第47卷 第10期 计 算 机 工 程 2021年10月
Vol.47 No.10 Computer Engineering October 2021

·网络空间安全·

文章编号：1000-3428(2021)10-0103-08 文献标志码：A 中图分类号：TP309

基于交互式流量回放的用户行为仿真技术

黄 宁，刘 渊，王晓锋

（江南大学 人工智能与计算机学院，江苏 无锡 214122）

摘 要：流量回放可为网络靶场提供逼真的流量数据并支持网络新技术验证与安全评测。面向复杂虚拟网络的交互式用户行为仿真需求，设计一种交互式流量链路的用户行为仿真架构。采用基于云平台的分布式流量仿真策略，以实现面向复杂虚拟网络用户的行为仿真多样化和可扩展加载。对交互式流量回放过程中延时修复与补偿策略进行研究，提升交互式用户行为仿真的时序逼真性。仿真实验结果表明，该仿真架构能够在保证流量时序准确性的前提下，实现交互式的大规模用户行为仿真，与传统的ITRM、Tcpreplay等方法相比，在仿真行为的多样性、规模性、逼真性上具有一定优势，可为安全评测提供有效支撑。

关键词：网络靶场；交互式流量回放；流量仿真；虚拟目标网络；大规模用户行为仿真

开放科学（资源服务）标志码（OSID）：

中文引用格式：黄宁，刘渊，王晓锋. 基于交互式流量回放的用户行为仿真技术[J]. 计算机工程，2021，47(10)：103-110.

英文引用格式：HUANG N，LIU Y，WANG X F. User behavior emulation technology based on interactive traffic replay[J]. Computer Engineering，2021，47(10)：103-110.

User Behavior Emulation Technology Based on Interactive Traffic Replay

HUANG Ning，LIU Yuan，WANG Xiaofeng

（School of Artificial Intelligence and Computer Science，Jiangnan University，Wuxi，Jiangsu 214122，China）

【Abstract】Traffic replay can provide realistic traffic data for the cyber range，and support new technology verification and network security evaluation. To meet the needs of interactive user behavior simulation for complex virtual networks，an architecture for user behavior emulation is designed based on interactive traffic links. The architecture adopts a distributed traffic emulation strategy based on cloud platform to achieve diversified and scalable loading of user behavior emulation for complex target networks. The delay repair and compensation strategy in the process of traffic replay is further studied to improve the timing fidelity of interactive user behavior emulation. Results of emulation experiments show that this method can realize interactive large-scale user behavior emulation with the accuracy of traffic timing ensured. It has certain advantages in the diversity，scale and fidelity of behavior emulation over traditional methods such as ITRM and Tcpreplay，providing effective support for security evaluation.

【Key words】cyber range；interactive traffic replay；traffic emulation；virtual target network；large-scale user behavior emulation

DOI：10. 19678/j. issn. 1000-3428. 0059355

0 概述

随着网络安全形势日趋严峻，网络靶场[1]已经成为支撑网络空间安全技术研究、攻防对抗试验的重要基础设施，为大规模网络用户行为仿真提供灵活逼真的仿真环境。同时网络用户行为仿真作为网络靶场的重要组成部分，也为靶场平台提供了注入背景流量、构建真实网络场景的有效手段，对于真实网络场景重现、网络安全态势评估、网络安全实验等方面研究具有重要意义。

网络用户行为仿真[2]的关键目标在于如何能够真实地还原现实网络用户行为，主要体现在仿真的逼真性及规模性。传统的网络用户行为主要通过对单个行为特征分析建立通用仿真模型，然后在NS3[3]、OPNET[4]等网络仿真软件上进行模拟，仿真的用户行为类型较为单一，无法体现网络用户行为多样性的特点，仿真的逼真度不高，且NS3等软件仿真的规模有限，无法支撑大规模高逼真的网络用户

基金项目：国家重点研发计划（2016YFB0800305）；国家自然科学基金（61672264，61972182）。

作者简介：黄 宁（1994—），男，硕士研究生，主研方向为网络安全、流量仿真；刘 渊，教授、博士生导师；王晓锋，副教授、博士。

收稿日期：2020-08-25 **修回日期：**2020-09-27 **E-mail：**6181611017@stu.jiangnan.edu.cn

行为仿真。流量回放近年来在网络安全、网络测试评估[5]等领域越发受到关注,通过流量回放的方式重现真实网络场景,对目标设备进行评估。随着云平台[6]兴起,可将回放目标扩大为整个虚拟网络,使得通过流量回放的方式在虚拟网路中复现真实网络场景,将网络流量作为用户行为仿真的载体,向虚拟目标网络注入背景流量的方案变得可行。

本文结合真实流量回放的方式进行大规模高逼真网络用户行为仿真,向虚拟目标网络注入现实网络中的用户行为流量,设计一个面向云平台的网络用户行为仿真系统。通过真实流量回放的网络注入多样化的网络用户行为,解决基于仿真模型模拟用户行为单一、逼真度低的问题,在此基础上,研究复杂目标网络回放时确保回放的交互性和逼真度,以及基于互动式流量回放驱动的大规模高逼真网络行为仿真的可行性与逼真性。

