计算机取证的相关法律技术问题研究*

丁丽萍[1,3,4+], 王永吉[1,2]

[1](中国科学院 软件研究所 互联网软件技术实验室,北京 100080)
[2](中国科学院 软件研究所 计算机科学重点实验室,北京 100080)
[3](北京人民警察学院,北京 100029)
[4](中国科学院 研究生院,北京 100039)

Study on Relevant Law and Technology Issues about Computer Forensics

DING Li-Ping[1,3,4+], WANG Yong-Ji[1,2]

[1](Laboratory for Internet Software Technologies, Institute of Software, The Chinese Academy of Sciences, Beijng 100080, China)
[2](Laboratory of Computer Science, Institute of Software, The Chinese Academy of Sciences, Beijng 100080, China)
[3](Beijing People's Police College, Beijing 100029, China)
[4](Graduate School, The Chinese Academy of Sciences, Beijing 100039, China)
+ Corresponding author: Phn: +86-10-82620803 ext 804, E-mail: dingliping@itechs.iscas.ac.cn, http://www.iscas.ac.cn
Received 2003-12-12; Accepted 2004-05-08

Ding LP, Wang YJ. Study on relevant law and technology issues about computer forensics. *Journal of Software*, 2005,16(2):260−275. http://www.jos.org.cn/1000-9825/16/260.htm

Abstract: Researchers of law investigate the relevant law features and identification of computer evidences, while computer scientists investigate the technological features and acquisition methods of computer evidences. As an interdiscipline of law and computer science, computer forensics must be studied from the view angle of law, computer science and their combination. The separation of them may result in the confusion of identification in law and technology in computer evidences. In this paper, relevant issues of computer forensics to criminal evidence, law and technology are jointly investigated. Since the technological methods and tools of computer forensics are important, a typical case study is presented and analyzed. Finally based on the analysis of the deficiency for the current work on law enforcement and technology, future work on the improvement of computer forensics is discussed from the viewpoint of both law and computer technology.

Key words: crime; law; evidence; computer evidence; computer forensics; computer forensics tools

摘 要: 法律界研究计算机证据的有关法律特性及其认定,而计算机科学领域的研究人员则从技术的角度研究计

* Supported by the National Natural Science Foundation of China under Grant No.60373053 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2003AA1Z2220 (国家高技术研究发展计划(863)); the Hundred Talents of the Chinese Academy of Sciences (中国科学院"百人计划"); the Chinese Academy of Sciences and the Royal Society of the United Kingdom under Grant Nos.20030389, 20032006 (中国科学院与英国皇家学会联合资助项目)

作者简介: 丁丽萍(1965－),女,山东青州人,博士生,副教授,主要研究领域为计算机取证学;王永吉(1962－),男,研究员,博士生导师,主要研究领域为实时系统,网络优化,计算机取证学,智能软件工程,优化理论,机器人,控制理论.

算证据的技术特征及其获取技术.由于这一学科是建立在法学和计算机科学之上的交叉学科,必须从这两个学科及其派生学科上体现出的特殊性的角度对其进行研究.在这一领域把法律和技术分离的做法会导致法律认定上的错误和技术上的无序性.通过将法律和计算机技术相结合对计算机取证进行了研究.阐明了计算机取证的相关法律问题,重点研究了计算机取证的技术方法和工具,并给出了一个计算机取证实验的技术过程.提出了目前计算机取证相关法律法规和计算机取证技术的不足,指出了今后法律法规的进一步健全、计算机取证工作的规范化和计算机取证技术的发展趋势.

关键词: 犯罪;法律;证据;计算机证据;计算机取证学;计算机取证工具

中图法分类号: TP391 **文献标识码**: A

打击犯罪的关键在于获得充分、可靠和强有力的证据.取证涉及到法律和技术两个方面.一般犯罪证据的提取目前已经有很多较为成熟的技术,例如,指纹提取和识别、法医鉴定、DNA 鉴定等等.这些技术的特点是针对性强,获取的证据符合证据的采用标准,而且取证工作有明确的法律依据.随着计算机应用的普及,计算机犯罪和其他犯罪的很多证据都以数字形式通过计算机或网络进行存储和传输,从而出现了计算机证据.由于计算机证据具有与一般犯罪证据不同的特点,例如,计算机证据的脆弱性、不可靠性、表现形式的多样性等,所以对其获取和可靠性的保证一直是计算机犯罪案件和其他与计算机有关的犯罪案件侦破工作的难点.因此,计算机取证(computer forensics)技术的研究变得越来越迫切.计算机取证学作为计算机科学和法学的交叉学科应运而生.美国十分重视计算机取证学的研究,至少有 70%的法律部门拥有自己的计算机取证实验室[1],经过资格认定的取证专家使用专门技术,通过网络或对从犯罪现场获取的计算机机器设备进行证据的提取和分析,目的在于提供非法活动的证据,并将这些证据提交给法庭,作为裁决的依据.近年来,我国法学家开始重视计算机证据这一特殊证据类型的研究,提出了一些新的法律观点,并呼吁学术界和立法界重视这方面的研究,否则,就会落在发达国家后面[2].另一方面,有些计算机科学家已在从事计算机取证技术的研究并开发出了一些应用系统.我国政府也已经意识到了计算机证据的重要性,在公安部门设置了网络安全监察机构,该机构的一部分职责是负责计算机证据的获取.然而,在我国,专门的取证机构还没有,计算机取证人员的认证也没有实行,律师界对计算机证据的认识还很模糊和肤浅.本文从法律和计算机技术两个角度对计算机取证进行了研究.阐明了计算机取证的相关法律问题,重点研究了计算机取证的技术方法和工具,并给出了一个计算机取证实验的技术过程.提出了目前计算机取证相关法律法规和计算机取证技术的不足,指出了今后法律法规的进一步健全、计算机取证工作的规范化和计算机取证技术的发展趋势.

本文第 1、第 2 节研究计算机取证的法律问题以及计算机取证应遵循的一般原则和步骤.第 3、第 4 节对计算机取证技术进行研究并给出一个取证案例实验的过程.第 5 节指出目前相关法律和技术上存在的问题以及计算机取证的发展趋势.第 6 节给出本文的结论.

1 计算机证据的法律认定

1.1 一般证据及计算机证据

所谓证据就是指能够证明案件真实情况的事实[3].我国的《刑事诉讼法》把证据分为 7 类: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;鉴定结论;勘验、检查笔录;视听资料.证据的基本功能就在于它具有证明一定事实存在与否的作用.对于司法和执法活动而言,这一作用具有特别重要的意义.司法和执法活动的两项基本任务是:(1) 准确认定案件事实或其他争议事实;(2) 正确使用有关的法律规定.其中,第 1 项任务无疑应排在首位.所以,准确认定案件事实,即获取证据,是正确使用法律的前提条件和基础.

在国内外证据法中,对于计算机证据的定义莫衷一是.我国证据法学家认为,计算机证据是存储有电子数据的电、磁、光记录物,这些电子数据是在计算机系统运行过程中产生的,并能够以其内容证明案件事实,它们可以转换为不同的输出表现形式[4].计算机证据本身的存在形式是数字形式,人们不能识别,只有计算机及其附属设备才能识别.在理解计算机证据的概念时应该特别注意不要将计算机证据的输出形式作为标准来界定计算

机证据,因为同样的计算机可能产生各种不同的输出形式,这将导致将同样的证据归入不同的证据种类.同时,这样做还容易将计算机证据与传统的证据方式混杂在一起,法律就难以同时对其特性及运用规则做出比较准确的规定.

1.2 计算机证据的特征[4,5]

数字性.计算机证据的物质载体是电子元件和磁性材料等.行为人蓄意操作、改变数据或程序,从物理表示上,只是集成电路的电子矩阵正负电平或磁性材料磁体等发生了变化.获取这些行为的证据需要特殊的手段,与其他证据的获取是完全不同的.

技术性.计算机证据的产生、储存和传输及其采集、分析和判断都必须借助于计算机科学中的计算技术、存储技术、网络通信技术等.

脆弱性.由于计算机信息容易被修改,并且对其进行真正意义上的修改(不可恢复的修改)后不留痕迹,从而使得计算机信息具有了脆弱而不可靠的一面.人为操纵数据和程序的破坏在某种程度上具有普遍性.计算机系统对数据的处理具有环节多、使用的技术和设备复杂的特点.而且由于计算机的处理速度越来越快,数据的修改都是在瞬间完成的.所以,计算机证据有时是不可靠的.

