网络空间靶场技术研究

方滨兴[1, 2, 3], 贾 焰[2], 李爱平[2], 张伟哲[3]

[1] 北京邮电大学 北京 中国 100876
[2] 国防科技大学计算机学院 长沙 中国 410073
[3] 哈尔滨工业大学计算机科学与技术学院 哈尔滨 中国 150001

摘要 网络靶场已经成为支撑网络空间安全技术验证、网络武器试验、攻防对抗演练和网络风险评估的重要手段。本文首先介绍了网络靶场国内外研究现状; 然后介绍了靶场相关技术的研究进展, 包括大规模网络仿真、网络流量/服务与用户行为模拟、试验数据采集与评估、系统安全与管理等方面; 最后阐述了网络靶场发展面临的挑战与发展趋势。

关键词 网络靶场; 网络仿真; 用户行为模拟; 数据采集与分析; 安全与管理
中图法分类号 TP309.2 DOI 号 10.19363/j.cnki.cn10-1380/tn.2016.03.001

Cyber Ranges: state-of-the-art and research challenges

FANG Binxing[1, 2, 3], JIA Yan [2], LI Aiping[2], ZHANG Weizhe[3]

[1]Beijing University of Posts and Telecommunications, Beijing 100876, China
[2]School of Computer, National University of Defense Technology, Changsha 410073, China
[3]School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

Abstract Cyber Range has become a very important means to support tasks such as network security technology validation, network weapon testing, training of network attack and defense, and network risk assessment etc. In this survey, we first give an overview of the current research works in the field of Cyber Range, including both domestic and international contributions; secondly, state-of-the-art techniques of Cyber Range is described, including large-scale network simulation, network traffic/service and user behavior simulation, acquisition and analysis of testing data, and system security and management etc.; finally, we concluded the paper by discussing the challenges and trends of Cyber Range.

Key words cyber range; network simulation; user behavior simulation; data acquisition and analysis; system security and management

1 引言

网络空间对抗形势日趋严峻, 网络攻防已成为各国网络攻防对抗的主要内容。网络环境已由单纯互联网发展到了泛在网络空间, 攻击方式也由单一模式朝着复杂的 APT 攻击方向发展。网络靶场是针对网络攻防演练和网络新技术评测的重要基础设施, 主要供政府和军队部门使用, 用来提高网络和信息系统的稳定性、安全性和性能[1]。世界各国均高度重视网络靶场建设, 将其作为支撑网络空间安全技术验证、网络武器试验、攻防对抗演练和网络风险评估的重要手段[2]。

网络靶场的主要功能包括: (1) 网络攻防武器评测验证: 新型网络攻防武器研制出来之后, 需要对其进行测试验证, 是否能有效攻破敌方防护系统, 以及是否能有效保护我方目标系统; 2)支持人员培训与竞演: 随着新型网络攻防武器的研发, 具体网络安全人员能否能有效掌握, 训练后, 谁掌握的技能更好; 3)科学试验和新技术验证: 网络空间科研人员研制出新的网络协议, 新型网络设备, 以及不同网络新技术, 在互联网上功能和性能如何, 也需要进行验证。

网络靶场已成为各国家进行网络空间安全研究、学习、测试、验证、演练等必不可少的网络空间安全核心基础设施。我国网络安全问题严重, 每年的经济损失达数百亿美元, 网络靶场研究成果如能很好的推广, 普惠网络安全相关企业及网民, 定可提高企业的核心竞争力及网民的安全意识与能力, 从而极大的减少经济损失。因此, 无论是从保证国家安全、维护社会稳定以及产业发展、减少经济损失

邀稿日期: 2016-03-28; 收稿日期: 2016-06-28

等网络靶场都具有广阔的应用前景, 具有很高的社会和经济效益。

目前, 关于网络靶场的研究已经取得了很多成果, 但仍处于探索阶段。对国内外关于网络靶场的研究进展进行较为全面的总结, 对未来网络靶场的深入研究具有重要意义。本文剩余部分组织入下, 第 2 节给出网络靶场国内外研究现状, 第 3 节给出网络靶相关技术研究进展, 第 4 节指出网络靶场面临的挑战与发展趋势, 第 5 节对全文进行总结。

2 靶场国内外研究现状

在网络靶场建设方面, 美国走在了世界的前列, 除了建成多个小型网络靶场外, 已开展国家级的网络靶场建设。2016 年 2 月 9 日, 美国白宫公布《网络安全国家行动计划》, 再次对美国网络基础设施、专业人才队伍、与企业合作等五个方面做全面提升, 提高美国在数字空间的安全。其他国家, 如英国等也正在建设自己的国家网络靶场。