1 相关工作

与本文相关的工作主要体现在基于模型驱动的流量生成方法、基于真实流量驱动的回放方法和用户行为建模与仿真方法3个方面。

在基于模型驱动的流量生成方法[7]方面,文献[8-10]基于流量模型的回放方法根据真实网络中流量的数学特征建立相关模型生成仿真流量,由于该方式回放的流量逼真性取决于建立的相关流量模型是否逼真准确。而真实网络场景复杂多变,几乎难以建立一个与真实网络场景完全一致的流量模型,因此该方式存在一定的局限性。

在基于真实流量驱动的回放方法方面,由于该回放方式生成的流量直接来自真实网络中,因此能够完整且精确地重现真实网络中的用户流量以及数据包层面的内容。相比基于流量模型回放的方式,该回放方式关注在回放过程中数据报文层面的应用有效载荷,更加符合网络用户行为仿真的高逼真需求。目前基于真实流量驱动的流量方式研究热点开始转向交互式流量回放,即更加注重回放过程中报文状态交互准确。文献[11]提出交互式流量回放这一概念,并开发了交互式回放系统TCPOpera,通过模拟TCP/IP协议栈维护请求与响应端的会话状态,基于状态判定控制报文的收发。文献[12]提出一种有状态流量回放方式,维护请求响应端的交互状态,重放应用层的流量用于测试应用代理服务器的安全性能。文献[13]引入收发平衡机制,提出一种基于收发平衡和状态判断相结合的TCP流量回放方法,在发送报文前通过优先收法平衡判定,减少状态判定的开销,提高回放性能。但是以上方法都只局限于物理回放设备,并且是单个DUT设备的回放场景,无法在虚拟网络中进行回放。针对现有回放方法难以在复杂的虚拟目标网络上实现回放问题,文献[14]论述了网络环境对回放的影响,通过计算物理网络与虚拟网络的相似度和IP映射的方法将流量回放到虚拟网络中,但该方法并没有实现交互式流量回放。文献[15]提出一种基于云平台的虚拟网络交互回放的方法ITRM,重点研究在一个缩小规模的虚拟网络场景下如何进行多节点的交互流量回放的问题,但该方法回放规模相对较小,前提条件需要保证各回放节点时钟高度同步,虚拟网络中时钟同步的精度无法满足此需求,而且该方法并未考虑到延时等网络环境对回放造成的影响,因此无法在复杂的目标网络场景下保证流量回放的交互性。

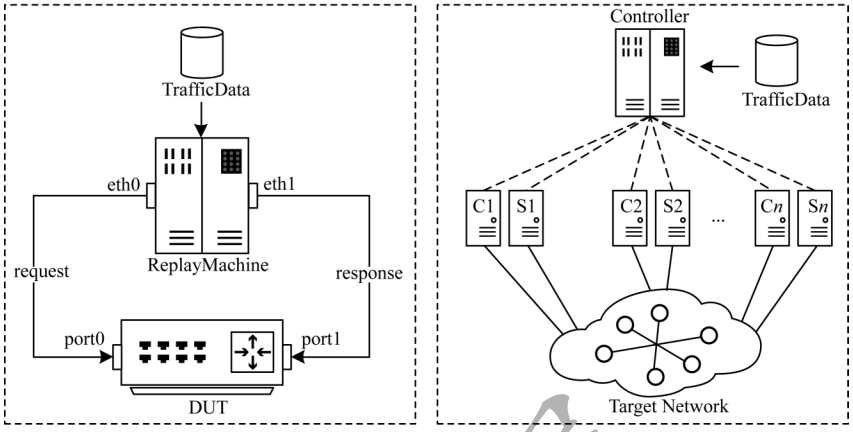
目前关于大规模多用户的行为仿真的研究较少,文献[16]基于目前用户网络行为研究现状,将目前研究方向细分为个人与群体网络行为,将用户网络行为定义为用户通过操作某个或某种应用程序与他人或服务进行交互的行为以及产生的网络流量。文献[17]对用户访问HTTP服务的行为进行了建模,并以此模型对HTTP协议流量进行模拟,且模拟的网络流量满足流量的自相似性。但上述方法都只是针对单个用户行为进行仿真研究,且模型无法灵活模拟真实用户的行为,仿真的逼真性不高。文献[18]提出一种基于云平台的多用户行行仿真的方法,针对天地一体化网络[19]中的多用户并发行为,通过模型驱动卫星用户行为仿真,但是该研究只停留在窄带用户行为。文献[20]提出基于“录制-回放”策略的网络桌面应用行为仿真方案,通过录制单个用户的应用操作行为,然后在虚拟机上进行回放,由此仿真出具有真实负载的交互流量,但该方法耗费的资源相对较高,仿真规模会受到一定的限制。

本文提出一种面向目标网络的流量回放方法,并设计一套多用户行为仿真架构,以解决无法在目标网络中进行精确交互流量回放以及目前用户行为仿真规模逼真度不高的问题。

2 体系结构

2.1 现有方法存在问题与设计思路

目前大部分的交互式回放工具为面向DUT设备(如路由器、防火墙)测试的应用需求而设计[21],因此回放场景通常是回放设备的端口与DUT设备串接测试,如图1(a)所示,无需考虑链路延时等网络环境对回放性能的影响,这显然与基于云平台的目标网络仿真场景需求不符:即链路回放的场景更为复杂。图1(b)所示为一个目标网络的回放场景,回放的对象为虚拟目标网络,即回放的流量需要经过整条链路,因此在面向云平台设计时,需要考虑延时导致的数据包乱序问题。同时,在进行大规模行为仿真实验时,设计的体系架构需满足多条链路并发交互式回放的需求。