多态性.是指计算机证据的表现形式是多种多样的,即不同形态的输出材料的证明力都来源于同一计算机存储的信息本身.虽然不同的证据表现形式并不能说明其在审查判断上有根本的区别,但是不同形态的证据材料的审查规则是不同的.

人机交互性.这是指计算机证据的形成,在不同的环节上有不同的计算机操作人员的参与,它们在不同程度上都可能影响计算机系统的运转.而且,这种影响的层次和程度与这些人员的工作性质有关.也就是说,计算机管理员、网络管理员、程序员、系统分析人员以及一般的计算机操作人员对数据信息的影响是不同的.所以,可能出现的问题也就存在于人、机两个方面.为了保证证据的可靠性和真实性,应该从技术和管理上严格控制人机系统.同时,在采集和获取计算机证据时应注意分析人和机器两个方面.

复合性.证据的复合性是指当证据以某种形式表现时,往往就具有了这种表现形式的证据特征.而计算机证据的表现形式是多种多样的,可以是打印在纸上的文字,可以是在显示器上输出的视频、图像、文字,也可以是声音设备输出的声音.在诉讼活动中采纳某一证据形式时,应当既考虑该证据生成过程的可靠程度如何,又考虑这一证据的表现形式是否被伪造、变造或剪辑、删改过.所以,计算机证据的采集和鉴别不仅技术含量高,而且其过程是十分复杂的.

基于计算机证据的以上特征,我国有的法学专家建议,把计算机证据作为一个独立的证据种类或者用“电、磁、光记录物”取代视听资料作为一个证据种类以涵盖视听资料和计算机证据.

1.3 计算机证据的采用标准[3]

计算机证据的采用标准和其他证据的采用标准相同,就是指什么样的证据可以被采用.这是研究计算机取证学必须弄清楚的法律问题.这些标准包括:

采用证据的客观性标准.证据的客观性是指证据应该具有客观存在的属性,或者说,证据应该是客观存在的东西.这一客观性包括两个方面,首先,证据的内容必须具有客观性,必须是客观存在事物的反映.其次,证据必须具有客观存在的形式,必须是人们能以某种方式感知的东西.电子证据一般存储在计算机中,这些数据信息毫无疑问是证据,而且也具有一定的客观性,但是因为它们仅存储在计算机中,犹如存在于人脑中的与案件有关的信息一样,没有以人们可以感知的形式表现出来,所以还不符合采用证据的标准.当这些数据信息被计算机显示器显示或打印机打印出来,呈现在法庭上时,电子证据就具备了证据的客观表现形式,也就可以采用了.

采用证据的关联性标准.证据必须与需要证明的案件事实或其他有争议的事实具有一定的联系,这就是证据的关联性或称相关性.人们对证据关联性的认识也受到科学技术水平等因素的影响.过去,人们没有找到工具恢复硬盘上删除的数据,这时删除的与某一案件事实有关的数据就不具备关联性.现在,我们已经可以很容易地将这些删除的证据恢复并呈现在法庭上,那么,由于技术的进步,就使得这些数据具有了证据的关联性.

采用证据的合法性标准.合法性是采用证据的重要标准之一,它包括主体、形式和程序 3 个方面的合法性:

第一,证据的主体必须符合有关法律的规定.这主要指各种“人证”,那些不具备证人能力的人提供的证据即使具备了客观性和关联性,也不能采用;同样,不具备鉴定资格的人做出的鉴定结论当然也不符合证据的合法性标准.所以,计算机取证技术的使用必须是那些具备取证资格的人,否则,即使获取了足够的具有客观性和关联性的证据,也是属于主体不合法的证据,不能采用.第二,证据的形式必须符合有关法律的规定.例如,我国刑事诉讼法规定,鉴定结论和勘验检查笔录上必须由鉴定人员或勘验检查人员签名盖章,因此那些没有上述人员签名盖章的鉴定结论和勘验笔录就属于不合法的证据,不能采用.第三,证据的提取方法和收集程序必须符合法律的有关规定.例如,我国行政诉讼法规定被告人在诉讼过程中不得自行向原告和证人收集证据,如果作为行政诉讼被告人的行政机关这样做了,那么收集的证据也因程序不合法而不能采用.

2 计算机取证的原则和步骤

2.1 计算机取证的原则

收集证据是指证明主体依法直接提取、采集并掌握证据的诉讼或非诉讼法律活动.在刑事诉讼活动中,收集证据的方法主要是现场勘验、检查、询问证人、询问被害人、讯问被告人、搜查、扣押和鉴定[3].计算机取证,亦即收集计算机证据,是指在现场勘查和搜查的过程中及时、准确、完整地获取或者扣押与案件有关的电子数据和相关的证据[3].计算机在相关的犯罪案件中(不仅仅是计算机犯罪案件)要么是入侵的目标,要么是作案的工具,要么是犯罪信息的存储器.无论作为哪种角色,机器中都会存留大量的与犯罪有关的数据信息.计算机取证就是对能够为法庭接受的、足够可靠和有说服性的、存在于计算机和相关外设中的电子证据加以确认、保护、提取和归档的过程.这一过程从某种意义上讲是一个重建犯罪事件的过程[1].

计算机证据的易受损特性对收集证据、审查判断证据都提出了严格的程序要求.犯罪分子大部分是通过网络而不需要到案发现场去作案.另外,由于网络的无国界性,不同国家在法律、道德和意识形态上是有差异的,可能会造成案件无法继续侦查的结果.根据这样的特点,计算机取证的原则是[4,6]:

及时性原则.这一原则要求计算机证据的获取有一定的时效性.

取证过程合法的原则.这一原则要求计算机取证过程必须按照法律的规定公开进行,得到第一手具有证明力的公正的证据信息.

多备份的原则.即对于含有计算机证据的媒体至少应制作两个副本,原始媒体应存放在专门的房间由专人保管,复制品可以用于计算机取证人员进行证据的提取和分析.

环境安全的原则.计算机证据应妥善保存,以备随时重组、试验或者展示.该原则是指存储计算机证据的媒体或介质应远离高磁场、高温、灰尘、积压、潮湿、腐蚀性化学试剂等.在包装计算机设备和元器件时尽量使用纸袋等不易产生静电的材料,以防止静电消磁.环境安全的原则还要求防止人为地损毁数据.

严格管理过程的原则.含有计算机证据的媒体的移交、保管、开封、拆卸的过程必须由侦查人员和保管人员共同完成,每一个环节都必须检查真实性和完整性,并拍照和制作详细的笔录,由行为人共同签名.

2.2 计算机取证的步骤

2.2.1 保护现场和现场勘查

现场勘查是获取证据的第 1 步,主要是物理证据的获取.这项工作可为下面的环节打下基础.包括封存目标计算机系统并避免发生任何数据破坏或病毒感染,绘制计算机犯罪现场图、网络拓扑图等,在移动或拆卸任何设备之前都要拍照存档,为今后模拟和还原犯罪现场提供直接依据.要特别注意保证“证据连续性”,即在证据被正式提交给法庭时,必须能够说明在证据从最初的获取状态到在法庭上出现的状态之间的任何变化,当然最好是没有任何变化.而且,整个检查、取证过程必须是受到监督的.也就是说,所有调查取证工作,都应该有其他方委派的专家的监督.

2.2.2 获取证据[6,7]

证据的获取从本质上说就是从众多的未知和不确定性中找到确定性的东西.而取证人员面对的是各种互不相同的案件,有些甚至不是刑法上定义的计算机犯罪而是其他犯罪.比如,北京市公安局与福建省公安厅联合

破获的某两高校爆炸案,从性质上看纯粹是一般刑事案件,但该案件的证据又确实是通过计算机获取的,即计算机取证应该不仅仅是计算机犯罪证据的获取.因此,证据的获取很难说有固定的、一成不变的方法和模式,应该具体问题具体分析.有的专家认为:检查一个系统并保持可靠证据的理想的方法是冻结现有的系统并分析原有数据的拷贝.但是这种做法并不总是可行的.例如,机器设备的冻结是否合法、是否会引起非议以及在停机时间难以确定时,是否会引起有关人员的反对,等等.所以,证据的获取方法首先要合法并不会造成太大的政治和经济上的损失.又比如,对于到达现场后,是否要立即切断电源的问题也应当具体分析,在理想状态下,任何一台需要分析的计算机都可以关掉电源,重新启动,做一个完整的镜像.这时,取证人员只需取下硬盘,进行各种操作即可.但是,这仅仅是个理想状态而已.计算机取证领域最具争议的话题之一就是:在取证时究竟是让一台计算机继续运行还是立即拔掉电源,或者进行正常的管理关机过程.大多数取证人员为了使计算机停留在当前状态,采取立即拔掉电源的做法,但是这一做法会毁掉入侵过程中的相关数据,并且可能毁坏硬盘上的数据.最好的做法应该是针对现场的具体情况迅速做出判断,采取最合适的方法获取证据.