我们将网络靶场的发展可以分为三个阶段, **第一阶段是以 21 世纪初期针对单独的木马类攻击武器而建立的实物高逼真型靶标时期**。在此阶段, 各国以敌方的靶标软硬件平台为目标, 建立尽可能逼真的靶标软硬件平台, 用于测试己方新研制的攻击武器能否成功绕过敌方的防护软件, 主要包括早期的蜜罐系统、木马测试系统等。**第二阶段是以 2005 年开始的小型虚拟化互联网靶场时期。**在此阶段, 云计算、软件定义网络等虚拟技术是该阶段的主流技术, 模拟真实的互联网攻防作战提供虚拟环境是各个国家的主要目标, 但模拟的互联网规模都比较小。主要包括: 2005 年美军联合参谋部组织建设的“联合信息作战靶场”(IOR) , 2009 年美国国防部国防高级研究计划局牵头建立的“国家网络靶场”(NCR)[3], 2010 年美军国防信息系统局组织建设的“国防部网络安全靶场”(GIG); 2010 年英国国防部正式启用了由诺•格公司研制的“网络安全试验靶场”; 此外, 日本的 StarBed 靶场系统, 加拿大的 CASELab 靶场系统, 英国的 SATURN 靶场系统以及台湾的 Testbed 测试平台等均属于这一阶段。**第三阶段是 2014 年开始的支撑泛在网的大型虚实结合网络空间靶场时期。**在此阶段, “震网”、“火焰”等针对工控网的新型网络攻击突现, 各国纷纷开始研究虚实结合的网络空间靶场技术。主要包括: 2014 年美国国家靶场增加了法拉第罩进行无线发射设备的测试, 并支持移动计算设备; 2014 年 6 月, 北大西洋公约组织在塔林建立 NATO 的网络靶场, 支持工控网的攻防测试; 2015 年 7 月, 欧洲防务署批准建立网络攻防测试靶场, 标志着 EDA 靶场工程的启动。为了保证国家安全, 为了加快向网络强国迈进的步伐, 为了在未来的网络战中占据有利的位置, 建立大型的网络靶场项目, 对网络靶场和网络战从理论和实践上进行深入研究, 具有重要的现实意义。

美国“国家网络靶场”项目(National Cyber Range, NCR)[4]。美军网络靶场是“曼哈顿计划”的五个组成部分之一。“国家网络靶场”项目由美国国防高级研究计划局(DARPA)负责组建, 通过构建可伸缩的互联网模型, 用来进行网络战争推演。“国家网络靶场”成为一种测试涉密与非涉密网络项目的国家资源。获得授权进行网络试验的政府及政府资助的测试组织可与“国家网络靶场”执行机构协调, 安排靶场时间与资源。该项目于 2009 年 1 月启动, 2011 年 10 月完成原型开发。这是一项多年计划, 由多个部门参加并分步骤实施, 其最终目的是保护美国的网络安全, 防止美国遭受敌对电子攻击, 并能对敌方展开在线攻击。具体包括网络攻防实验床 Emulab[5]、DETERlab[6]及 PlanetLab[7]。Emulab 是由犹他大学计算机学院 Flux 研究团队开发的一个网络实验床。该网络实验床由一定数量的计算机、服务器和路由器等硬件设施和一套专用的管理运营的软件系统组成。基于 Emulab 进行改进的 DETERlab, 将不同地理位置的 Emulab 平台进行网络集成的 PlanetLab[8], 此外, 还有美国惠普公司、英特尔公司和雅虎公司正在联合开发“全球云计算试验平台”等。

英国的网络靶场主要包括联邦网络实验靶场及 Breaking point 系统。英国联邦网络实验靶场(federated cyber test range)[9]由 Northrop Grumman 公司于 2010 年 10 月建立, 是英国第一个商业的网络实验平台。联邦网络实验靶场是将已有的网络靶场联邦而成的网络实验平台, 被用来模拟大型复杂网络, 并在安全可控的试验环境下进行基础设施生存能力和可靠性方面的网络试验及评估, 以评价它们对网络攻击的承受能力。Breaking point 是英国 Ixia 公司的网络靶场系统。它支持流量生成和仿真, 以创建一个互联网规模的网络靶场环境。Breaking point 中对互联网环境仿真包括目标仿真、漏洞仿真、逃避仿真和流量仿真。并且还包括互联网 IPV4 和 IPV6 基础架构、企业和 IT 服务、人口和国家用户群, 用于数据丢失预防(DLP)的相关数据、移动用户群等仿真。

日本提出了“星平台(StarBed)[10]”系统规划, 由日本情报通信研究机构(NICT)于 2002 年主导研制。StarBed 主要提供大规模的网络试验环境用于评估真

实场景下的新技术。目前已经发展到 StarBed 的第三个版本。第三代 StarBed 将研究范围扩大, 研究领域扩展到了安全性和服务质量, 复杂的有线无线网络的扩展和构建信息安全物理系统的方法, 提供软件实现以实现大规模网络仿真。截止到 2014 年, 在用的实验组总节点数达到 1398 个/11.7K 个核, 以及 60TB 的存储。

加拿大国家仿真实验室(CASELab)也建立了相应的项目开展类似的工作, 由维多利亚大学计划开发。其建立的试验平台提供云计算、大规模网络安全和保密等领域的核心研究能力, 并为研究人员提供系统分析和仿真工具, 使他们在完全可重复的实验条件下对真实世界大规模网络系统的行为建模, 从而支持对互联网的新技术(武器)的鉴定和评估。

国内在网络靶场建设工作方面, 科研院所方面主要有国防科技大学、中科院计算所、CNCERT/CC、中科院信工所、中国电子科技集团、哈尔滨工业大学、北京邮电大学均建设了自己的网络靶场, 在大规模网络仿真、大规模网络攻击行为场景仿真, 网络攻击数据采集与安全效果评估, 以及系统安全管理等关键技术方面进行了突破和技术积累, 公司方面, 合天网安实验室和国内 i 春秋也建设了网络试验培训平台和平台。其中, 合天网安实验室开发的互联网教学靶场, 用户遍及 300 余高校 3 万余人, 并举办了 XP 挑战赛、暴恐音视频挑战赛、强网杯挑战赛等全国性系列网络安全竞赛等。

