网络取证技术研究

段 玲[1,2],王 锋[1,2]
(1.昆明理工大学 云南省计算机技术应用重点实验室,云南 昆明 650051;
2.昆明理工大学 信息工程与自动化学院,云南 昆明 650051)

摘 要:介绍了网络取证的概念、网络证据的特点、与传统静态取证的比较及研究现状,详细分析了数据捕获、数据分析技术、专家系统和数据挖掘技术等在网络取证中的应用,并分析了目前网络取证存在的问题和发展趋势。

关键词:实时;网络取证;计算机取证
中图分类号:TP393.1 **文献标识码**:A

Research on real-time network forensics techniques

DUAN Ling[1,2], WANG Feng[1,2]
(1.Yunnan Key Laboratory of Computer Technology Applications, Kunming University of Science and Technology, Kunming 650051, China;
2.College of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650051, China)

Abstract: This paper introduces the concept of network forensics, the characteristics of network evidence, comparison with the traditional static computer forensic and current research situation. Data capture, data analysis techniques, expert system and data mining techniques were presented in detail in the paper. Current network forensics problems and trends were analyzed.

Key words: real-time; network forensics; computer forensics

随着计算机网络的迅速发展和普及,网络安全变得日益重要和复杂。网络入侵、计算机犯罪问题也越来越严重,攻击和入侵事件甚至直接威胁到国家的安全,信息安全问题日益严峻,在犯罪事件发生后对犯罪行为进行事后的取证,存在着证据的真实性、有效性和及时性等问题。如何及时准确地从网络中获取计算机证据,打击和遏制网络犯罪正日益成为计算机网络安全的研究热点。计算机网络取证是计算机科学与法学交叉的一门新兴学科,是伴随着计算机技术的发展和普及以及计算机犯罪这一新的犯罪形式出现而出现的。

1 网络取证概述

1.1 概念

网络取证(network forensics)概念最早是在20世纪90年代美国防火墙专家Marcus Ranum[1]提出的,借用了法律和犯罪学领域中用来表示犯罪调查的词汇“forensics”,网络取证是指捕获、记录和分析网络事件以发现安全攻击或其他的问题事件的来源。在2001年数字取证研究工作组DFRWS[2](Digital Forensic Research Workshop)的会议上,明确将网络取证作为会议的4个主题之一进行讨论,并给出了网络取证的定义:“为了揭示与阴谋相关的事实,或者为了成功地检测出那些意在破坏、误用或危及系统构成的未授权行为,使用科学的技术,对来自各种活动事件和传输实体的数字证据进行收集、融合、识别、检查、关联、分析和归档等活动过程”。

按照Simson Garfinkel[3]的观点,网络取证包括2种方式:(1)“catch it as you can”(尽可能地捕捉)。这种方式中所有的包都经过一定的节点来捕获,将报文全部保存下来,形成一个完整的网络流量记录,把分析的结果按照批量方式写入存储器。这种方式能保证系统不丢失任何潜在的信息,最大限度地恢复黑客攻击时的现场,但对系统存储容量的要求非常高,通常会用到独立冗余磁盘阵列(RAID)系统。(2)“stop look and listen”(停、看、听)。这种方式中每个包都经过基本的分析,为以后的分析留下基本信息,对存储的要求比较小,但需要一个较

快的处理器,以便速度能跟得上输入的网络流。这种方式能减少系统存储容量的需求,但有可能丢失一些潜在的信息,同时过滤进程还会增加系统负荷。

中科院学者指出:“网络取证技术指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、数据过滤技术及漏洞扫描技术等”[4]。

目前还没有网络取证的统一标准定义,综上所述,笔者认为,网络取证是指通过计算机网络技术,按照符合法律规范的方式,对网络事件进行可靠的分析和记录,并以此作为法律证据的过程。

