

云环境隐私侵犯取证研究

李维奉, 羌卫中, 李伟明, 邹德清

(华中科技大学计算机科学与技术学院, 湖北 武汉 430074)

摘要: 云计算是主流的信息系统技术, 但是云计算环境下的用户隐私保护和隐私侵犯追踪取证一直是一个挑战。当前主流的云计算取证技术体系和标准并不成熟和完备, 距离计算机取证的目标相去甚远。针对云环境下隐私侵犯取证, 首先, 给出了云取证和云隐私侵犯的研究现状和挑战; 其次, 着重分析了云环境下隐私取证的研究内容和技术路线; 最后, 提出了一个具体的面向云环境的隐私侵犯取证系统。

关键词: 云计算; 云取证; 云证据分析; 隐私侵犯

中图分类号: TP393

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2018004

Research on forensics of privacy violations in cloud environment

LI Weifeng, QIANG Weizhong, LI Weiming, ZOU Deqing

School of Computer Science & Technology, Huazhong University of Science and Technology, Wuhan 430074, China

Abstract: Cloud computing is the mainstream information system technology, however, privacy protection of user and privacy violation tracking and forensics in the cloud computing environment have always been a challenge. The current mainstream cloud computing forensics system and standard are not mature and complete, and are far from the target of computer forensics. With the focus on privacy violation and forensics in the cloud environment, status and challenges of cloud forensics and cloud privacy violation were reported firstly. Then, the research content and technical route of privacy forensics in the cloud environments were analyzed. Finally, a specific privacy violation forensics system for cloud environments was presented.

Key words: cloud computing, cloud forensics, cloud evidence analysis, privacy violation

1 引言

过去三十年来, 数字取证在计算机犯罪调查中得到广泛应用。云计算是一种新的计算模式, 可实现普遍存在的、按需的、弹性的网络访问, 提供可共享、可配置的资源池(如网络、服务器、存储、应用程序和服务)。但是, 信息系统和数据在云计算平台上的广泛部署和存储给数字取证带

来了特殊的挑战。

云取证(cloud forensics)是一种尝试调查和分析云安全威胁的方法。它作为一种威慑方法, 可以确保攻击者更加谨慎, 以避免因非法行为被取证, 从而减少网络犯罪, 提高云计算系统的安全性。云取证中通过3个来源进行证据提取: 客户端系统、网络层和云服务提供商的管理服务器。美国国家标准技术研究(NIST, National Institute

收稿日期: 2017-12-04; 修回日期: 2018-01-02

通信作者: 羌卫中, wzqiang@hust.edu.cn

基金项目: 国家重点研发计划项目(No.2017YFB0802205)

Foundation Item: The National Key Research & Development (R&D) Plan of China (No.2017YFB0802205)

of Standards and Technology)^[1]指出,“云计算取证科学是应用科学原理、技术实践及其衍生而形成的方法,它通过识别、收集、保存、检验和解释数字证据来重建过去的云计算事件”。云取证是云计算和计算机取证的交叉学科,云环境下存在着多用户证据混杂、资源可动态配置、虚拟化隔离环境、数据易丢失等特性,使传统的计算机取证框架、取证工具不再适用。云环境下取证工具应具备更细致的分工、动态性高、隔离环境下可用、具有时间戳的数据物理定位等新的特点^[2]。云安全联盟(CSA, Cloud Security Alliance)^[3]于2013年提供了云取证的潜在证据来源清单,其中列出了Web服务器日志、应用服务器日志、数据库日志、客户端操作系统日志、主机访问日志、虚拟化平台日志、网络捕获、计费记录、管理门户日志、DNS服务器日志、虚拟机监视器日志、主机操作系统日志和API日志等证据。

本文先调研云计算的取证问题和云环境中隐私侵犯问题,研究云取证和云环境中隐私侵犯问题中存在的挑战和机遇,提出云环境下隐私侵犯取证的关键研究内容和技术路线,致力于解决隐私侵犯行为的精准取证与追踪,并提出面向云计算环境的追踪取证系统,支撑云环境下隐私侵犯行为及时发现、联动追踪、精准取证。

2 研究现状和挑战

2.1 云环境取证

在云取证定义方面,文献[4]提出:云取证是数字取证科学在云环境中的应用。从技术层面上讲,它包括具有远程、虚拟化、实时、大规模等特点的混合环境下产生数字证据的取证方法;从组织结构层面上讲,它涉及云行为者(即云提供商、云消费者、云代理商、云运营商和云审计师等)之间的交互;从法律层面上讲,它往往意味着多域和多租户的场景,涉及多个管辖区的证据获取。文献[5]在云取证现有挑战和问题基础上,讨论了存储在云上数据的可访问性、多个司法管辖区证据的可靠性和所有权、整个调查期间监管链的保持等问题。

在云基础设施取证方面,文献[6]提出了多级别的云环境取证分析模型,将取证层级分为:云

管理级、云平台级、云应用级和云客户级。在鉴定、获取、分析和呈现的传统取证基础上,提出了云环境下取证的4个阶段:确定潜在的证据来源、保证证据的完整性、分析现场事件的关系,以及合理使用取证的工具。常用取证工具包括EnCase^[7]、FTK (forensic toolkit)^[8]等,由于云环境存在远程Web访问、虚拟化平台、第三方物理位置获取、可疑操作确定等新的取证场景,EnCase和FTK等传统计算机取证工具无法用于执行云环境下所有的取证功能。

