基于信元序列重建的暗网流量远程指纹识别技术

李泽禹[1],黄文涛[1],王学宾[2],刘 凯[1],时金桥[1*]
(1. 北京邮电大学网络空间安全学院,北京 100876;2. 中国科学院信息工程研究所,北京 100085)

摘 要: 流量指纹攻击可以应用于第二代洋葱网络(The second onion router,Tor)的隐藏服务溯源。由于Tor网络协议设计的限制,隐藏服务侧流量难以获得,因此需要利用电路中其他位置的流量作为隐藏服务侧流量的替代来开展远程指纹攻击。然而,指纹模型面临Tor中继中数据缓存、加解密与重封装带来的报文序列的结构性变化,导致指纹模型在位点迁移时存在性能衰减。已有的方法在一定程度上可以从模型层面和数据层面缓解随机噪声,却难以精准拟合由协议机制导致的结构性差异。本文深入剖析了暗网流量的转发机理,提出了上下文感知的远程信元序列重建(Remote Cell Estimator,RemoCellEst)。该方法引入了跨报文的残留补偿机制,动态修正报文起始处的载荷偏移,从而重建成跨位点一致的信元序列,使模型具备更稳定的位置迁移能力。实验结果表明,在隐藏服务溯源场景下,基于RemoCellEst方法重建的信元序列构建的远程指纹模型可达94.85%的准确率,相较于基于桶分割方法或直接使用传输控制协议(Transmission Control Protocol,TCP)/传输层安全协议(Transport Layer Security protocol,TLS)层流量构建的模型提升了6%~16%的准确率。本文将远程指纹攻击的位点迁移问题归因到可解释、可建模的报文序列结构性差异,可为后续加密流量分析、流关联等相关研究提供新的视角。

关键词: 远程指纹攻击;信元序列重建;隐藏服务溯源;网站指纹攻击;流量指纹攻击

基金项目: 国家重点研发计划(No.2023YFB3106600);国家自然科学基金(No.62372056);海南省方滨兴院士工作站资金资助(No.YSGZZ2023003);海南省院士创新平台科研项目资金资助(No.YSPTZX202506)

中图分类号: TP393.0 **文献标识码:** A **文章编号:** 0372-2112(2026)04-1529-19

电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20251035

Remote Traffic Fingerprinting Attack Based on Cell Sequence Reconstruction

LI Zeyu[1], HUANG Wentao[1], WANG Xuebin[2], LIU Kai[1], SHI Jinqiao[1*]
(1. *School of Cyberspace Security*, *Beijing University of Posts and Telecommunications*, *Beijing* 100876, *China*;
2. *Institute of Information Engineering*, *Chinese Academy of Sciences*, *Beijing* 100085, *China*)

Abstract: Traffic fingerprinting attacks can be applied to hidden services de-anonymization in the Tor (The Second Onion Router) network. Due to the limitations of the Tor network protocol, traffic on the hidden service-side is difficult to obtain. Therefore, traffic from other positions is used as a substitute for hidden service-side traffic to train remote fingerprinting models. However, fingerprinting models face structural changes in packet sequences caused by data caching, encryption/decryption, and repackaging at Tor relays, leading to performance degradation during position migration. Existing methods, whether at the model or data level, can partially mitigate random noise but struggle to accurately fit structural differences induced by protocol mechanisms. This paper delves into the forwarding mechanisms of Tor traffic and proposes a context-aware remote cell sequence estimator method, called remote cell estimator (RemoCellEst). This approach introduces a cross-packet residual compensation mechanism to dynamically adjust payload offsets at boundaries of each packet, thereby reconstructing consistent cell sequences across position and enhancing the model's capability for positional migration. Experimental results demonstrate that remote fingerprinting attacks based on cell sequences reconstructed using the RemoCellEst method achieve an accuracy of 94.85%, representing a 6% to 16% improvement over which using bucket-based estimator methods or directly using TCP (Transmission Control Protocol)/TLS (Transport Layer Security Protocol) layer traffic. This paper attributes the position migration challange in remote fingerprinting attacks to interpretable and modelable structural differences in packet sequences, offering a novel perspective for subsequent research such as encrypted traffic analysis and flow correlation.

收稿日期:2026-01-26;录用日期:2026-03-09;责任编委:沈蒙
*通信作者:时金桥

Keywords: remote fingerprinting attack; cell-layer reconstruction; hidden service de-anonymization; website fingerprinting attack; traffic fingerprinting attack

Foundation Item(s): National Key Research and Development Program (No.2023YFB3106600); National Natural Science Foundation of China (No.62372056); Academician Fang Binxing Workstation in Hainan Province (No.YSGZZ2023003); Specific Research Fund of The Innovation Platform for Academicians of Hainan Province (No.YSPTZX202506)

0 引言

第二代洋葱网络(The second onion router, Tor)通过多层加密、多跳转发,对搭建在其中的隐藏服务提供匿名保护,可以在不暴露隐藏服务位置的情况下,向匿名用户发布网络服务,由此带来滥用风险与治理挑战[1-2]。因此,开展隐藏服务溯源具有重要安全价值与现实意义。隐藏服务溯源包括主动[3-4]和被动[5-6]的方法,其中流量指纹识别是实现隐藏服务溯源的一种可行的方法[7-12]。在训练阶段,敌手通过访问目标隐藏服务捕获其对应的加密流量并进行标注,训练流量指纹模型。此后,敌手在Tor网络中捕获加密流量,通过流量指纹模型提取其指纹特征并判断是否是目标隐藏服务对应的流量。当敌手的流量监控位置为目标隐藏服务的入口节点时,即可确定隐藏服务和观测到的互联网协议(Internet Protocol, IP)地址的对应关系,成功实现溯源。

然而,和传统的网站指纹攻击不同,在隐藏服务溯源场景下,标注流量的捕获位置与实际监测流量的位置往往是不同的。敌手无法获取目标隐藏服务接入暗网时的流量,其根本原因是目标隐藏服务的入口地址是敌手未知且非可控的,因此,使用目标隐藏服务入口侧的流量用于指纹模型训练仅是一种理想场景。为解决目标隐藏服务入口侧流量难以获取的问题,已有方法通过在Tor网络中部署可控的客户端[13]、汇聚节点[14]或隐藏服务代理[15]来捕获带域名标签的替代流量,并据此训练指纹模型,随后监控未知流量并进行预测,实现隐藏服务与IP地址的关联。这种训练位置和预测位置不同场景下的流量指纹攻击称为远程指纹攻击[13]。然而,由于训练阶段的替代流量与预测时捕获的流量之间跨越多个转发中继,加密数据的逐跳缓冲、解密与重封装会使同一会话流量在不同监测位置呈现的可见长度、数量组合与上下行序列形态上产生结构性差异[15-16],从而导致指纹模型在位点迁移时出现明显的性能衰减。

为提升指纹模型在训练与测试条件不一致时的鲁棒性[17],相关工作可分为两类思路:一类从模型层面开展环境无关特征学习,通过架构设计或特征工程使模型主动忽略噪声。例如,在架构设计上,Sirinam等人[18]和Wang等人[13]利用三元组来开展对比学习;Zhao等人[14]在此基础上进一步引入四元组[14],强迫模型提取跨环境的不变指纹嵌入。在特征层面,Gong等人[8]和Shen等人[19]提出了流量对齐矩阵(Traffic Alignment Matrix, TAM);Mitseva等人[20]构建了流量图(Traffic Graph),这些方法将流量转化为更稳健的宏观表示,将网络拥塞或抖动引起的噪声过滤掉或平滑化,使模型能够专注于观察流量中不易改变的整体模式,从而提升在不同网络环境下的稳定性。另一类从数据层面开展样本空间扩展,通过生成多样化的数据来覆盖潜在的噪声分布漂移。例如,Bahramali等人[21]利用数据增强技术,通过模拟带宽波动、链路拥塞等引起流量形变;Gong等人[8]和Wang等人[13]利用不同会话样本的线性组合来扩展数据;Li等人[15]利用流量整形技术强制对齐训练与测试数据的载荷大小分布;而Jansen等人[22]则尝试通过构建高保真匿名网络仿真环境生成跨位点流量。上述方法能够提升模型鲁棒性,在少样本训练[23]、概念漂移[24]、流量防御[19,25]等方面取得了较好的效果。然而,无论是模型层面的噪声容忍策略,还是数据层面的样本空间增强,其本质都是基于观测流量或特征的统计性归纳。在面对远程指纹攻击场景时,Tor中继对数据的缓存、加解密与重封装带来的报文序列的结构性变化,本质上是由协议机理驱动的动态演变过程。上述基于统计性分析的方法虽然能缓解随机噪声,却难以精准拟合由协议机制导致的结构性差异,在远程指纹攻击场景下具有局限性。