一般地,在防止远程攻击的同时保护现场证据的方法有如下几种:

在取证中心设置专门的取证计算机来进行硬盘的检查和镜像.这一方法不必担心可疑主机上软硬件环境的有效性.产生的证据在法庭上很容易得到认可.但是这样做很不方便,要取走可疑主机的硬盘,比较费时,容易丢失数据.

将可疑主机关闭后,用经过验证的写保护的软盘或者光盘启动被检查的系统.这种做法方便、快捷.如果可以将硬盘以只读方式装载的话,产生的证据是比较有证明力的.但是,这一方法容易使可疑主机的硬件系统受损害,也容易丢失数据.

使用经过验证的软件的外部介质来检查原有的系统.这种方法方便、快捷,能够检查易失信息.但当系统内核受到损害时,产生的结果可能是错误的.同时,外部介质可能不具备所有需要的工具.

首先验证可疑系统上的软件,并使用经验证的本地软件来进行检查.这需要很少的前期准备,可以检查易失信息,能够进行远程的检查.然而,由于这样做缺乏对可疑硬盘的写保护,使得产生的证据很难具备可靠性.而且,需要的时间较长.

不经验证地使用可疑系统上的软件检查可疑系统.这样的做法所需要的准备时间最少,可以检查易失信息,能够进行远程检查.但是,这样做是最不可靠的,入侵者也最希望取证人员采用这种技术以便他们发现后采取反取证措施.所以,在某种情况下,这种方法完全是浪费时间.

因为计算机证据必须是真实、可靠、完整和符合法律规定的[7],所以,取证人员在开始取证阶段所采取的行动对整个取证工作是至关重要的,如果这一阶段采取的方法不正确或者程序不正确都会导致证据可靠性的丧失,甚至一无所获.

2.2.3 鉴定证据[6,9]

计算机证据的鉴定主要是解决证据的完整性验证.计算机取证工作的难点之一是证明取证人员所搜集到的证据没有被修改过.而计算机证据又恰恰具有易改变和易损毁的特点.例如,腐蚀、强磁场的作用、人为的破坏等等都会造成原始证据的改变和消失.所以,取证过程中应注重采取保护证据的措施.例如,可以采用形成所谓的证据监督链的技术和方法.电子指纹技术是常用的技术,它也被称为数字指纹,其对象可以是单个的文件,也可以是整张软盘或者硬盘.其原理是:如果一方的身份“签名”未与任何应签署的报文(message)本身相联系,就留下了篡改、冒充或抵赖的可能性.我们需要从报文中提取一种格式确定的、符号性的摘要,以将千差万别的报文与数字签名不可分割地结合起来,这种“报文摘要(message digest)”就是“数字指纹(digital fingerprint)”.在开始取证时就使用数字指纹技术,而且,每做一个分析动作,都要再生成数字指纹,与分析前进行对比以保证所收集到的证据是可靠的[10].

时间戳也是取证工作中非常有用的技术,必将成为一种有效的证据鉴定方法.它是对数字对象进行登记来提供注册后特定事物存在于特定日期的时间和证据.时间戳对于收集和保存数字证据非常有效.因为它提供了数字证据在特定的时间和日期里是存在的,并且从该时刻到出庭这段时间里不曾被修改过.

2.2.4 分析证据

这是计算机取证的核心和关键.证据分析的内容包括:分析计算机的类型、采用的操作系统是否为多操作系统或有无隐藏的分区;有无可疑外设;有无远程控制、木马程序及当前计算机系统的网络环境.注意,分析过程的开机、关机过程,尽可能地避免正在运行的进程数据丢失或存在的不可逆转的删除程序.分析在磁盘的特殊区域中发现的所有相关数据.利用磁盘存储空闲空间的数据分析技术进行数据恢复,获得文件被增、删、改、复制前的痕迹.通过将收集的程序、数据和备份与当前运行的程序数据进行对比,从中发现篡改痕迹.可以通过该计算机的所有者,或电子签名、密码、交易记录、回邮信箱、邮件发送服务器的日志、上网 IP 等计算机特有信息识别体,结合全案其他证据进行综合审查.注意,该计算机证据要与其他证据相互印证、相互联系起来综合分析.同时,要注意计算机证据能否为侦破该案提供其他线索或确定可能的作案时间和罪犯.

例如,在单机情况下,证据的分析过程是:首先,使用 Image MaSSter(这是专门为取证设计的工具,有两个版本,这是桌面版.可移动的便携式版本的体积比硬盘稍大,适合室外使用,称为 Solo Forensic)制作两份原始数据的备份[6],每次取出原始证据都必须在报告和证据监督链记录中作相应的记录.并且要严格按照操作准则(尽管目前没有这种准则)进行.在处理证据时,采取的步骤越少,发生错误的可能性就越小.分析过程是基于原始证据的数字拷贝进行的.一份用于取证的备份必须是对原始数据的每一比特的精确克隆.同时,在进行分析之前,一定要为被分析的数据生成数字指纹.接下来就可以正式开始分析工作了.第 1 步,分析硬盘的分区表.分区表的内容是我们最后提交法庭的报告的一项重要内容,同时也决定着下一步分析工具的采用.假定是 NTFS 格式的分区,则只能采用支持这种格式的工具,例如,“诺顿反删除”工具就不支持这种文件系统.第 2 步,浏览文件系统的目录树并将其打印出来.在分析过程中还要注意分析犯罪嫌疑人的技术实力.假如他经常使用加密程序或解密程序的话,他可能就是一位狡猾的高手.第 3 步,进行关键字搜索.使用特制的取证程序检查主引导区记录和引导扇区.要特别注意那些标记为已损坏的簇,要使用工具仔细检查,因为其中可能藏有有效的证据.关键字要进行多种可能的变换,如张三,可以用张三、章三等进行多种尝试.第 4 步,使用数据恢复工具找回那些已经被删除的文件.第 5 步,使用专门的工具软件,检查文件系统中的未分配空间和闲散空间以寻找残留的数据.最后,对找到的证据做多份备份,并制作成具有可读性的文件.

2.2.5 进行追踪

上面提到的计算机取证步骤是静态的,即事件发生后对目标系统的静态分析.随着计算机犯罪技术手段的升级,这种静态的分析已经无法满足要求,发展趋势是将计算机取证与入侵检测等网络安全工具和网络体系结构技术相结合,进行动态取证.整个取证过程将更加系统化并具有智能性,也将更加灵活多样.对某些特定案件,如网络遭受黑客攻击,应收集的证据包括:系统登录文件、应用登录文件、AAA 登录文件(比如 RADIUS 登录)、网络单元登录(network element logs)、防火墙登录、HIDS 事件、NIDS 事件、磁盘驱动器、文件备份、电话记录等等.对于在取证期间犯罪还在不断进行的计算机系统,采用入侵检测系统对网络攻击进行监测是十分必要的.也可以通过采用相关的设备或设置陷阱跟踪捕捉犯罪嫌疑人.

2.2.6 提交结果

打印对目标计算机系统的全面分析和追踪结果,然后给出分析结论:系统的整体情况,发现的文件结构、数据、作者的信息,对信息的任何隐藏、删除、保护、加密企图,以及在调查中发现的其他相关信息.标明提取时间、地点、机器、提取人及见证人.然后以证据的形式按照合法的程序提交给司法机关.

3 计算机取证的技术方法和工具

取证专家 Reith Clint Mark 认为:计算机取证(computer forensics)可以认为是“从计算机中收集和发现证据的技术和工具[11]”.以计算机证据的来源为标准,计算机取证技术可分为:单机取证技术、网络取证技术和相关设备取证技术.

3.1 基于单机和设备的计算机取证技术

单机取证技术是针对一台可能含有证据的非在线计算机进行证据获取的技术.包括存储设备的数据恢复

技术、隐藏数据的再现技术、加密数据的解密技术和数据挖掘技术等等.

