一种自适应的动态取证机制

陈 琳[1] 李之棠[1,2] 高翠霞[1]

(华中科技大学计算机学院 武汉 430074)[1] (华中科技大学网络中心 武汉 430074)[2]

摘 要 随着网络入侵技术和计算机犯罪技术的发展,动态取证变得越来越重要。利用入侵检测系统和蜜罐来实现入侵取证的方法在取证的实时性方面有很大优势,但这些方法没有过多考虑系统被入侵时证据可靠性以及系统可靠性的问题,而且取证的时机难以掌握。提出了一种自适应的动态取证方法,该方法采用入侵检测系统作为取证触发器,利用影子蜜罐对疑似攻击进行确认和进一步观察分析,自适应调整取证过程,获取关键证据,最后采用有限状态机对该机制进行建模,并对该机制中的状态转换时机、影子蜜罐、证据安全存储等关键技术进行描述。利用该机制来实现动态取证,可以使得取证过程更可控,可以减少不必要的证据量,并增强系统的容侵性。

关键词 动态取证,影子蜜罐,自适应,有限状态机

中图法分类号 TP393 **文献标识码** A

Self-adaptive Mechanism of Dynamic Forensics

CHEN Lin[1] LI Zhi-tang[1,2] GAO Cui-xia[1]

(School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)[1]

(Network Center, Huazhong University of Science and Technology, Wuhan 430074, China)[2]

Abstract With the development of intrusion and computer crime technologies, dynamic forensics is becoming more and more important. Dynamic forensics based on intrusion detection and honeypot technologies has great advantage in real-time performance, whereas these methods are defective in overcoming the difficulty of evidence and system reliability, and hard to seize the opportunity of investigation. A self-adaptive mechanwasm was proposed which used intrusion detection system as forensics trigger and shadow honeypot was used to verify the suspicious attack, observe and analyze the attack activities further more to gather key evidences. And then the finite state machine model of this mechanism was illuminated and key technologies such as shadow honeypot, state transition opportunity and evidence security storage method were described. The dynamic forensics system with this mechanism can tolerate intrusion in a certain degree and get the investigation process under control. Moreover, the amount of unnecessary evidences can be reduced obviously.

Keywords Dynamic forensics, Shadow honeypot, Self-adaptive, Finite state machine

1 引言

随着网络入侵技术和计算机犯罪技术的不断升级,静态的事后取证方法已不再满足网络取证的需求,实时的动态取证变得越来越重要。

动态取证主要通过对网络数据流、审计迹、主机系统日志等进行实时监控和分析,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,要求能够从海量的、动态的数据中及时分析出有效的数据,减小误判率,解决实时性、有效性、可适应性和可扩展性问题。相对于传统的取证方法来说,动态取证强调对网络动态信息的收集和网络安全的主动防御。从法律的角度来看,证据的可靠性和可信性应该引起足够的重视。

2 动态取证中的常用技术

2.1 入侵检测技术

当前动态取证的一个研究热点就是利用入侵检测系统IDS来检测非法的入侵或恶意行为,对这些行为记录日志,并将日志作为网络取证的主要证据源。很多IDS工具,如Snort,就是先捕获网络流量,对流量进行分析,发现攻击活动后就记录告警日志。

Sommer[1]认为IDS的一个更进一步的目标应该是为计算机犯罪提供法律上的证据。Stephenson指出,可以利用IDS的告警日志作为实时或接近实时的网络取证的证据源[2]。Stephenson还提出了一个入侵管理模型(Intrusion Management Model)[3],对取证和入侵检测的有效结合做出了重要的理论贡献。文献[4]介绍了第一个实时入侵取证原型系统的实现。将IDS应用到网络取证中,可以使得整个取证过程具有更好的实时性,将多个部署在不同位置的不同类型IDS结合起来,可以获得更为全面的证据,有助于及时做出响应。