对于云环境的各种云服务类型,文献[9]检验现有解决方案的安全性,对基础设施即服务(IaaS, infrastructure as a service)、平台即服务(PaaS, platform as a service)和软件即服务(SaaS, software as a service)的建模方法进行分析,但未详细讨论扩大取证场景后的建模方法。文献[4]对云取证分析的关键标准做了详细的讨论。根据257个数字取证专家和从业人员对云取证基础问题的回应,得出需要建立广泛的取证即服务(FaaS, forensics as a service)体系结构的结论。文献[10]提出一种三层架构的FaaS,并在OpenStack上实现了该模型,该模型支持自动取证分析,并可提供以下服务:1) 取证分析服务;2) 定制化的取证分析,即允许客户对任何已购买的服务(IaaS、PaaS或SaaS)进行定制的取证分析;3) SaaS模型中的高级取证工具。针对IaaS云方面,文献[11]提出一种云取证框架,该框架包括在虚拟服务器中的证据抓取器、Xen Hypervisor中的实时取证模块和一个取证虚拟机。实时取证模块会收集所有虚拟机的证据信息,证据抓取器会自动抓取所需的信息并发送给取证虚拟机,取证虚拟机对这些信息进行分析并保存。该框架可在Windows及Linux平台的虚拟机中正常工作,能够有效且快速地获取云环境中的证据数据,并保障证据数据的完整性及机密性。文献[12]提出了一种云环境下现场迁移的取证技术,其将待取证虚拟机实例迁移至本地,并在迁移过程中对虚拟机实例的内存映射、网络连接等易失性数据进行保全。迁移虚拟机内存在本地加载成功后,再利用传统的取证工具在本地虚拟机实例中进行取证。

针对云环境中取证工具不易部署,并且取证工具之间缺乏交互等问题,文献[13]提出了一种基于云环境的高性能证据数据搜索服务,将证据数据进行索引并存储在 Hbase 数据库中,以实现大规模证据数据的高效查询;文献[14]提出一个领域特定的云环境来处理大量的取证数据,通过规范取证工具之间的交互接口,并提供创建和定制取证数据处理工作流程的方法,尝试解决取证工具之间的交互问题。

在云取证的其他方面,文献[15]设计了一个云环境下的日志记录模型,通过第三方在云服务端和客户端构建日志模块,每次云服务器端动作发生后,同步云服务端和客户端的日志记录。该模型在客户端可以始终维护一份更新的日志,不需要云服务提供商协助调查取证,方便取证者获得日志。文献[16]基于 VMWare 的 Vsphere 架构,提出了一种 Vsphere 私有云电子取证技术框架。该框架通过结合基于虚拟机硬盘的常规取证和基于虚拟机重启的动态取证,获得虚拟机的电子证据,并结合现场笔录生成证据文书。该框架可应用到案件侦查实战中,获取的电子证据在现行刑事诉讼法的体系下具有法律效力。文献[17]设计并实现了一种高效的基于 Hadoop 的分布式取证系统。它通过调度控制服务,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,很大程度上提高了云环境取证的效率。

2.2 云环境隐私侵犯

在云计算系统中,用户的隐私数据和业务的私密信息会被云服务提供商存储和管理,信息的所有权和使用权的分离,会导致信息隐私的泄露。针对传统隐私问题,已有许多法案和相关技术方案,但对于云环境这种存在于多用户与多服务提供者的新情景,这些法案和技术方案已变得不再适用。

在隐私保护的相关法案方面,文献[18]列出了可能无法保护用户隐私信息的几种法案。1986年提出的电子通信隐私法(ECPA, Electronic Communications Privacy Act)^[19]是为了防止政府访问存储在服务提供商的存储设备中的电子信

息,包括电子邮件和其他计算机信息等。然而云计算中应用众多、交互频繁、行为难预测,该法案无法准确识别所有行为,所以难以有效地防止云环境中的隐私侵权行为。公平信用报告法(FCRA, Fair Credit Reporting Act)^[20]规定用户的信用报告被允许的用途,如果债权人向云提供商存储信用报告,而第三方(其他机构、组织、用户等)从云提供商处获取报告,则可能会违反法定使用限制。视频隐私保护法案(VPPA, Video Privacy Protection Act)^[21]限制了客户数据的披露,如果云提供商的服务条款允许提供商观看、使用或披露信息,则云提供商的行为可能会导致违法。与隐私有关的法案不仅颁布的时间较早,而且初衷只是保护双方之间的隐私。所以在云环境中,第三方(其他机构、组织、用户等)也许能够更轻易地从云提供商处获取信息。

在隐私保护相关的技术方面,文献[22]讨论了云环境下的数据隐私的风险和需要遵从的法律。文献[23]强调需要开发健全的数字身份基础设施,为云环境下隐私和安全问题的解决提供有力的支撑。文献[24,25]分析了云环境下现有的隐私问题,提出关于设计隐私保护云服务的全面指导方针,并强调在云服务设计阶段把隐私保护考虑进去的重要性。文献[24]还总结了几种推荐的隐私保护模式:1) 发送和存储在云中的客户个人信息最小化;2) 保护云中敏感的客户信息;3) 用户控制最大化;4) 允许用户选择;5) 指定和限制数据使用的目的;6) 向客户提供隐私反馈。