3.1.1 数据恢复技术

数据恢复技术主要用于把犯罪嫌疑人删除或者通过格式化磁盘擦除的数字证据恢复出来.由于磁盘的格式化只不过是对用于访问文件系统的各种表进行了重新构造,因此,如果格式化之前的硬盘上有数据存在,则格式化操作后这些数据仍然存放在磁盘上,同时,格式化操作会创建一个新的空索引列表,指向未分配数据块.删除文件的操作也不能真正删除文件,只不过把构成这些文件的数据簇放回到系统中,对于通常的读写操作而言,这些簇不可见.这些簇可以从空闲块列表中得到,而从目录项(或inode项)中访问不到.可能包含已删除数据的文件系统区域有:应用程序产生的数据文件可能包含文件系统中的游离数据;文件的最后一个簇通常会因为没有被完全使用而使得上次写进这个簇的数据没有被全部覆盖;不在使用中的文件系统的未分配数据块(或簇);计算机的当前配置可能没有使用硬盘上的所有空间,但以前的配置可能使用了,则这些空间就可能含有隐藏数据[12];分区表和引导信息所在的磁道也可能有证据信息.我们可以使用国际取证专家普遍看好的取证软件TCT(the Coronor's toolkit)和Encase等把这些数据恢复出来.即使使用安全删除工具删除的数据也会留有痕迹,擦除一个磁道的数据时留下的边缘数据和被覆盖后仍留下的痕迹称为阴影数据(shadow data)[13],我们可以使用特殊的电子显微镜一比特一比特地恢复写过多次的磁道.一个有经验的技术人员,可以使用适当的设备恢复被覆盖过7次以上的数据.在国外,这种数据恢复服务公司或机构已经存在,如Ontrack公司、Ibas实验室等[6].

3.1.2 加密解密技术和口令获取

取证在很多情况下都面临如何将加密的数据进行解密的问题.目前的加密解密算法及工具很多[14],计算机取证中使用的密码破解技术和方法主要有:

- 密码分析技术.这种技术需要取证专家具有密码学专业领域的知识,目前的软件工具也并不实用.
- 密码破解技术.包括口令字典、重点猜测、穷举破解等技术.其中口令字典一般是基于软件的,而且已经有了多种字典可供使用.目前基于字典的口令破解软件很多,如专门用于Office文件的破解工具AOPR(advanced office password recovery)等,这些软件的破解效率很高.
- 口令搜索.包括物理搜索(在计算机四周搜查可能有口令的地方)、逻辑搜索(在文档或电子邮件中搜索明文的口令)和网络窃听(从网络中捕获明文口令).
- 口令提取.许多Windows的口令都以明文的形式存储在注册表或其他指定的地方,我们可以从注册表中提取口令.
- 口令恢复.使用密钥恢复机制可以从高级管理员那里获得口令.

3.1.3 信息搜索与过滤技术

在计算机取证的分析阶段往往使用搜索技术进行相关数据信息的查找.这些信息可以是文本、图片、音频或视频.这方面的技术主要有:数据过滤技术、数据挖掘技术等.目前这方面的软件种类繁多,既有基于单机的,又有基于网络的,例如美国NTI(New Technologies Inc.) (http://www.forensics-intl.com)公司的系列取证工具中的Filter、中软通用产品研发中心信息安全实验室的NetMonitor网络信息监控与取证系统等.

3.1.4 磁盘映像拷贝技术[15]

由于证据的提取和分析工作不能直接在被攻击机器的磁盘上进行,所以,磁盘的映像拷贝技术就显得十分重要和必要了.而且,这一技术可以实现磁盘数据的逐字节拷贝.

3.1.5 反向工程技术[16]

反向工程技术用于分析目标主机上可疑程序的作用,从而获取证据.但目前这方面的工具还很少.

3.2 基于网络的计算机取证技术

所谓网络取证技术就是在网上跟踪犯罪分子或通过网络通信的数据信息资料获取证据的技术.包括IP地址和MAC地址的获取和识别技术、身份认证技术、电子邮件的取证和鉴定技术、网络侦听和监视技术、数据过滤技术、漏洞扫描技术等[6,15].应该说,在基于网络的犯罪日益猖獗的今天,网络取证技术在计算机取证技术中占有举足轻重的地位.

3.2.1 IP 地址获取技术

使用 ping 或 traceroute 命令.ping 是一个简单而又非常有用的程序,它使用了因特网信息控制协议的 ECHO_REQUEST 数据报.该数据报向目标主机发送请求并监听 ICMP 应答.我们可以使用 ping 这样的命令来判断一台机器是否在线,然后使用 What's Up Gold 这样的程序继续检查这台运行着的机器.但是,ping 的这种行为很容易被原段的系统发现.假设对方是一个比较高明的罪犯,它会对到自己机器的任何连接进行监视,从而发现你在调查他[1].

混乱地址的恢复[1].很多高明的罪犯使用伪造或混乱的地址发送信息,这种混乱的地址一般是一个 10 位长度的整数,例如http://2280853951.我们可以采用数学方法变换成正常格式:把 2280853951 转换成 16 进制为:87F311BF.然后依次把每两位 16 进制数都转换成 10 进制数并加上点号.不足两位在后边加 0.从而,得到正常的 IP:135. 243. 17.191.反过来也可以把正常的 IP 转化成 10 进制数.

当然,也可以使用 ping 命令:C:\>ping 2280853951 得到结果 Pinging 135. 243. 17.191 with 32 bytes of data.

使用 IP 扫描工具程序,就是利用特定的软件来获取 IP 地址.一般这种软件都有较好的用户界面,使用简单、方便.例如 solarwins 2001 engineers edition 就有扫描 IP 的功能.

由 DNS 获取 IP 地址.大多数的域名服务器都支持逆向查询,也可以使用工具软件进行手工解析.一流的工具就是 nslookup,它可以基于 Windows,UNIX 和 NT 使用,可以进行正向和逆向查询[1].

MAC 地址的获取.MAC 地址只能使用在硬件层上,IP 地址和 MAC 地址的转化是通过查找 ARP 表来实现的,该表是由地址解析协议(address resolution protocol,简称 ARP)自动创建的.需要指出的是,MAC 地址也不能绝对信赖,因为现在已经有了很多软件可以对 MAC 地址进行修改,在 UNIX 中,就可以通过命令改变 MAC 地址.所以,有时虽然找到的 MAC 地址与我们所掌握的并不一致,但也不能说明这台主机不是嫌疑主机[1].

在 ISP 的支持下获取 IP.一个互联网服务提供商一般通过 RADIUS 协议支持拨号路由器和中央用户目录之间的验证请求,这一协议可以用于用户身份认证,也可以用于记帐.RADIUS 服务器通常是互联网服务提供商可以为跟踪罪犯提供记录的唯一设备,所以它对取证工作非常重要,该服务器通常会将每一个注册请求的记录保存一年以上.一般情况下,ISP 都愿意提供这种日志,因为他们也不想让人利用他们的系统从事非法活动.

3.2.2 针对电子邮件和新闻组的取证技术[6]

电子邮件和新闻组的共同特征是,都是用简单的应用协议和文本存储转发,只允许信息在多个中间系统上穿过,信息的主体由可打印的字符构成,头信息中包含了从发送者到接收者之间的路径.所以,可以对信息发送路径上的痕迹进行分析以获取证据.

