电子取证技术基础与实践

取证，司法解释是具有调查取证权的国家机关对于立案处理的案件，为查明案情，收集证据。电子取证，顾名思义可理解为基于计算机的证据收集。

取证相关介绍

对于电子取证的介绍，百度是这样回答的：电子取证是指利用计算机软硬件技术，以符合法律规范的方式对计算机入侵、破坏、欺诈、攻击等犯罪行为进行证据获取、保存、分析和出示的过程。从技术方面看，计算机犯罪取证是一个对受侵计算机系统进行扫描和破解，对入侵事件进行重建的过程。具体而言，是指把计算机看作犯罪现场，运用先进的辨析技术，对计算机犯罪行为进行解剖，搜寻罪犯及其犯罪证据。

电子证据概念

电子证据在很多年前已经作为一种新兴证据被列入法律，但我实在是没找到这个法律条文……

现在的中华民族共和国刑事诉讼法，中华人民共和国刑法里的第四十八条里，电子证据是在第八条，这八条依此是：

- 物证
- 书证
- 证人证言
- 被害人陈述
- 犯罪嫌疑人、被告人供述和辩解
- 鉴定意见
- 勘探、检查、辨认、侦查实验等笔录
- 视听资料、电子数据

证据必须经过查证属实，才能作为定案的根据。那么，怎么去取证，下面我们来讲讲正确的取证姿势。

电子取证行业标准

电子数据取证行业标准

GB/T 29360-2012 电子物证数据恢复检验规程	GA/T 978-2012 网络游戏私服检验技术方法
GB/T 29361-2012 电子物证文件一致性检验规程	GA/T 1069-2013法庭科学电子物证手机检验技术规范
GB/T 29362-2012 电子物证数据搜索检验规程	GA/T 1070-2013法庭科学计算机开关机时间检验技术规范
GA/T 754-2008 电子数据存储介质复制工具要求及检测方法	GA/T 1071-2013法庭科学电子物证Windows操作系统日志检验技术规范
GA/T 755-2008 电子数据存储介质写保护设备检测方法	GA/T 1770-2014 《移动终端取证检验方法》
GA/T 756-2008 数字化设备证据数据发现提取固定方法	GA/T 1771-2014 《芯片相似性比对检验方法》
GA/T 757-2008 程序功能检验方法	GA/T 1772-2014 《电子邮件检验技术方法》
GA/T 825-2009 电子物证数据搜索检验技术规范	GA/T 1773-2014 《即时通讯记录检验技术方法》
GA/T 826-2009 电子物证数据恢复检验技术规范	GA/T 1774-2014 《电子证据数据现场获取通用方法》
GA/T 827-2009 电子物证文件一致性检验技术规范	GA/T 1775-2014 《软件相似性检验技术方法》
GA/T 828-2009 电子物证软件功能检验技术规范	GA/T 1776-2014 《网页浏览器历史数据检验技术方法》
GA/T 829-2009 电子物证软件一致性检验技术规范	《计算机犯罪现场勘验与电子证据检查规则》（公信安〔2005〕161号）
GA/T 976-2012 电子数据法庭科学鉴定通用方法	《公安机关电子数据鉴定规则》（公信安〔2005〕281号）
GA/T 977-2012 取证与鉴定文书电子签名	

上图这个表，就是在取证里，怎么去取证操作才能是合法的，必须遵守公安机关电子取证鉴定规则，基于这条规定，你的取证才能合法合理的。

现场取证技术

现场取证，涉及到计算机取证硬件、软件、移动智能取证工具，这里我们分开详讲。

计算机取证硬件

- 硬盘复制机，他的功能就是对硬盘进行一个打镜像，进行复制然后取证的工具，因为取证过程是不允许对原硬盘进行直接操作，都需要打镜像。如下图：

这就是一个正在运行的复制机，结构根据厂商都有稍不同。

- 硬盘只读锁

  硬盘只读锁，来看看这是个什么玩意。

  硬盘只读锁的功能就是，给硬盘加上一块锁，阻止写入通道，有效的保储存介质中的数据在获取和分析过程里不会被修改，从而保证取证工作的司法有效性于数据完整性。简单说就是，确保我拿到手的硬盘数据是原生态，没有被二次修改过的一个设备。