信元序列重建是一种基于协议感知的数据特征重构方法,该方法回归Tor协议机理,利用TCP/TLS元数据将结构差异的跨位点流量直接重构为结构一致的信元(Cell)序列。由于Cell是Tor网络传输过程中上层应用使用的固定长度的最小单元,这种直接重构的方法不易受到数据封装带来的报文结构性差异的影响。信元序列重建的典型代表是Wang等人[26]在复杂网络环境下,通过TLS记录载荷大小推断其中包含的Cell数量,将访问站点页面的流量重建为带方向的信元序列。相比于TCP/TLS数据报文训练的指纹模型,利用底层传输语义单元训练的模型具备更好的分类准确性[27-30],对网络条件变化引起的报文结构形变具有一定的鲁棒性。

然而,现有的协议感知信元序列重建方法[26]在隐藏服务溯源场景中仍存在一定的挑战。该方法是

一种“上下文无关”的静态映射方法,仅关注单个TLS记录的载荷长度并孤立地推导信元数量。事实上,Tor在数据传输中可能存在将Cell分片并在两个连续的报文中发送的现象。中继在转发流量时的缓存、解密与重封装行为,会进一步打乱信元分片的位置分布,导致报文序列在传输过程中发生非线性变化,这种动态变化不仅破坏了报文大小与信元数量之间的线性对应关系,还导致重构出的信元序列具有显著的估计误差,无法有效解决指纹模型性的跨位点性能衰减问题。

针对上述问题,本文提出远程信元重建方法(Remote Cell Estimator, RemoCellEst)。RemoCellEst深入分析Tor中继转发过程中数据重封装、缓冲调度引发的聚合与分片等机理,创新性地引入了跨报文的残留补偿机制。该方法不再对单个报文采取孤立分析,转而动态修正报文起始处的载荷偏移,从而能够将跨报文边界的信元分片进行逻辑对齐与整合,使其重建成跨位点一致的信元序列,在长路径、多跳转发的隐藏服务溯源场景下使模型具备更稳定的位置迁移能力。实验证明,RemoCellEst重建的信元序列训练的远程指纹模型准确率在隐藏服务溯源场景上达到了94.85%,显著优于传统重建方法训练的指纹模型。此外,该方法训练的指纹模型不会随着中继跳数的增加而产生性能削弱,解决了指纹模型对监测位点的强依赖问题,为多跳环境下的隐藏服务溯源、远程指纹攻击提供了稳健的理论和技术支撑。表1展示了RemoCellEst与不同提升指纹模型鲁棒性的方法对隐藏服务溯源能力的对比分析。

表1　隐藏服务溯源场景下不同鲁棒性方法对比

Table 1　Comparison of different robustness methods in hidden service De-anonymization

技术分类	代表工作	核心逻辑	隐藏服务溯源/位点迁移能力
环境无关特征学习	三元组/四元组[13-14,18] TAM矩阵[8,19]、流量图[20]	从模型层面,通过架构设计或特征对齐,使模型主动忽略或平滑噪声	较低,难以应对协议驱动的结构变化
样本扩展	数据增强[8,13,21] 流量整形[15] 网络仿真[22]	从数据层面,模拟网络波动或线性组合数据样本,以生成多样化样本	中,受限于仿真保真度或样本多样性
流量重建	Wang等人[26]	基于TLS载荷长度静态推断Cell数量,执行“上下文无关”映射	中,重建序列存在跨位点的估计误差差异
	RemoCellEst(本文)	模拟Tor中继转发机理,基于残留补偿机制,实现“上下文感知”的动态重建	高,还原跨位点一致的信元序列

因此,本文的核心贡献是解决了远程指纹攻击场景下跨位点一致性流量特征的直接估计问题。具体如下:

(1)深入刻画了Tor中继在转发过程中因解密重封装引发的信元分片与聚合机理,揭示了该机制下数据传输过程中导致的TCP/TLS流量报文结构动态变化。

(2)提出了上下文感知的远程信元序列重建框架RemoCellEst。该方法动态追踪并传递报文边界处的残留载荷,通过跨报文传递残留载荷长度,对齐了报文中被切断的信元边界,解决了因分片导致的信元估计偏差。RemoCellEst实现了跨位点流量向一致性信元序列的高保真还原,显著抑制了指纹模型在远程攻击场景下的性能衰减,增强了特征表征的跨位点泛化能力。

(3)构建并公开了两个多层级元数据数据集RePHST(面向6跳隐藏服务溯源)与ReSurface(面向短跳跨位点差异分析),填补了现有研究缺乏多观测点对齐数据的空白。实验表明,RemoCellEst重建的序列使指纹模型在隐藏服务溯源任务中达到了94.85%的准确率,能够有效缓解中继跳数累积带来的性能削弱,解决了指纹模型对监测和测试位点的强依赖问题,为复杂匿名网络环境下的远程指纹攻击提供了稳健的理论与技术支撑。

1　相关工作

1.1　隐藏服务机制

隐藏服务使用自认证的顶级域名地址(56字符base32的.onion),由服务的公钥派生并内嵌校验与版本信息,这使得客户端无需DNS解析即可验证服务身份,并使服务在不暴露其IP地址的前提下提供可达性。如图1所示,隐藏服务部署与客户端访问两个阶段中,分别与以下关键节点之间建立一条由多个志愿者节点组成的加密逻辑传输路径,称之为电路。同一电路中的所有信元均沿此路径进行多层加密转发,以实现通信双方的匿名性。

在隐藏服务接入Tor网络时,隐藏服务会:(1)选择若干引入节点(Introduction Points, IPs),对每个引入节点建立长期3跳HS-IP电路,并发送特殊信元ESTABLISH_INTRO;(2)生成并签名服务描述符,通过另一条3跳HS-HSDir电路上传至隐藏服务目录(Hidden Service Directory, HSDir)。随后,隐藏服务部

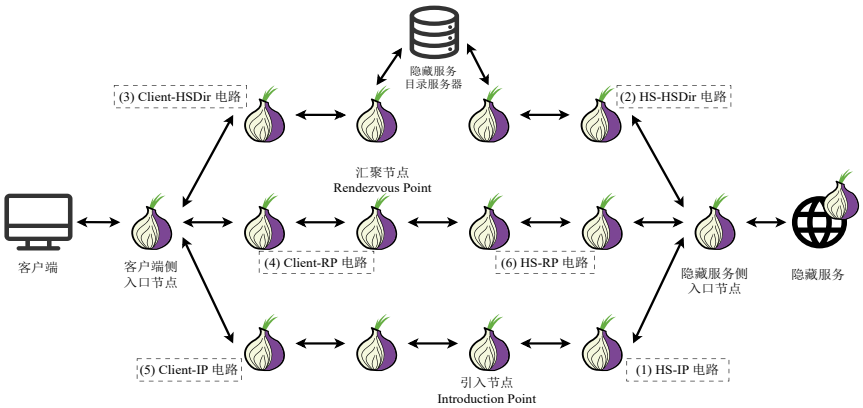


图1 隐藏服务协议

Figure 1 Hidden service protocol

署成功,可由客户端进行访问。

当客户端访问该隐藏服务时:(3)客户端据.onion地址构建3跳Client-HSDir电路获取隐藏服务的引入节点并拉取描述符;(4)客户端随机选定汇聚节点(Rendezvous Points,RPs),建立3跳Client-RP电路,发送ESTABLISH_RENDEZVOUS;(5)客户端再建立到某一引入节点的3跳Client-IP电路,通过Client-IP-HS电路向隐藏服务递交携带RP节点信息与INTRODUCE消息;(6)隐藏服务收到INTRODUCE后,建立3跳HS-RP电路至同一RP节点,发送RENDEZVOUS完成握手与密钥协商。