安全外包计算可以较好地处理隐私数据,但依然存在较大的效率问题,其难点主要在于对隐私信息的加密。文献[26]认为基于完全同态和可验证加密的纯密码隐私解决方案延迟较高,建议将受信任的硬件令牌与安全功能评估相结合,以便在仍处于加密形式的计算数据上执行任意功能。计算不泄露任何信息且可验证,该工作的重点是尽量减少计算延迟。文献[27]研究大规模线性优化问题的外包计算高效验证问题,结合线性优化问题特性,并利用云端返回的额外验证结构,实现了对计算结果的高效验证。在此基础上,文献[28]研究了符合一定条件的凸优化外包计算问题的高效验证问题,该方案将符合特定性质的凸

优化验证问题转化为线性优化问题,从而达到高效的验证结构。

对出于隐私安全考虑的云计算架构和模型,文献[29]介绍了隐私即服务(PaaS, privacy as a service),一种确保云环境中用户隐私数据不被侵犯的安全协议。为在云计算中提供可靠和隔离的执行环境,该解决方案依赖于安全的加密协处理器,可通过利用密码管理员的防篡改功能,实现用户数据的安全存储和计算,反过来保护用户数据不被未经授权的访问。

在云环境隐私侵犯的其他方面,文献[30]探讨云环境下数据索引引起的隐私问题,为解决数据索引导致的信息泄露,提出了一个三层数据保护架构,为云客户提供不同级别的隐私保护。文献[31,32]提出了依赖混淆技术的隐私管理器,可以提供模糊和去混淆服务,以减少存储在云环境中敏感信息的数量。由于该方案客户端私有数据的加密形式存储在云端,数据处理直接在加密数据上执行,该方案的一个限制是云供应商可能不愿意为保护隐私而实施其他服务。

2.3 存在的问题与挑战

云计算是一种主流的信息系统技术,但云环境下的用户隐私保护和隐私侵犯追踪取证一直存在着巨大的挑战。在云环境下,最突出的特征是用户隐私泄露痕迹易销毁、虚拟机迁移频繁、难定位和难溯源,最本质的问题是云平台虚拟化环境动态性高的特点,最大的问题是虚拟化环境回滚、迁移导致很难保证虚拟机取证数据的时空行为状态一致性以及虚拟化环境下攻击行为多元化导致攻击事件重构困难等。

依据云计算取证和隐私侵犯现状的调研可以明确,当前主流的云计算环境中取证技术体系和标准体系并不成熟和完备。而典型的互联网基础设施仅具备最基础的账号、访问控制体系和非规范化的日志系统,而这些机制的设计初衷也仅是面向资源的使用和业务调试,距离计算机取证的目标相去甚远。

一方面,云环境隐私保护取证需要可信的取证技术方案来支撑证据的获取、保存、分析、出示和审计等关键环节,该技术方案需要覆盖敏感云环境下数据全生命周期的各个环节:存储、传

输、计算、共享、发布和销毁;且能够无缝融入当前主流的云环境中。另一方面,云环境隐私保护取证需要处理的是海量、多维、异构的证据源,因此,同样需要使用云计算存储、计算架构和挖掘技术进行智能化、自动化的取证分析。

总体来说,云环境下隐私侵犯取证带来了新的挑战,需要从以下几个方面进行关键技术突破:云环境下隐私侵犯线索捕获;隐私侵犯线索的归一化融合分析;基于语义分析的隐私侵犯事件识别;分布式虚拟节点隐私数据标记;隐私数据标记的监测自适应覆盖;隐私数据访问与计算流转全流程追踪。突破以上技术挑战后,便可形成一整套覆盖敏感数据处理全流程的分布式、智能化取证方案,支撑研制分布式隐私数据访问、流转、计算监控系统,分布式多元证据采集系统,并在当前主流的云计算平台进行验证,实现云环境下的证据采集和智能取证分析系统。

3 云环境隐私侵犯取证

3.1 研究内容

根据云环境隐私侵犯取证存在的问题和挑战,针对云计算的复杂环境,云环境隐私侵犯取证的主要研究内容有以下3个方面(如图1所示)。

3.1.1 多元隐私侵犯行为归一化融合分析

基于虚拟机时空行为状态一致模型。研究虚拟机多内存镜像比对分析取证技术,提出增量内存镜像取证方法,通过对增量内存镜像还原,实现多镜像比对分析和数据挖掘,还原内存中的攻击痕迹,解决云计算环境下用户隐私的窃取具有动态易消失性问题;研究面向取证的虚拟机迁移技术,结合用户的虚拟机位置隐私策略,实现对虚拟机迁移路径和迁移环境的取证,解决用户虚拟机环境迁移导致隐私泄露难定位的问题。研究面向第三方的可信云监控技术,通过对云管理员和用户及其行为进行准确追踪和定位,实现云环境下虚拟身份的追踪和取证,解决用户隐私窃取的难溯源问题。在多元证据的基础上研究面向云环境的多维事件重构,将搜集的多种证据重构为隐私侵犯线索。

针对云计算环境下用户隐私数据分布在多层次、多节点,泄露痕迹易销毁等问题,基于虚拟

机内存连续增量镜像,结合虚拟机在 Host、Guest、network 等环境下的多种上下文数据,提出关键对象变化连续跟踪方法、将虚拟机在某段时间内的变化重新还原;具体突破快速增量信息镜像、关键对象在内存镜像中准确定位和多种数据关联分析等机制;解决多种数据融合问题,并为多维事件重构提供支撑。

3.1.2 面向虚拟机的隐私侵犯行为检测与取证

在归一化融合分析的基础上,根据隐私侵犯行为线索,进一步降低误报和漏报实现对隐私侵犯线索的准确捕获,并结合云环境研究基于语义分析的隐私侵犯事件识别,将多个隐私侵犯线索进行分析得到更高层的隐私侵犯时间。为了克服虚拟机独有的回滚行为对取证的干扰,通过构建细粒度抗回滚监测取证域,实现对窃密攻击行为的全生命周期取证,解决虚拟机回滚后取证数据碎片化问题。