虽然采用IDS作为取证工具可以很好地解决实时性、智

到稿日期:2008-12-24 返修日期:2009-03-05 本文受国家自然科学基金(60573120)资助。

陈 琳(1976-),女,博士生,讲师,主要研究方向为网络安全,E-mail:chenlin@mail.hust.edu.cn;李之棠(1951-),男,教授,博士生导师,主要研究方向为网络技术、网络安全等;高翠霞(1974-),女,博士生,主要研究方向为网络安全。

能性等问题，但 IDS 的日志作为证据存在以下问题：

1）日志量非常庞大，格式不统一且存在冗余，不适合直接作为证据来使用；

2）存在较多误报和漏报，不能直接适用于取证。

2.2 诱骗技术

取证同诱骗技术相结合是动态取证的另一个研究方向。利用蜜罐等诱骗系统来收集证据，可以减小取证量，直接将取证者的注意力集中在入侵活动上，并且可以利用蜜罐拖延攻击者对真实目标的攻击。Alec[5]提出利用蜜阱(honeytrap)来监视攻击者的活动和策略，自动收集电子证据，设计了串联和并联两种取证框架，以实现取证技术与诱骗技术的结合。文献[6]指出利用蜜罐的取证可以帮助分析者识别攻击的指纹。

蜜罐等诱骗系统的架设和使用首先应该获得法律上的肯定。另外，作为诱骗系统的蜜罐或蜜阱，并没有向外界提供真正有价值的服务，提供的所谓敏感信息也都是虚假的，不具有实际价值。对这样的系统或者信息的攻击进行取证调查后获得的信息是否能作为法律认可的证据，仍是一个有争议的问题。高明的黑客甚至可能反过来控制蜜罐，把它作为进一步攻击的跳板，这也是利用蜜罐来进行取证所必须要考虑的问题。

2.3 Agent 技术

Agent 技术在网络安全领域有着很广泛的应用，特别是在入侵检测领域已有很多研究成果。

利用 Agent 技术来实现网络动态取证，可以实时从多个网络取证源上获取证据，从更全面的视角来分析入侵。ForNet[7]是一个用于数字取证的分布式网络日志系统，可以用在广域网上，利用安装在其他网络设备上的代理 SynApp 来收集网络事件。Ren 介绍了一种基于 Agent 的分布式主动实时网络取证系统的框架[8]。

3 自适应动态取证的思想

采用入侵检测、诱骗系统等技术来实现动态取证，有助于获得实时的证据，并有助于及时地对网络安全风险进行评估。但是，除了应用在取证方面具有一些特定的缺陷之外，这些方法都是基于一个共同的假设：系统在遭受入侵时，仍然可以正常工作，所以可以比较完整地获得证据。而实际上，系统被攻击时，已经处于一种不安全、不可靠的状态，进程可能已被入侵者劫持，所获得的证据很有可能是已被入侵者篡改过的。因此，我们希望有一种机制可以使得系统即使在遭受到入侵的时候仍然处于一种可控的状态或者是一种自适应的状态，对入侵行为进行观察、容忍并在入侵被确认的情况下获取必要的证据。整个过程应该是可控且可信任的。

入侵容忍是近年来的一个研究热点，它关注的是入侵造成的影响而不是入侵的原因，能够实现自我诊断、故障隔离。我们可以利用入侵容忍的思想构造一个更为健壮的动态取证系统，并能够比较准确地把握取证时机，进一步消除入侵检测的误报，减少证据量。