(a)DUT串接回放场景　　(b)目标网络链路回放场景

图1　DUT串接回放场景和目标网络链路回放场景对比

Fig.1　Comparison of DUT serial replay scene and target network link replay scene

2.2　基于交互式流量回放的用户行为仿真体系

本文提出的用户行为仿真架构面向OpenStack作为仿真平台,采用分布式架构可按需灵活地将用户行为仿真节点加载到任意目标网络实验中,较好地满足用户行为仿真对多样性和灵活性的需求。该仿真体系如图2所示。其中仿真控制端部署在控制节点上,仿真用户节点由KVM或Docker生成。

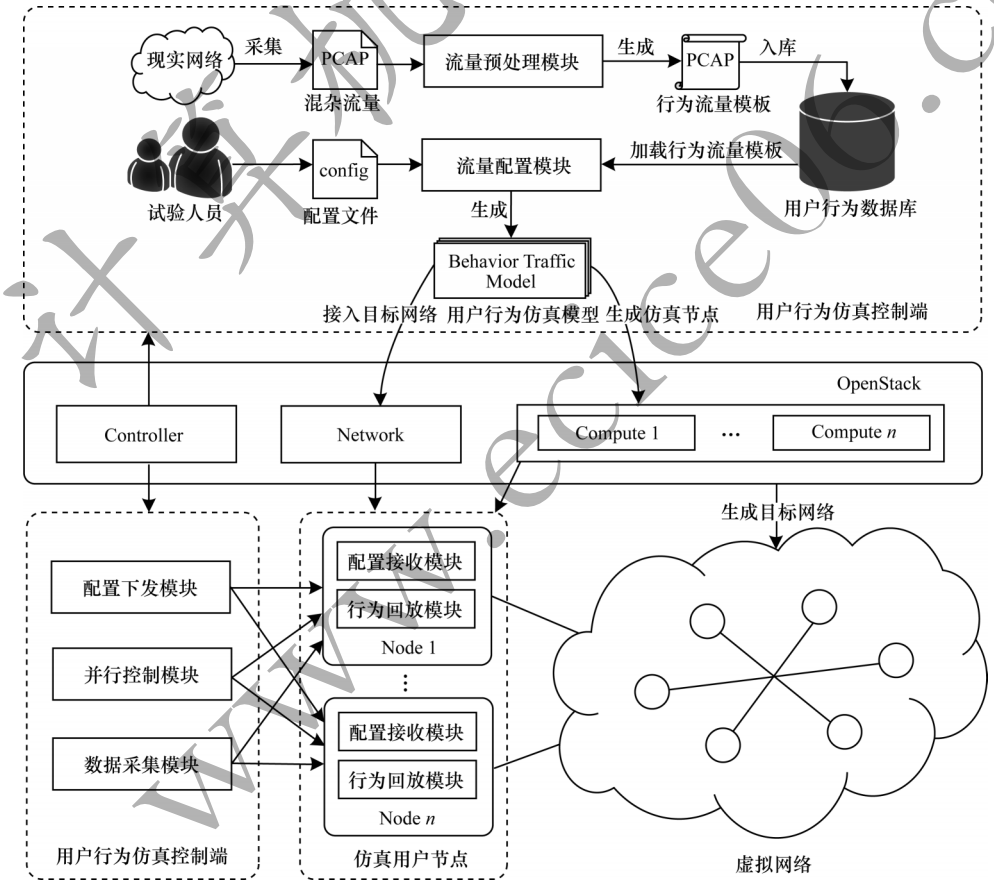


图2　基于流量回放的网络用户行为仿真体系架构

Fig.2　Network user behavior emulation architecture based on traffic replay

具体仿真流程如下:

1)用户行为流量模板生成。控制端负责用户行为仿真模板的生成、配置和下发,并且具有一个用户行为数据库,用于存放各种类型的用户行为流量模板。试验人员可以将现实网络中采集的混杂流量导入仿真控制端,仿真控制端首先按照输入配置对输入流量过滤清洗,然后匹配流量中各个报文的五元组(源/目的IP地址、源/目的MAC地址、协议类型)

识别特征行为流量,再根据协议的端口号、状态码、载荷抽取出单个完整的会话流,按照请求-响应标记分类,最后生成行为流量模板存入用户行为数据库中。

2)行为仿真任务配置与下发。在一次行为仿真任务过程中,首先仿真控制端解析试验人员的配置文件,获取目标网络部署信息,创建所需的回放节点并接入到目标网络中,并根据配置文件在目标网络中自动化配置路由。同时,从用户行为数据库中加载对应类型的用户行为流量模板进行处理,修改数据包MAC地址等字段保证用户行为流量能够在目标网络中转发。然后建立用户行为仿真模型,该模型包括目标网络链路信息以及用户规模数、行为类型、仿真时间等具体的仿真运行参数,将生成的仿真模型以及流量模板通过配置下发模块下发至对应的仿真节点中。

3)行为仿真任务并行控制。在所有实例模型以及流量模板下发到回放节点后,仿真控制端通过消息队列控制各仿真节点启动运行,按照时间线实现对不同节点不同用户行为仿真事件的集群调度。仿真节点在收到仿真指令后,调用回放模块报文发送函数将用户行为流量注入到虚拟目标网络中,并实时采集仿真运行状态。