3 相关技术研究进展

网络靶场涉及大规模网络仿真、网络流量/服务与用户行为模拟、试验数据采集与评估、系统安全与管理等多项复杂的理论和技术, 是一个复杂的综合系统。

3.1 大规模网络仿真

在大规模网络仿真方面, 主要包括模型模拟和虚拟化两种方法。在模型模拟方面, 代表性工作有 UC Berkeley 大学开发的基于并行离散事件的网络模拟器(Network Simulator, version 2, NS2), 尽管能实现超大规模网络的构建[11], 但难以保证网络节点的逼真度以及用户行为复制的逼真度。因此, 以虚拟化为基础的网络仿真成为了主流, 虚拟化技术又分为节点虚拟化和链路虚拟化两方面。在节点虚拟化方面, 作为云计算平台中最具代表性的 Openstack[15], 其基于虚拟网桥实现宿主机内部的链路仿真, 实现虚拟机间的互联互通。在轻量级的节点虚拟化方面, 最具有代表性的是 docker[30], 这是一种基于 linux container(LXC)的技术, 一个容器就相当于一个拥有一个应用的虚拟机, 开发者可以在上面操作而不会影响到整个下层系统。美国空军技术学院基于操作系统级虚拟化以及全虚拟化技术实现了大规模、高逼真度网络节点仿真平台并用于网络安全训练[14]。在链路虚拟化两方面, 作为网络仿真平台的代表 Emulab[4], 其基于 Dummynet, 通过协议栈的方式拦截数据包, 并通过一个或多个管道模拟带宽、传播时延、丢包率等链路特性, 具有较高宿主机内部的链路仿真逼真度。网络功能虚拟化技术(NVF)[16]通过通用性硬件以及虚拟化技术实现网元(路由器、交换机等)虚拟化以及网元间连接的虚拟化, 但缺乏对网络链路参数的仿真。软件定义网络技术(SDN)[17]主要是基于数据层与控制层的分离, 在整个网络架构上提供网络虚拟化和自动化的配置, 为新的网络服务提供快速部署, 其网络的灵活构建与快速部署可为网络仿真提供基础, 但 SDN 的研究目标并不是网络仿真, 对传统网络的链路参数、路由路径的仿真有待研究。基于网络模拟和虚拟化技术各自的优缺点, 美国伊利诺伊大学香槟分校和佛罗里达国际大学整合了两种技术, 形成了基于虚拟机以及模拟器的融合仿真[12, 13]。

在大规模虚拟网络快速部署方面, 分为主要包括 3 类[18]: 基于镜像启动的方法, 基于内存拷贝的方法和轻量级的虚拟化技术部署。基于镜像启动是虚拟机部署方法中最普遍的一种方法, 主要工作集中在镜像管理, 镜像格式的升级, 镜像传输优化, 镜像存储等方面。普渡大学的 K.R.Jayaram 等人[19]在 2011 年提出来在 IaaS 环境下, 底层虚拟机镜像的相似度很高, 它们很多的数据块之间的内容都是相似的。加利福尼亚大学的 Peng[20]在基于镜像相似问题基础上, 对镜像进行了块划分, 提出了一个基于块级的镜像分布系统 VDN。Zhang 等人[18]提出了一个镜像管理部署系统 VMThunder 中对于镜像之间的相似性, 提出了一种按需获取镜像内容, 而不是整块镜像的传输。镜像启动的另一个关键技术就是镜像格式的优化。威廉玛丽学院的 Duy Le[21]对虚拟环境下的镜像文件系统进行细致分析, 对于原始镜像 Raw 格式, 它保留了物理磁盘或文件上的比特图, 从而不需要进行地址翻译等工作。典型的基于增量的镜像是 qcow 和 qcow2[22], 它通过不断占用磁盘剩余空间来提供需求容量。利用了“copy on write”策略[23], 为多重访问提供不同的版本, 且提供回滚操作, 初始磁盘大小也比较小。国防科技大学的陈斌[24]等人提出了一种虚拟机镜像按需获取技术, 对镜像进行了细粒度的分割。高效的传输机制可以减少部署的时间,

云计算早期, 亚马逊 EC2 toolbox rocks 来传输镜像[25]。Peer-to-peer[26]方法在镜像传输方面就比较高效, 而且镜像在传输之前也可以被压缩或者分割。东田纳西州立大学的 THOMAS MORGAN JR 等人利用一些网络协议来满足无磁盘的远程启动与部署, 通过共享的一个存储池[27], 只需要一个网络连接就可以利用里面的存储资源而不需要传输镜像。内存拷贝的方法多用于正在运行的虚拟机。多伦多大学的 H. Andrés Lagar-Cavilla 等人[28]提出了基于内存拷贝方法的虚拟机快速部署 SnowFlock, 达到了在秒级部署的任务。北京大学信息科学技术学院的 Zhu 等人[29]提出了一种基于之前存储的虚拟机快照来实现快速启动的 Twinkle, 从而实现秒级启动。