4 自适应动态取证的机制

4.1 系统模型架构

自适应动态取证机制的系统架构如图 1 所示。

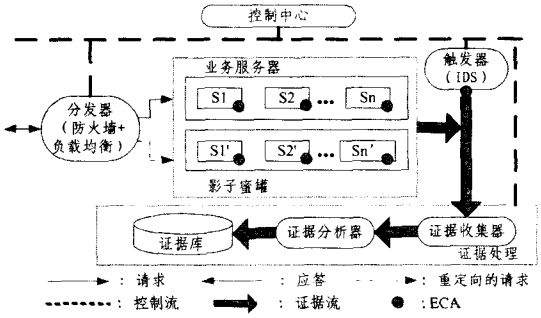

图 1 自适应动态取证机制的系统架构

1）ECA(证据收集代理)：ECA 是集成在各个证据源中的代理，这些证据源包括 IDS、服务器以及蜜罐。ECA 从证据源收集证据，然后把证据传送给证据接收器。IDS 告警日志、服务器系统日志、访问日志、安全日志以及蜜罐上所有通信日志、击键记录的分析报告都是非常有用的、需要收集的证据。

2）触发器：IDS 可以作为启动入侵容忍以及取证的触发器。NIDS 监视网络中的通信，HIDS 对主机的文件、进程以及网络连接进行检测，如果发现有攻击活动就产生告警。常用的 IDS(如 snort)是基于规则来检测入侵的，规则中有一个字段“priority”用来表示攻击活动的级别，匹配该条规则的攻击活动所触发的告警可以相应反映出该攻击活动的威胁级别。攻击的威胁级别是触发取证以及系统状态转换的重要参数之一。

3）影子蜜罐：服务器配备有影子服务器。影子服务器本质上是蜜罐系统，其数据都与真实服务器一致，但从容忍的多样性角度来考虑，其结构、实现都与真实服务器有所区别。影子蜜罐用来监视可疑流量，通过安装 Sebek，Nepenthes 等工具来进一步检测攻击。IDS 判断为疑似攻击的流量被重定向到影子蜜罐。影子蜜罐对所有进入的流量都进行分析和记录，一旦确认为攻击，则通知更新入侵防御策略。蜜罐系统增加了强访问控制措施，防止入侵者利用蜜罐实施跳板攻击。

4）分发器：分发器可以由带有负载均衡功能的路由器或防火墙来充当，在需要的时候对流量进行重定向。正常情况下，请求被送给真实的服务器。如果 IDS 发现有不能确定的疑似攻击流量，则这种流量被同时发送给真实的服务器和影子蜜罐，由蜜罐进行观察和确认。一旦可疑流量被确定为攻击，则切断阻断恶意流量到真实服务器的连接，将恶意流量全部重定向到影子蜜罐，由蜜罐进一步监视和记录证据。

5）证据处理：证据处理模块包括证据接收器、证据分析器和证据库 3 个部分。证据接收器负责接收由 ECA 传送来的事件和日志记录，证据分析器对这些事件和日志进行分析。分析产生的结果是最后用来保存和出示的证据，存储在证据库中。

6）控制中心：控制中心是整个系统的核心，控制系统状态的转换，并通知相关的模块。

4.2 系统的有限状态机

有限状态机(FSM)的模型可以用一个五元组来表示：$M=(Q,\Sigma,\delta,q_0,Z)$，其中 Q 是一个有限的状态集合；q_0 是初始状态，$q_0\in Q$；Σ 是输入字母集，可以表示系统接收的所有事件的集合；δ 是状态转移函数，$\delta:Q\times\Sigma\rightarrow Q$，它描述了系统中每个状态转换到其他状态的可能性。常用定义式 $\delta(q_i,e)=q_j$ 表示在 q_i 状态下接收事件 e 之后，转入指定的新状态 q_j；Z

是结束状态的集合，$Z \subseteq Q$。

自适应动态取证系统的 FSM 表示如下：

$M=(Q,\Sigma,\delta,q_0,Z)$

其中，$Q=\{q_0,q_1,q_2,q_3,q_4,q\}$；$q_0$ 是初始状态；$Z=\{q\}$ 是结束状态集合；$\Sigma=\{E_1,E_2,E_3,E_4,E_5,E_6\}$，$E_i$ 是系统接收的可能的事件或动作。