3 交互式流量回放关键技术

交互式流量回放最为关键的目标就是保证回放流量过程中数据包的时序状态以及报文内容与原始流量一致,该目标也是评估回放逼真性的重要指标之一。为实现该目标,现有主流的交互式回放方法普遍将请求和响应端放在同一服务器上,以便于更加精确地协同控制请求和响应。本文采用此架构,但是回放的对象为目标网络,即回放的流量需要经过整条链路,因此,需要考虑延时导致的数据包乱序问题:

1)原始流量文件中可能出现部分相邻报文间隔极短的特殊情况,造成此情况的原因之一可能是采集过程中采集点没有选取请求与响应端链路的中点,或者是在一个超低延时的网络环境下进行流量采集。因此,有必要针对该情况进行时间戳校正。

2)回放过程中链路延时、抖动以及回放节点处理延时对回放效果的影响。面向云平台的流量回放相比于DUT测试场景更为复杂,延时和抖动会随着链路的复杂度增大而增加,尤其是大规模的网络试验场景。不同于实物回放设备,在时间精度方面,实物网卡往往可以达到微秒级时钟控制,而虚拟机的时钟无法达到此精度,因此,还需要考虑报文处理阶段虚拟机的I/O性能,包处理能力等造成的回放误差。

基于上述问题,本文设计一种基于RTT校正的时间戳修复算法和延时补偿策略,通过对原始流量文件进行RTT时间戳修正,链路延时补偿,解决因延时造成的误差。在保证报文顺序准确的前提下,尽可能将回放时间误差降至最低。

此外,在基于流量回放的用户行为仿真过程中,保证回放数据包的顺序准确性优先级比保证时序的精确性更高。因为采集的数据包的时间戳是按网卡接收数据包的时间记录的,从而无法获取数据包真实发送时间,而且有的协议例如TCP协议具有延时确认机制。较小的时间误差(微秒级)通常对一个会话流状态的影响很小,但是请求与响应顺序会决定会话流的协议状态。因此,设计回放算法的原则应在保证回放数据包顺序一致的前提下尽可能将时间精度误差降低,以维持重放的会话状态,保证用户行为仿真的逼真度。为此,本文采用的报文发送函数基于相邻报文时间间隔进行发送,而非距离第1个报文的时间间隔进行发送,该方式能够有效确保报文顺序的精确性。

3.1 RTT时间戳修复算法

回放时需要维持整个会话的状态,即数据报文交互的顺序需与原始流量文件保持一致,否则将会导致协议状态错乱。图3所示为一个TCP会话建立连接过程,当采集点不是客户/服务端之间的中间节点,或者是采集网络延迟非常低时,流量会话中SYN请求报文与SYN-ACK响应报文的时间间隔非常短,则回放时就有可能因为传输延时出现SYN-ACK响应报文比SYN请求报文先到达目标节点的状态错乱情况。因此,有必要对采集的用户行为流量进行延时修复处理,保证其在回放过程中顺序的一致性,以维持正确的用户行为协议状态。

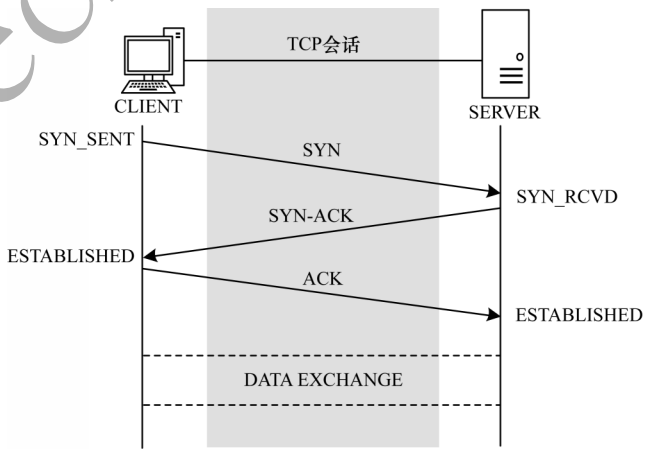


图3 TCP会话3次握手过程

Fig.3 Process of TCP 3-way handshake

本文提出一个基于RTT修复算法解决流量采集时采集点不在中间节点的问题。其中心思想是:根据TCP会话3次握手中第1次握手SYN报文与第3次握手SYN-ACK时间戳可计算出原网络RTT往返延时,再根据RTT将报文修复为在中心节点采集

时的时间戳,具体的伪代码描述如算法1所示。

算法1 基于RTT校正的时间戳修复算法

```
1.first_syn_pkt ← False
2.for packet in origin_pcap do
3. if first_syn_pkt is False and packet_p == TCP and TCP_flag == SYN
4.first_syn_pkt ← True
5.ts_syn ← packet_ts
6.client_tuple ← packet_tuple
7.server_tuple ← swap(client_tuple)
8. if first_syn_flag is True and packet_tuple == server_tuple and tcp_flag == ACK
9.ts_syn_ack ← packet_ts
10.delta ← ts_syn_ack–ts_syn
11. if first_syn_flag is True and tuple == client_tuple and tcp_flag == ACK
12.rtt ← packet_ts – ts_syn
13.break
14.end for
15.tc ← delta/2-rtt/4
16.pcap_fix ← create a new pcap file
17. client_IP, server_IP ← get source IP and destion IP from the client_tuple
18.for packet in origin_pcap do
19.if src_IP == client_IP then
20.packet_ts ← packet_ts + tc
21.else
22.packet_ts ← packet_ts – tc
23.add packet to fix_pcap
24.end if
25.end for
26.return fix_pcap
```