至此,Client-RP电路与HS-RP在RP处拼接,形成Client-RP-HS电路,应用数据随后沿该6跳路径双向传输,直至会话终止。

1.2 流量指纹攻击

1.2.1 网站指纹攻击

网站指纹攻击是一种流量分析技术,在不解析加密内容的情况下,分析浏览活动的网络流量中的元数据和模式,来识别Tor用户浏览的网页。网站指纹攻击的有效性高度依赖于从网络流量中提取的流量特征。最初,研究主要通过提取传输控制协议(Transmission Control Protocol,TCP)数据包的元数据作为流量特征,包括数据包长度、方向、时间、数量等。然而,这些基本特征易受网络拥塞、性能差异等因素影响,导致同一网站下流量模式不一致,从而降低了网站指纹攻击的有效性。为了克服这些问题,后续研究开始探索更复杂、更不易受到网络状态影响的特征,例如:相邻数据包方向一致的序列;30个非重叠数据包中的数据包数量;单方向前20个数据包的大小[31]。

网站指纹攻击最初依赖于手动特征选择,这需要研究人员根据经验和领域知识精心挑选最佳特征。然而,手动选择可能限制潜在的准确性。随着深度学习的发展,自动化特征提取算法(如通过人工神经网络)逐渐被引入。例如:Abe最早使用堆叠去噪自编码器(Stacked Denoising AutoEncoders,SDAE)[32];Rimmer引入了自动化网站指纹识别框架,包含SDAE、卷积神经网络(Convolutional Neural Network,CNN)和长短期记忆网络(Long Short-Term Memory Network,LSTM)[33];Sirinam提出了深度指纹识别(Deep Fingerprinting,DF)[34],扩展了CNN算法。

由于Tor网络流量对数据的封装,导致网站指纹攻击可以在不同的流量层进行,其中包括:TCP层,直接分析TCP/IP数据包的元数据;传输层安全协议(Transport Layer Security protocol,TLS)层,分析TLS记录的元数据;Cell层[35-36],Tor中传输的Cell的时间和方向。为了克服TCP/IP数据包的局限性,Wang等人[26]在2013年提出了一种基于桶分割(Bucket-divided)的Cell层流量重建方法,旨在通过深入解析TLS记录的载荷大小,从而估算出Tor网络中的实际Cell序列,以克服传统上直接依赖TCP数据包序列进行流量分析的局限性。此方法的核心在于深刻理解Tor网络的数据传输机制:Tor数据以固定大小的514字节Cell为基本单位进行传输,并始终进行填充以达到Cell边界,同时在传输层使用TLS协议进行加密。相较于易受TCP/IP协议层面影响导致不一致的变长数据包,Tor的Cell作为更小的传输单元,能够提供更为稳定和具有代表性的流量指纹特征。在后续研究中,这类方法被广泛适用于网站指纹攻击中[37]。

1.2.2 远程指纹攻击

随着网站指纹攻击的不断发展,部分研究假设攻击者可以在Tor出口中收集真实的Tor流量,并利用这些数据来训练网站指纹模型。因为运行出口中继的攻击者可以持续获得标记的流量,可以减轻“概念漂移(Concept Drift)”的负面影响。然而,研究发现当

监控位点从出口转移到客户端侧时,网站指纹模型的性能会急剧下降。例如,当监控50个网站时,性能从76.2%(出口训练,出口测试)下降到65.1%(出口训练,入口测试),研究猜测网络效应(如延迟)可能导致流量痕迹在入口和出口之间存在失真[38]。

近期的研究转向利用流量指纹攻击实现隐藏服务溯源,这是由于流量指纹模型可以根据加密流量识别流量来自隐藏服务域名。当敌手将流量指纹模型部署到隐藏服务入口侧(后续称为隐藏服务侧)时,预测的隐藏服务与加密流量中的IP地址相互关联,实现溯源。然而,在隐藏服务侧无法获得大量真实、有标签的隐藏流量来训练模型,在难以直接获取监控隐藏服务流量作为训练数据的情况下,如何通过可观测的"替代流量"来训练指纹模型,从而实现对隐藏服务溯源,是远程指纹攻击的挑战。

在远程隐藏服务溯源的研究中,通常以长跳作为主要威胁模型与评估条件,其根本原因在于敌手能够控制并获取带标注训练流量的观测位置,往往与目标隐藏服务入口侧的观测位置相隔多跳中继。例如,采用客户端侧作为替代位置对应跨6跳迁移[13],在基于隐藏服务代理的溯源攻击中也需要跨6跳迁移[15]。相对而言,在靠近隐藏服务的短跳(如1~2跳)条件下,中继节点通常由目标隐藏服务自主选择,敌手难以稳定控制并获取带标注训练流量,因此,短跳设置难以作为隐藏服务溯源中可稳定成立的训练与测试的位点迁移场景。

为了获取用于指纹模型训练的替代流量,Wang等人[13]假设了一个节点级敌手,在电路中客户端侧捕获Cell层流量作为替代数据,以训练模型。其关键在于将特征提取与分类解耦,允许训练模型使用的隐藏服务侧流量与被监控隐藏服务完全不同,且特征提取器只需训练一次。通过采用三元组网络(Triplet Network)学习客户端侧和隐藏服务侧数据的共同特征,并结合数据增强(随机混合锚点和客户端样本),增强模型的泛化能力,从而实现对隐藏服务的去匿名化。

Zhao等人[14]同样假设了一个节点级敌手,提出了基于四元组网络(Quadruplet Network)的指纹识别(Quadruplet Fingerprinting,QF)方法。该方法创新性地选择RP节点作为远程观测点,因为RP节点比客户端更接近洋葱服务,在特征上与隐藏服务侧流量更为相似,从而在数据可访问性和特征相似性之间取得平衡。此外,QF引入了四元组损失函数(Quadruplet Loss),通过引入额外的负样本来考虑正负样本之间的绝对距离,结合边界样本挖掘损失,以实现更小的类内差异和更大的类间差异,显著提升了模型的泛化能力和去匿名化精度。

Li等人[15]提出了一种代理指纹识别(Proxied Fingerprinting,PF)方法。该方法通过部署隐藏服务代理来捕获隐藏服务代理侧流量,解决了传统替代方案(如镜像隐藏服务流量和客户端侧流量)在响应和协议一致性方面的缺陷,使隐藏服务代理侧流量在特征上更接近隐藏服务流量。为进一步弥合跨位点流量之间因路由导致的非MTU数据包分布和额外数据包差异,PF引入了突发重塑(Burst Reshaping)技术,包含突发重构(处理数据包大小分布差异)和伪标签学习(利用未标记的隐藏服务流量迭代优化模型),从而克服路由和协议差异,最终实现对隐藏服务的有效溯源。

综上所述,虽然现有研究对比学习、流量整形来弥合位点间的分布差异,但这些方法本质上仍停留于对TCP报文的统计特征处理。由于加密报文在多跳转发过程中受中继重封装机理影响,导致流量结构差异较大,单纯的流量整形难以从根本上提取跨位点一致性的流量特征。相比之下,信元作为Tor协议的基本语义单位,其组成的信元序列比网络报文在跨位点流量上具备更高的结构一致性与位置无关性。因此,在不控制中继节点的限制下,如何通过网络报文构建跨位点结构一致的信元序列,是突破位点迁移难题、实现稳健隐藏服务溯源的关键研究方向。

1.3　数据封装协议及信元重建

Tor在中继之间先建立基于TLS-over-TCP的链路(Channel),随后双方在该链路上连续发送固定长度的Cell。Cell是Tor电路内的基本传输单元,典型长度为514字节。由于TLS记录只是外层承载容器,Cell可以嵌入在任意大小的TLS记录中发送,也可跨TLS记录分片传输,但TLS记录的封装结构不得泄露单元类型或内容信息。在每一跳中继处,TLS记录会被解密以取出Cell并完成转发处理,然后再重新封装为新的TLS记录发送到下一跳,因此,网络级敌手通常只能看到TCP/TLS层的长度与方向等元数据,而难以直接获得稳定的Cell边界与Cell序列。