1.2 网络证据的特点

网络证据是正在网上传输的计算机证据,其实质是网络流[5]。网络证据的获取属于事中取证,即在犯罪事件进行或证据数据的传输途中进行截获。网络流的存在形式依赖于网络传输协议,采用不同的传输协议,网络流的格式也不相同。网络取证的目标就是对网络流进行正确地提取和分析,真实全面地将发生在网络上的所有事件记录下来,为事后的追查提供完整准确的资料,将证据提交给法庭。网络证据具有以下特点:

(1)动态:不同于存储在硬盘等存储设备中的数据,网络数据是正在网上传输的数据,是“流动”的数据。

(2)实时:就网络上传输的一个数据包而言,其传输的过程是有时间限制的,从源地址经由传输介质到达目的地址后就不再属于网络流了。

(3)海量:随着网络带宽的不断增加,网上传输的数据越来越多,形成了海量数据。

(4)多态:网络上传输的数据流有的是文本,有的是视频,有的是音频,其表现形式呈多态性。

1.3 与传统静态取证的比较

计算机取证就是对计算机犯罪的证据进行获取保存分析和出示,它实质上是一个详细扫描计算机系统以及重建入侵事件的过程[6]。传统的计算机取证主要是静态取证,即在计算机犯罪发生后,有时甚至时间间隔很长。主要集中在计算机磁盘的分析,通过克隆存储介质生成镜像文件、恢复删除的文件和关键字查找有关证据。由于事后的静态取证处于被动状态,取证工作受计算机犯罪分子留下现场的制约,电子证据的证明力相对较弱,对于证据的提取很不充分,而且不能保证所获得的证据一定有效。另一方面,静态取证过多地依赖人为经验,从海量的数据中根据经验来提取分析。静态在事件发生后借助手工分析网络数据包和主机日志,并利用这些数据重建攻击事件,所涉及的数据量大,数据包和日志的格式复杂,收集和分析这些网络信息非常困难。因此,这种手工分析方式的效率通常很低。另外,这种获取和分析证据的方法通常不严格,从而导致所获得的证据的可信度不高。

网络取证专门针对网络证据的获取与分析,是计算机静态取证的补充。网络取证对所有可能的计算机网络犯罪行为进行实时数据获取和分析,在确保系统安全的情况下获取最大量的证据,并将证据保全、分析和提交的过程。网络取证的数据源是全面、真实地反映客观事实的、海量的且不断更新的。网络取证主要通过对网络流、审计迹、主机系统日志等的实时监控和分析,发现对网络系统的入侵行为,自动记录犯罪证据,并阻止对网络系统的进一步入侵。

1.4 研究现状

在学术界,近几年每年都会有讨论计算机取证为主题的学术会议召开:应急响应和安全小组论坛 FIRST[7](Forum of Incident Response and Security Teams)年会、数字取证研究工作组 DFRWS[2](Digital Forensic Research Workshop)会议、e-forensics 电子取证会议[8]。参考文献[9]讨论了自动网络取证分析技术,参考文献[10]提出了基于证据图形推理方法的网络取证技术,参考文献[11]提出了几种高速有效的网络取证分析方法。

国外打击计算机犯罪有着 20~30 年的历史,有许多专门的计算机取证部门、实验室和咨询服务公司,开发了许多非常实用的取证产品。虽然我国在计算机取证方面的研究起步较晚,但是经过国家有关部门的不懈努力,我国在计算机取证方面也取得了一定成果,开发出了不少计算机取证工具和系统。中科院高能物理研究所提出了网络取证与分析系统模型[12]、参考文献[13]提出了一种基于模糊决策树的网络取证分析方法、浙江大学和复旦大学在取证技术、吉林大学在网络逆向追踪、电子科技大学在网络诱骗、北京航空航天大学在入侵诱骗模型等方面展开了研究工作。

2 网络取证技术

网络取证作为全新的信息安全技术,不同于已有的网络安全系统的实现技术,网络取证过程充满了复杂性和多样性,使得相关技术既复杂又多样。主要包括:

(1)捕获网络数据:高效截包技术

Libpcap 是网络截包最通用的函数库,适合于多种操作系统平台。目前有许多著名的截包分析程序都建立在 Libpcap 的基础上,如 tcpdump、wireshark、snort 等。这种基于网络数据包的取证系统,如果丢包率太高,后面的会话重建、协议分析将无法进行,取证将失去意义。所以,能否高效截取所有数据包是整个网络取证的基础。

(2)分析获取的数据:会话重建技术

会话重建是网络取证中的重要环节。分析数据包的特征,并基于会话对数据包进行重组,去除协商、应答、重传、包头等网络信息,以获取一条基于完整会话的记录。基于 Libpcap 的截包应用程序截获原始的网络数据包,把捕获到的数据包分离,逐层分析协议和内容,在传输层将其组装起来,在重新组合的过程中可以发现很多有用的证据。例如,数据传输错误,数据丢失,网络的联

结方式等等。

(3)专家系统

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。有专家提出了 NFAT(Network Forensics Analysis Tools)[14]的设计框架和标准,核心是开发专家系统 ES(Expert System)并配合入侵检测系统或防火墙,对网络流进行实时提取和分析,对发现的异常情况进行可视化报告。

典型的专家系统包括知识库和推理机,将有关的证据知识转化为 if-then 结构的规则,当检测到当前的数据包符合知识库中的 1 个或几个条件时,就将该数据包存入原始证据库以备进一步处理和分析。采用专家系统的优势在于灵活性和可扩充性,在基于规则的系统中很容易使系统的性能和正确性得到持续地检查。专家系统的难点在于知识库的建立,很难全面地从各种犯罪手段中抽象出能够规则化的知识。

(4)数据挖掘技术

实时网络取证不同于静态取证的关键在于它事前就进行实时数据获取,这要求实时网络取证要从海量的数据中及时分析出具有网络犯罪特征的数据,并对具有新特征的数据进行分析判断。这一阶段需要将数据挖掘技术引入数据的分析中。数据挖掘技术主要包括关联规则分析、分类和联系分析等。运用关联规则分析方法可以提取犯罪行为之间的关联特征,挖掘不同犯罪形式的特征、同一事件的不同证据之间的联系;运用分类方法可以从获取的海量数据中找出可能的非法行为,发现各种事件在时间上的先后关系。

3 存在的问题和发展趋势

3.1 存在的问题

网络取证目前在国内刚刚兴起,人们的网络安全意识和法律观念较淡薄,取证技术相对滞后,使得网络取证还面临着一些困难和挑战,主要表现在:

(1)海量的计算机数据给证据收集和分析带来困难

计算机系统和计算机网络中每天产生的数据复杂而庞大,如何及时获取和保存这些海量数据给人们提出了巨大的挑战。如何在海量的数据中审查判断出与案件关联的、反映案件客观事实的计算机证据更是一项艰巨任务。由于计算机数据自身容易修改且不留任何痕迹的特性,使得信息的完整性保护变得非常困难。

(2)没有标准的网络取证流程

由于计算机取证是一门新兴的学科,发展还不完善,还没有统一的、科学的计算机取证过程和步骤,没有统一的、科学的计算机取证工具的评判标准和评判机构;商用和个人的计算机取证工具侧重于事后取证。

(3)相关的法律法规还不完善

现在美国至少有 70%的法律部门拥有自己的计算机取证实验室[6]。而目前我国还没有设立计算机取证方面专门的法律法规,还没有专门的取证机构,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。还需要继续加强计算机和网络安全相关的立法工作。

(4)用户网络安全意识和法律观念淡薄

目前大部分政府部门、企事业单位、学校、中小型公司等联网单位没有经过专业培训的计算机安全技术人员,用户网络安全观念非常淡薄。另外,一些单位为了自身的声誉,在遭受网络攻击或者其他网络违法犯罪侵犯时,没有及时向公安机关报案,使得违法犯罪分子逃避打击,导致更多的网络犯罪发生。