3.2 低延时补偿策略

因为RTT修复算法只有在原始流量中包含TCP会话时才有效,当原始流量中不包含完整TCP会话则无法预估采集时的RTT延时,所以需要对其他低延时的情况进行延时补偿。假设原始流量相邻报文间隔为T_d,回放时链路的抖动为T_j,当$T_j>T_d$时,则有可能发生数据包乱序的情况,为尽可能避免这一问题,本文提出一种低时延补偿策略。通过对报文较小的包进行延时补偿,尽可能消除抖动对回放造成的误差。

本文具体策略如下:在仿真实验前,设定一组等同数量具有固定间隔的流量文件进行回放。比较原始流量与回放流量中每对相邻报文的时间间隔Δo、Δr,得到一个误差均值。计算公式如下:

$$E_{\text{avg}}^{\text{Error}}=\frac{\sum_{i=1}^{n}(|\Delta r_i-\Delta o_i|)}{n},\ i=1,2,\cdots,n$$

将得到的误差均值对原始流量进行补偿,具体操作如下:从原始流量中第2个报文开始顺序遍历,当报文满足$\Delta o_i<E_{\text{avg}}^{\text{Error}}$且与上一条报文不同侧时,则对当前报文的时间戳$Ts_i$进行误差补偿:$Ts_i=Ts_i-1+E_{\text{avg}}^{\text{Error}}$,并更新报文时间戳,直至最后一个报文。使得每个报文间隔始终不小于平均误差,以此减小延时在回放延时中造成的误差。

4 实验验证与评估

本节对基于交互式流量回放的用户行为仿真方法进行实验验证。

4.1 实验环境

该用户行为仿真体系部署在基于Openstack Mitaka版本搭建的云平台上,控制节点采用Intel Xeon E5-2620 v2×4机架式处理器,内存为32 GB;网络节点处理器为Intel Xeon E5-4607 v2×8,内存为32 GB;计算节点1处理器为Intel Xeon E5-2620 v3×4,内存为32 GB;计算节点2、3、4处理器均为IntelXeon E5-2620 v3×4,内存为32 GB;采用OVS提供虚拟网络;所有节点的操作系统均为CentOS 7.5。

实验场景如图4所示,目标网络为一个普通的多尺度虚拟网络拓扑,图中灰色节点均为虚拟路由器,实验场景两端为本仿真方法生成的用户节点,中心黑色路由节点为实验采集点。共设计了3个验证实验分别对链路交互流量回放方法的逼真性与精准度,以及用户行为仿真的可行性与逼真性进行验证。

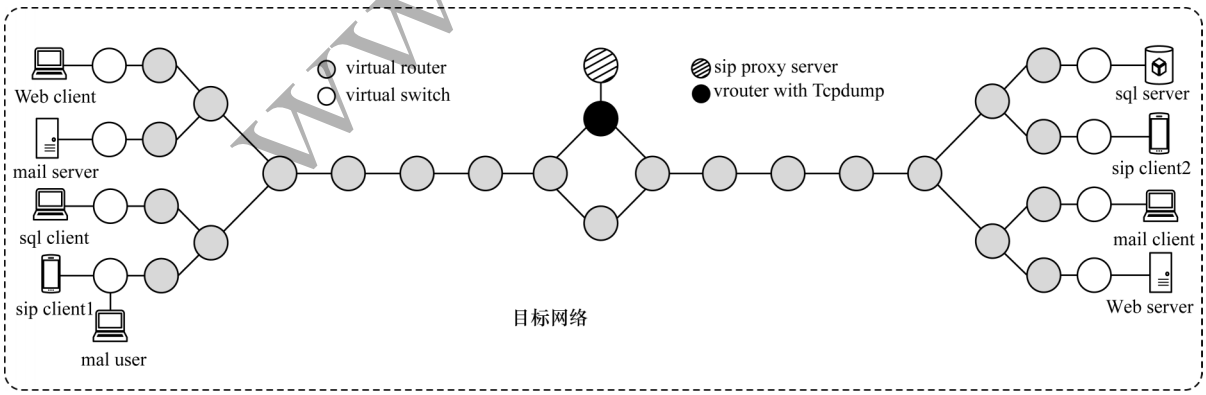


图4 仿真实验场景

Fig.4 Simulaton experiment scene

4.2 实验结果

4.2.1 交互式流量链路回放方法逼真性验证实验

为验证本文所提出的链路交互式流量回放方法可满足大规模高逼真用户行为仿真的需求,对该方法进行逼真性实验评估。在图4的目标网络中,创建一个回放节点接入目标网络中,仿真一对Web请求响应会话;链路延时为1.21 ms,在整条链路上回放了5组采集自互联网的真实的TCP流量文件(数据报文个数依次为2万、5万、10万、20万、30万个),对本文所设计的交互式流量链路回放方法在顺序的准确性以及时间的精确度2个方面进行评估,以验证该回放方法的逼真性。基于该评估方法对比直接回放方法,具体的评估方法如下:

通过对回放后的流量文件与原始流量文件进行遍历,比较两组在同一位置的每个报文内容是否相同,统计错误的报文数以验证该方法的顺序准确性。通过计算回放与原始流量中每个相邻报文的时间间隔的误差验证时间精确度,即设Δr为回放流量中2个相邻报文的时间间隔,Δo为原始流量中对应的2个相邻报文的时间间隔,可以计算出:

$$E_{\text{avg}}^{\text{Error}}=\frac{\sum_{i=1}^{n}(|\Delta r_i-\Delta o_i|)}{n}$$

在5组回放实验中,本文方法产生的乱序报文数分别为2、4、4、6、4个。如图5所示,随着回放数据包的增多,直接回放方法产生的乱序报文数明显增多,在回放30万个数据包时,直接回放共产生乱序报文5 900个,而本文方法仅为4个。

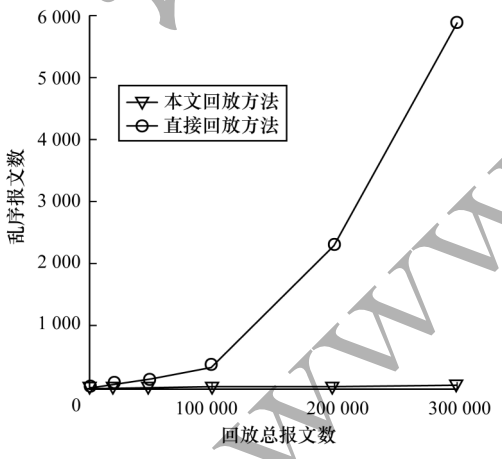


图5 本文方法与直接回放方法的乱序报文数对比

Fig.5 Comparison of the out-of-order packets number between the method in this paper and direct replay method

本文所提出延迟补偿策略极大地减少了回放中产生错误数据包的数量,在回放的顺序准确性上具有一定的优势,能够保证流量回放不乱序这一基本目标。相比未经过延时补偿的直接回放方法,本文方法经过延时补偿后回放产生的乱序报文数并不会随着回放报文数增加而大幅增加,回放的准确率稳定在99.99%。

由于直接回放法在报文的顺序准确率低,无法准确验证其回放误差,因此本文通过对比经典单机流量回放软件Tcpreplay回放效果进行时间精确度验证。对比结果如图6所示,本文所设计的链路交互式流量回放方法平均误差均小于0.026 ms,与流量回放软件Tcpreplay基本持平,但是Tcpreplay为单机回放,无需考虑链路延时问题,且无法实现在链路中交互回放。

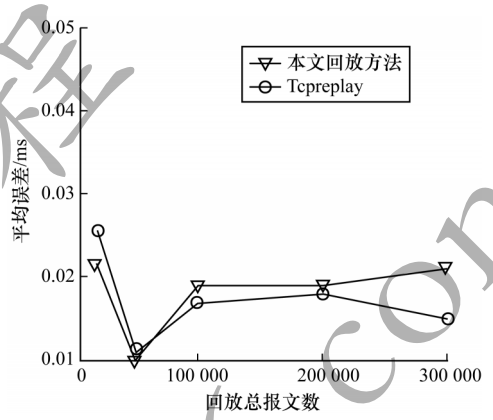


图6 本文方法与Tcpreplay单机回放方法平均误差对比

Fig.6 Average error comparison of the this paper method and Tcpreplay stand-alone replay method

表1为本文方法与目前主流的回放方法的对比分析。在回放的逼真性方面,Tcpreplay由于是单机回放,在时序的准确性和精确度上要略优于本文方法,但是其无法面向目标网络回放,ITRM和本文方法都是面向虚拟网络回放的,但是其无法保证各回放节点的高精度的同步时钟,且其回放机制是根据第1个报文间隔进行回放,没有考虑到复杂网络环境产生乱序报文的问题。因此,本文设计的链路交互式流量回放方法在回放时间和顺序的精确性上具有一定的逼真性,特别是在复杂网络的情况下,能够满足用户行为仿真对于回放逼真度的需求。

表1 不同回放方法的对比分析

Table 1 Comparative analysis of different playback methods

回放方法	是否支持交互回放	是否支持复杂场景	顺序准确性	时序准确性/ms
本文方法	是	是	99.99%	<0.20
ITRM[14]	是	否	无法保证	—
Tcpreplay[21]	否	否	100%	<0.20

4.2.2 交互式流量链路回放方法多样性验证实验

真实网络中多用户行为具有多样化、大规模并发的特点。因此,多用户行为仿真也应具备大规模实验的能力。为验证本文仿真方法的多样性与规模,在图4网络场景下设计了基于Web请求、数据库

操作、邮件操作3种行为的多用户并发仿真实验,目标网络两端为本文方法创建的3组仿真节点,通过消息队列下发用户仿真配置参数,并进行分布式协同控制。本次仿真规模为36 000个用户,仿真时间设置为315 s,共回放了917 000个报文。具体的行为参数及回放结果如表2所示。

表2　并发行为仿真流量回放结果

Table 2　Replay result of concurrent behavior simulation traffic

行为类型	仿真用户数	协议	报文数/占比	乱序报文数	回放准确率/%
Web请求	15 000	HTTP	315 000/34.4%	28	99.98
数据库操作	1 000	TNS	302 000/32.9%	30	99.99
邮件操作	20 000	SMTP	300 000/32.7%	8	99.99
所有行为	36 000	TCP	917 000/100%	66	99.99