针对云计算平台租户通过回滚虚拟机销毁攻击痕迹问题,制定细粒度的监测数据同步策略,构建轻量级抗回滚监测取证域,解决隐私信息窃取者通过回滚虚拟系统导致的取证记录碎片化的问题,确保取证时,能准确还原证据数据,保障监测取证软件不受恶意回滚影响,对窃密攻击行为的全生命周期取证;解决虚拟机时间状态一致

性问题,为云环境虚拟机时空行为状态一致性模型提供支撑。

3.1.3 云计算处理过程中的隐私窃取取证

云计算环境中数据复杂、异构,数据流转和计算模型繁多,隐私窃取证据难以完整提取等特点,研究云计算存储和计算架构中的访问控制、异构模块间的互访特点、计算拓扑的抽象等内容;提出基于三层级的隐私数据取证模型;突破在云计算环境下隐私数据的访问取证、计算流转取证、计算语义逻辑取证等关键技术;实现云计算环境中隐私数据处理全流程的完整取证。

针对云计算证据多维、海量、异构、不完备等造成的有效关键证据难提取、高效取证分析难等问题,提出云计算环境下的多元证据分析模型;突破基于云计算平台的集群化分析,智能化取证分析和内容抽样等关键技术;实现在云计算环境下对海量、异构证据的自动化、智能化取证分析;解决云计算环境下对隐私数据生命周期全环节的取证问题,并为基于云计算平台的隐私数据智能化取证分析系统提供支撑。