现有的信元重建方法通常从捕获的TCP/IP流量中进行传输层重组,恢复连续字节流,并进一步解析TLS记录边界以得到完整的TLS记录序列。为保留方向信息,方法将客户端到服务器方向的TLS记录记为正号,将服务器到客户端方向记为负号,从而把一次会话表示为带符号的TLS记录载荷长度序列。在此基础上,该方法利用Tor的数据封装语义完成从TLS记录到Cell的映射。由于Tor在应用层以固定长度Cell作为协议语义单元,并通过TLS/TCP进行封装传输,因此TLS记录载荷长度在一定程度上反映了其

所承载的Cell数量。基于这一数据封装机理，信元重建方法对每条TLS记录的载荷长度进行量化：将其向下取整到最接近的514字节的倍数，相当于把连续长度归一到以514字节为单位的桶中（在后续称为“桶分割”方法）。完成量化后，桶分割将规整化的载荷长度除以514得到该TLS记录估计承载的Cell数，最终形成Cell方向序列作为指纹模型输入。在流量指纹攻击的研究中，为了将捕获的原始报文转化为深度学习模型可处理的序列，通常利用正负号来表征数据传输的方向。正数代表上行报文，即从客户端发往隐藏服务的方向；负数代表下行报文，即从隐藏服务返回客户端的方向。以一组TLS负载长度序列[+544, −1 088, +1 088]为例，在桶分割方法下，TLS估计对应的Cell计数为[+1, −2, +2]，最终展开得到Cell序列[+1, −1, −1, +1, +1]。

然而，需要注意的是，向下取整会产生不可整除的余数部分。桶分割方法将其视为TLS记录载荷中通常包含与Cell字节流无关的记录层开销，例如初始化向量（Initialization Vector）、认证标签等。为此，该方法通常直接丢弃这部分剩余字节，以防止错误计数。基于上述步骤，桶分割方法在一定程度上能削弱数据缓存、报文聚合与分片带来的报文序列的结构性变化，可以从环境不同导致的报文序列结构差异的情况下提取结构一致的信元序列，从而提高模型鲁棒性。

2 威胁模型

本文假设一个网络级敌手，例如，互联网服务提供商（Internet Service Provider，ISP）、自治系统（Autonomous System，AS）运营商或骨干网提供商，通过被动监控多个非受控节点，来实现对未知加密流量的观测和捕获，进而实现隐藏服务溯源。与需要攻陷或控制Tor中继节点的节点级对手不同，网络级对手依托其在网络拓扑中的有利位置，作为远程流量指纹的监控位置。

如图2所示，敌手选择暗网中部分隐藏服务作为监控隐藏服务。为了获取标注流量来训练远程指纹模型，该敌手根据PF[15]提出的训练数据的流量捕获方法，在Tor网络中部署隐藏服务代理，并将其作为自己的可控位点。之后，敌手通过客户端访问隐藏服务代理，捕获访问过程中隐藏服务代理与其入口节点之间的加密流量，并根据隐藏服务代理设置的监控HS域名作为该加密流量的标签。敌手从这些训练数据中提取TCP层的时间戳、方向、载荷大小等元信息，并用信元重建方法将TCP层或TLS层流量重建为Cell层流量，基于已有的流量（网站）指纹模型架构来训练远程指纹模型。

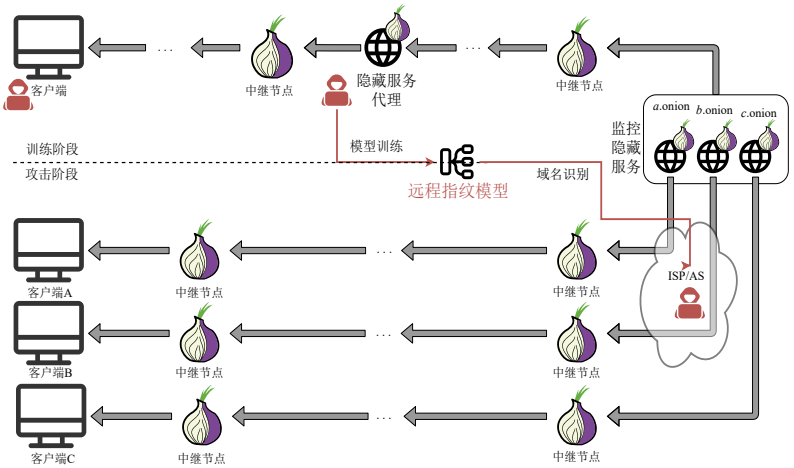


图2 威胁模型

Figure 2 Threat model

在攻击阶段，敌手在其所处的ISP/AS中被动观察流量，并将捕获到的流量利用与训练阶段相同的方法，交由远程指纹模型预测流量来自哪个隐藏服务。当敌手监控的隐藏服务与其中继节点的流量经过敌手所处的ISP/AS域时（如a.onion，b.oinon），敌手可将模型预测的隐藏服务域名与流量中隐藏服务的IP地址相互关联。

本文提出的威胁模型具备可行性，网络级对手无需控制大量Tor中继节点（这对节点级对手往往难以实现），而是通过利用大规模网络流量可见性，同时监控众多Tor流量。因此，敌手面临的核心挑战在于如何从可观测的TCP/TLS流量中，恢复结构一致且语义稳定的Cell层流量，从而实现有效的远程流量攻击与隐藏服务溯源。

3 Tor跨位点流量序列结构性差异分析

为了刻画Tor加密流量在电路中的不同位置下，TCP/TLS/Cell层的流量差异及其成因，本文从Tor协

议设计、中继传输等方面给出机理解释,分析中继数据传输引起的报文序列结构变化及该变化在不同流量层下表现出的差异性。

3.1 Chunk机制及Cell分片

Tor的数据封装协议表明,Cell可以跨TLS记录分片传输。通常情况下,Tor采用分块(Chunk)机制限制单条TLS记录中的字节数[39,15],默认分块阈值为4 065字节,因此单次写入至TCP缓冲区的总字节数不超过4 065字节。如图3所示,一个最大的TLS记录能够聚合7个完整的Cell(总计3 598字节)。为降低等待时延,记录中常追加下一枚Cell的部分字节以填充至阈值,形成Cell分片。随后,该TLS记录在TCP层按最大报文长度或最大传输单元进一步分段并发送。

3.2 Tor中继解密与重封装

图4(a)展示了单向流量下,在电路中不同位置TCP、TLS、Cell层的流量差异。Tor中继在接收入站

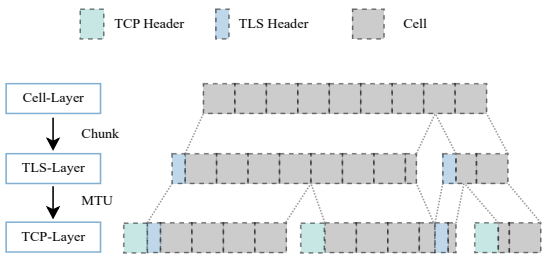
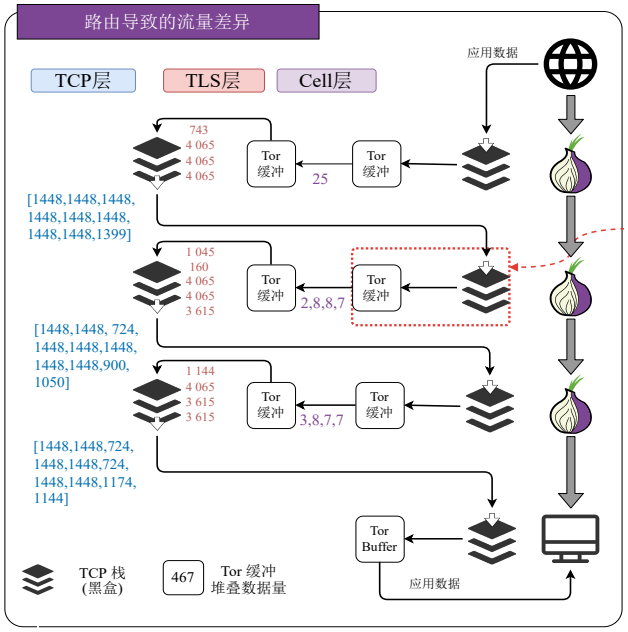