状态转换函数 $\delta: Q\times\Sigma\rightarrow Q$，如表 1 所列。

状态 q_i 具体描述如下：

q_0：初始状态“正常”；

q_1：被探测；

q_2：可能被攻击，进行实时取证；

q_3：被恶意攻击，进行实时取证；

q_4：重定向或者防御失败，进行事后取证；

q：结束状态“切断所有连接，进行事后取证”。

事件 E_i 具体描述如下：

E_1：检测到探测或扫描活动；

E_2：检测到异常流量；

E_3：确认无害；

E_4：确认是恶意流量，重定向该流量；

E_5：重定向或者防御失败；

E_6：切断所有连接。

表 1　自适应动态取证机制的状态转换表

$\delta(q_0,E_1)=q_1$	$\delta(q_0,E_2)=q_2$
$\delta(q_0,E_6)=q$	$\delta(q_1,E_2)=q_2$
$\delta(q_1,E_4)=q_3$	$\delta(q_1,E_5)=q_4$
$\delta(q_2,E_3)=q_0$	$\delta(q_2,E_4)=q_3$
$\delta(q_2,E_5)=q_4$	$\delta(q_3,E_3)=q_0$
$\delta(q_4,E_6)=q$	

状态转换图如图 2 所示。

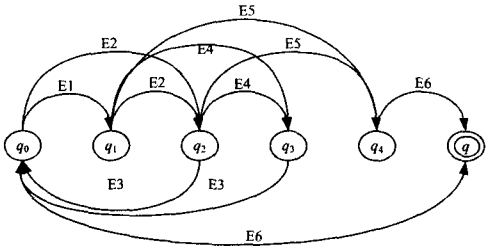

图 2　自适应动态取证机制的状态转换图

系统初始处于 q_0 状态。如果 IDS 检测到探测或扫描活动，这意味着可能有人正在对系统进行入侵前的踩点扫描，了解系统开放端口和服务的情况。一般情况下，探测和扫描的告警日志威胁度比较低，此时，系统的状态从 q_0 转为 q_1。

系统处于状态 q_0 或 q_1 时，收到被 IDS 判断为不正常的流量，入侵可能已经在发起了，这种情况被 IDS 报告给控制中心，系统状态转换为 q_2。由于此时流量尚属于疑似攻击，控制中心会通知负载均衡模块将疑似攻击的数据包发送给真实的服务器和影子蜜罐，同时启动实时取证，ECA 开始收集 IDS 以及影子蜜罐上的有关证据。影子蜜罐上的安全工具进一步分析可疑流量，观察这些流量对蜜罐造成的影响，从而确定其是不是真的恶意流量。

当系统处于状态 q_1 时，如果 IDS 能确认某些流量的确是入侵，或者某些攻击行为达到一定的威胁级别，系统状态就根据重定向成功与否直接转换到 q_3 或 q_4。

当系统处于状态 q_2 时，影子蜜罐会根据可疑流量在蜜罐中造成的后果来判断到底是正常流量被误报还是的确为恶意入侵，并把分析结果报告给控制中心。如果结果是前者，系统状态就从 q_2 转换回初始状态 q_0，真实服务器仍然继续提供服务，反之，控制中心会立即通知访问控制模块阻止恶意源与真实服务器之间的连接，通知负载均衡模块将所有相关的数据包都重定向到影子蜜罐，由影子蜜罐进一步观察和深入分析，系统状态也相应转换为 q_3。如果访问控制或者重定向失败，控制中心将会收到告警，系统状态随之转换为 q_4。此时，控制中心将通知访问控制模块切断流向被保护系统的所有流量，取证模块从系统各个相关模块收集证据，开始事后取证，系统转换到结束状态 q。