3.2 网络流量/服务和用户行为模拟

在网络流量行为模拟方面, 主要集中在流量模型的建立、预测与回放等方面。1997 年, 贝尔实验室的 Willinger 提出了具有重尾分布周期的 ON/OFF 模型[31]。1998 年, 美国马里兰大学的 Krunz 提出了服务时间分布无穷方差的 M/G/排队模型等[32]。2011 年, 瑞典乌普萨拉大学 Dombry 等人研究了高速通信链路中数据流量的传输模式, 结合重尾分布与 ON/OFF 模型, 证明了 ON/OFF 源的数量与时间尺度均趋于无穷时, 流量数据的行为模式近似于分数泊松运动[33]。2001 年, 美国莱斯大学 Sarvotham 等人给出了一种α-β-ON/OFF 模型, 解释了网络流量具有突发性和长相关性的原因, 将网络流量的特性与用户的行为联系起来[34]。2004 年, 加利福尼亚大学 Cheng 和 Google 共同开发出来用于测试服务器端性能的流量回放系统 Monkey, Monkey see 用于在服务器端一侧捕获服务器与客户端交互的流量, 而 Monkey do 用于模拟客户端和传输网络行为[35]。2009 年, 日本早稻田大学 PHAM 研究了大规模网络流量并行回放中的流分割和回放质量评估问题, 但局限于对捕获流量进行单向回放研究[36]。2013 年, 南加州大信息科学研究所 Hussain 使用流量分析工具 LANDER 对真实网络中的流量进行了捕获的分析, 并通过在 DeterLab 实验床中开发了一个代理, 实现了真实网络攻击流量的回放。实验中, 作者使用随机 Web 访问流量作为非恶意流量, 并与真实网络攻击流量合成, 完成了小规模的网络模拟实验[37]。2008 年, 马来西亚多媒体大学的 Lim S C 等人从柯西过程角度讨论了网络流量建模方法, 给出了一种十分灵活的描述多重分形性质的模型, 该模型可以同时精确的刻画流量的短相关和长相关性质[38]。2008 年, 加拿大魁北克大学蒙特利尔分校的 Zhani 等人通过对实际网络流量数据的分析, 给出了一种 Training-based 模型[39], 并研究了模型性能与模型参数间的关系, 预测结果比较令人满意。

在网络用户行为模拟方面, 2011 年, 智利大学工业工程系学者 Loyola 利用蚁群优化算法对 web 用户的浏览行为进行建模, 解决了传统 web 挖掘方法与模型适应性不强的问题[40], 该方法缺点是偏好模型比较单一, 训练过程较慢, 不适合大规模网络用户行为分析。在 2013 年, 哥伦比亚大学学者 Song 采用生物特征提出了一种基于机器学习的用户行为生物识别方法, 该论文是在系统级别用户行为生物特征识别方面最早的论文, 通过提炼典型特征并采用高斯混合模型对每个用户的特征进行训练，实验结果与 SVM 相比提高了 17.6%[41]。, 其缺点是仅在 windows 测试环境下进行了测试, 且未给出误识率和拒识率。在 2015 年, 德国哈索•普拉特纳研究所学者 Amirkhanyan 研究了一种用户行为状态图对用户行为进行描述和表示方法, 实现了通过设计目标场景和人工合成活动产生真实用户行为数据的方法[42]。但该文献未给出用户行为来源, 且作者提出的构建方法仅限于模拟简单的用户行为。

3.3 试验平台采集与效果评估

试验数据采集分成物理数据采集和虚拟化数据采集。因为前者的方法、技术、工具都相对成熟, 所以本节主要分析虚拟化数据采集技术。即在体系结构栈的硬件层和操作系统层之间加入一个新层次--虚拟层(hypervisor), 为单一物理机上提供同时运行多个相互独立的操作系统, 并已成为当今云计算和数据中心的支撑。总结起来, 虚拟化数据采集可进一步分成带内数据采集和带外数据采集两种技术路线。在带内数据采集方式中, 典型的做法是以基于主机的入侵检测系统为主[43], 由中心采集程序和植入虚拟机的代理程序组成。该方式具有被攻击的风险, 抗破坏能力差。而一个理想的数据采集系统应当既具备对监控对象全面彻底的观察能力, 也具备健壮的自身保护机制, 因此带外方式被广泛认可是网络空间安全试验靶场数据采集的有效路线。

在带外数据采集方式中, 2003 年, 美国斯坦福大学的 Garfinkel 等[44]首次提出了虚拟机内省 VMI (virtual machine introspection)技术, 将数据采集方式移到了带外, 该方式减轻了直接攻击 IDS 的风险, 但是数据采集引入的性能代价较大。2007 年, 美国佐治亚理工学院的 Payne 等[45]设计的 libvmi, 不需要对虚拟机监视器进行修改, 而是直接利用虚拟机监视器提供的接口, 这种方式需要虚拟机监视器的支持,

引入的性能代价小于 5%。但是如果虚拟机操作系统内核数据结构发生变化, 这种方法采集的数据即会出错。2008 年, 美国佐治亚理工学院的 Dinaburg 等[46]设计的 Ether, 用于恶意软件分析, 这种方式能够较好地拦截内核数据。2013 年, 美国犹他大学的 Burtsev[47]设计了记录和重放系统 XenTT, 该系统支持透明的虚拟机记录方式, 而且可以对系统执行历史和系统状态进行分析。记录程序位于虚拟机监视器中, 记录虚拟机内发生的中断和异常事件、CPU 状态、指令等底层二进制数据。但是该方式的数据采集需要将底层二进制重构成高层语义, 实现较为困难[48]。2014 年美国国防部 DARPA 将 VMI 技术作为 Cyber Fast Track program 项目一部分, 并以美国 MIT Lincoln 实验室为载体, 集结相关技术专家和工程师, 形成 Panda 等为代表的系统, 从而标志带外数据采集技术在网络靶场开始得到实际使用。