- 取证一体机

一体机这个比较好理解，基于拷贝克隆、读取、销毁于一体的取证设备。

- 取证塔，和一体机的一样的，不过是多了一些ID分析接口和集合的大平台。

- 介质修复设备

这个设备的作用就是个修复设备，可理解为“医疗兵”，硬盘磁道坏了，就用这个设备结合软件进行修复。工作环境必须无尘。

设备就介绍完了，接下来讲讲计算机取证软件，计算机取证软件分为国内外，值得一提的是，硬盘隐藏数据以上设备也是能读取出来的，加不加锁形同虚设。

计算机取证软件

国外取证分析软件：

- ENcase
- X-Ways
- FTK

国外取证分析软件：

- 取证大师
- 盘石介质取证分析系统

以上设备均可取证，比如一木马制作者制作的木马威胁用户很多，在中途过程中他在浏览器IE里搜索过：怎么写入注册表实现开机自启过杀毒。那么以上软件均可在读取到你的搜索记录。

分布式取证系统

分布式，怎么理解，建立在网路之上的软件系统。

研究与开发

基于 Hadoop 的高效分布式取证:原理与方法 *

吴松洋 [1],张熙哲 [1],王旭鹏 [1],李祥学 [2]
(1. 公安部第三研究所 上海 201204;2. 华东师范大学 上海 200241)

摘 要:随着信息技术的发展以及各种智能设备的普及,设备的平台多样化使得现有电子数据勘查取证分析装备已不能满足网络和存储技术所需要的高速数据镜像存储和海量数据相关性分析等要求,并表现出操作复杂、效率低等缺陷。设计并实现了一种高效的基于 Hadoop 的分布式取证系统,它能够支持多介质并行取证的工作场景,并通过调度控制服务将不同的证据介质中的数据存储到不同的分布式数据存储服务器上,每个取证任务运行时都可以独占一个取证介质,从而实现多介质的并行取证分析。实验数据显示,搜索一个 2~4 GB 的文本数据的响应时间可以达到仅 0.1 s。

关键词:Hadoop;分布式系统;取证;海量数据;多介质
doi: 10.3969/j.issn.1000-0801.2014.01.005

An Efficient Distributed Forensic System Based on Hadoop: Principle and Method

Wu Songyang[1], Zhang Xizhe[1], Wang Xupeng[1], Li Xiangxue[2]
(1.The Third Research Institute of Ministry of Public Security, Shanghai 201204, China;
2.East China Normal University, Shanghai 200241, China)

Abstract: With the development and popularization of information technology and intelligence device, the diversity of different device making forensic analysis of existing equipment cannot meet today's networking and storage technology requirements, and exhibit complex operation, low efficiency, on high speed disk image storage and massive data correlation. An efficient distributed forensics system based on Hadoop technique, which can support multiple concurrent media scene forensics work, was designed and implemented, and through the dispatch control services would be evidence of different data storage media to a different distributed data storage server, each forensic task runtime could monopolize a forensic medium to achieve a parallel multiple media forensic analysis. Data show that responsible acknowledge duration will be 0.1 s for a 2~4 GB text file.