从电子邮件中获取证据.电子邮件中获取证据的关键是要了解电子邮件协议中的邮件信息的存储位置,例如接收邮件时,对于仅存储收到信息的 POP3 协议,我们必须访问工作站才能对邮件进行跟踪.基于网页进行发送和接收的 HTTP 协议将发送和接收到的信息存储到服务器上,但可以手工下载到本地,有助于鉴别假冒行为.微软的邮件应用编程接口(mail API,简称 MAPI)能够存储所有信息.发送邮件时,采用的协议是 SMTP(简单邮件传输协议),网络黑客最先学会的技术之一就是如何通过 telnet 到 SMTP 服务器的 25 端口手工发送邮件信息.他们可以插入任何信息到你要发送的邮件的头文件中——包括伪造的源地址和目标地址.他们也可能在配置邮箱时选择手工输入发信人地址,由于 SMTP 没有强壮的认证机制,在没有使用 PGP 或者 S/MIME(安全的多功能互联网邮件扩展)来添加数字签名的情况下,邮件信息的可信度非常低.跟踪伪造的电子邮件的主要方法是请求 ISP 的帮助.而取证人员必须学会解读邮件的头文件,一般的邮件服务程序,如 FOXMAIL,OUTLOOK,OUTLOOK EXPRESS 等都可以通过选中某邮件再执行菜单命令(如文件菜单中的属性或查看菜单中的选项)来查看详细的头信息.跟踪电子邮件还可以使用专用工具,如 NetScanTools.在使用 Eudora 作为客户端软件时要想看到头文件只需在"Blah Blah Blah"按钮上单击鼠标即可.从电子邮件中获取证据的一个例子就是某两所高校爆炸案的破获.案件发生后,公安机关立即对互联网特别是校园网进行严密监控.3 月 6 日 11 时许,其中一高校宣传部电子信箱收到一用户名为 huanglaoxie0225@yahoo.com.cn 发来的邮件,发件人在邮件中自称是两校爆炸案的制造者,并在该邮件中详细叙述了爆炸装置的构成情况,从该人叙述内容看与公安机关现场勘查情况基本相符,他声称第 3 次爆炸在所难免,并约定收件人于 3 月 7 日~10 日在搜狐网站"校园文化"聊天室中以"水木清

华”的昵称与其讨论有关两校爆炸的相关情况.3 月 6 日 14 时,公安机关对高校宣传部接收邮件的计算机进行了现场勘查,获取了邮件信头等相关信息,并立即针对 huanglaoxie0225@yahoo.com.cn 开展调查取证工作.经查,该邮箱于 2003 年 3 月 5 日 9 时在雅虎中国网站申请注册,注册 IP 地址为 218.104.239.141,经使用取证技术定位,其物理位置在福建省福州市,由中国网通福建分公司分配给某网吧使用.该邮箱曾于 3 月 5 日凌晨 3 时 32 分、6 时 23 分相继向两所高校、新华网和中华网发送了内容相同的电子邮件.后经一系列的侦查取证确定了犯罪嫌疑人并在确认的网吧将其抓获[17].

跟踪新闻组.跟踪的方式和跟踪电子邮件一样,不仅需要与所跟踪的信息有关的所有新闻服务器管理员协作,而且还需要这些管理员提供足够的日志文件.但由于繁忙的 USENET 站点每天都更新日志文件,所以必须在 24 小时内完成取证工作.跟踪的方法也是从接收点开始往回查找,对路径中的每个主机进行验证.解读分析头文件信息也是关键.

3.2.3 网络输入输出系统取证技术[6]

可以使用 NetBIOS 的 nbtstat 命令来跟踪嫌疑人.该命令可以获取嫌疑人的机器所在的域名和 MAC 地址等,并且可以将获得的信息打印出来.最常用的技术是入侵检测技术(IDS).IDS 一般有两种使用方式:基于网络的和基于主机的.基于网络的 IDS 一般安置在一个网段内,检查网段内每台主机的流量,寻找未授权活动的证据.例如,可以使用一台单独的集中式的入侵检测引擎,记录所有的日志信息,并基于多个运程传感器所提供的数据报警,这些传感器位于多个不同的局域网段中.基于主机的 IDS 在单个主机上执行监测功能.IDS 一般分为检测特定事件的和检测模式变化的.检测特定事件的 IDS 需要定期更新特征数据库,因为它无法对数据库以外的事件发出警报.检测模式变化的模型使用人工智能技术,创建一个系统,用来监测反常行为,当超出了正常模式时,就可以报警.IDS 对取证的最大帮助是它提供的日志或记录功能可以被用来监视和记录犯罪嫌疑人的行为.

3.2.4 网络入侵追踪技术[18]

入侵追踪的最终目标是能够定位攻击源的位置,推断出攻击报文在网络中的串行路线,从而找到攻击者.IP 报文入侵追踪技术包括连接检测、日志记录、ICMP 追踪法、标记报文法等,可以追溯到发送带有假冒源地址报文的攻击者的真实位置,还可以发现在网络连接链中“前一跳”的信息,特点是需要利用路由器作为中间媒介.基于主机的入侵追踪技术的代表是“身份识别系统(caller identification system in the internet environment,简称 CISIE)”,它是为了在多台“跳板”主机追踪入侵者而设计的,其目标是向处于“扩展连接”的主机报警,提供前若干跳的连接信息.此外,“指纹技术(thumbprint)”、网络入侵源点的安全管理框架(decentralized source identification for network-based intrusions,简称 DecIDUouS 系统)以及协同的入侵追踪和相应框架(CITRA)、基于主动网络的入侵检测框架 SWT(sleepy watermark tracing)都是具有代表性的入侵追踪技术.

3.2.5 人工智能和数据挖掘技术

计算机的存储容量越来越大,网络的传输速度越来越快.对于计算机内存储的和网络中传输的大量数据,可以应用数据挖掘技术以发现与特定的犯罪有关的数据.有的专家提出了 NFAT(network forensics analysis tools)[19]的设计框架和标准,核心是开发专家系统(expert system,简称 ES)并配合入侵检测系统或者防火墙,对网络数据流进行实时提取和分析,对于发现的异常情况进行可视化报告.

3.2.6 网络取证应用案例分析:计算机运行环境勘查取证系统[17]

该系统融合了网络扫描技术、系统扫描技术和网络侦听技术,运行在 Windows NT 操作系统之上,采用汉字图形界面,操作方便、快捷.在保护系统(网络)环境原有运行状态的前提下,通过多层次的网络协议实现跨平台的网络通信;利用动态漏洞知识库实现了系统和网络的漏洞检查;利用模型检查,实现了安全协议分析,在现有的 ICMP 协议工具的基础上,利用有效的算法,实现了网络拓扑信息的搜索;采用 SNMP,DNS 等网络协议实现了网络隐藏信息的自动挖掘和关联搜索,并最终生成对涉案的全部系统的网络描绘和证据的提取.该系统采用的关键技术包括:网络设备;主机配置的自动扫描与分析技术;基于网络管理协议(SNMP)的分析技术;网络拓扑结构的自动发掘与分析技术;网络中应用和服务等信息的自动采集与表达技术;运行环境与拓扑结构的可视化技术;利用模型检验技术以实现对系统的漏洞分析.其技术上的创新点主要有:利用 TCP/IP 协议栈指纹识别技术进行操作系统类型探测;多种端口扫描机制使得扫描更加隐蔽;利用优化的多线程并发扫描机制可使扫描效率大大

提高;利用漏洞规则库结合动态链接库技术进行漏洞扫描,便于系统升级和扩充;基于插件技术的系统扫描技术统一解决系统扫描的多版本问题.其系统的结构如图 1 所示.

但是该系统主要是针对网络证据进行勘查和定位,缺乏对单机和计算机相关设备的详细取证分析功能,仍然难以获得法庭所希望的证据资料.

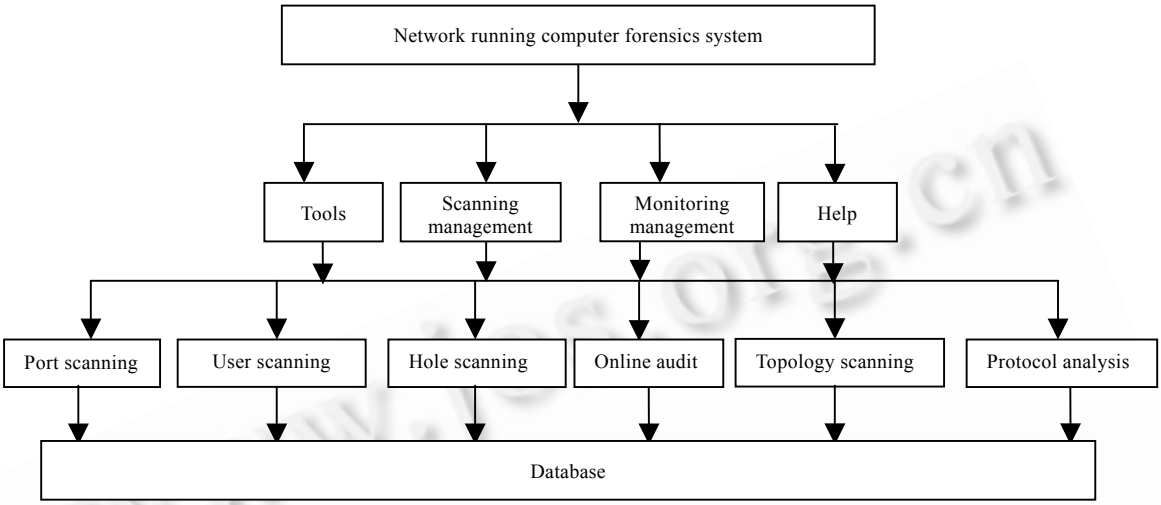

Fig.1 Structure of the network running computer forensics system

图 1 计算机运行环境勘查取证系统结构示意图

3.3 计算机取证软件工具

3.3.1 一般工具软件

一般工具软件包括用于检测分区的 PartitionMagic、杀毒软件、各种压缩工具软件等.