在试验数据分析评估方面, 主要是基于试验运行采集到的数据, 根据一定的评估标准和模型, 对被测的攻防武器或技术进行定量与定性相结合的效果评估, 以及网络攻防对抗态势评估分析与可视化, 并尽可能保证评估的可操作性和客观性。分析评估方法主要有三类, 即基于数学模型的方法、基于指标体系的方法和基于知识推理的方法。在基于数学模型的方法中, Tim Bass 于 2000 年提出的基于多传感器的入侵检测框架[49]是对基于网络安全态势感知的分析评估模型; 2002 年美国国防部联合指挥实验室提出了 JDL 数据融合处理模型[50], 将试验数据进行预处理、融合、精炼和评估; 2012 年 SA Technologies 的 M.R.Endsley 提出了态势感知理论 SA 模型[51], 对影响网络安全态势的要素进行理解、评估和预测; 基于指标体系的方法中, 以 20 世纪 70 年代美国运筹学家 T.L.Saaty 教授提出的层次式指标体系分析法[52]最为广泛, 是安全态势评估领域最常用的评估方法, 其评估函数通常由网络安全指标及其重要性权重共同确定; 在基于知识推理的方法中, 2006 年挪威约维克大学 Arnes 等人提出了基于隐马尔科夫推理的网络安全态势评估模型[53]; 2007 年挪威科学技术大学的 Mehta 等人[54]提出了一种基于攻击图状态的排序方案, 通过对状态的排序反映出安全状态的重要性进行推理, 实现对网络安全态势的评估; 2010 年美国乔治梅森大学的 Noel 等人[55]利用攻击图来理解攻击者如何借助网络漏洞一步步来实施攻击推理, 通过模拟增量式的网络渗透攻击和攻击在网络中传播的可能性来衡量这个网络系统的安全性。

3.4 试验平台安全及管理

在试验平台安全技术方面, 主要包括虚拟机安全隔离、虚拟网络隔离和试验平台隔离方面, 从虚拟机内核、内存、存储、监控器、网络流量、系统平台等各个层次研究试验平台的安全技术。虚拟机安全方面 XEN 和 KVM 有不同的方法, XEN 通过修改操作系统特权级、内存分段保护机制、分离设备驱动模型等实现安全隔离[56], KVM 通过 CPU 的绑定设置、修改、优化 KQEMU 源代码、影子页表法、硬件辅助的虚拟化内存等实现安全隔离[57]。根据现有的带宽隔离策略是否基于本地交换机或链路, 可以分为本地策略和端对端策略。VLANs[58]和 802.1p[59]服务类型标签(CoS Tags)是以太网提供的分割不同用户和类型流量的机制, 属于本地策略。端对端策略在端节点维护速率控制状态, 因而更加灵活、扩展性更强。端对端策略还可以针对单个流进行调整, 而不影响其他流, 因而更加准确。试验平台隔离的相关研究主要涉及虚拟机用户恶意行为监控与记录、基于平台配置的安全管理、恶意行为安全取证、安全审计和追责四项关键技术, 2010 年美国北卡罗莱纳州立大学的 Jiang X 等设计了 VMwatcher[60], 可以实现对文件系统和进程等虚拟机状态进行行为监测。在工业界, VMware 依据其虚拟化平台 vSphere 的配置选项, 设计 vSphere 云平台安全配置的安全加固文档, 以确保 VMware vCenter Server 和 VMware ESXi 的 vSphere 环境安全。2008 年美国阿拉斯加大学 Nance 开发出 VIX 工具包[61], 将监控系统部署在 Xen 的特权域 domain0, domain0 拥有对所有资源的访问权限。2002 年美国密歇根大学研发的 Revirt 系统[62]在半虚拟化环境下实现, 通过在虚拟机前端和 Domain0 后端作中介获得共享请求数据, 记录下来以重放时间和外部中断等不确定性事件。2012 年美国雪域大学 Yan 等人设计的 V2E [63]把虚拟机系统分为两个范围(main realm 和 recording realm)。恶意软件运行在 recording realm 中, 虚拟机系统的剩下部分仍然在 main realm 中。通过从记录器获得的记录日志, 然后放入重放器进行重放。然而, V2E 需要进行指令级记录和翻译, 所以性能开销很大。

在试验任务运行控制与管理方面, 主要研究集中于试验任务的自动化配置、试验运行控制以及网络仿真系统协同融合控制。在试验运行控制方面, 传统的并行离散事件模拟技术(PDES) [64]基于同步技术, 可实现试验任务运行时钟的可控性与因果性, 但仅限于离散事件模拟; 美国伊利诺伊大学香槟分校提出了一种离散事件模拟与虚拟机的时钟同步与控制

技术[12], 可为试验运行的灵活时钟控制提供支撑。在试验任务自动化配置方面, 以 NS3[65]、Emulab[66]为代表的开源网络模拟与仿真软件根据用户提交的网络配置文件, 可自动构建仿真环境; 以 OPNET[67]、QualNet[68]为代表的商用网络模拟与仿真软件可为用户提供可视化配置界面。

