高职高专教育法律类专业教学改革试点与推广教材

计算机数据分析技术与应用

主　编　凌　彦
副主编　潘大四　孙培梁

清华大学出版社

華中科技大學出版社
http://www.hustp.com

高职高专教育法律类专业教学改革试点与推广教材 | 总主编 金川

计算机数据分析技术与应用

主 编 凌 彦
副主编 潘大四 孙培梁

清华大学出版社
北京

華中科技大学出版社
http://www.hustp.com
中国 · 武汉

内容简介

面对日趋严重的计算机犯罪，计算机数据分析技术正变得越来越重要。本书系统地阐述了计算机数字证据分析的基本概念、原理及方法，全书由12章组成。基础部分的内容包括了计算机系统基础知识、数据存储原理。数据分析理论部分包括了计算机犯罪概述、数据分析内容和工作流程、调查现场的处理、数字证据固定技术。数据分析技术部分主要以EnCase为例介绍了数据分析技术的实际应用以及如何编制计算机调查分析报告。

本书面向那些想深入了解计算机系统的工作原理，对计算机数据分析工作有兴趣的读者，也适用于计算机系统管理员、开发人员、安全专家等查阅参考。

图书在版编目（CIP）数据

计算机数据分析技术与应用/凌彦主编．—武汉：华中科技大学出版社，2010.11
ISBN 978-7-5609-6613-7
Ⅰ.①计… Ⅱ.①凌… Ⅲ.①电子计算机—数据处理—高等学校—教材 Ⅳ.①TP309

中国版本图书馆CIP数据核字（2010）第193534号

计算机数据分析技术与应用 凌 彦 主编

策划编辑：王京图
责任编辑：王京图
封面设计：傅瑞学
责任校对：北京书林瀚海文化发展有限公司
责任监印：周治超
出版发行：华中科技大学出版社（中国·武汉）
武汉喻家山 邮编：430074 电话：（027）87557437
录 排：北京楠竹文化发展有限公司
印 刷：武汉中远印务有限公司
开 本：710mm×1000mm 1/16
印 张：25.25
字 数：524千字
版 次：2010年11月第1版第1次印刷
定 价：39.00元

总　序

我国高等职业教育已进入了一个以内涵式发展为主要特征的新的发展时期。高等法律职业教育作为高等职业教育的重要组成部分，也正经历着一个不断探索、不断创新、不断发展的过程。

2004年10月，教育部颁布《普通高等学校高职高专教育指导性专业目录（试行）》，将法律类专业作为一大独立的专业门类，正式确立了高等法律职业教育在我国高等职业教育中的重要地位。2005年12月，受教育部委托，司法部牵头组建了全国高职高专教育法律类专业教学指导委员会，大力推进高等法律职业教育的发展。

为了进一步推动和深化高等法律职业教育的改革，促进我国高等法律职业教育的类型转型、质量提升和协调发展，全国高职高专教育法律类专业教学指导委员会于2007年6月，确定浙江警官职业学院为全国高等法律职业教育改革试点与推广单位，要求该校不断深化法律类专业教育教学改革，勇于创新并及时总结经验，在全国高职法律教育中发挥示范和辐射带动作用。为了更好地满足政法系统和社会其他行业部门对高等法律职业人才的需求，适应高职高专教育法律类专业教育教学改革的需要，该校经过反复调研、论证、修改，根据重新确定的法律类专业人才培养目标及其培养模式要求，以先进的课程开发理念为指导，联合有关高职院校，组织授课教师和相关行业专家，合作共同编写了“高职高专教育法律类专业教学改革试点与推广教材”。这批教材紧密联系与各专业相对应的一线职业岗位（群）之任职要求（标准）及工作过程，对教学内容进行了全新的整合，即从预设职业岗位（群）之就业者的学习主体需求视角，以所应完成的主要任务及所需具备的工作能力要求来取舍所需学习的基本理论知识和实践操作技能，并尽量按照工作过程或执法工作环节及其工作流程，以典型案件、执法项目、技术应用项目、工程项目、管理现场等为载体，重新构建各课程学习内容、设计相关学习情境、安排相应教学进程，突出培养学生一线职业岗位所必需的应用能力，体现了课程学习的理论必需性、职业针对性和实践操作性要求。