3.3.2 取证专用工具软件

文件浏览器.这类工具是专门用来查看数据文件的阅读工具.只用于查看而没有编辑和恢复功能,从而体积较小并可以防止证据的破坏.比较好的软件是 Quick View Plus(http://www.jasc.com).它可以识别 200 种以上文件类型,可以浏览各种电子邮件文档.比起 WordPerfect 的频繁转换要方便得多.Conversion Plus 可以用于在 Windows 系统下浏览 Macintosh 文件.

图片检查工具.ThumbsPlus 是一个功能很全面的进行图片检查的工具.

反删除工具.这方面的取证分析工具中最主要的是诺顿工具,虽然这是一个老式的工具,但在有些时候是很有用的.

CD-ROM 工具.使用 CD-R Diagnostics 可以看到在一般情况下看不到的数据.

文本搜索工具.dtSearch 是一个很好的用于文本搜索的工具,特别是具有搜索 Outlook 的.pst 文件的能力.

驱动器映像程序.可以满足取证分析(即逐位拷贝以建立整个驱动器的映像)的磁盘映像软件,包括 :SafeBack(http://www.forensics-intl.com),SnapBack(http://www.cdp.com),Ghost(http://symantec.com),dd(UNIX 中的标准工具)等.

磁盘擦除工具.这类工具主要用在使用取证分析之前,为了确保分析机器的驱动器中不包含残余数据,显然,只是简单的格式化肯定不行.从软盘启动后运行 NTI 公司的 DiskScrub 程序即可把硬盘上的每一扇区的数据都清除掉.

取证程序.取证软件的效能倾向于同时拥有收集及分析数据的功能.目前,国际上的主流产品有:

Forensic Toolkit 是一系列基于命令行的工具,可以帮助推断 Windows NT 文件系统中的访问行为.这些程序包括的命令有:AFind(根据最后访问时间给出文件列表,而这并不改变目录的访问时间)、HFind(扫描磁盘中有

隐藏属性的文件)、SFind(扫描整个磁盘寻找隐藏的数据流)、FileStat(报告所有单独文件的属性)、NTLast(提供标准的 GUI 事件浏览器之外对每一个会话都记录了登录及登出时间,并且它能够指出登录是远程的还是本地的).

The Coroner's Toolkit(TCT)主要用来调查被"黑"的 Unix 主机,它提供了强大的调查能力,它的特点是可以对运行着的主机的活动进行分析,并捕获目前的状态信息.其中的 grove-robber 可以收集大量的正在运行的进程、网络连接以及硬盘驱动器方面的信息.数据基本上以挥发性顺序收集,收集所有的数据是个很缓慢的过程,要花上几个小时的时间.TCT 还包括数据恢复和浏览工具 unrm&lazarus、获取 MAC 时间的工具 mactime.另外,还包括一些小工具,如 ils(用来显示被删除的索引节点的原始资料)、icat(用于取得特定的索引节点对应的文件的内容)等等.

EnCase 自称是唯一一个完全集成的基于 Windows 界面的取证应用程序,其功能包括:数据浏览、搜索、磁盘浏览、数据预览、建立案例、建立证据文件、保存案例等.

ForensicX 主要运行于 Linux 环境,是一个以收集数据及分析数据为主要目的的工具.它与配套的硬件组成专门工作平台.它利用了 Linux 支持多种文件系统的特点,提供在不同的文件系统里自动装配映像的能力,能够发现分散空间里的数据,可以分析 Unix 系统是否含有木马程序.其中的 Webtrace 可以自动搜索互联网上的域名,为网络取证进行必要的收集工作,新版本具有识别隐藏文件的工具.

New Technologies Incorporated(NTI, http://www.forensics-intl.com)是取证软件最为固定的商家之一.NTI 以命令的形式执行软件,所以速度很快,软件包的体积小,适合于在软盘上使用.该公司提供的取证工具包括:

CRCMD5:可以验证一个或多个文件内容的 CRC 工具.

DiskScrub:用于清除硬盘驱动器中所有数据的工具.

DiskSig:CRC 程序,用于验证映像备份的精确性.

FileList:磁盘目录工具,用来建立用户在该系统上的行为时间表.

Filter_we:一种用于周围环境数据的智能模糊逻辑过滤器.

GetSlack:一种周围环境数据收集工具,用于捕获未分配的数据.

GetFile:一种周围环境数据收集工具,用于捕获分散的文件.

Net Threat Analyzer:网络取证分析软件,用于识别公司互联网络账号滥用.

M-Sweep:一种周围环境数据清除工具.

NTI-DOC:一种文件程序,用于记录文件的日期、时间以及属性.

PTable:用于分析及证明硬盘驱动器分区的工具.

Seized:一种用于对证据计算机上锁及保护的程序.

ShowFL:用于分析文件输出清单的程序.

TextSearch Plus:用来定位文本或图形文件中的字符串的工具.

4 基于 OpenBSD,Linux 或者 Solaris 系统的技术取证过程实验案例[20]

基于以上讨论,下面我们研究一个使用基于 OpenBSD,Linux 或者 Solaris 的取证工具和有关命令进行取证的实验过程.

4.1 有关命令和软件工具的功能

vmstat 是一个可以快速查看内存、CPU 和磁盘子系统的统计信息的命令.它的执行时间短,以便可以查看子系统利用的趋势.vmstat 经常可以在系统性能有一些问题的时候帮助我们知道哪些地方应该去深究.mpstat 命令在 Linux 和 Solaris 上都有,可以用它查看处理器利用情况的统计数据.mpstat 提供一个选项,允许在多处理器系统中查看指定 CPU 的统计数据,vmstat 则没有这个功能.iostat 显示比 vmstat 更详细的与子系统相关的统计信息,主要用来监控 IO 设备、磁盘的负载、平均传输速度及相关活动记录.sar,sa,lastcomm,last 是用来检查历史数据和一些近来的系统事件的工具.sar 是一个 Solaris 和 Linux 的系统性能分析工具.这些可以用于检查的性能

数据类似于 vmstat,mpstat 和 iostat 的显示.sar 的数据是一段时间保存的内容,因此可以查看过去的信息.lastcomm 可以再现系统最近被执行的命令.这些可以用在系统审计中.sa 可以在*BSD 和 Linux 中找到,它在系统审计中给用户更多的选项来收集信息.ps 立足于进程状态,用于显示系统执行的进程和他们的信息.top 显示的信息与 ps 接近,但是 top 可以了解到 CPU 消耗,可以根据用户指定的时间来更新显示.lsof 列举打开的文件,显示系统当前打开的所有文件.Unix 系统的所有东西几乎都可以看作是文件,因此,lsof 也显示了系统的状态中有重要意义的内容.file 判断文件是什么,不同的文件格式可以 16 进制的形式显示文件的内容.readelf 显示二进制文件的 ELF(可执行链接和格式)头的细节.这些内容可以判断可执行文件提供的函数.od 以用户指定的格式输出文件内容.od 对于在文件内容中有一些解释的情况下查看原始内容是有帮助的.ldd 读取 ELF 头的内容,并显示可执行文件依赖的对象库.string 显示文件中的 ASCII 字符串.对于在二进制文件中查找其中可读的字符串是非常有用处的.find 用于在文件系统中查找指定的对象.strace 这个工具位于一个当前运行的进程的开始或者附加到一个当前运行的进程中,显示这个进程所作的所有系统调用.这可以用来判断程序运行的行为,并且用来决定是否是合适的程序.strace 存在于 Linux.在 Solaris 中是 truss,*BSD 提供的 ktrace 可以达到相似的功能效果.sudo 可以让管理员给用户以其他用户的权限执行命令的能力,而不用给该用户密码.grep 用于按照用户指定的模式查询.grep 使用匹配规则.less 页面调度程序,用来按页显示文本.