4 面临的挑战与发展趋势

综上所述, 网络空间靶场的关键技术面临的挑战与发展趋势具体如下: (1)大规模网络仿真方面, 面临的挑战为含人、物、信息的虚实互联网络靶场的灵活快速构建问题, 主要发展趋势为物理集群网络拓扑透明的大规模任意拓扑及特征的虚拟网络生成、高逼真度的数据报文转发与链路复现及自适配的大规模虚拟网络快速构建等技术, 通过镜像文件的存储优化、传输优化以及网络感知的大规模资源调度等方法, 以实现万级规模网络拓扑、特征的快速复现及自动配置; (2)在网络流量/服务和用户行为模拟方面, 面临的挑战为如何解决面向攻击的自适应的网络空间环境的逼真模拟仿真问题, 主要发展趋势为场景化的网络行为逼真模拟技术, 通过多层级融合网络流量行为模拟、基于时序确保的网络应用逼真模拟、大规模服务交互行为模拟、网络终端用户行为模拟等技术, 以达到场景化的多层级、全方位综合互联网行为逼真模拟效果; (3)在试验数据采集与评估方面, 面临的挑战为低损、实时、准确的网络攻防评估和分析问题, 主要发展趋势为低损耗的靶场信息实时采集技术, 主要包括虚拟机与虚拟机监视器配合的低损实时采集、大规模试验数据的订阅/分发、多模态试验数据流的存储与管理等技术, 以实现对网络安全试验过程信息的低损、实时、准确采集; 在试验效果评估方面, 其发展趋势研究可量化的攻防效果评估指标体系、可伸缩的实时绩效评估计算模型、支持可反馈的攻防武器量化评估自适应机制等; (4)在试验平台安全及管理方面, 面临的挑战为如何保证整个靶场平台的安全及试验安全隔离问题, 发展趋势为多层次动态隔离的安全管控体系, 发展趋势为高效、灵活、可控的虚实资源分配与隔离管控机制, 实现高安全、高可靠的联合试验环境。

5 结束语

网络靶场还可为我国军队及相关安全部门研究网络攻防技术、进行网络攻防试验、验证攻防工具效果、攻防演练对抗等提供平台支持; 同时可为我国培养网络安全的高层次技术人才提供学习平台、技术支撑平台、培训平台、攻防技术课程体系等, 从而为国家重大网络安全决策提供重要依据, 为维护我国网络主权提供技术和能力支撑。因此, 为了保证国家安全, 为了加快向网络强国迈进的步伐, 为了在未来的网络战中占据有利的位置, 建立大型的网络靶场项目, 对网络靶场和网络战从理论和实践上进行深入研究, 具有重要的现实意义。

致 谢 感谢在本文成文过程中, 哈尔滨工业大学(威海)的王佰玲, 江南大学的王晓峰, 哈尔滨工业大学(深圳)的刘川意, 西安电子科技大学的朱辉和国防科技大学的江荣所作出的贡献。

参考文献

[1] What is a Cyber Range? - Definition from Techopedia https://www.techopedia.com/definition/28613/cyber-range.

[2] Davis J, Magrath S. A survey of cyber ranges and testbeds[R]. DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION EDINBURGH (AUSTRALIA) CYBER AND ELECTRONIC WARFARE DIV, 2013.

[3] Ranka J. National Cyber Range[R]. DEFENSE ADVANCED RESEARCH PROJECTS AGENCY ARLINGTON VA STRATEGIC TECHNOLOGY OFFICE (STO), 2011.

[4] Pridmore L, Lardieri P, Hollister R. National Cyber Range (NCR) automated test tools: Implications and application to network-centric support tools[C]//AUTOTESTCON, 2010 IEEE. IEEE, 2010: 1-4.

[5] Hibler M, Ricci R, Stoller L, et al. Large-scale Virtualization in the Emulab Network Testbed[C]//USENIX Annual Technical Conference. 2008: 113-128.

[6] Mirkovic J, Benzel T. Teaching cybersecurity with DeterLab[J]. Security & Privacy, IEEE, 2012, 10(1): 73-76.

[7] Chun B, Culler D, Roscoe T, et al. Planetlab: an overlay testbed for broad-coverage services[J]. ACM SIGCOMM Computer Communication Review, 2003, 33(3): 3-12.

[8] Laih C S, Li J S, Lin M J, et al. The Development and Operation of Testbed@ TWISC[C]//Proceedings of the 3rd Joint Workshop on Information Security. 2008: 532-546.

[9] Winter H. System security assessment using a cyber range[C]// System Safety, incorporating the Cyber Security Conference 2012, 7th IET International Conference on. IET, 2012: 1-5.

[10] MIYACHI T, MIWA S, HASEGAWA S, et al. Hands-on Environments for Network Technologies on StarBED[J]. Educational technology research, 2011, 34(1): 107-118.

[11] Liu N, Carothers C, Cope J, et al. Model and simulation of exascale communication networks[J]. Journal of Simulation, 2012, 6(4): 227-236.

[12] D Jin, Y Zheng, DM Nicol. A parallel network simulation and virtual time-based network emulation testbed [J]. Journal of Simulation. 2014, 8(8): 206-214.

[13] Miguel A. Erazo, Jason Liu. Leveraging symbiotic relationship between simulation and emulation for scalable network experimentation [A]. ACM SIGSIM Conference on Principles of Advanced Discrete Simulation[C]. 2013: 79-90.

[14] Todd R. Andel, Kyle E. Stewart, Jeffrey W. Humphries. Using Virtualization for Cyber Security Education and Experimentation [C]. 14th Colloquium for Information Systems Security Education. 2010: 130-136.

[15] OpenStack Community. http: //www.openstack.org/

[16] Mijumbi R, Serrat J, Gorricho J L, et al. Network function virtualization: State-of-the-art and research challenges[J]. IEEE Communications Surveys & Tutorials. 2016, 18(1): 236-262.