4.3　关键技术

4.3.1　状态转换的时机

不同状态之间转换需要有特定事件触发，但在具体实现中需要进行量化。可以定义一个状态转换参数 SP，SP 主要由以下几个参数来确定：

1）攻击威胁度 A；

2）服务器访问延迟时间 T；

3）服务器/蜜罐的服务进程数 P；

4）服务器/蜜罐的 CPU 利用率 C；

5）服务器/蜜罐的内存利用率 M；

6）重定向处理结果 R。

其中，攻击威胁度是主要因素。当 SP 超过某个特定阈值时，控制中心开始转换系统状态。

4.3.2　影子蜜罐

影子蜜罐是整个系统中非常重要的一个部分，其主要作用是作为真实服务器的冗余实现部分程度的容侵，吸引攻击流量，利用安装在蜜罐中的安全工具实现对疑似攻击的确认，其产生的所有日志记录均作为取证分析的重要证据。

影子蜜罐的多样性是保证系统安全的一个基本要求，影子蜜罐与真实服务器具有不同的入侵难度，入侵者必须更换攻击方法才能转而对影子蜜罐进行进一步入侵，避免了在入侵者对真实服务器进行了一段时间的初步攻击之后，影子蜜罐才被相同的方法在短时间内迅速攻陷而不能及时通知到访问控制模块。系统的多样性可以通过使用不同类型的操作系统、多版本程序设计(N-version programming)[9]等方法实现。

4.3.3　证据的收集与分析

为减少误报证据量，提高分析的效率，需要提交分析的证据应该是针对确为攻击的活动的日志记录。因此，证据接收模块采用两步式接收，设置两个证据队列：疑似队列 Q_s 和确认队列 Q_r，分别存放疑似证据和确认证据。在状态 q_2 时开始接收证据，证据先存放在 Q_s 中，如果状态转换为 q_0，则 Q_s 中的信息无需继续处理，直接清空或转入历史库中，以待后查；如果状态转换为 q_3 或 q_4，Q_s 中的数据转入 Q_r，接收模块再继续接收的数据都直接进入 Q_r，分析模块直接对 Q_r 中的证据进行在线分析。

对证据的分析包括去除冗余、证据融合及关联分析，可以采用的方法有模式匹配、数据挖掘、因果关联等。

4.3.4　证据的安全存储

为防止证据被篡改、伪造，需要对证据的完整性、真实性

（下转第 156 页）

息过滤、信息感知和显示感知信息。从社区规模、个人中心性和网络集中性3个方面讨论了社交网络环境感知，在社交网络模型的基础上研究了基于关系的群体感知。最后，介绍了面向科研工作者的社交网络——学术社区，给出了它的系统结构，介绍了各个部分的主要功能，并在学术社区内应用了感知技术，帮助社区内的研究者发现科研热点或某一领域的研究群体，促进了学术交流和创新。在未来的工作中，将结合用户体验，考虑用户意见，做深入的理论分析和实验验证，进一步完善感知理论，提高感知的应用效果。

参 考 文 献

[1] Boyd D M, Ellison N B. Social network sites: definition, history, and scholarship[J]. Journal of Computer-Mediated Communication, 2007, 13(1): 210-230

[2] Acquisti A, Gross R. Imagined communities: Awareness, information sharing, and privacy on the Facebook[J]. Privacy Enhancing Technologies, 2006: 36-58

[3] Schmidt A, Gellersen H-W. Visitor Awareness in the Web[C]// Proceedings of the 10th International Conference on World Wide Web(www2001). Hong Kong, China, 2001: 745-753

[4] Steiny D, Oinas-Kukkonen H. Network awareness: social network Search, innovation and productivity in organisations[J]. International Journal of Networking and Virtual Organizations, 2007, 4(4): 413-430

[5] Steiny D. Network Awareness, Social Context and Persuasion [C]// PERSUASIVE2008. LNCS 5033. 2008: 58-70

[6] Heo M. Relationship-based Social Awareness Disclosure[C]// Proceedings of the Sixth International Conference on Advanced Learning Technologies (ICALT'06). 2006: 852-853

[7] Susanne B, Ellen C. Computer support for social awareness in flexible work [J]. Computer Supported Cooperative Work (CSCW'06), 2006, 15(1): 1-28