图3　Tor数据封装机理

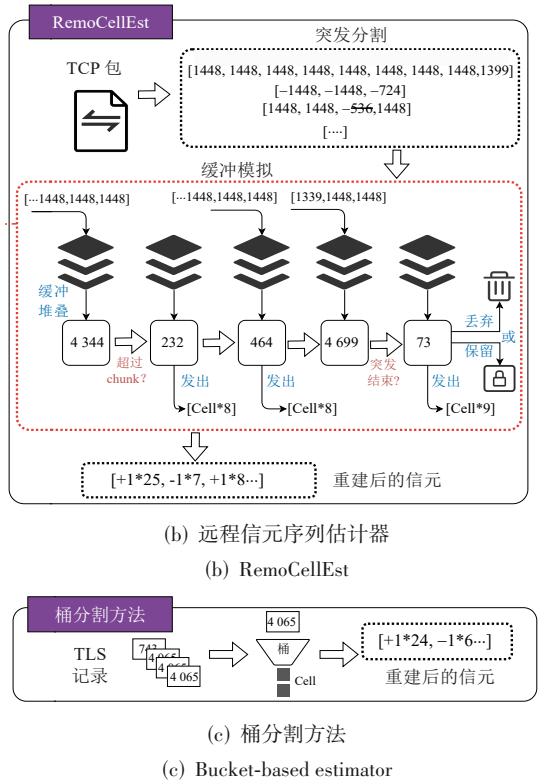
Figure 3　Encapsulation mechanism of Tor data

TCP报文后,由TCP协议栈先对入站报文去重、重排和按序重组,将同一连接内快速到达的全最大报文(Maximum Segment Size,MSS)段长度堆叠在重组缓冲中,形成短时"到达突发(Burst)"。应用层读取一条连续的字节流,并按记录边界解析并完成解密与完整性校验。只有校验通过的TLS记录才会上交给Tor进程,在这个过程中,TLS只维护应用层的完整性,不感知也不约束Cell的对齐与完整性。



(a) TCP/TLS/Cell层跨位点流量表现在不同Tor中继下的差异

(a) Differences in TCP/TLS/Cell-layer cross-position traffic performance on different Tor relays



(b) 远程信元序列估计器

(b) RemoCellEst

(c) 桶分割方法

(c) Bucket-based estimator

图4　RemoCellEst机理及其与桶分割方法的对比

Figure 4　The mechanism of RemoCellEst and its comparison with bucket-based estimator

Tor进程在解密后的字节流上以514字节为粒度读取Cell,并对Cell的完整性进行校验。若记录末尾不足514字节,则作为缓冲残留,待下一条TLS补齐后再交付处理,从而完成跨TLS记录、跨TCP分段的分片重组。对还原出的完整Cell,Tor会执行解密与协议处理,部分Cell还会校验其端到端摘要/识别字段以检测篡改或错位。

Tor将待发送Cell写入发送缓冲,将多个Cell聚

合为TLS记录。如果Cell载荷未超过默认Chunk大小,则会直接聚合成1条TLS记录;若Cell载荷超过了默认Chunk大小,则会将这些Cell聚合为多条TLS记录。随后,这些TLS会下交TCP,TCP再根据MSS、拥塞控制与重传策略切分并发送,这常表现为连续的全MSS报文"发送突发"。接收侧依次完成记录级完整性校验与Cell粒度的重组,确保即便存在跨TLS的Cell分片,仍可按514字节正确复原Cell。

在传输过程中,Cell分片带来的缓冲残留是造成跨位点流量差异的主要原因之一。例如,当一个4 065字节大小的TLS记录在经过Tor中继后,其中7个完整的Cell会被Tor立即处理并释放,由TCP发送给下一个中继。然而,Cell分片无法解析,会被作为残留缓冲被Tor保留,造成TLS记录大小差异(从4 065字节变成3 615字节)。TLS结构改变进一步传导到TCP层的分段,从而造成TCP层流量差异(从[1 448,1 448,1 448]变成[1 448,1 448,724])。尽管TCP/TLS层的表示发生漂移,Cell层的语义顺序与计数保持一致;即使Cell分片被中继暂存,后续TLS记录的接收与解析会将其与后续字节拼接补齐,复原为完整Cell并按原电路时序转发。

3.3 Tor中继解密与重封装

Tor的TLS记录中不可避免地包含Cell分片,在正常解析过程中,这些分片会被写入Tor的残留缓冲,等待后续记录补齐后再还原为完整Cell。相比之下,现有的桶分割方法在重建Cell层流量时不断丢弃TLS记录中的残留字节,也会将Cell分片直接丢弃掉,导致重建序列相对真实序列出现单调累积的计数偏差。本文将这种单调累积的计数偏差称为累计估计误差。

在面向隐藏服务溯源场景下,累计估计误差与被丢弃的残留缓冲数量直接相关,表现为:

(1)累计估计误差与传输数据量相关。传输数据越多,需要构建TLS记录的次数越多,TLS中承载的Cell分片就越多,缓冲残留随总传输数据规模逐渐增长,累计估计误差会不断变大。

(2)数据发起方一侧误差最大。隐藏服务作为响应的主要发送方,下行数据量大,缓冲区堆叠过多,Tor往往以分块阈值构建TLS记录。在这个过程中,产生的Cell分片与残留缓冲也最多。在桶分割方法下,隐藏服务侧的累计估计误差最大。

(3)跨位点的累计估计误差具有差异。随着流量在各中继间转发,即使TLS记录中存在Cell分片,其会在每一跳上都会被Tor中继重新拼接,还原为完整Cell,并在下一个TLS记录中转发到后续中继中。越远离隐藏服务,上一跳遗留的缓冲残留可能会被Tor中继拼接为完整的Cell而逐渐消除。因此,流量观测点之间间隔的跳数越多,缓冲残留规模相差越大。基于桶分割重建出的信元序列在跨位点下的累计估计误差不一致,导致指纹模型在位点迁移时性能削弱严重。

上述分析结果表明,在Tor协议引导下,中继在数据转发过程中会对TLS记录进行解析和重封装,进而会改变报文的结构形态,导致报文的长度序列在传输过程中发生了复杂的非线性演变。如果仅依赖现有的桶分割方法,仍然难以有效实现指纹模型的位点迁移。如何基于Tor的转发机理,显式建模并跟踪缓冲残留,降低重建信元与真实信元之间的累计估计误差,是提升远程流量指纹模型鲁棒性,实现隐藏服务溯源的关键挑战之一。

4 远程信元序列估计RemoCellEst

本文提出一种上下文感知的远程信元序列估计RemoCellEst,用于从加密TCP报文中重建跨位点一致性的Tor信元序列。该框架不是传统方法中对单个报文的孤立分析,而是通过动态追踪中继转发过程中的缓冲残留,识别和重建报文中的信元分片以实现上下文关联。这一机制能有效缓解由加解密与重封装引发的缓冲残留及累计估计误差,从而为复杂多跳环境下的隐藏服务溯源提供稳健的技术支撑。

RemoCellEst的处理流程如图3(b)所示:首先,该方法将双向TCP报文按方向划分为一系列突发段;随后,在每个突发段内通过缓冲模拟估计该段的Cell数目;最后,依据突发段在原始TCP流中的时间与方向顺序,串接各段的估计结果,得到完整的Cell序列。该方法建立在Tor协议端到端语义一致的基础之上,通过上下文感知的缓冲残留追踪,能够还原中继转发过程中的信元分片的重组逻辑,使得无论经过多少次中继,可以在跨监控位置捕获到的结构差异的报文序列中重建出一致的信元序列,从而实现了对跳数的鲁棒性。

4.1 突发分割

突发分割将双向TCP流细分为方向一致、时间连续的突发段,目的在于降低段尾Cell碎片对估计的干扰,并抑制误差在跨位置观测中的传播。突发(Burst)本身由报文方向进行分割,被描述为同一方向的一串连续的TCP报文。在方向作为分割规则之外,采用两条切分规则:

(1)包含完整Cell的TCP报文。在同向连续包序列中,若出现同向的非MTU包,其有效负载大小落在536±4字节或1 045±4字节区间,则该TCP包很可能承载了1个或2个完整Cell。此时在该包之前切分,

使其作为新突发段的起始。这样可避免突发段过大,减少TLS头部聚合超过Cell大小导致的额外Cell计数偏差。

(2)忽略反向的单Cell包。若一串同向的满MTU包被一个反向的非MTU包(大小在536±4字节区间)打断,则该反向包很可能为单个控制类Cell(如relay_sendme)。这类信元为中继间信元,不会随着电路传输,因此不以该反向包作为切分点,保持前后同向突发的连续性。原因在于前一突发往往包含未闭合的Cell碎片,强行切分会增加估计误差的风险。

遵循上述规则,可将单条双向TCP流划分为多段方向一致、时间连续的突发段,随后对每个突发段独立执行缓冲区仿真。

4.2　缓冲模拟

缓冲区仿真在突发段内模拟Tor发送缓冲的行为。首先,将该段内同向的TCP包按到达顺序重组为连续字节流,顺次写入模拟的Tor缓冲区。当累计字节达到或超过分块阈值(默认4 056字节)时,触发一次发出事件(Emit):从缓冲区中释放尽可能多的完整Cell(每个514字节),将不足514字节的剩余记为残留。

例如,若该突发段内连续到达的TCP有效负载为[1 448, 1 448, 1 448, 1 448],当写入第3个1 448字节后,累计达到4 344字节并首次越过阈值,触发一次Emit,释放8个Cell,并在缓冲区中保留232字节残留。随着后续报文继续到达,缓冲区再次累积并在越阈时触发新的Emit事件(例如,再次释放8个Cell)。

在突发段结束时,若缓冲区仍有不足1个Cell的残留(例如73字节),则依据最后一个包的形态进行处理:若段末为满MTU包(如以太网MTU为1 500时常见的1 448字节TCP负载),则将该残留作为同向下一突发段的初始缓冲,避免因人为截断导致的低估;否则丢弃残留,以防跨突发传递造成的误计数。

4.3　时间戳估计

由于网络级敌手无法直接获得中继内部逐Cell的精确释放时刻,因此本文对被释放Cell的时间戳采用受限可观测条件下的线性插值近似。对每个被释放的Cell赋予一个线性插值的时间戳:若一次Emit由时间窗口$[T_i, T_{i+1}]$内的触发报文所致,则将k个释放的Cell的时间戳均匀分布在该窗口内,记为Linear(T_i, T_{i+1}, k)。

RemoCellEst对时间戳的处理属于一种时间近似。然而,本文并非在任意长度的双向序列上直接插值时间戳,而是利用RemoCellEst的结构化重建过程,尽可能压缩时间近似误差的来源与传播范围,从而使时间戳分配在统计意义上更接近真实释放过程。

首先,突发分割将原始双向TCP流切分为方向一致、时间连续的单向突发段,并在方向切换及可能承载完整Cell的非MTU包等位置进行分段,使一次Emit对应的触发窗口通常落在单一方向的连续到达区间内。该设计降低了上下行报文交织造成的时间噪声,减少上下行Cell时间戳相互穿插、顺序颠倒等典型错误,使时间误差主要表现为突发内部的细粒度抖动,而非跨方向的结构性错序。

其次,缓冲模拟以分块阈值触发Emit事件,并仅在每次Emit的触发窗口时间内对被释放的Cell进行时间分配。该方式将时间估计误差限制在单次释放事件内部。相比于直接按全序列均匀分配时间戳的做法,RemoCellEst的事件边界为时间近似提供了更强的约束。

最后,RemoCellEst在段末残留处理上避免了不必要的截断与错配:当段末为满MTU包时保留残留并传递至同向下一段,使由残留补齐触发的后续Emit不会被人为提前或延后到错误的时间窗口内,从而减少由于分段边界选择不当引入的错位。尽管Cell的精确时刻不可观测,但本文基于缓冲追踪设计的信元估计方法将时间估计限制在Emit触发窗口段中,以降低时间估计误差对指纹模型的干扰风险。

4.4　信元序列重建

缓冲区仿真结束后,每个突发段输出该段内释放的Cell数与方向的重建片段,例如[+1]*25、[−1]*7、[+1]*8,其中"+1"与"−1"分别表示两个传输方向。依据各突发段在原TCP流中的先后与方向,将这些重建片段按时间顺序串接,得到全局的Cell层重建流量。

算法1展示了RemoCellEst的整体执行流程。该过程可归纳为3个关键阶段,突发分割将连续的TCP流分段成为具备完整性信元的观测区间,限制缓冲残留带来的估计误差。缓冲区模拟维护跨报文传递的残留载荷来实现信元的上下文感知。信元发送阶段通过分块阈值触发Emit事件,从而将结构差异的报文序列还原为定长信元序列,并利用受限窗口下的线性插值对信元赋予时间戳。最后,根据原始报文序列的时序关系,将各段重建结果线性串接,输出重建信元序列。

5　数据集

为评估RemoCellEst重建方法的有效性,本文构建并采集了RePHST和ReSurface两个数据集,覆盖不同类型的Tor远程流量分析场景。两个数据集均在真实Tor网络中采集,除用于数据采集与控制实验流程的客户端/代理等关键节点由本文部署外,其余转发路径均由Tor网络中的志愿者中继节点承担,流量在实际网络负载、链路拥塞等条件下逐跳传输与封

算法 1 RemoCellEst:远程信元序列估计算法

输入:TCP数据包序列 $P=\{(p_i, t_i)\}$,其中,p_i为有效载荷大小,t_i为时间戳

输出:重建信元序列 $C=\{(d_j, \tau_j)\}$

初始化 $B\leftarrow\varnothing$ //当前突发缓冲区
初始化 $d_{\text{curr}}\leftarrow$ None //当前方向
foreach $(p_i, t_i)\in P$ do
  $d_i\leftarrow \text{sign}(p_i)$;
  if d_{curr} is None then
    $d_{\text{curr}}\leftarrow d_i$;
  end
  // 检测完整TLS记录作为分割点
  if $d_i=d_{\text{curr}}$ and $|p_i|\in\{536\pm4, 1\,045\pm4\}$ then
    $C\leftarrow C\cup$EstimateCells(B, d_{curr});
    $B\leftarrow\varnothing$;
  end
  // 方向切换并跳过控制信元
  else if $d_i\neq d_{\text{curr}}$ then
    if $|p_i|\in[536\pm4]$ and $\text{sign}(p_{i+1})=d_{\text{curr}}$ then
      $C\leftarrow C\cup\{(d_i, t_i)\}$; // 独立控制信元
      continue;
    end
    else
      $C\leftarrow C\cup$EstimateCells(B, d_{curr});
      $B\leftarrow\varnothing$, $d_{\text{curr}}\leftarrow d_i$;
    end
  end
  $B\leftarrow B\cup\{(p_i, t_i)\}$;
end
return $C\cup$EstimateCells(B, d_{curr});

装,从而天然包含真实环境中由于中继转发带来的报文结构差异。其中,RePHST数据集中包含隐藏服务代理侧流量与隐藏服务侧流量,两侧流量分别相距6个中继节点。ReSurface数据集包含客户端-入口节点、入口节点-中间节点、中间节点-出口节点3个关键位置的Tor流量,两个位置的流量间相距1~2个中继节点。

5.1 RePHST数据集

由于Tor网络中的监控隐藏服务不受控制,直接获取其服务端侧流量在技术上较为困难,且出于伦理与合规考虑,主动对真实在线隐藏服务实施溯源也是不可接受的。鉴于此,本研究基于Li等人[15]的思路部署可控的隐藏服务镜像作为监控隐藏服务,并在其前端配置隐藏服务代理,用于转发客户端与隐藏服务之间的通信。基于该架构,采集到包含多流量层(TCP/TLS/Cell)的远程指纹攻击数据集,从而在不干扰真实Tor隐藏服务的前提下,对所提方法在隐藏服务溯源场景下的有效性进行系统评估。