在保证不丢包和回放的数据包内容与原始报文一致的前提下,表2结果显示在回放的917 000个数据报文中共产生了66个乱序报文,3种行为仿真准确率均维持在99.99%,回放的精确度能够满足多用户并发行为仿真的逼真需求。

4.2.3　恶意用户行为仿真逼真性验证安全实验

恶意用户行为仿真对于网络安全研究有着重要意义,本文通过重放复现攻击场景进行恶意用户为行为分析。本文用户行为仿真方法亦可构建恶意用户行为模型,根据基于重放攻击的安全实验验证恶意用户行为模型的逼真性。在图4目标网络中设置了一组通话节点sip_client1与sip_client2,并在中间路由节点添加了一个通信代理服务节点sip_server,该节点集成开源通信代理软件SIPp,对通话进行验证。采用本文行为仿真方法创建恶意用户节点mal_user进行重放攻击,具体攻击流程如下:首先截获sip_client1到sip_server的通话请求流量,构造通话用户行为模型,具体的通话用户模型如表3所示。然后在恶意用户节点mal_user上进行重放攻击,即在链路2(mal_user->sip_server)中重放链路1中(sip_client1->sip_server)的通话请求行为。

表3　通话行为仿真流量模型

Table 3　Simulation traffic model of call behavior

序号	ID	被叫ID	IP	所属节点
1	185000000	185000500	111.1.1.1	sip_client1
2	185000500	185000000	111.2.1.1	sip_client2
⋮	⋮	⋮	⋮	⋮
999	185000499	185000999	111.1.2.246	sip_client1
1000	185000999	185000499	111.2.2.246	sip_client2

本次实验共重放500组用户通话行为,图7为代理节点sip_server实时统计的通话成功数,图8为通话节点sip_client2最终显示的通话成功用户数。在图7中,曲线呈线性增长的趋势符合预设的正常通话速率。实验结果证明,通过恶意节点重放的通话流量能够通过代理服务器的交互验证,并实现与SIP_c2节点上的用户成功通话对正常通话行为进行干扰。该行为仿真方法具有一定的逼真性。

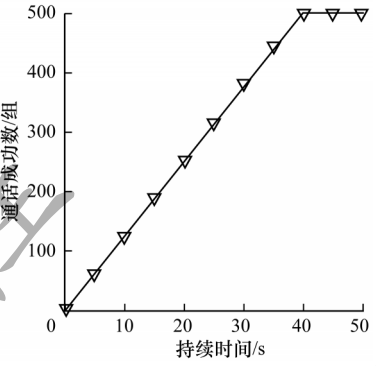


图7　代理服务器sip_server节点实时统计的通话成功数

Fig.7　Number of successful calls counted by the call proxy server sip_server node in real time

```
------------------------------ Statistics Screen ------- [1-9]: Change Screen --
Start Time              | 2020-06-08   10:43:04.957619 1591612984.957619
Last Reset Time         | 2020-06-08   10:44:11.774610 1591613051.774610
Current Time            | 2020-06-08   10:44:12.698345 1591613052.698345
-------------------------+----------------------------+----------------------------
Counter Name            | Periodic value             | Cumulative value
-------------------------+----------------------------+----------------------------
Elapsed Time            | 00:00:00:923000            | 00:01:07:740000
Call Rate               |    0.000 cps               |    7.381 cps
-------------------------+----------------------------+----------------------------
Incoming call created   |       0                    |      500
OutGoing call created   |       0                    |        0
Total Call created      |                            |      500
Current Call            |       0                    |
-------------------------+----------------------------+----------------------------
Successful call         |       0                    |      500
Failed call             |       0                    |        0
-------------------------+----------------------------+----------------------------
Response Time 1         | 00:00:00:000000            | 00:00:00:002000
Call Length             | 00:00:00:000000            | 00:00:15:013000
------------------------------ Test Terminated --------------------------------
```

图8　通话节点sip_client2最终通话成功数

Fig.8　Number of successful final calls of call node sip_client2

5　结束语

本文通过流量回放的方式进行网络用户行为仿真,提出一种交互式流量链路回放的方法。将回放对象从单个节点扩大到整个目标网络链路,并基于该方法设计一个高逼真大规模的网络用户行为仿真体系。实验结果表明,该方法能够在保证会话顺序与内容的准确性前提下,达到与现有单机回放方法同等的精确性及逼真性。基于该回放方法的网络安全仿真体系具有可行性,其仿真体系可实现灵活构建大规模高逼真的用户行为仿真目标,从而满足基于云平台构建的目标网络注入网络用户行为背景流量的需求。下一步将融合虚实互连技术,研究面向大规模目标网络的高逼真用户行为仿真方法。

参考文献

[1] 方滨兴,贾焰,李爱平,等. 网络空间靶场技术研究[J]. 信息安全学报,2016,1(3):1-8.
FANG B X,JIA Y,LI A P. Cyber ranges:state-of-the-art and research challenges[J]. Journal of Cyber Security,2016,1(3):1-8. (in Chinese)

[2] 单晓红,王宁,蔡培. 网络社区用户行为仿真与分析[J]. 计算机系统应用,2014,23(7):17-23.
SHAN X H,WANG N,CAI P. Simulation and analysis of virtual community users' behavior[J]. Computer Systems & Applications,2014,23(7):17-23. (in Chinese)

[3] 茹新宇,刘渊,陈伟. 新网络仿真器NS3的研究综述[J]. 微型机与应用,2017,36(20):14-16.
RU X Y,LIU Y,CHEN W. A survey of new network simulator NS3[J]. Microcomputer & Its Applications,2017,36(20):14-16. (in Chinese)