这批教材无论是形式还是内容，都以崭新的面目呈现在大家面前，它在不同层面上代表了我国高等法律职业教育教材改革的最新成果，也从一个角度集中反映了当前我国高职高专教育法律类专业人才培养模式、教学模式及其教材建设改革的新趋势。我们深知，我国高等法律职业教育举办的时间不长，可资

借鉴的经验和成果还不多，教育教学改革任务艰巨；我们深信，任何一项改革都是一种探索、一种担当、一种奉献，改革的成果值得我们大家去珍惜和分享；我们期待，会有越来越多的院校能选用这批教材，在使用中及时提出建议和意见，同时也能借鉴并继续深化各院校的教育教学改革，在教材建设等方面不断取得新的突破、获得新的成果、作出新的贡献。

全国高职高专教育法律类专业教学指导委员

2008 年 9 月

前　言

自 1994 年 3 月中国加入互联网始，计算机和互联网开始慢慢融入我们的生活。经过近十五年的发展，计算机和互联网已经成为中国数以亿计网民生活中不可或缺的部分。每天，我们都有数小时甚至更多的时间是在计算机前度过的，如通过互联网收发电子邮件、查询最新资讯、跟世界各地的朋友愉快交流以及进行其他的活动。

近年来，随着电子商务的发展成熟，电子货币广泛应用，计算机和互联网开始进入我们的实际生活，如电子银行、电子货币、虚拟货币等等。然而，与方便快捷的计算机生活一同到来的是越来越多的计算机问题，比如，侵犯个人隐私、盗窃虚拟财产、病毒和木马攻击等等，我们用计算机犯罪来形容这类活动。计算机犯罪是一种在虚拟空间、使用虚拟身份、通过虚拟手段来完成，但却可以在现实空间获得利益的犯罪方式。正是由于计算机犯罪具有犯罪主体的专业化、犯罪行为的智能化、犯罪客体的复杂化、犯罪对象的多样化、犯罪后果的隐蔽化等特征，使其有别于传统的刑事犯罪。此外，一旦发生计算机犯罪事件，对计算机犯罪证据的提取和展示也成为亟须解决的问题。

在对计算机犯罪进行司法调查的过程中，计算机数据分析技术起到了至关重要的作用。计算机数据分析技术是全世界的计算机安全专家在长期同计算机犯罪分子进行信息安全较量的过程中，集思广益，由实战经验总结出来的计算机安全领域的一个全新分支。计算机数据分析技术在中国已经发展了近十年的历史，在实际的案件调查过程中已经形成一套符合中国国情的计算机数据分析方法和流程。但是，由于计算机数据分析技术形成晚、技术新，因此在涉及计算机犯罪案件分析调查的过程中，广大信息安全领域的技术人员还是存在诸多疑问，涉及调查的工具、调查的软件、调查的方法、调查的规范等等。本书通过对计算机数据分析技术的简要介绍，结合实际生活中一些典型的计算机犯罪案例，对计算机数据分析技术在计算机司法调查分析过程中的应用做了阐述。

如何阅读本书：

第 1 章和第 2 章用简要的篇幅，回顾了一下计算机硬件以及操作系统的结构，供刚参加计算机调查的人员参考。

第 3 章介绍了计算机犯罪概述以及一些理论知识，作为计算机数据分析的入门理论基础。

第 4 章介绍了计算机数据分析工作的定义，由此引入计算机数据分析的概

念以及一些常见的数据分析内容和流程知识。

第5章介绍计算机犯罪调查过程中的计算机犯罪现场证据固定、应急响应准备以及在前往计算机犯罪现场之前的一些准备工作。

第6章以及后面的章节介绍了计算机数据分析的实际应用，包括证据固定技术和计算机数据分析软硬件的介绍和使用。

在本书的附录部分，我们增加了一些计算机数据分析技术名词的解释、一些常用的知识、法律法规汇总等，供计算机数据分析人员查阅。

本书提供配套的电子教案，有助于教师上课备课，同时也便于学生课后复习及自学。

本书的第1、3、4、12章由凌彦编写，第2、7、8章由孙培梁编写，第5、9、10章由潘大四编写，彭辉和来自企业一线的赵庸、王跃聪参加了第6章、第11章和附录的编写工作，全书由凌彦统稿。