Key words: Hadoop, distributed system, forensic, massive data, multiple media

1 引言

随着社会信息化的快速推进,越来越多的数据以电子

信息资源存储媒介的比例仍呈上升趋势。因此,存储于计算机及其他信息设备中的电子数据逐步成为针对和利用计算机网络犯罪案件的重要证据和诉讼依据。随着社会信

这是一份文档，提供个链接给大家：

http://www.chinacloud.cn/upload/2014-04/14041406572100.pdf?WebShieldDRSessionVerify=Izq4UwNheLLIWpOvgpD6

作用就是，采集单个设备的数据，上传云，在系统里搜索关键词进行比对发现问题，一般用于公安局省厅。

智能终端取证

人工分析

智能终端也跟的比较快，刚开始是人工分析的，比如拍照或者手抄分析，效率也普遍较低。(人工分析)

逻辑分析

人才有逻辑，计算机逻辑也是人工编的逻辑，一成不变，也不知变通。逻辑分析是用取证软件对照手机电话本、QQ、微信、邮箱等信息的获取，进行人为逻辑分析，也是等于

给嫌疑人画像，即使你已经见过他的样子了。

JTAG分析

也就是十六制镜像。在电影里有这样的情节，一个犯罪团伙老大，联系了他的杀手后，就把手机砸碎，放厕所里，或者是丢水里。以前看刘德华的一部电影，华仔是反派还是间谍身份，公安组在手机上安装了GPS定位，团伙老大等华仔接完电话后就让华仔把车扔对面水池里，与此同时，公安也失去的定位。

讲道理的话，手机砸碎和丢水里都是能够取证的，如果利用JTAG芯片分析技术，还是能做数据提取的，因为手机芯片上是记录了运动轨迹的。

iPhone，iPod也都属于移动智能终端，遇到人为损坏的，特殊情况下，照样是能取的。

动态仿真

这个就比较前沿了，动态，仿真，有没有人能想到？安卓模拟器用过吧？这就是比较接近的，类似于模拟器进行手机仿真，比如你在微信群里讨论一些政治敏感的内容，如果拿到了你的手机，我是能够还原你聊天的一幕幕，对你聊天的场景进行模拟，你做了什么，发了什么图，就好像是我自身在进行对话。

还有就是，群里发送黄色视频这种，如果影响大，公安指定取证的话，是对你的微信QQ提取文件分析，转发了多少，影响范围，甚至还有播放数量，查到源头，也不是完全不可能。

远程勘验技术

远程勘验取证技术包括也比较多，有网络取证和现场取证，也有Web网页取证和服务器取证技术等，——详解。

现场取证与网络取证

现场取证与网络取证

现场取证	网络取证
静态取证	动态取证
事后取证	事中取证
证据链的发端	证据链的构成
软硬件的恢复技术	数据抓取技术
数据格式分析与检索技术	海量数据与协议分析技术

这个图就已经说明是现场和网络取证的不同和相关技术。

现场取证

分为静态取证，事后取证。证据链的发端，软硬件的恢复技术，数据格式分析于检索技术。现场一般是静态取证，也就是事后取证，证据面的开始， 就会接触到一些软硬件的恢复，对软件进行数据的分析和检索。

假设，到现场 发现一台电脑，如果是开机状态，你就不能关机的。只能对电脑进行开凿然后进行数据提取，如果是关机状态，就只能保持关机，直接对硬盘进行复制，打镜像。做到开机不关，关机不开即可。

服务器如果取证的话，取证的话，是不能正常关机的，只能直接拔掉电源。你也可以理解为，直接拔掉电源，防止有人远程连接服务器进行数据篡改。

静态取证，是电脑已经摆在面前了，直接对数据进行分析。事后取证，是已经案发了，这个事情已经发生了，我们在进行一个取证调查过程。软硬件恢复技术，是对数据进行恢复，因为你不知道他的硬盘是否进行了一次格盘操作，当你找不到相关的信息，得到允许后，你可对当前的硬盘进行一次数据恢复。

有个案子，是一个制作外挂的，非法牟利百万，严重影响了游戏厂商正常运营，无奈选择报案，当公安局抓获嫌疑人后，在他的硬盘里没有找到外挂源代码或者是其他信息，取证方就可以思考，是否在我们来之前，嫌疑人就已经把硬盘数据删掉了，正常的取证你是取不到什么，所以只能对硬盘做一次数据恢复然后尝试取证。