3.2 技术路线

基于 3.1 节的 3 点关键研究内容,为突破云环境下隐私侵犯取证困难、分析效率低、隐私侵犯难跟踪等问题,可以采用如图 2 所示的技术路线。

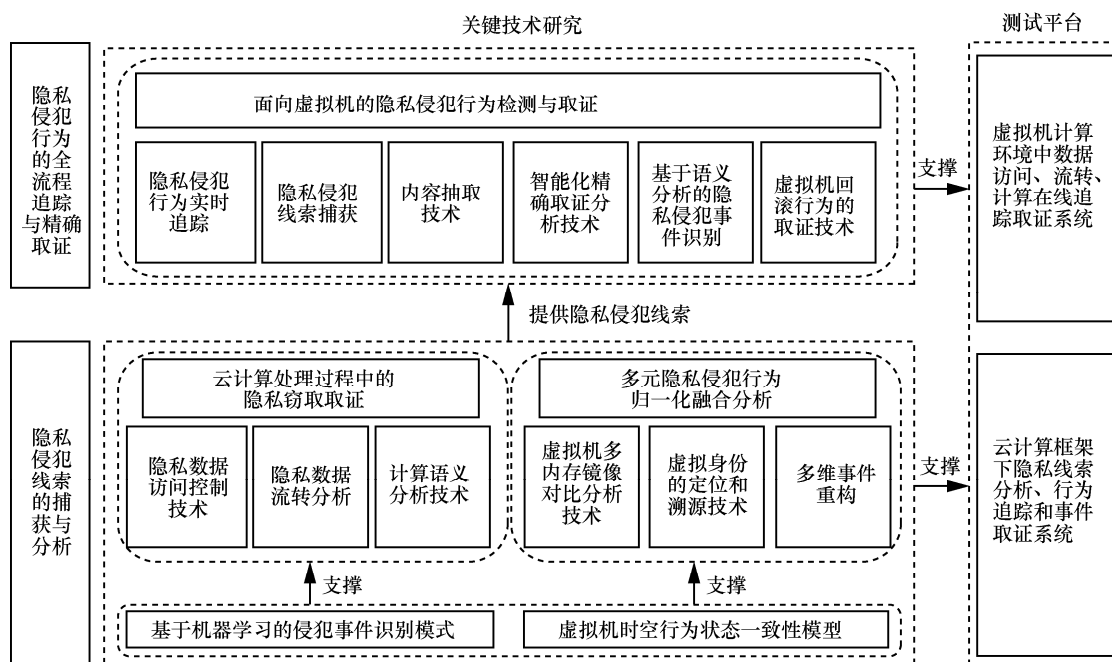


图1 研究内容示意

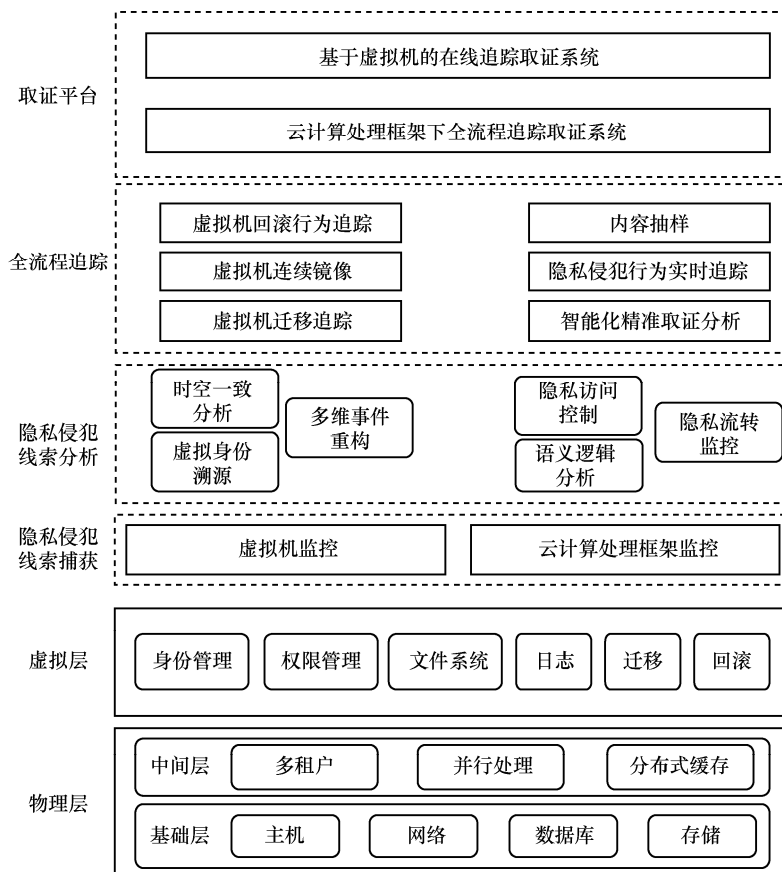


图2 技术路线示意

证据的采集应包括虚拟机监控和云计算处理框架监控，监控点分布在物理层和虚拟层的多个节点和组件上，通过建立数据通道对证据进行可靠的采集，建立控制通道对证据采集进行控制和反馈，使用采集策略采集（通过分布式的采集代理把多元证据采集）并上报到隐私侵犯线索分析层。

隐私侵犯线索分析层应对云计算环境下的证据，通过时空一致分析、多维事件重构、虚拟身份溯源等技术将采集的数据进一步语义化，从而使多种证据关联在一起，并对证据采集形成反馈和控制闭环；同时对云环境中采集到的数据采取以下操作。1) 对隐私数据访问取证。基于隐私数据的定级分类技术与分布式访问控制技术，实现云计算存储与计算中涉及隐私数据访问的细粒度取证；做到对隐私数据的 who-when-where-what-how 访问取证。2) 对隐私数据计算流转取证。在操作 1) 的基础上，基于云计算数据流追踪技术、数据抽样技术、差分隐私等技术，对云计算分布式计算的全过程中涉及隐私数据的各个

环节进行取证；做到对数据源、数据流经各个节点产生的中间结果，以及各类数据输出中涉及隐私泄露的取证。3) 隐私数据计算语义逻辑取证。在操作 1) 和操作 2) 的基础上，基于云计算语义分析技术，构建顶层云计算运算关系的逻辑拓扑图，推断敏感数据在计算过程中的逻辑关系，实现正向/反向隐私数据逻辑走向追踪取证，云计算语义隐私泄露的取证。

关联后的证据进一步提交给全流程追踪层，这个层次针对云计算环境采用虚拟机回滚行为追踪、虚拟机连续镜像分析、虚拟机迁移追踪多维事件重构和时空一致性分析、内容抽样技术，有针对性地具有隐私泄露的相关行为提取出来，形成有效的隐私侵犯证据。其中多维事件重构中虚拟机内存连续镜像比对分析技术，将跟踪虚拟机内存中的隐私数据，结合虚拟机在 Host、Guest、Network 等环境下的多种上下文数据，提出关键对象变化连续跟踪方法，检测隐私侵犯行为。虚拟机回滚行为的取证技术，制定细粒度的监测数

据同步策略,构建轻量级抗回滚监测取证域,解决隐私信息窃取者通过回滚虚拟系统导致的取证记录碎片化的问题,确保取证时,能准确还原证据数据。全流程追踪层针对云计算处理框架的证据通过进一步的内容抽样、隐私侵犯行为实时追踪、智能化精准取证分析。

4 面向云环境的隐私侵犯取证系统

4.1 系统架构

在云环境中存在着利用迁移回滚技术抹去隐私侵犯证据的行为。利用传统或者一般的方法,很难获取证据并分析其隐私侵犯行为。为此,本文提出了基于虚拟机时空状态一致性的面向云环境的隐私侵犯取证系统。该系统合理有效地解决了云环境取证、分析、数据多元化等难点问题,为云环境隐私侵犯提供有效的解决方案。

面向云计算环境的追踪取证系统整体架构如图3所示,主要分为3部分。

1) 改造现有的虚拟机监控器(VMM, virtual machine manager),加入可信的云环境监控模块,这个模块的主要功能包括:隐私数据标记、内存镜像、迁移回滚标记。核心的功能是监控客户虚拟机中隐私数据的流转,保证在云计算环境这种

碎片化、传播路径随机情况下能够正确地跟踪隐私侵犯的线索。

2) 构建在云环境外部的隐私侵犯分析系统。这个模块主要的功能包括:面向虚拟机迁移/回滚取证,克服虚拟机迁移或回滚操作对取证的影响,保证能够正确地跟踪隐私数据;通过对不同虚拟机的时空一致性分析,克服隐私数据在不同虚拟机中分布和时序不一致问题;多元隐私线索归一化,能够将不同虚拟机中的多个线索进行管理分析,得到统一的隐私侵犯场景。

3) 隐私侵犯证据库。包含了隐私侵犯行为的全过程,包括隐私数据标记位置、隐私侵犯行为以及影响的操作系统对象、隐私侵犯在云计算环境中的上下文。

整个系统设计的关键技术主要是对现有云环境中VMM的改造,使之更加适合隐私侵犯取证,涉及的主要技术包括隐私标记和隐私侵犯行为追踪、隐私侵犯行为多元化分析、隐私侵犯取证保存3个部分。