[8] Buder J, Bodemer D. Supporting controversial CSCL discussions with augmented group awareness tools[J]. International Journal of Computer-Supported Collaborative Learning, 2008, 3(2): 123-139

[9] Newman M E J. The structure of scientific collaboration networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2001, 98(2): 404-409

[10] Miki T, Nomura S. Semantic web link analysis to discover social relationships in academic community[C]// Proceedings of the 2005 symposium on applications and the internet (SAINT'05). 2005: 38-45

[11] Matsuoa Y, Morib J, Hamasaki M. POLYPHONET: An Advanced Social Network Extraction System from the Web[C]// Proceedings of the 15th international conference on World Wide Web (www'06). 2006: 397-406

[12] Mika P. Flink: Semantic web technology for the extraction and analysis of social networks[J]. Journal of Web Semantics, 2005, 3(2): 211-223

[13] Kimmerle J, Cress U. Group awareness and self-presentation in computer-supported information exchange [J]. International Journal of Computer-Supported Collaborative Learning, 2008, 3(1): 85-97

[14] Valdis K. Social Network Analysis: A Brief Introduction[EB/OL]. http://www.orgnet.com/sna.html, 2006

(上接第67页)

加以保护。通常的方法是对证据计算报文摘要，增加数字签名，附加时间戳等，以便一定程度地保护证据的完整性和真实性。除此之外，利用Shamir秘密共享可以使得在证据遭到部分破坏时仍能恢复证据信息[10]，而采用强访问控制机制将取证服务器和证据数据库与业务系统强制隔离[11]，可以更有力地对证据加以安全保护。

结束语 动态取证最重要的是对网络的动态信息收集和网络安全的主动防御，如果人工来控制取证时机，很容易导致过早开始取证，从而获取大量无关的信息；或者过晚取证导致不能及时加以防御。本文提出的自适应的动态取证机制，可以通过系统组件之间的协作来适时触发取证，并采用影子蜜罐来保护真实服务器并进一步观察、确认和分析攻击过程，提供更完整的证据，提高取证过程的健壮性和智能性，并对证据的可靠性和可信性提供保护。

参 考 文 献

[1] Sommer P. Intrusion detection system as evidence [Z]. Recent Advances in Intrusion Detection-RAID 98

[2] Stephenson P. The Application of Intrusion Detection Systems in a Forensic Environment [C]// Proceedings of the RAID 2000 Conference. Toulouse, France, 2000

[3] Stephenson P. Intrusion Management: A Top Level Model for Securing Information Assets in an Enterprise Environment [C]// Proceedings of EICAR 2000. Brussels, Belgium, March 2000

[4] Payer U. Realtime Intrusion-forensics, a First Prototype Implementation [C]// TERENA Networking Conference. 2004

[5] Yasinsac A, Manzano Y. Honeytraps, a Network Forensic Tool [C]// Sixth Multi-conference on Systemics, Cybernetics and Informatics. Orlando, Florida, USA, July 2002

[6] Berthier R F, Biondi Y, Kaminsky P, et al. Forensics[C]// Proceedings from the Fifth Annual IEEE SMC. June 2004: 22-29

[7] Shanmugasundaram K, Memon N, Savant A, et al. ForNet: A Distributed Forensics Network [C]// Proceedings of the Second International Workshop on Mathematical Methods, Models and Architectures for Computer Networks Security. St. Petersburg, Russia, 2003

[8] Ren Wei, Jin Hai. A framework of distributed agent-based active and real time network forensics system [Z]. DFRWS, 2004

[9] Dugan J B, LyuBell M R. System Reliability Analysis of an N-version Programming Application[J]. IEEE Transactions on Reliability, 1994, 43(4): 513-519

[10] 杨晓元，季称利，秦晴，等. 基于Sharmir秘密共享的安全取证服务器方案[J]. 计算机工程与应用，2005，22：147-149

[11] 孙波，纪建敏，孙玉芳，等. 电子数据证据收集系统保护机制及其发展趋势[J]. 计算机科学，2004，31(7)：9-11