本书在编写过程中，得到了厦门市美亚柏科资讯科技有限公司董事长刘祥南教授、总经理滕达先生的大力支持，在此表示感谢。厦门美亚公司在本书的编写过程中，提供软硬件设备供实践操作，保证了本书关于各种软硬件操作介绍的正确性和严谨性。

由于时间仓促，书中错误或不妥之处在所难免，衷心希望广大读者、同行以及使用本教科书的师生给我们提出宝贵的意见。

作者

2010年10月

目　录

第 1 章　计算机系统基础知识

本章内容包括：

- 计算机硬件系统组成
- 计算机硬件系统启动过程
- 系统分区原理
- 计算机文件系统

计算机数据分析人员要分析的对象，就是各种各样的电子数据，电子数据的范畴非常广泛，存在的形式多种多样，甚至可以说跟我们的工作和生活密不可分。比如，人们最经常接触的计算机中，就存在着大量的电子数据，以电子表格、文本文档、数据库、各种日志文件、电子邮件、电子图片、视频文件、音频文件等形式存在。除计算机以外，还有很多地方存在着电子数据，比如手机、掌上电脑、传真机、打印机、摄像机等。也就是说，凡是能存储电子数据的介质，其中运行产生的和保存的都是电子数据。

在日常工作和生活中，人们能接触到的存储介质有很多，包括软盘、硬盘、ZIP 盘、JAZ 盘、磁带、磁光盘、CD・ROM、CD・R、CD・RW、DVD、CF 卡、MS 卡、SD 卡、MMC 卡等。在所有存储介质里面，使用最广泛的，存储数据量最大的，也是我们接触最多的，就是计算机里面的硬盘。因此，充分了解组成计算机的各种组件以及存储介质等，在进行调查分析的时候可以快速识别出分析的对象类型，是一个计算机数据分析人员所必须具备的基本知识。

另外，作为一个计算机数据分析人员，需要在法庭上向法官或者其他人解释计算机的各种相关术语，包括软件和硬件方面的功能或者作用。但是要作出这类的解释，不仅需要从一个技术人员的角度来解释计算机的各种相关术语，而且要能够把这些技术概念解释为非计算机人员可以听懂并接受的技术名词。

注：我们把易于听懂或接受的名词称为 TWAIN，即 Technology Without An Interesting Name，一般来说，对于非计算机人员，很难向其解释计算机的工作原理等一些特有的名词和技术，而解释这些名词和技术往往又是必需的。因此，作为计算机数据分析人员，不只要掌握各项计算机分析技术、掌握各种分析名词，还需要掌握把各种技术、名词用通俗易懂方式表达出来的能力。

当计算机数据分析人员在分析平台上对电子数据进行调查分析的时候，通常会遇到各种各样的问题，比如硬件故障、系统故障、信息加密、信息隐藏、系统不能运行等，因此掌握计算机底层的基础知识，可以让数据分析人员在遇上这类问题的时候，能够更快地发现和解决问题，保证分析工作顺利进行。

本章的主要内容是向计算机数据分析人员介绍计算机硬件系统的各个组件，以及了解这些硬件组件是如何在一个完整的启动过程中扮演各式各样的角色的。除此以外，分析人员通过本章还可以学习硬盘分区以及文件系统的基础知识。

1.1 计算机硬件系统组成

每一个行业，总是有一些核心的名词，它们被用于该行业实施人员进行交流时或分享经验时使用。计算机数据分析行业也不例外，在本章里，将要讨论各种计算机系统的组件。

一个计算机硬件系统主要由以下几部分组成：

- 机箱。
- ROM（Read Only Memory，只读存储器）。

注：这里所指的计算机硬件系统，指的是个人计算机系统，与小型机、服务器等大型计算机系统在硬件结构上是有较大区别的。

只读存储器是一种永久地保存数据的设备，或者几乎是永久的，而且它的本质是几乎不可写的。ROM 的另一个重要的属性是非易失性，意味着数据将保存在 ROM 里，即使关机或者电源出现问题也是不丢失的。因为有了只读和非易失性，使 ROM 成为一个理想的包含了数据文件的初始启动设置的连接器，一般把它用于计算机系统启动的初步阶段，如 ROM BIOS。

