件就内置有加密文档功能。有两种方法可以处理类似于 Office 这种加密程序：密钥搜索和分布式破解[7]。

4.2.1 密钥搜索技术

以 Word®/Excel® 97/2000 为例，如果加密文件使用的 RC4 算法，如果文件保护被使用，最简单的办法是利用字典或者暴力方式破解密码。然而这种方法往往只能对简单密码有效。比如如果密码有 10 位长，并且包含大小写和数字，首先无法找到包含这些字符的字典，只能用暴力破解。如果使用暴力破解，则需要尝试的口令个数为：

$$(26 + 26 + 10)^{10} = 839,299,365,868,340,224$$

假如采用 P4 的电脑，每秒尝试 1000,000 个密码，也需要 26614 年时间完全尝试完。即使平均也需要 13307 年才能破解，相当于无法解密。密钥搜索是一种新的密码破解方法。这种方法并不是尝试去恢复密码。以 40 位的 RC4 算法为例，这意味着所需要尝试的密钥数量为：

$$2^{40} = 1,099,511,627,776$$

这样可以替代尝试所有密码的方式，通过测试所有的加密密钥，一旦密钥被发现，就可以解密文档，而无需口令就可以打开文档。以一个解密节点为例，如果速度为每秒测试 1000000 个密钥，仅需要 305 个小时，即 13 天即可尝试完所有的密钥。如果采用多个解密计算节点同时进行破解，则解密时间可缩短为几天甚至是几个小时。

4.2.2 多节点分布式破解技术

多节点分布式密码破解是一种新的密码破解技术，它能够将非常庞大繁重的密码破解计算问题分解成许许多多小的密码破解运算任务，并分散至多个计算节点上进行。然后通过这些分布式的计算机节点将这些密码破解问题逐个解决。该方法可以利用多个计算节点同时进行字典搜索、暴力破解和密钥搜索破解，大大提高破解成功率和缩短密码破解时间。

4.3 数据的相关性分析

当识别案件中各类数据关联时，可以用节点来表示在他们曾经逗留的地点、所使用过的电子邮件和 IP 地址、财务交易、用过的电话号码，这有助于确定节点之间是否存在值得关注的联系[8]。例如在一个大规模的诈骗案调查中，通过把个人与组织之间的活动关系进行连线，可以显示出资金转帐关系，从而揭露出诈骗案件中最活跃的实体。同样，通过在大量

相互交换的消息中描绘嫌疑人发送或接收的电子邮件消息，可以帮助分析员发现可能的同谋。在分析计算机入侵案时，画一个计算机之间的关系图，可以提供对案件的概况，并且可以帮助确定先前被忽视的数字证据源的位置。

5 结束语

虽然计算机技术正在飞速发展,但是其基本构建和操作却是相对稳定的,因此取证程序、数据获取和分析的过程也同样保持相当的稳定。本文分析了目前计算机取证过程中的关键要素。随着对于计算机取证研究的不断深入,计算机取证技术的发展也会不断加快,在打击计算机犯罪中发挥越来越重要的作用。

参考文献:

- [1] Forensic Examination of Digital Evidence: A Guide for Law Enforcement, U.S. Department of Justice, 1994
- [2] Disk Imaging Specification, National Institute of Standards and Technology, 2001
- [3] Digital Data Acquisition Tool Specification, National Institute of Standards and Technology, 2004
- [4] Carrier, File System Analysis Techniques: Sleuth Kit Reference Document, 2003
- [5] Crane, Linux Ext2fs Undeletion mini-HOWTO, www.tldp.org/HOWTO/Ext2fs-Undeletion.html
- [6] Erdelsky, A Description of the DOS File System, 1993
- [7] Access Data Corp, Forensic Toolkit, <http://www.accessdata.com/products/utk/>
- [8] Belgrade G. Javnosti, Croatia Using Advanced US-Installed Intelligence Technology, 2002

作者:

金波 Jin Bo, 公安部第三研究所, 国家反计算机入侵与防病毒研究中心, 总工, 博士, 高级工程师/副研究员; jinbo@trimps.ac.cn

陶明明 Tao Mingming, 上海金诺网络安全技术发展股份有限公司, taomingming@kingnet.biz