数据格式分析检索技术，这个比较好理解，数据格式，比如是TXT，word，PPT，EXE都属于一个格式，取证过程对你想要的数据进行格式检索，提高效率之用。

网络取证

网络取证，他是区别于现场取证的，网络取证都是动态的一种取证方式，现有大多数案子都是网络取证的。

数据抓取技术：利用抓包工具（wireshark等），对信息日子进行分析，过滤IP等等。

海量数据与协议分析：有关海量的，必定是基于大数据平台，海量数据与协议分析就是基于大数据平台来获取相关信息。

网络取证的内容也比较多，首先你得对来源进行取证，也就是犯罪嫌疑人所在的位置，取证的内容包括 IP地址，MAC地址，电子邮件（邮件头有IP地址），一些软件或者的互联网的账号等。

有个案例是，一黑客对手机用户进行木马植入，取证方对木马APK进行反编译，从而发现黑客做数据提交的一个IP地址，查阅后发现是某大型云服务器平台，从而提交到相关人员，锁定此人。

事实取证：确定犯罪事实的具体内容和过程。

这个取证的内容包括网络状态和数据包分析、日志文件分析、然后对文件内容进行调查、使用痕迹调查，软件的功能分析等。

数据包分析就和数据抓取一样了，日志文件的分析，像WIN系统都有运行日志，像伪基站系统，他也会记录日志，什么时候向什么人发送了什么信息，这些都是可以在日志里得到的。文件内容调查，取证时对相关文件进行调查分析，比如一个商业机密的Word或者是合同。

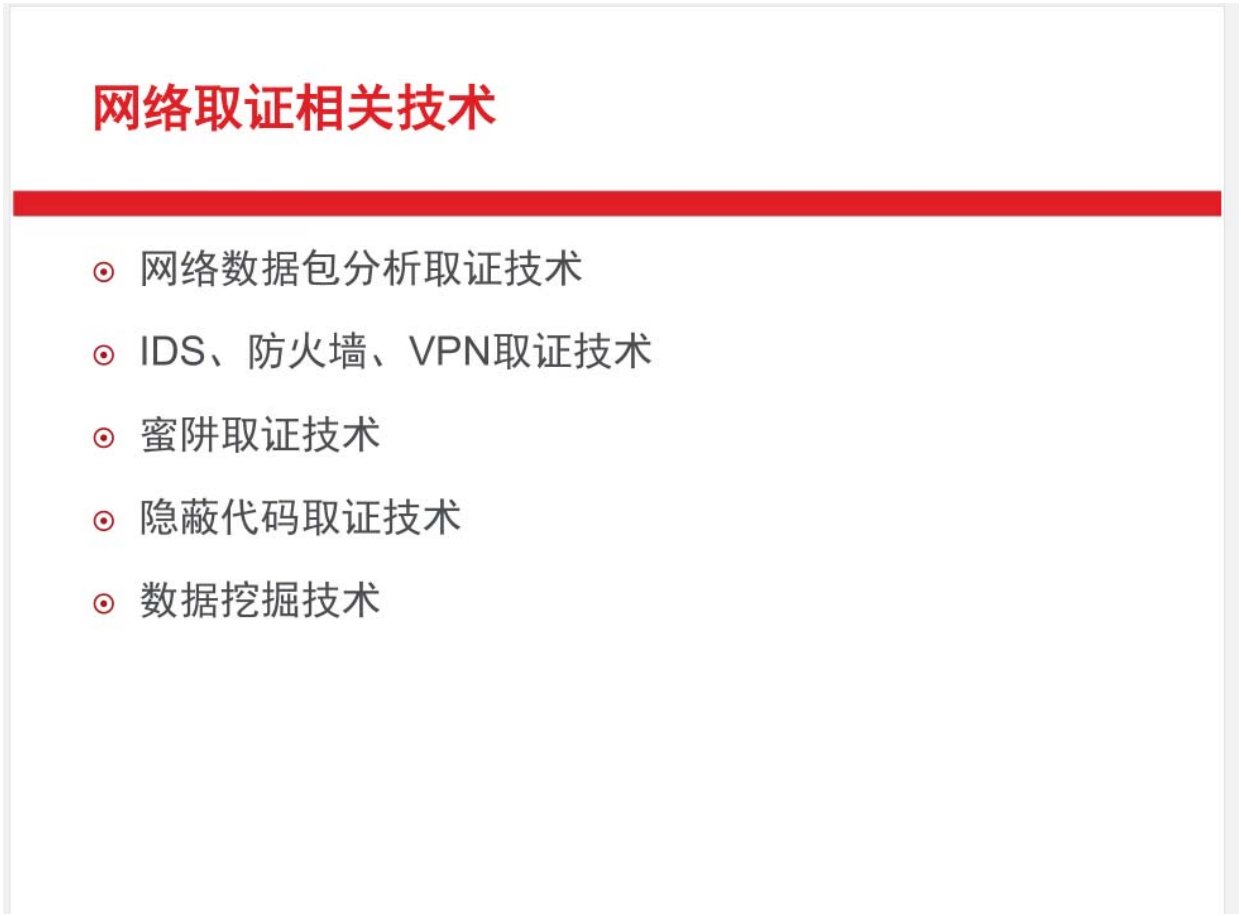
使用痕迹的分析，就好比什么用户，什么时候做了什么事。比如我在我电脑上插上U盘，拔下，拷贝录入和删除行为，都属于使用痕迹。