[17] Jammal M, Singh T, Shami A, et al. Software defined networking: State of the art and research challenges[J]. Computer Networks, 2014, 72: 74-98.

[18] Zhang Z, Li D, Wu K. Large-scale virtual machines provisioning in clouds: challenges and approaches[J]. Frontiers of Computer Science, 2016, 10(1): 2-18.

[19] Jayaram K R, Peng C, Zhang Z, et al. An empirical analysis of similarity in virtual machine images[C]//Proceedings of the Middleware 2011 Industry Track Workshop. ACM, 2011: 6.

[20] Peng C, Kim M, Zhang Z, et al. VDN: Virtual machine image distribution network for cloud data centers[C]//INFOCOM, 2012 Proceedings IEEE. IEEE, 2012: 181-189.

[21] Le D, Huang H, Wang H. Understanding performance implications of nested file systems in a virtualized environment[C]//FAST. 2012: 8.

[22] Bellard F. QEMU, a Fast and Portable Dynamic Translator[C]// USENIX Annual Technical Conference, FREENIX Track. 2005: 41-46.

[23] Xiao W, Liu Y, Yang Q, et al. Implementation and performance evaluation of two snapshot methods on iSCSI target storages[C]//Proc. of NASA/IEEE Conference on Mass Storage Systems and Technologies. 2006.

[24] 陈彬. 分布环境下虚拟机按需部署关键技术研究[D].国防科学技术大学, 2010.

[25] Papadopoulos P M. Extending clusters to Amazon EC2 using the Rocks toolkit[J]. International Journal of High Performance Computing Applications, 2011, 25(3): 317-327.

[26] Li D, Cao J, Lu X, et al. Efficient range query processing in peer-to-peer systems[J]. Knowledge and Data Engineering, IEEE Transactions on, 2009, 21(1): 78-91.

[27] Morgan Jr T. DRBL: Diskless Remote Boot in Linux[J]. NETWORK, 2006, 192: 100.0.

[28] Lagar-Cavilla H A, Whitney J A, Scannell A M, et al. SnowFlock: rapid virtual machine cloning for cloud computing[C]//Proceedings of the 4th ACM European conference on Computer systems. ACM, 2009: 1-12.

[29] Zhu J, Jiang Z, Xiao Z. Twinkle: A fast resource provisioning mechanism for internet services[C]//INFOCOM, 2011 Proceedings IEEE. IEEE, 2011: 802-810.

[30] Merkel D. Docker: lightweight linux containers for consistent development and deployment[J]. Linux Journal, 2014, 2014(239): 2.

[31] WILLINGER W, TAQQU M S, SHERMAN R, et al. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level[J]. IEEE/ACM Transactions on Networking (ToN), 1997, 5(1): 71-86.

[32] KRUNZ M M, MAKOWSKI A M. Modeling video traffic using M/G/∞ input processes: a compromise between Markovian and LRD models[J]. Selected Areas in Communications, IEEE Journal on, 1998, 16(5): 733-748.

[33] DOMBRY C, KAJ I. The on–off network traffic model under intermediate scaling[J]. Queueing systems, 2011, 69(1): 29-44.

[34] SARVOTHAM S, RIEDI R, BARANIUK R. Network traffic analysis and modeling at the connection level[C]// Proceedings IEEE/ACM SIGCOMM Internet Measurement Workshop, c2001.

[35] CHENG Y, HÖLZLE U, CARDWELL N, et al. Monkey See, Monkey Do: A Tool for TCP Tracing and Replaying[C]//USENIX Annual Technical Conference, General Track. c2004: 87-98.

[36] PHAM VAN D, ZHANIKEEV M, TANAKA Y. Effective high speed traffic replay based on IP space[C]//Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on. IEEE, c2009, 1: 151-156.

[37] A HUSSAIN, Y PRADKIN, J HEIDEMANN, Replay of malicious traffic in network testbeds[C]//Technologies for Homeland Security (HST), 2013 IEEE International Conference on. IEEE, c2013: 322-327

[38] LI M, LIM S C. Modeling network traffic using generalized Cauchy process[J]. Physica A: Statistical Mechanics and its Applications, 2008, 387(11): 2584-2594.

[39] ZHANI M F, ELBIAZE H, KAMOUN F. Analysis of prediction performance of training-based models using real network traffic.[J]. International Journal of Computer Applications in Technology, 2008, 37(1): 10-19.

[40] LOYOLA P, ROMÁN P E, VELÁSQUEZ J D. Clustering-based learning approach for ant colony optimization model to simulate web user behavior[C]//Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01. IEEE Computer Society, c 2011: 457-464.

[41] SONG Y, BEN SALEM M, HERSHKOP S, et al. System level

user behavior biometrics using fisher features and gaussian mixture models[C]//Security and Privacy Workshops (SPW). IEEE, c2013: 52-59.

[42] AMIRKHANYAN A, SAPEGIN A, GAWRON M, et al. Simulation user behavior on a security testbed using user behavior states graph[C]//Proceedings of the 8th International Conference on Security of Information and Networks. ACM, c2015: 217-223.

[43] Daniels T E, Spafford E H. A Network Audit System for Host-based Intrusion Detection (NASHID) in Linux [A]. // Proceedings of the Computer Security Applications, ACSAC '00, 2000: 178-187.

[44] Garfinkel T, Rosenblum M. A virtual machine introspection based architecture for intrusion detection [A]. // Proceedings of the Network and Distributed System Security Symposium [C]. San Diego, USA, 2003.