(5)取证专业技术人才严重匮乏

网络取证是一门新兴的交叉学科,涉及法律和计算机专业知识,这就造成网络取证专业人严重匮乏。因此,必须培养这方面的人才,建立相关的资格认证机制,对该类人员进行审核和认证。目前还没有机构对计算机取证人员的资质进行认证,使得取证结果的权威性受到质疑。

(6)反取证技术的发展

目前有不少的黑客从事反取证技术的研究,通过删除、篡改或隐藏证据使得取证失效。例如用数据隐藏技术、数据擦除技术和数据加密等技术削弱取证工作的效果。

3.2 发展趋势

未来网络取证技术的发展主要有以下几个方向:

(1)网络取证方法与工具的研究

数据获取技术、数据恢复技术、数据分析技术、数据保存技术、基于数据挖掘的海量数据取证技术。

(2)网络取证规范和标准的研究与制定

研究和制定科学的网络取证步骤和流程,为了保证数字证据在收集、保存、检查和转移等过程中的准确性和可靠性,法律实施组织和取证组织必须建立并维护一个高质量的系统,建立规范的文档、使用广为认可的设备和材料、取证人员的资质认定等。

(3)网络取证相关法律法规的健全

研究网络取证的合法性标准,制定和健全与信息安全、计算机基础设施保护等相关的法律法规,为计算机取证和电子证据的应用打下法律基础。

(4)网络取证机构的设立和认证

取证机构的配置、取证机构的管理、取证机构的资质认证、取证人员的培训和认证等。

(5)无线网络取证技术

随着手机上网、无线上网、3G 网络越来越普及,利用手机和无线局域网、CDMA 无线上网犯罪的案件逐年上升,目前涉及无线网络取证的技术正在开发和完善中,已取得了阶段性的成果。

网络取证是一个迅速发展的研究领域,是一门有待标准化和探讨、不断发展的学科,它在网络信息安全和犯罪调查方面有着重要的应用前景。我国在网络取证研

究上刚刚起步，这更加显示了网络取证研究的紧迫性，也为今后的研究提供了机遇和空间。

参考文献

[1] RANUM M J. Network forensics and traffic monitoring[J]. Computer Security Journal,1997,13(2):35-39.

[2] PALMER G. A road map for digital forensic research[R]. Report from the First Digital Forensic Research Workshop (DFRWS), Utica, New York, August 7-8. 2001: 27-30.

[3] GARFINKEL S. Network forensics:tapping the internet [EB/OL].http://www.oreillynet.com/pub/a/network/2002/04/26/nettap.html. 2002.

[4] 丁丽萍,王永吉.计算机取证的相关法律技术问题研究[J].软件学报,2005,16(2):260-275.

[5] 丁丽萍.基于网络数据流的计算机取证技术[J].信息网络安全,2005(6):74-76.

[6] 王玲,钱华林.计算机取证技术及其发展趋势[J].软件学报,2003,14(9):1635-1644.

[7] The forum of incident response and security team[Z]. http://www.first.org. 2009.

[8] e-forensics conference[Z].http://www.e-forensics.eu/.2009.

[9] MERKLE L D. Automated network forensics [C]. Atlanta, GA, USA: Proceedings of the 2008 GECCO conference companion on Genetic and evolutionary computation. ACM, 2008.

[10] WANG W, DANIELS T E. A graph based approach toward network forensics analysis [J]. ACM Trans Inf Syst Secur 2008,12(1):1-33.

[11] PONEC M, HERV G P. Highly efficient techniques for network forensics[C]. Alexandria, Virginia, USA: Proceedings of the 14th ACM conference on Computer and communications security. ACM, 2007.

[12] 杨泽明,许榕生,曹爱娟.网络取证与分析系统的设计与实现[J].计算机工程,2004,30(13):72-74.

[13] 刘在强,林东岱,冯登国.一种用于网络取证分析的模糊决策树推理方法[J].软件学报,2007,18(10):2635-2644.

[14] COREY V, PETERMAN C, SHEARINS S, et al. Network forensics analysis[J]. IEEE Internet Computing, 2002, 6(6):60-66.

(收稿日期:2009-09-10)