本文使用TorSocks和HTTrack从Tor网络中爬取了部分隐藏服务托管的内容,以此作为监控隐藏服务的托管内容。之后,本文在真实的Tor网络中部署监控隐藏服务,每个监控隐藏服务托管不同的资源,例如HTML页面、JPG图片等。之后,在Tor网络中部署隐藏服务代理,用于转发客户端与监控隐藏服务之间的通信。本文控制客户端分别访问隐藏服务代理和隐藏服务,在访问过程中捕获隐藏服务代理侧和隐藏服务侧流量。在每次访问过程中,本文在两个观测点均记录双向流,并从原始流量中提取每个TCP报文和TLS记录的时间戳与载荷大小,从而得到每次访问对应的网络层(TCP与TLS)流量轨迹。之后,本文根据电路ID从Tor日志中提取真实的Cell层流量元数据,包括时间戳、方向和Cell类型。

最终的RePHST数据集如表2所示,共包含58个监控隐藏服务,每个类别包含200个在不同会话下采集的双向流量样本。该数据集在提供可控实验环境以评估隐藏服务追踪攻击的同时,包含了此前缺失的细粒度TLS层和Cell层流量信息。

表2 RePHST数据集规模

Table 2 The scale of RePHST dataset

流量捕获位置	流量层	数据量	监控隐藏服务数量	流量踪迹
隐藏服务代理(PHST)	TCP/TLS/Cell	11 600	58	200
隐藏服务侧(HS)	TCP/TLS/Cell	11 600		

5.2 ReSurface数据集

Cherubin[19]首次展示了远程流量指纹攻击,其使用出口流量训练流量指纹模型,并在客户端侧流量进行测试,揭示了远程流量指纹的限制[38]。受该工作启发,本文设计并构建了访问明网的跨位点流量数据集,称为ReSurface数据集。具体而言,通过在受控条件下运行一个中间中继节点,并引导客户端访问公共网络服务,在3个关键位置采集流量轨迹:客户端到入口节点流量Client-Entry、入口节点到中间节点流量Entry-Middle以及中间节点到出口节点流量Middle-Exit。该设置构建了一个可用于系统评估远程指纹识别场景的数据集,在Cherubin[19]的研究基础上进一步扩展,并为跨观测点流量相关性分析提供了更可靠的数据支撑。

客户端通过Stem从Tor共识中随机选择不受控的入口和出口中继,并与受控的中间中继组成3跳电路。每次访问过程中,在3个观测点采集双向流量。同时,从采集的流量中提取每个TCP报文和TLS记录的时间戳与载荷大小。最终得到的数据集ReSurface如表3所示,包含60个不同的监控明网服务,每个类

别至少150个流量样本。

表3　ReSurface数据集规模

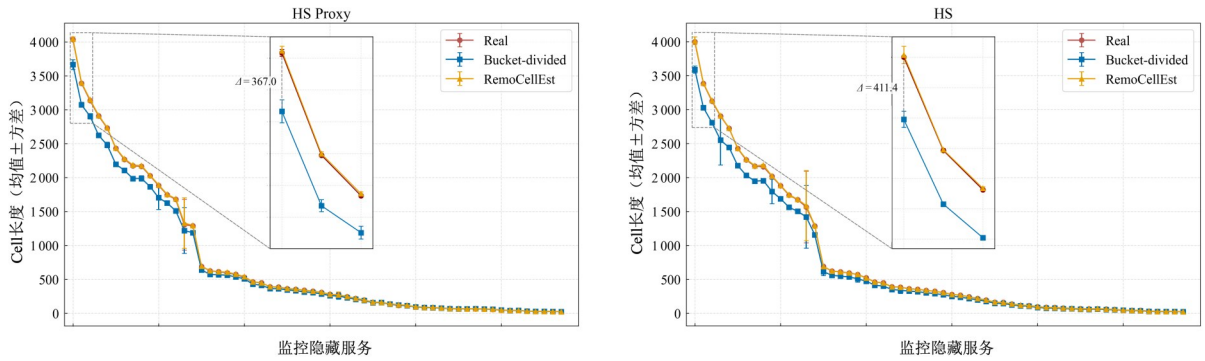
Table 3　The scale of ReSurface dataset

流量位置	流量层	数据量	监控明网服务数量	流量踪迹
Client-Entry	TCP/TLS	11 226	60	150+
Entry-Middle	TCP/TLS	11 226		
Middle-Exit	TCP/TLS	11 226		

5.3　RemoCellEst与桶分割下重建信元序列对比

为直观评估RemoCellEst与桶分割方法在Cell层重建上的差异,本文在RePHST数据集下,以真实Cell序列为基准,对两种方法重建得到的Cell序列进行定量对比。具体而言,本文在捕获的RePHST数据集下,统计隐藏服务侧流量样本和隐藏服务代理侧流量样本的重建序列与真实序列在Cell数量上的偏差,并将该偏差作为估计误差指标。考虑到不同隐藏服务承载的内容与交互模式存在差异,本文按隐藏服务维度进行分组统计,对每个隐藏服务独立汇总50条访问轨迹的误差分布,从而考察估计误差在不同服务之间的稳定性与可变性,以及两种方法在各隐藏服务上的相对优劣。

图5展示了隐藏服务代理侧和隐藏服务侧跨位点流量下基于Bucket-divided和RemoCellEst重建后的Cell计数差异。从两处观测点可以看出,随着隐藏服务流量规模增大,Bucket-divided的Cell数量估计误差总体呈上升趋势,尤其在流量规模较大的服务上出现明显差异;而RemoCellEst曲线始终紧贴真实Cell计数基线,误差明显更小。此外,图5(a)与图5(b)中RemoCellEst的表现高度一致,表明RemoCellEst在代理侧与隐藏服务侧均能提供稳定的Cell层重建精度,具有较好的鲁棒性。



(a) 隐藏服务代理侧流量　　(b) 隐藏服务侧流量

(a) Proxy hidden service-side traffic　　(b) Target hidden service-side traffic

注:Real为真实Cell计数,Bucket-divided为桶分割方法。

图5　不同隐藏服务下Cell数量估计误差对比结果

Figure 5　Comparison of the estimation errors on Cell counts across different hidden services

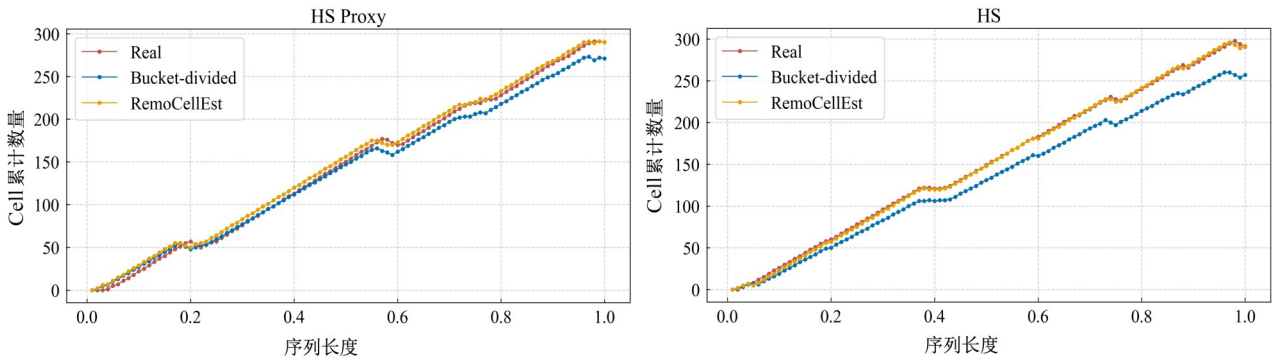
进一步地,本文抽取了隐藏服务代理侧流量与隐藏服务侧流量的真实Cell序列与两种方法在累积Cell数量上的对比。从图6可以看出,两处观测点的曲线形态表明,RemoCellEst在整个序列范围内几乎与真实曲线重合,无论在隐藏服务代理侧还是隐藏服务侧,其累计Cell数量的增长斜率和细节变化都能准确跟踪真实Cell序列。相比之下,Bucket-divided方法在中后段开始出现明显偏离,累计曲线整体偏低,展现出一定程度的累计估计误差,且偏差随序列增长而逐步放大。这一现象在隐藏服务侧更为明显,表明Bucket-divided方法下在跨位点场景下误差具有明显的累积效应。