隐私数据标记分为手工标记和自动化标记。手工标记是提供一组隐私数据标记API,由云程序设计者调用API标记隐私数据位置,通过VMM的监控系统进行维护。自动化标记是VMM监控

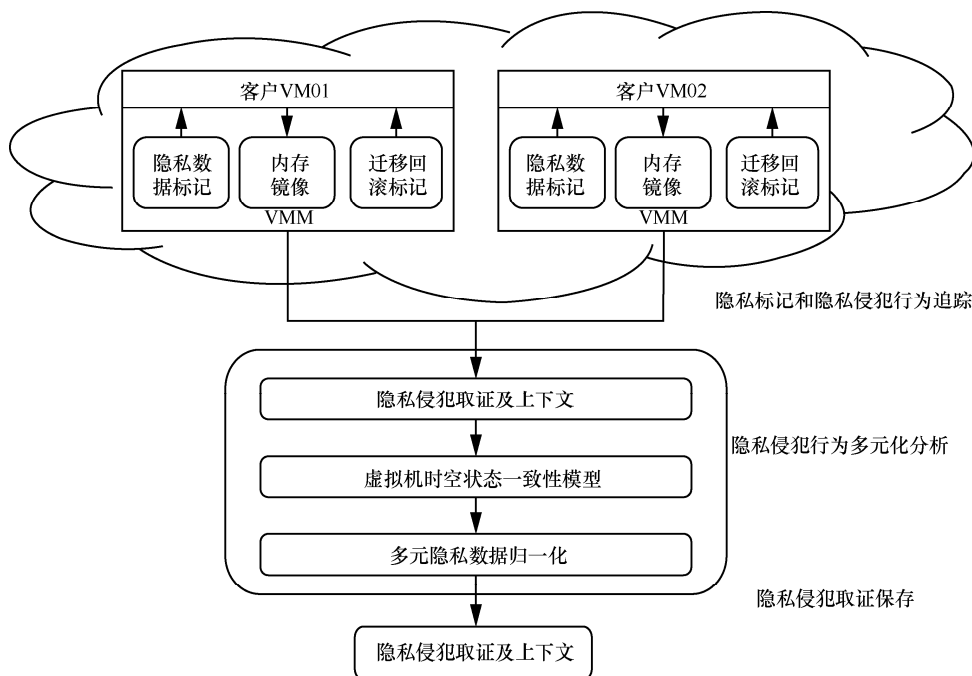


图3 面向云计算环境的追踪取证系统整体架构

虚拟机的执行，将加密关键函数访问的内存区域自动标记为敏感数据，为了提高标记精度，需要采用智能化的分析。

迁移回滚标记需要跨多个物理主机和虚拟机系统，对迁移的进程、虚拟机进行准确定位，保持隐私数据的唯一性，以便下一步的分析系统能够整合所有的线索，因此迁移回滚标记是一个分布式的协作模块。自适应的内存镜像模块可以根据需要自动化地对虚拟机内存进行保存和分析，能够监控虚拟机中内存变化和操作系统核心对象的创建、修改、删除，跟踪各种隐私侵犯行为，作为隐私侵犯线索保存。最终，整个系统的设计和实现需要充分考虑对云计算环境的性能影响，保证不影响云计算环境正常的业务处理。

隐私侵犯行为多元分析的主要作用是在云环境中收集可能存在隐私侵犯行为的信息，通过多维度 and 长时间跨度范围的综合分析，找出隐私侵犯行为并取证。隐私侵犯行为多元分析系统首先使用 Bro IDS 获取云环境中客户机中的可执行文件，然后通过 Yara 分析该可执行程序的行为，并通过恶意软件样本判断是否为可疑应用。若判断为可疑应用后，再获取客户虚拟机的系统版本等信息，并利用虚拟机自省技术（VMI，virtual machine introspection）模拟出该客户机同样环境的一个虚拟机。接下来，在此模拟环境中对应用的行为做进一步细致的分析，其中包括连续内存镜像分析、隐私泄露路径监测等过程，并将其存在的隐私侵犯行为记录下来。通过这样的方法和过程，可以有效地避免破坏客户机原有的运行环

境和正常任务，同时可以抵抗虚拟机回滚等抹去隐私侵犯记录的行为。为提高效率，在某些过程中保持粗粒度的检测，同时，经过进一步细粒度的检测，准确地验证隐私侵犯的证据。整体架构如图4所示。

4.2 内存隐私数据标记实例

在云环境中，存在把内存中的隐私数据通过文件或者网络的方式泄露出去的问题。由于文件读写和网络请求十分频繁，很难从大量的文件和网络请求中分辨是否存在泄露隐私数据的行为。而通过内存隐私数据标记和污点追踪的方法，则可以很容易地分辨出隐私数据并进一步对该数据实施污点追踪，存在泄露隐私数据行为的证据也将轻易被获取到。

内存标记和污点追踪的关键点在于如何标记关键的内存数据，需要通过特征字符串定位、关键系统调用、SSL 密钥等数据定位隐私数据。例如，在本节实例中，定义 Firefox 的历史记录（'0x06 0x25 0x08','0x06 0x25 0x09'等）、cookie（'0x04 0x06 0x06','0x05 0x06 0x06'等）等关键特征字符串，并在物理内存中搜索该字符串，再通过 hook 系统截取信息关键函数（GetComputerName、GetUserName、GetSystemInfo 等），从而完成内存标记并做进一步的污点追踪。