- RAM（随机访问存储器）。

计算机数据分析人员在现场进行数据处理，比如进行电子数据固定的时候，经常会遇到对象——计算机没有关机的情况，比如 IDC 机房。操作系统在运行的时候，很多的交换数据、临时数据、进程信息、登录账号，都驻留于 RAM 里面，使系统的运行速度与效率得到提高，而这些数据对我们的调查分析工作十分重要甚至是关键的，驻留在 RAM 里面的数据在系统重启或者关机或者电源出现问题的时候，就会丢失，因此我们称这些数据为易失性数据。计算机数据分析人员要具备在 RAM 数据丢失之前，尽可能固定 RAM 里面的数据，但有时候，RAM 里面的数据也会被写入硬盘的交换文件

中。该交换文件可以被系统压缩存储，甚至有可能该数据就是位于磁盘的未分配簇里。

注：在实际的工作中，我们发现易失性数据里面隐藏着至关重要的数据，因此提取易失性数据是计算机数据分析工作中非常重要的一项工作，但是这项工作需要依赖特定的设备和软件，常用的工具有“美亚网警”DC-8600H 特定数据获取设备等。

- 电源。
- 母板或主板。
- 微处理器或 CPU。
- CPU 或者系统散热风扇。
- 硬盘。

计算机数据分析人员遇到最多的就是各种各样接口的硬盘，一般我们要求计算机数据分析人员能够快速识别出硬盘的各种接口类型和品牌、容量大小等等硬盘的主要参数。正确识别出硬盘后，分析人员才能决定使用什么样的设备、准备什么样的备份硬盘以及使用哪些调查工具等。

目前常见的硬盘接口类型有：

IDE（Integrated Drive Electronics）接口

SATA（Serial Advanced Technology Attachment）接口

SCSI（小型计算机系统接口）

SAS（Serial Attached SCSI）串行 SCSI 接口

RAID（冗余廉价磁盘阵列）

注：RAID 冗余廉价磁盘阵列，包含 RAID 0/1/5/6 等多种多样的模式，每种模式的调查和分析方法都不一样。每种阵列可以由 SATA、SCSI、SAS 等各种接口的硬盘来组成，分析人员一般很难直接从阵列以及硬盘本身判断该阵列属于哪一种 RAID，而要从阵列本身的配置数据上来获得。因此，当分析人员在工作过程中遇上磁盘阵列的时候，要十分注意收集阵列的信息，比如阵列模式、硬盘的顺序、阵列卡驱动程序以及各种连接线等，因为一旦阵列受到破坏或者被初始化，其数据的毁坏几乎是不可恢复或者需要付出昂贵代价的。

- CD-ROM（Compact Disc Read-Only Memory）or CD-RW（Compact Disc-Read/Write）驱动器。
- DVD-ROM（Digital Versatile Disc Read-Only Memory）or DVD-RW（Digital Versatile Disc Read/Write）驱动器。
- USB 端口。
- IEEE 1394 A/B 端口。

- 扩展槽（ISA，MCA，EISA，VL-Bus，PCI，AGP，PCI Express）。
- 声卡。
- 显卡（PCI，AGP，PCI Express）。
- RTC（实时时钟）。
- CMOS、CMOS电池。

对于CMOS和CMOS电池，一个需要注意的问题是关于CMOS的密码，CMOS密码经常作为计算机的第一重保护措施，虽然简单但确有很多人使用它。在计算机数据分析工作中，遇上CMOS密码是常有的事，常用的一个办法就是移除CMOS的电池，然后把它的设置复位到出厂的默认的没有密码的设置就可以了。

注：尽管移除CMOS电池是移除CMOS启动密码的一种非常有效方式，但是CMOS恢复到出厂的默认设置后，也意味着其他所有存储在CMOS的信息全部被清除了，其中就包括时间信息。