[45] Payne B D, Carbone M D P de A, Lee W. Secure and flexible monitoring of virtual machines [A]. // Proceedings of the 23rd Annual Computer Security Applications Conference [C], 2007: 385-397.

[46] Dinaburg A, Royal P, Sharif M, et al. Ether: Malware Analysis via Hardware Virtualization Extensions [A]. // Proceedings of the CCS'08 [C], 2008: 51-62.

[47] Burtsev, A. Deterministic systems analysis. Doctoral dissertation [D], The University of Utah, 2013.

[48] Bauman E, Ayoade G, Lin Z. A Survey on Hypervisor-Based Monitoring: Approaches, Applications, and Evolutions [J]. ACM Computing Surveys, Vol. 48, No. 1, 2015: Article 10: 1-33.

[49] Bass T. Multisensor Data Fusion for Next Generation Distributed Intrusion Detection Systems [C]. In Proceedings of the IRIS National Symposium on Sensor and Data Fusion, May 24-28, . 1999: 24-27.

[50] Blasch E P, Plano S. JDL level 5 fusion model: user refinement issues and applications in group tracking [J]. Proceedings of SPIE. 2002, 4729 (May 2012): 270-279.

[51] Endsley M. Situation awareness global assessment technique (SAGAT) [C]. In Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, May 23-27, Dayton, OH. 1988: 789-795.

[52] Dagdeviren M, Yüksel . Developing a fuzzy analytic hierarchy process (AHP) model for behavior-based safety management [J]. Information Sciences. 2008, 178 (6): 1717–1733.

[53] Årnes A, Valeur F, Vigna G, et al. Using hidden markov models to evaluate the risks of intrusions [C]. In Recent Advances in Intrusion Detection: 145–164. Recent Advances in Intrusion Detection.

[54] Mehta V, Bartzis C, Zhu H, et al. Ranking Attack Graphs [C]. In Proceedings of the 9th International Symposium On Recent Advances in Intrusion Detection (RAID), September 20-22, Hamburg, Germany. 2006: 127–144.

[55] Noel S, Jajodia S, Wang L, et al. Measuring security risk of networks using attack graphs [J]. International Journal of Next-Generation Computing. 2010, 1 (1).

[56] 王雅超, 黄泽刚. 云计算中 XEN 虚拟机安全隔离相关技术综述[J]. 信息安全与通信保密, 2015(6): 85-87.

[57] 黄煜, 罗省贤. KVM 虚拟化技术中处理器隔离的实现[J]. 计算机系统应用, 2012, 21(1): 179-182.

[58] 802.1Q - Virtual LANs [EB/OL]. http://www.ieee802.org/1/pages/802.1Q.html.

[59] Ek N. IEEE 802.1 P, Q - Qo S on the MAC level [EB/OL]. http://www.tml.tkk.fi/Opinnot/Tik-110.551/1999/papers/08IEEE802.1Qos In MAC/qos.html.

[60] Jiang X, Wang X, Xu D. Stealthy malware detection and monitoring through VMM-based "out-of-the-box" semantic view reconstruction.[J]. Acm Transactions on Information & System Security, 2010, 13(2): 128-138.

[61] Nance K, Hay B, Bishop M. Virtual Machine Introspection: Observation or Interference? [J]. IEEE Security & Privacy, 2008: 32-37.

[62] Dunlap G W, King S T, Cinar S, et al. ReVirt: Enabling Intrusion Analysis through Virtual-Machine Logging and Replay [A]. // Proceedings of the 5th Symposium on Operating Systems Design and Implementation [C], 2002.

[63] Yan L, Jayachandra M, Zhang M, et al. V2E: Combining Hardware Virtualization and Software Emulation for Transparent and Extensible Malware Analysis [A]. // Proceedings of the VEE'12 [C], 2012: 227-237.

[64] R. M. Fujimoto. Parallel Discrete Event Simulation. Communications of ACM. 1990, 33(10): 30~53

[65] NS3 https: //www.nsnam.org/

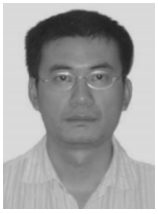
[66] Brian White, Jay Lepreau, Leigh Stoller, Robert Ricci, Shashi Guruprasad, Mac Newbold, Mike Hibler, Chad Barb, Abhijeet Joglekar. An integrated experimental environment for distributed systems and networks[J]. ACM SIGOPS Operating Systems Review. 2002, 36(SI): 255-270.

[67] OPNET http: //www.opnet.com/

[68] QualNet http://scalable-networks.com

方滨兴 于 1989 年在哈尔滨工业大学计算机系获得博士学位, 现任中国网络空间安全协会理事长, 中国工程院院士。研究领域为大数据、计算机网络和信息安全。Email: fangbx@cae.cn

贾焰 于 2000 年在国防科学技术大学获得计算机软件与理论博士学位, 现任国防科学技术大学教授, 研究领域为大数据、网络信息安全和社交网络。Email: jiayanjy@vip.sina.com



李爱平 于 2004 年在国防科技大学计算机软件与理论专业获得博士学位。现任国防科技大学计算机学院研究员。研究领域为网络安全、大数据分析等领域。Email: liaiping@nudt.edu.cn

张伟哲 于 2006 年在哈尔滨工业大学计算机系统结构专业获得博士学位。现任哈尔滨工业大学计算机网络与信息安全研究中心任教授。研究领域为网络空间安全、并行与分布式系统。研究兴趣包括: 虚拟化、资源管理。Email: wzzhang@hit.edu.cn