图5为利用本系统实现的一个内存隐私数据标记的大致流程，在该例子中，通过对虚拟机内存中 Firefox 历史记录、cookie 等隐私数据关键字字符串的搜索，标记相应隐私数据为污点源。通过 API hook 技术监控 GetSystemInfo 等系统关键信

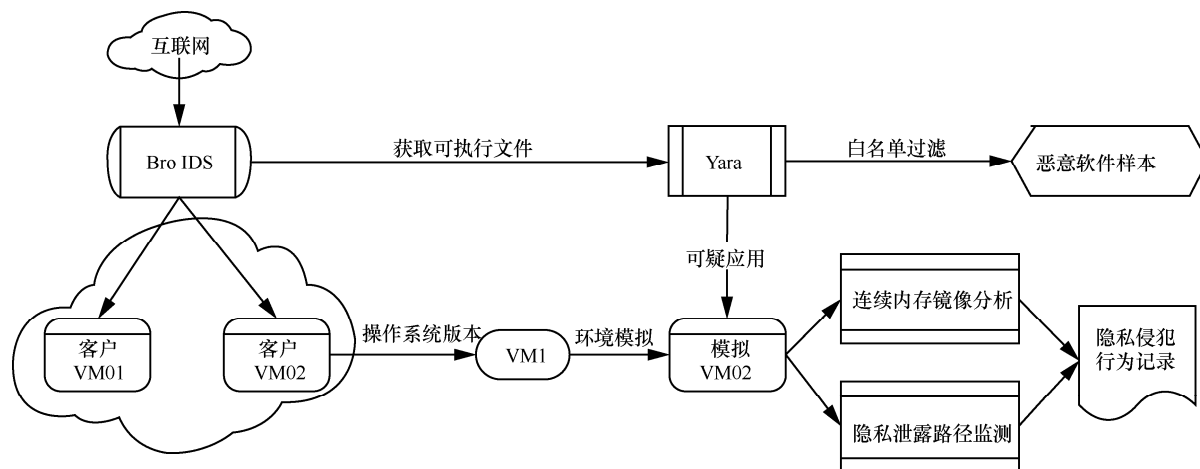


图4 隐私侵犯行为多元分析架构

息函数的调用, 将其参数标记为污点源。通过对污点内存数据的追踪, 分析污点数据在操作系统中的流向。若应用程序通过网络套接字和文件相关接口进行操作时, 检查是否存在被标记的污点隐私数据有写入文件或通过数据分组发送至网络的操作, 达到标记并追踪隐私数据的目的, 为整个面向云计算环境的追踪取证系统提供标记隐私数据的基础。

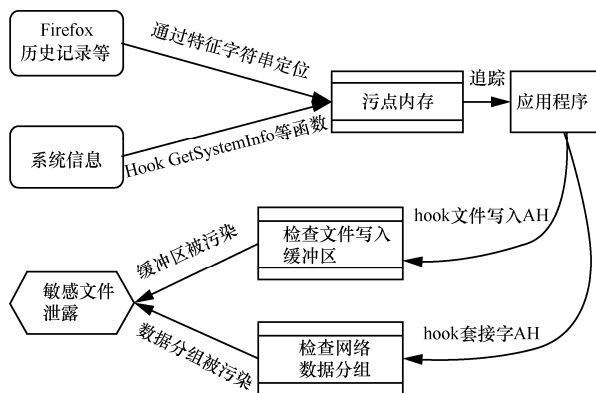


图5 内存隐私数据标记实例

5 结束语

随着云计算的进一步发展, 云环境下取证分析的重要性不言而喻。本文主要从现有的研究状况入手, 分析云环境下隐私取证存在的问题和挑战, 然后提出了其关键的研究内容和技术路线, 最后提出一个面向云环境隐私取证的系统。随着云环境取证技术的不断完善和进步, 云取证必将推动计算机其他领域的技术发展。

参考文献:

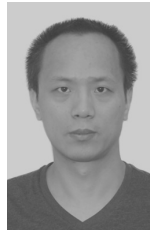
- [1] NIST cloud computing forensic science challenges[EB/OL]. https://csrc.nist.gov/csrc/media/publications/nistir/8006/draft/documents/draft_nistir_8006.pdf.
- [2] RUAN K, CARTHY J, KECHADI T, et al. Cloud forensics[C]// IFIP International Conference on Digital Forensics. 2011: 35-46.
- [3] Mapping the forensic standard ISO/IEC 27037 to Cloud Computing[EB/OL]. <https://downloads.cloudsecurityalliance.org/initiatives/imf/Mapping-the-Forensic-Standard-ISO-IEC-27037-to-Cloud-Computing.pdf>.
- [4] RUAN K, CARTHY J, KECHADI T, et al. Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results[J]. Digital Investigation, 2013, 10(1): 34-43.
- [5] ZARGARI S, BENFORD D. Cloud forensics: concepts, issues, and challenges[C]//The third Conference on Emerging Intelligent Data and Web Technologies (EIDWT). 2012: 236-243.
- [6] TAYLOR M, HAGGERTY J, GREESTY D, et al. Forensic investigation of cloud computing systems[J]. Network Security, 2011(3): 4-10.
- [7] [EB/OL]. <http://www.guidancesoftware.com/encase-forensic>.
- [8] [EB/OL]. <http://www.accessdata.com/solutions/digital-forensics/forensic-toolkit-ftk>.
- [9] ZAWOAO S, HASAN R. Cloud forensics: a meta-study of challenges, approaches, and open problems[EB/OL]. <http://www.cryptome.org/2013/02/cloud-forensics.pdf>.
- [10] NANDA S, HANSEN R A. Forensics as a service: three-tier architecture for cloud based forensic analysis[C]//The 15th International Symposium on Parallel and Distributed Computing (ISPDC). 2016: 178-183.
- [11] 谢亚龙, 丁丽萍, 林渝淇, 等. ICFF: 一种 IaaS 模式下的云取证框架[J]. 通信学报, 2013, 34(5): 200-206.
- [12] XIE Y L, DING L P, LIN Y Q, et al. ICFF: a cloud forensics framework under the IaaS model[J]. Journal on Communications, 2013, 34(5): 200-206.
- [12] 周刚. 云计算环境中面向取证的现场迁移技术研究[D]. 武汉: 华中科技大学, 2011.
- [12] ZHOU G. Research on forensics-oriented site migration technology in cloud computing environment[D]. Wuhan: Huazhong University of Science and Technology, 2011.
- [13] LEE J, UN S. Digital forensics as a service: a case study of forensic indexed search[C]//The International Conference on ICT Convergence (ICTC). 2012: 499-503.
- [14] WEN Y, MAN X, LE K, et al. Forensics-as-a-service (faas): computer forensic workflow management and processing using cloud[C]//The fifth International Conferences on Pervasive Patterns and Applications. 2013: 1-7.
- [15] SANG T. A log based approach to make digital forensics easier on cloud computing[C]//Intelligent System Design and Engineering Applications (ISDEA). 2013: 91-94.
- [16] 田庆宜, 王琳, 黄道丽. Vsphere 架构私有云取证方法研究及其实例[J]. 信息安全, 2011, 9: 44-46.
- [16] TIAN Q Y, WANG L, HUANG D L. Research and practice of VSPHERE architecture private cloud forensics methods[J]. Netinfo Security, 2017, 30(1): 31-38.
- [17] 吴松洋, 张熙哲, 王旭鹏, 等. 基于 Hadoop 的高效分布式取证: 原理与方法[J]. 电信科学, 2017, 30(1): 31-38.
- [17] WU S Y, ZHANG X Z, WANG X P, et al. Highly distributed distributed forensics based on hadoop: principles and methods [J]. Telecommunications Science, 2017, 30(1): 31-38.
- [18] ZHOU M, ZHANG R, XIE W, et al. Security and privacy in cloud computing: a survey[C]//The sixth International Conference on Semantics Knowledge and Grid (SKG). 2010: 105-112.
- [19] BURNSIDE R S. The electronic communications privacy act of 1986: the challenge of applying ambiguous statutory language to intricate telecommunication technologies[J]. Rutgers Computer & Technology, 1987, 13: 451.
- [20] SHELDON J. Fair credit reporting act[J]. National Consumer Law Center Consumer Credit & Sales Legal Practice, 1994, 65(2): 595-609.

- [21] EVideo Privacy Protection Act[EB/OL]. <http://epic.org/privacy/vppa/>.
- [22] GELLMAN R. Privacy in the clouds: risks to privacy and confidentiality from cloud computing[C]//The World privacy forum. 2012.
- [23] CAVOUKIAN A. Privacy in the clouds[J]. Identity in the Information Society, 2008, 1(1): 89-108.
- [24] PEARSON S. Taking account of privacy when designing cloud computing services[C]//The 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing. 2009: 44-52.
- [25] PEARSON S, CHARLESWORTH A. Accountability as a way forward for privacy protection in the cloud[C]//IEEE International Conference on Cloud Computing. 2009: 131-144.
- [26] SADEGHI A R, SCHNEIDER T, WINANDY M. Token-based cloud computing[C]//The International Conference on Trust and Trustworthy Computing. 2010: 417-429.
- [27] WANG C, REN K, WANG J. Secure and practical outsourcing of linear programming in cloud computing[C]//INFOCOM. 2011: 820-828.
- [28] XU Z, WANG C, REN K, et al. Proof-carrying cloud computation: the case of convex optimization[J]. IEEE Transactions on Information Forensics and Security, 2014, 9(11): 1790-1803.
- [29] ITANI W, KAYSSI A, CHEHAB A. Privacy as a service: privacy-aware data storage and processing in cloud computing architectures[C]//Dependable, Autonomic and Secure Computing. 2009: 711-716.
- [30] SQUICCIARINI A, SUNDARESWARAN S, LIN D. Preventing information leakage from indexing in the cloud[C]//The 3rd International Conference on Cloud Computing (CLOUD). 2010: 188-195.
- [31] PEARSON S, SHEN Y, MOWBRAY M. A privacy manager for cloud computing[C]//IEEE International Conference on Cloud Computing. 2009: 90-106.
- [32] MOWBRAY M, PEARSON S. A client-based privacy manager for cloud computing[C]//The Fourth International ICST Conference on Communication System Software and Middleware. 2009: 5.

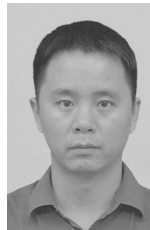
[作者简介]



李维奉（1991-），男，江西南昌人，华中科技大学硕士生，主要研究方向为信息安全。



姜卫中（1977-），男，江苏南通人，博士，华中科技大学副教授，主要研究方向为信息系统安全。



李伟明（1975-），男，湖南株洲人，博士，华中科技大学副教授，主要研究方向为信息安全。



邹德清（1975-），男，湖南长沙人，博士，华中科技大学教授，主要研究方向为信息安全。