4.2 实验案例研究

假如突然发现某个 Web 服务器上 CPU 负载很高,而这个服务器正常情况下每天的浏览量仅仅几千而已.从这种迹像可以初步断定存在问题.以下是分析过程及其系统的显示.

vmstat 和 mpstat(只在*BSD 系统上)表明 CPU 被用户空间的一个或者多个进程消耗掉了,但是内存和 I/O 子系统还没有大量使用,iostat 也显示出磁盘系统不正常,如图 2 所示.

```
$ vmstat  1 4
procs                  memory    swap          io     system         cpu
r  b  w   swpd   free   buff  cache  si  so    bi    bo   in    cs  us  sy  id
1  0  0    376   7756  29772 570960   0   0     7     3  441   397  87   5   8
5  0  0    376   6728  29772 570960   0   0     0     0  498  1249 100   0   0
6  0  0    376   7240  29772 570960   0   0     0     0  652  1563  97   3   0
6  0  0    376   7604  29772 570960   0   0     0     0  536  1323  97   3   0
$ mpstat 1 4
20:51:21     CPU   %user   %nice %system   %idle    intr/s
20:51:22     all  100.00    0.00    0.00    0.00    479.00
20:51:23     all  100.00    0.00    0.00    0.00    496.00
20:51:24     all  100.00    0.00    0.00    0.00    499.00
20:51:25     all   97.00    0.00    3.00    0.00    481.00
Average:     all   98.00    0.60    1.40    0.00    486.60
$ iostat -dk 1 4
Device:            tps   Blk_read/s   Blk_wrtn/s   Blk_read   Blk_wrtn
dev3-0            0.00         0.00         0.01         73       1296
dev3-0            0.00         0.00         0.00          0          0
dev3-0            0.00         0.00         0.00          0          0
```

Fig.2 Experimental result of vmstat, mpstat and iostat

图 2 vmstat,mpstat(只在*BSD 系统上)和 iostat 的运行结果

同时,sar(*BSD 的 sa)显示 CPU 从凌晨 03:17 就开始被使用.lastcomm 显示 FTP 客户端在 03:17 以 root 运行了几次,last 这个命令显示最近的登录,没有显示出这个时间段有 root 从什么地方登录进来.另外,这个服务器

又使用 sudo 来管理 root 权限.这里有两个系统管理员账号可以用 root 登录,但是他们并没有登录过,特别是在凌晨 3:00 这个时间附近.根据这一点,决定使用 CD 上准备好的二进制工具.现在运行 top-d1,并且发现 apache 这个进程占用了 100%的 CPU.

用 grep 检查 apche 的错误日志,这个 GNU 工具可以有更多有用的参数,-A,-B 和-C 可以指定想从哪一个匹配的行的开始、之前或者周围查看.grep 可以通过-E 来使用扩展的模式匹配表达式.可以检查系统日志.但是并没有在这些日志里面检查到什么奇怪的地方.

继续进行检查.运行 ps-eflcyL(Solaris9),ps-eflcym-headers(Deb3.0,RH8.0)或者 psauwxhkwvl(OBSD3.2),找到进程 apache,发现了问题.这里只有一个 apache 进程和多个 httpd 进程.httpd 是在 apachectl 文件执行的,运行的 apache 是之前派生的,因此,如果二进制文件被重命名为“apache”,那么也应该有多个进程存在.这就是问题所在.

使用 lsof-p⟨pid⟩,发现这个 apache 进程打开了一个叫 john.pot 的文件.立刻查询到这是一个叫“John the Ripper”的密码破解工具,正是它让 CPU 满负荷了.现在来检查这个所谓的 apache 二进制程序.通过运行 file apache,表明这个文件是 ELF 可执行的.现在进行进一步的挖掘,执行 readelf-a apache,确信这是一个 ELF 可执行程序.od-xc apache|less 在文件的开始显示 ELF 数字(7f454c46),ldd 显示 apache 连接的共享对象库比实际的 httpd 进程要小,这表明一定存在问题,如图 3 所示.执行 strings|grep-ijohn,显示如图 4 所示.进一步用 strace-fp'pgrepapache'(Deb3.0,RH8.0),truss-fp'pgrepapache'(Sol9),或者 ktrace-dip⟨pid⟩(OBSD3.2),来查看 apache 进程到底在做什么,如图 5 所示.

```
$ ldd ./apache
        libc.so.6 => /lib/libc.so.6 (0x4001f000)
        /lib/ld-linux.so.2 => /lib/ld-linux.so.2 (0x40000000)
$ ldd /usr/sbin/httpd
        libm.so.6 => /lib/libm.so.6 (0x7002c000)
        libcrypt.so.1 => /lib/libcrypt.so.1 (0x700c4000)
        libdb.so.2 => /lib/libdb.so.2 (0x70100000)
        libdb2.so.2 => /lib/libdb2.so.2 (0x70120000)
        libexpat.so.1 => /usr/lib/libexpat.so.1 (0x70180000)
        libdl.so.2 => /lib/libdl.so.2 (0x701b4000)
        libc.so.6 => /lib/libc.so.6 (0x701c8000)
        /lib/ld-linux.so.2 => /lib/ld-linux.so.2 (0x70000)
```

Fig.3 Checking of the binary program of apache

图 3 检查 apache 二进制程序

```
$ strings apache | grep -i john
/usr/share/john/password.lst
/etc/john.ini
~/john.pot
/etc/john.ini
john
John the Ripper  Version 1.6  Copyright (c)
1996-98 by Solar Designer
/etc/john.ini
/etc/john.ini
/etc/john.ini
```

Fig.4 Experimental result of strings|grep-ijohn

图 4 运行 strings|grep-ijohn 的结果

用 grep-n 在 John the Ripper 的源代码中找到在 crk_password_loop()函数中有一个叫 sig_timer_emu_tick()的函数,它在 timer_emu_max 到达之后产生 SIGALRM.这就是 strace 显示的 SIGALRM.通过 lsof –p 'pgrepapache'以及在读取数据的文件描述器(fd)4 发现那个文件的文件描述符 fd4,以及 fd4 相关联的叫 all.chr 的文件.查看 john.ini 文件,找到在负担增大的区域所引用的 all.chr 文件.

所有证据表明,这个 John the Ripper 的确是一个密码破解工具.

至此,技术取证过程的前几个步骤就完成了,接下来就要进行证据的分析、犯罪分子的追踪,最后形成证据结果.

```
# strace -fp `pgrep apache`
--- SIGALRM (Alarm clock) ---
sigreturn()                                 = ? (mask now [])
--- SIGALRM (Alarm clock) ---
sigreturn()                                 = ? (mask now [])
--- SIGALRM (Alarm clock) ---
sigreturn()                                 = ? (mask now [])
_llseek(4, 65536, [65536], SEEK_SET)        = 0
read(4, "\2YUz\0\2Z\0241\0\2ZA1dkmnr\0\2ZDa\0\2ZElnrd"..., 4096) = 4096
read(4, "\2\5Ma\0\2\5Te\0\2\6\n(\0\2\6>%\0\2\10\n&\0\2\10\0210\0"..., 4096) = 4096
read(4, "IPVm\0\2(%L\0\2((H\0\2(/CMRBNTUWcr\0\2(5"..., 4096) = 4096
read(4, "fgq\0\2BCh26o\0\2BDaikmoy\0\2BEanreltc"..., 4096) = 4096
read(4, "ag\0\2N\23l\0\2N\25j\0\2N\0269\0\2N\30f\0\2N!ds\0\2N)"..., 4096) = 4096
read(4, "\0\2_9LRMTes\0\2_:6Z\0\2_>&%\0\2_?)Cw\0\2_"..., 4096) = 4096
read(4, "%Antv\0\2%Cr\0\2%Dd\0\2%Gg\0\2%Ke\0\2%Lls\0"..., 4096) = 4096
read(4, "rtlpbdgiuv\0\2GBeam\0\2GDuy\0\2GEnert\0"..., 4096) = 4096
read(4, "We6hilw\0\2RYasdpcgilmno\0\2S\0205c\0\2S\21"..., 4096) = 4096
read(4, "68jDxMEBTIRNLAGSKCPOqVHFUZWJ$%#^"..., 4096) = 4096
read(4, "\0022In\0\0022Ke\0\0022LiEe\0\0022Ma\0\0022NEe\0\0022P"..., 4096) = 4096
read(4, "I5y\0\2I6e\0\2I7ae\0\2IAnmrs9t67dhlpuz"..., 4096) = 4096
read(4, "Eprylsenbu9ak013ft!Lcox\0\2SFi9aey"..., 4096) = 4096
read(4, "o149dlytwMbckr0u\0\2_Xixae01537926"..., 4096) = 4096
read(4, "2\0\2-!NCLS\0\2-%RS\0\2-)ETDNS\0\2--M\0\2-"..., 4096) = 4096
read(4, "!*bfgk\0\2GJu\0\2GLeiayYs\0\2GMaAoegu\0"..., 4096) = 4096
read(4, "Utsdbgmpy\0\2PWa\0\2PXy\0\2PY1!27sy.05"..., 4096) = 4096
read(4, "w!9bAjp5clgkr068EFSYf$.?CJKLW]n\0"..., 4096) = 4096
--- SIGALRM (Alarm clock) ---
sigreturn()                                 = ? (mask now [])
```