软件功能分析，有可能涉及到对软件进行OD反编译，如果有源代码就对源代码分析，没有就只能分析软件运行后的行为，一步一步调试。

这个一是自己搭建环境，在环境里对功能进行分析。有点类似于分析一个病毒，他运行后是调用了系统的什么进程。

二就是对源代码分析了，也需要对编程有所了解，得看懂源码才行。

网络取证相关技术



蜜罐取证技术，也就是挖一个坑让你跳，跳了我就知道你的行为。也包括上面所讲的网络数据包分析取证技术。还有一个数据挖掘技术，也就是对数据进行恢复、提取、深入的分析。

网络数据包分析取证工具：

通用网络数据包分析取证工具

工具名称	使用环境	特性
TcpDump&Windump	Unix & Windows	采集过滤
Wireshark	Unix & Windows	采集过滤
Sleuth Kit	Unix	采集过滤、流重组、数据关联
Argus	Unix	采集过滤、日志分析
SNORT	Windows /Unix	采集过滤

功能和优势都列举出来了，需要体验的请自行百度下载。

- 网页取证相关技术

  查看网页使用语言，网站信息，利用爬虫来获取网站相关数据。

- 网站/服务器取证技术

  远程链接登陆服务器（如有需要，获取服务器账号密码的手段不限）查看日志，截获快照，但是全程都必须屏幕录制，是必须。

- 新型取证技术

  新型取证技术，内存取证，芯片取证，云取证，物联网取证，边信道于量子计算。

- 内存取证技术
  - 内存证据以及和硬盘证据成为打击网络违法犯罪的重要依据。
  - 内存证据分析可被用于发现系统的各种关键信息和用户行为特征。
  - 内存取证技术也可被用户恶意代码检测的分析。

内存取证实践：

- 虚拟内存文件

- 休眠文件
- 内存转储
- DMA
- 冷启动

这是一个内存取证的完整过程，1.2名词请百度一下。

DMA的原理就是，数据传输要经过CPU然后再传给电脑，这时候直接转到DMA，不经过CPU，数据传输非常快，但是保存数据量比较小。

冷启动，按住电源键强制启动或者关机，就是冷启动。但是可能会造成数据的丢失，这些数据都保存在内存，冷启动取证就是对机器进行冷却，尽快的恢复数据。

以上就是对内存进行喷冰处理。

内存转储文件：内存转储是用于系统崩溃时，将内存中的数据转储保存在转储文件中，供给有关人员进行排错分析用途。而它所保存生成的文件就叫做内存转储文件。

芯片取证技术

这个是针对现有手机取证工具无法解决的问题，才会搬出芯片取证技术，多涉及硬件。

1. 已损坏的手机（恶意破坏）可以用这种情况。
2. 掩埋，水浸等原因无法开机。
3. 数据接口（USB等）无法使用的手机，但屏幕可正常亮起。
4. 设置密码且没有办法解锁的手机。