[4] 张磊. 基于OPNET的智慧家庭异构网络的仿真研究[D]. 北京:北京邮电大学,2019.
ZHANG L. Opnet-based simulation research on heterogeneous network for smart-home[D]. Beijing:Beijing University of Posts and Telecommunications,2019. (in Chinese)

[5] 林秀,基于TCPCopy的在线引流压测通用架构设计[J]. 电信技术,2014(11):30-33.
LIN X. General architecture design of online drainage pressure measurement based on TCPCopy [J]. Telecommunications Technology,2014(11):30-33. (in Chinese)

[6] 李春艳,张学杰. 基于高性能计算的开源云平台性能评估[J]. 计算机应用,2013,33(12):3580-3585.
LI C Y,ZHANG X J. Performance evaluation on open source cloud platform for high performance computing[J]. Journal of Computer Applications,2013,33(12):3580-3585. (in Chinese)

[7] 方熙,曾剑平,吴承荣. 背景流量生成模型综述[J]. 计算机应用,2019,39(S1):124-131.
FANG X,ZENG J P,WU C R. Journal of computer applications[J]. Journal of Computer Applications,2019,39(S1):124-131. (in Chinese)

[8] 焦久隆. 基于流量矩阵的背景流量建模与生成技术研究[D]. 哈尔滨:哈尔滨工业大学,2016.
JIAO J L. Research on background traffic modeling and generation method based on traffic matrix[D]. Harbin:Harbin Institute of Technology,2016. (in Chinese)

[9] 张华川,田杰,许静. 自相似网络流量模拟的分布式系统的设计与实现[J]. 电子学报,2009,37(4):31-35.
ZHANG H C,TIAN J,XU J. The design and implementation of distributed system for self-similar network traffic simulation[J]. Acta Electronica Sinic,2009,37(4):30,31-35. (in Chinese)

[10] 储伟,燕亚兰. 5G应用场景业务流量模型仿真平台研究[J]. 通信与广播电视,2020(2):6-12.
CHU W,YAN Y L. Research on simulation platform of 5G application scenario traffic model[J]. Communication & Audio and Video,2020(2):6-12. (in Chinese)

[11] HONG S S,WU S F. On interactive internet traffic replay[C]//Proceedings of the 8th International Conference on Recent Advances in Intrusion Detection. Berlin,Germany:Springer,2005:257-264.

[12] HUANG C Y,LIN Y D,LIAO P Y,et al. Stateful traffic replay for Web application proxies[J]. Security and Communication Networks,2015,8(6):970-981.

[13] 褚伟波,蔡忠闽,管晓宏,等. 基于收发平衡判定的TCP流量回放方法[J]. 计算机学报,2009,32(4):835-846.
CHU W B,CAI Z M,GUAN X H,et al. A new method for interactive TCP traffic replay based on balance-checking between transmitted and received packets[J]. Chinese Journal of Computers,2009,32(4):835-846. (in Chinese)

[14] LI Z,HAO Y,ZHANG Z,et al. Traffic replay in virtual network based on IP-mapping [C]//Proceedings of International Conference on Algorithms and Architectures for Parallel Processing. Berlin,Germany:Springer,2015:697-713.

[15] LIU H,AN L,REN J,et al. An interactive traffic replay method in a scaled-down environment[J]. IEEE Access,2019,7:149373-149386.

[16] 李晶晶,刘红日,王佰玲,等. 用户网络行为模拟技术研究[J]. 信息技术与网络安全,2018,37(1):36-39.
LI J J,LIU H R,WANG B L,et al. A survey of user network behavior simulation technology[J]. Information Technology and Network Security,2018,37(1):36-39. (in Chinese)

[17] LEUNG K Y,YEUNG K H. The design and implementation of a WWW traffic generator [C]// Proceedings of IEEE Parallel and Distributed Systems. Washington D. C.,USA:IEEE Press,2000:509-514.

[18] 张光杰,叶海洋,王晓锋. 基于多尺度虚拟化的卫星终端用户行为仿真[J]. 计算机工程,2019,45(8):165-172.
ZHANG G J,YE H Y,WANG X F. Emulation of satellite terminal user behavior based on multi-scale virtualization [J]. Computer Engineering,2019,45(8):165-172. (in Chinese)

[19] 李凤华,殷丽华,吴巍,等. 天地一体化信息网络安全保障技术研究进展及发展趋势[J]. 通信学报,2016,37(11):156-168.
LI F H,YIN L H,WU W,et al. Research status and development trends of security assurance for space ground integration information network[J]. Journal on Communications,2016,37(11):156-168. (in Chinese)

[20] 张新星. 基于虚拟化的网络流量模拟系统设计与实现[D]. 哈尔滨:哈尔滨工业大学,2017.
ZHANG X X. Design and implementation of network traffic simulation system based on virtualization[D]. Harbin:Harbin Institute of Technology,2017. (in Chinese)

[21] 吕平,董春雷,刘冬培,等. 基于FPGA的软件定义流量发生器[J]. 通信学报,2018,39(S2):66-71.
LÜ P,DONG C L,LIU D P,et al. Implementation of software defined traffic generator based on FPGA[J]. Journal on Communications,2018,39(S2):66-71. (in Chinese)

编辑 索书志