Fig.5 Further check of apache process

图 5 进一步查看 apache 进程

5 存在的问题及其发展趋势

5.1 法律法规的健全和取证工作的规范化

目前,我国还没有设立计算机取证方面专门的法律法规,计算机证据也没有作为一种单独的证据类型加以确认.所以,立法势在必行.相关的法律法规有《刑法》、《刑事诉讼法》、有关的网络法律法规以及和被调查案件相关的其他法律法规.计算机取证工作的程序没有标准和规范,取证工作随意性太大,获取证据的证明力差.法庭上律师和审判人员对电子证据的质疑很少,律师和其他司法工作人员缺乏计算机取证的起码常识.获取证据的过程没有严格的规定,使计算机取证工作不具备权威性和科学性.没有制定取证人员的认证和培训机制.任何具备一定技能的人员,都可以取证,由于取证人员水平偏低,或缺乏经验或法律知识匮乏,取得的证据不具备可靠性.所以,取证工作的标准和规范化以及取证人员的资格认定是当务之急.

计算机取证工具已有很多,但缺乏评价机制和标准.什么样的证据应该适用什么样的取证工具,进行怎样的

操作过程才能使获取的证据具有证明力,这些都是需要通过制定标准来确定的.由于取证软件的特殊性,不能像一般软件那样,任何软件公司都可以制作,然后拿到市场上去销售.应当制定行业规范,只有符合资质要求的软件企业才可以从事取证软件的开发研制.所以行业规范也是需要解决的问题.

5.2 计算机取证技术的完善和发展

现在的计算机取证,很大程度上是手工操作硬件或者使用工具软件,能够在作案的同时或一定的时限内取得证据的可能性几乎没有.取证工作的成败主要取决于技术人员的经验和智力[1],证据的自动获取技术还没有.所以,取证技术的发展方向之一就是取证技术的自动化.

由于对计算机证据数据进行各种分析操作的过程越复杂、越频繁,就越容易损坏证据,从而降低其证明力,所以计算机取证技术的另一个发展趋势就是设计对证据数据的操作尽可能少的取证工具.

取证技术还不能完全击败反取证技术[1].反取证技术也在不断发展.例如,数据隐藏技术、数据擦除技术等.反取证工具 Runefs 就利用了取证工具 TCT 不能检查磁盘坏块的缺陷,把敏感文件的数据块标记为坏块来逃避检查[1].所以,如同病毒和反病毒软件的发展一样,取证技术必须在研究反取证技术的基础上进一步发展.

在未来的几年里,计算机取证技术将充分应用人工智能、数据挖掘、实时系统、反向工程技术、软件水印技术[21],使用更加安全的操作系统.计算机取证学将会作为一门新兴学科飞速发展,形成一套系统的理论,并会研制出大量的专门用于取证的自动化程度较高的工具,培养一批“电子法医”(取证人员),为打击犯罪获取强有力的证据.

6 结 论

计算机取证学的研究是从 20 世纪 80 年代开始的,并随着计算机和网络技术的普及而变得越来越重要.本文研究了计算机证据的法律认定、计算机取证的步骤和原则、计算机取证应用的技术,概括了目前的计算机取证软件工具,分析研究了一个计算机取证的技术实验过程.提出计算机取证学应该从法律和技术两个方面进行研究并指出了这两个方面的不足和今后的发展趋势.对计算机取证的法律和技术问题的深入探讨和研究,建立适合我国国情的计算机取证专业实验室和技术平台,将有助于澄清这一领域中的模糊认识,促进计算机取证法律法规的健全和计算机取证技术的进一步完善.

References:

[1] Wang L, Qian HL. Computer forensics and its future trend. Journal of Software, 2003,14(9):1635−1644 (in Chinese with English with abstract). http://www.jos.org.cn/1000-9825/14/1635.htm

[2] Zhang P. Internet Law Review. Vol.3, Beijing: Law Press China, 2003 (in Chinese).

[3] He JH. The New Issue of Evidence Law. Beijing: Law Press China, 2000 (in Chinese).

[4] Teaching Material Writing and Editing Council in Ministry of Public Security. Security Supervise on Network Information. Beijing: The Mass Press, 2000 (in Chinese).

[5] Yang CG. Electronic evidences in e-business and its law status. 2001. http://cooltoy.yesky.com/20010118/155824_1.shtml

[6] KruseII WG, Heiser JG. Computer Forensics: Incident Response Essentials. Pearson Education, Inc. 2002.

[7] Anderson MR. Computer evidence processing: The important first step-safe seizure of the computer. http://www.forensics-intl.com

[8] Sommer P. Computer forensics: An introduction. In: Proc. of the Computer Forensics'92—the 9th World Conf. on Computer Security Audit and Control. 1992. 82−96. http://www.virtualcity.co.uk/vcaforens.htm

[9] Anderson MR. Electronic fingerprints——Computer evidence comes of age. http://www.forensics-intl.com

[10] He M. New focus on the subject of computer security——The subject of computer forensics. System Security, 2002,7:42−43.

[11] Reith M, Carr C. An examination of digital forensics models. Int'l Journal of Digital Evidence, 2002,1(3):1−12.

[12] Bryson C, Anderson MR. Shadow Data——The 5th Dimension of Data Security Risk. http://www.forensics-intl.com

[13] Liang JH, Jiang JC, Dai FY, Qiang SH. Research on technology of computer forensic. Computer Engineering, 2002,28(8):12−14.

[14] Anderson MR. Internet security——Firewalls & encryption: the cyber cop's perspective. http://www.forensics-intl.com/art1.html

[15] Qian GQ, Yang ZM, Xu RS. Study and design of computer forensics. Computer Engineering, 2002,28(6):56−58.

[16] Armouring the ELF: Binary encryption on the UNIX platform. Phrack #58 article5. 2001. http://www.phrack.org

[17] Zhang YJ. Network Security and Detection Technologies on Computer Crime. Beijing: Tsinghua University Press, 2003 (in Chinese).

[18] Zhang J, Gong J. A summary of network intrusion traceback. Computer Science, 2003,30(10):155−166 (in Chinese with English abstract).

[19] Corey V, Peterman C, Shearins S, Greenberg MS, Van Bokkelen J. Network forensics analysis. IEEE Internet Computing, 2002 6(6):60−66.

[20] Sorenson H. Incident response tools for Unix, part one: System tools. 2003. http://www.securityfocus.com/infocus/1679

[21] Zhang LH, Yang YX, Niu XY,Niu SZ. A survey on software watermarking. Journal of Software, 2003,14(2):268−277 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/14/268.htm

附中文参考文献:

[1] 王玲,钱华林.计算机取证技术及其发展趋势.软件学报,2003,14(9):1635−1644. http://www.jos.org.cn/1000-9825/14/1635.htm

[2] 张平.网络法律评论.北京:法律出版社,2003.

[3] 何家弘.新编证据法学.北京:法律出版社,2000.

[4] 公安部教材编审委员会.信息网络安全监察.北京:群众出版社,2000.

[5] 杨晨光.电子商务中的电子证据及其法律地位. http://cooltoy.yesky.com/20010118/155824_1.shtml

[10] 何明.计算机安全学的新焦点——计算机取证学.系统安全,2002,7:42−43.

[13] 梁锦华,蒋建春,戴飞雁,卿斯汉.计算机取证技术研究.计算机工程,2002,28(8):12−14.

[15] 钱桂琼,杨泽明,许榕生.计算机取证的研究与设计.计算机工程,2002,28(6):56−58.

[17] 张越今.网络安全与计算机犯罪勘查技术学.北京:清华大学出版社,2003.

[18] 张静,龚剑.网络入侵追踪研究综述.计算机科学,2003,30(10):155−166.

[21] 张立和,杨义先,钮心忻,牛少彰.软件水印综述.软件学报,2003,14(2):268−277. http://www.jos.org.cn/1000-9825/14/268.htm