以上情况都可用于芯片取证技术。

现在来介绍一下芯片取证的相关流程:

这是个被摔坏的手机。将它拆解，如下图红框里所示的就是芯片所在位置。

拆下芯片后，连接取证工具，选择芯片类型，设置芯片镜像文件，继续对芯片镜像进行读取。

镜像读取完以后就可操作解析恢复，以及仿真。仿真他是模拟手机的实际运行环境。如下图微信红包所示。

一个芯片取证到这里就算结束了。

IOS取证实践：

- 取证工具
- 硬件/系统漏洞
- iCloud
- 社会工程学
- 暴力破解

IOS取证国内外都有工具，但是是比较难的，在没有越狱的情况下。限制比较多。线下取证的话，可使用IOS的备份功能，将应用数据一起备份出来。然后再解析数据。原理实现就是iCloud的备份功能，用过iPhone的都知道，连上了就会问你备不备份 可下载到本地。

漏洞/硬件漏洞，iPhone能被越狱。

iCloud，用户名和密码。

社会工程学：这方面用的很少，主要是利用欺骗手段进行活动。

暴力破解：暴力破解现只支持IOS7以下的版本，高于7 都是无法破解的。

IOS取证流程

1. 吹下iPhone flash芯片并使用复制设备制作副本。
2. 讲副本使用测试架桥接至iPhone主板 然后开机。

3. 穷举密码，进行暴力破解额，5-10次为一轮。

4. 如果触发了IOS安全机制，被锁定或者是数据抹除，那么更换副本继续测试。

5. 直到解锁。

网络取证上，IOS比较难，因为要越狱。线下取证，越不越狱没关系的，都可直接取证。

云取证的相关技术

云端数据的保全和迁移：把数据从一个云端转移到另一个云端，保证数据的完整性。

云端服务重现：对他的服务里找到相关数据。

云数据恢复：数据被删了，想恢复 就必须找到他的云服务商，找到服务器进行数据的恢复，涉及到硬盘恢复等技术。

在线取证：远程提取数据进行取证。

客户端现场取证：和在线取证一个意思。

物联网取证

1. 物理安全。

2. 物理俘获固件器件等修改。

3. 接入层安全。

4. 网络层安全。

5. 应用层安全（数据安全、隐私保护）。

有兴趣的朋友可自行查阅资料，物联网没怎么接触过。

物联网取证相关技术

1. 物联网黑匣子技术。

2. 物联网分布式IDS技术。

3. 物联网嗅探取证技术。

物联网黑匣子技术，是通过访问他的日志，固件更新日志，操作日志，用入侵检测（分布式IDS）的方式进行操作取证。

物联网嗅探取证技术，大多指摄像头监控，什么人在什么地点做什么事，进行取证。

边信道攻击与取证

新手段，边信道攻击简称SCA，是针对加密电子设备在运行过程中的时间消耗或者功率消耗之类、电磁辐射造成的信息泄露来进行攻击。如果想得到你的数据，是根据这几种来实现攻击。说明白一点 就是窃听监听你的数据然后进行破译。感兴趣的朋友可以百度了解一下。

量子计算取证

还没普及，就跳过吧……

到此电子取证技术基础就结束了，硬盘数据恢复这个，军方的条件是格式化37次，基本上无法提取数据了，最简单的方法是格式化后用大文件覆盖硬盘空间，大文件也就指垃圾文件。