计算机操作系统的时间是以CMOS时间作为基准的，CMOS的时间发生变化，会对操作系统时间造成影响，而在某些对时间的准确性依赖较大如网络攻击、黑客入侵案件的调查工作中，系统时间信息对案件的分析结论非常重要，有时候会影响结论甚至得出错误的结论。因此要求分析人员在做移除CMOS这个操作之前，必须确保该操作的必要性，如果确实有必要，也要详细记录CMOS内原始的时间信息，并且做正确的记录处理。

- BIOS（Basic Input Output System）。
- 鼠标端口。
- 键盘端口。
- 网卡接口。

注：MAC（Media Access Control，介质访问控制）地址是烧录在网卡里面的。MAC地址也叫硬件地址，是由48位长的十六进制数字组成的，0～23位叫做组织唯一标志符（Organizationally Unique，是识别局域网节点的标识）；24～47位是由厂家自己分配的。其中第40位是组播地址标志位。网卡的MAC地址通常由网卡生产厂家烧入网卡的EPROM（一种闪存芯片，通常可以通过程序擦写）中，它存储的是传输数据时真正赖以标识发出数据的主机和接收数据的主机的地址。也就是说，在网络底层的物理传输过程中，是通过物理地址来识别主机的，具有全球唯一性。

正是由于网卡的这个特性，因此，在某些网络犯罪案件中，MAC地址成为了唯一确定主机身份的方法，计算机数据分析人员在对嫌疑机器进行检查的时候，MAC地址是需要特别关注的对象。尽管现在有很多程序可以修改MAC地址，但是仍然值得关注。

- MODEM 调制解调器。
- 串口。

1.2　计算机硬件系统启动过程

介绍完计算机的各种硬件组件之后，接下来，来看一下计算机硬件系统的启动过程。计算机硬件系统组件都是一些芯片和电子元器件，它们的启动过程可以通过进行各种测试，组装在一起。计算机的启动过程就是把这些元器件组织在一起，然后让一台计算机激活。作为计算机数据分析人员要理解并能描述硬件系统启动过程，才能够对计算机系统进行深入的分析工作。

当用户按下电源按钮并启动系统的时候，不管它是哪一种系统，都会发生如下步骤的操作：

当按下计算机电源开关的时候，将会开始 POST（Power On Self Test，加电自检）这一处理过程。它首先检查电压，确保电压和电流强度是可以接受的。电源将会按照预定义的路径抵达 CPU，任何驻留在 CPU 里的数据都会被擦除。该信号会把 CPU 的寄存器设为程序计数器。

ROM BIOS 里的启动程序（有时称为自举程序）将会启动一系列的系统自检程序。第一步是运行一系列的处理过程以检查 CPU 或者 POST 处理过程，并将它与存储在 BIOS 芯片里的信息进行对比。只要该值是匹配的，将继续下一个处理过程。

信号由 CPU 发送到系统总线，必须确保系统总线是正常工作的。如果测试通过，POST 将继续下一步的处理过程。

CPU 将检测 RTC，或者称为系统时钟。这个时钟用来保证内部的电子信号是同步的。如果 RTC 通过检查，POST 将进入下一步的处理工作。

POST 接下来需要测试系统的视频组件，我们有时候也将视频组件称之为显卡。视频的内存需要进行测试，由设备送出的信号也需要进行测试。显卡的 BIOS 将被添加到全局的系统的 BIOS 里，存储在 RAM 里。到这一步骤后，在屏幕上将开始可以看到所有的启动处理过程。

在接下来的处理过程中，POST 将对系统的主内存 RAM 进行测试，数据将被写入 RAM，然后读出来与原始的数据进行对比。匹配则通过，不匹配则不通过。有些系统可以看到一个不停增加的计数器来表示测试的内存的数量。如果所有的 RAM 都通过了该测试，POST 将继续进行下一步的处理工作。

CPU 将检查键盘是否存在，并且是否有键正在被按下。如果不小心把一本书或者一串钥匙放在键盘上，将听到扬声器发出“嘀嘀”声。如果成功地完