上述结果表明,Bucket-divided方法重建得到的信元序列存在显著的累计估计误差,且该误差随传输数据量的增加而持续放大。对于隐藏服务下行流量,由于在多跳中继处反复经历解密和重封装,在到达隐藏服务代理侧时累计误差有所缓解,这一现象与前文关于残留缓冲作用机制的分析相吻合。相比之下,在RemoCellEst方法下,由于缓冲模拟跟踪并利用缓冲区残留,基于上下文相关的重建Cell序列在计数与时序上均与真实序列高度一致,几乎不出现随流量规模累积的系统性偏差,这充分体现了所提方法在跨位置Cell层重建精度与稳定性方面的优势。

6　实验

6.1　RemoCellEst在不同指纹模型架构下的隐藏服务溯源能力对比

为了评估不同流量层和指纹模型在远程指纹攻击中的有效性,本节实验使用RePHST数据集的6跳场景下,验证了各种模型的隐藏服务溯源能力。正如



(a) 隐藏服务代理侧流量 (b) 隐藏服务侧流量

(a) Proxy hidden service-side traffic (b) Target hidden service-side traffic

注:横轴为序列进度(例如0.1表示取该序列前10%的元素),纵轴为从序列起点累加到该位置的Cell数量。其中,Real为真实Cell计数,Bucket-divided为桶分割方法。

图6 隐藏服务代理侧与隐藏服务侧流量样本下真实Cell序列与两种方法在累积Cell数量上的对比

Figure 6 Comparison of cumulative Cell sequences between real Cell sequence and two estimation methods under the hidden service proxy side and the hidden service side traces

在威胁模型中所述,网络级攻击者只能观察到TCP和TLS层的流量,并且可以通过桶分割或RemoCellEst来重建Cell层的流量。因此,本文使用上述4种不同的流量元数据,并基于多种最先进的指纹识别分类器,包括网站指纹模型(DF[34]、Var-CNN[40]、WF-Transformer[7])和远程指纹识别模型方法(PF[15]),实现隐藏服务溯源。

在这些方法中,DF通过输入流量元数据实现端到端指纹模型;Var-CNN优化了模型输入,在流量元数据的基础上,输入了部分统计特征(如包数量、总传输字节等);WF-Transformer在Cell层上统计得到流量特征,输入到模型中进行训练。在远程流量指纹攻击中,PF利用伪标签学习,在训练过程中逐渐观测和学习测试数据分布。

从表4的实验结果可以观察到,在隐藏服务溯源的跨位点识别场景下,本文提出的RemoCellEst机制表现出了显著的优越性。基于原始TCP层和TLS层的指纹模型识别准确率普遍较低,例如DF模型在TLS层下的准确率仅为0.679 3。这证实了在远程溯源场景中,网络抖动及中继节点的重封装行为使报文的时空统计特征发生了剧烈形变,导致模型在跨位点测试时性能大幅下滑。相比之下,RemoCellEst在所有测试模型中均取得了最高的准确率,均超过了0.92(除WF-Transformer外)。以DF模型为例,其准确率达到0.948 5,相较于效果次优的特征层提升了15.23%;在Var-CNN上,性能提升更是达到了16.76%。这一数据有力证明基于Cell这一底层传输语义能提供更具鲁棒性的指纹表示。

进一步对比发现,现有的基础重建方法(Bucket-based)虽然在一定程度上缓解了报文结构差异,但其准确率(最高0.827 3)仍远低于本文方法。这是因为Bucket-based方法属于“上下文无关”的静态映射,忽略了Tor转发过程中的缓冲残留机理,导致重建序列存在严重的累计估计误差。RemoCellEst通过模拟Tor中继的残留补偿机制,实现了上下文相关的动态重建,能够有效消除由协议机理驱动的报文结构差异。即使在识别难度较大的WF-Transformer模型上,RemoCellEst依然带来了7.42%的性能提升。实验结果表明,RemoCellEst能够有效校正非线性结构演变,还原出跨位点一致的信元序列原貌,在多种指纹识别架构下均展现出极强的跨位点迁移能力与普适性。

表4 隐藏服务溯源场景下,TCP、TLS及重建信元下不同流量指纹模型的准确率

Table 4 Accuracy of different traffic fingerprinting models under TCP, TLS, and reconstructed cells

模型	TCP	TLS	Bucket-based	RemoCellEst
DF[34]	0.796 2	0.679 3	0.750 9	**0.948 5** (+0.152 3)
Var-CNN[40]	0.714 1	0.610 2	0.752 6	**0.920 2** (+0.167 6)
WF-Transformer[7]	—	—	0.649 8	**0.724 0** (+0.074 2)
PF[15]	0.858 1	0.637 8	0.827 3	**0.923 0** (+0.064 9)

注:加粗代表溯源效果最优的流量层,下划线代表溯源效果次优的流量层,括号内为RemoCellEst相较于效果次好的流量层的性能提升。

6.2 不同训练范式在隐藏服务远程迁移场景下的有效性分析

现有提升鲁棒性的代表性方法多在可观测的网络报文特征层面进行优化,通过训练策略或样本构造缓解噪声与分布偏移,但对Tor中继逐跳解密与重封装引发的结构性演化缺乏显式刻画,因此,在远程位点迁移的隐藏服务溯源场景中,其收益可能受限。为此,本文选取并实现了覆盖不同范式的代表方法进行对比:包括采用三元训练的TF[18],用于代表训练策略驱动的跨条件一致性学习;两类数据增强方法NetAugment[21]与WFCAT[8],用于代表通过样本构造扩展训练分布的思路;结合三元训练与数据增强的Wang-Mix[13],用于评估两类训练期鲁棒化叠加后的效果。所有方法在相同的远程迁移设置下训练与测试,以分析在跨位点结构差异显著的条件下,不同训练鲁棒化策略后模型准确率。

上述鲁棒化方法在原始工作中可能基于不同的报文层级或输入结构提出,但其共同目标是学习对训练和测试样本之间更不敏感的判别表示,理论上应具备一定的跨输入结构迁移能力。基于这一认识,本文在同一威胁模型与同一远程迁移设置下,将这些方法分别作用于3种可获得的输入结构,并报告其在不同输入下的性能差异:其一,为远程敌手直接可观测的TCP层流量;其二,为现有协议感知重建中采用的桶分割方法;其三,为本文提出的RemoCellEst重建Cell序列。另一方面,由于隐藏服务溯源的敌手无法直接获得隐藏服务侧的带标注流量,部分数据增强策略在迁移到本文场景时需要做相应约束:对于WFCAT与Wang-Mix,其增强样本由训练集内部的会话混合/合成产生,因此混淆与组合操作仅在隐藏服务代理侧的带标注流量之间进行,而不涉及隐藏服务侧流量;对于NetAugment,其通过对网络条件扰动进行模拟来生成新样本,本文在训练阶段对每条隐藏服务代理侧流量额外生成5条增强流量,以保证增强强度在可控成本下与原方法设定保持一致。

从表5中结果可以看到,在远程位点迁移的隐藏服务溯源场景下,部分鲁棒化方法在可直接观测的TCP层流量上能够带来一定收益,但仍然难以精准拟合跨位点结构性差异。例如:TF在TCP层的准确率与端到端基线DF相近,说明仅依赖训练策略并不能稳定获得显著优势;WFCAT在TCP层进一步下降,也反映出部分增强/混淆策略在该威胁模型下并不适配。更值得注意的是,当这些方法迁移到桶分割重建的Cell序列输入时,性能出现明显不稳定甚至大幅退化,尤其是TF与Wang-Mix在桶分割下显著失效,表明在远程多跳条件下,桶分割带来的系统性重建偏差会主导模型性能,使训练期鲁棒化难以发挥作用。

表5　不同训练范式对隐藏服务溯源准确率的影响

Table 5　The accuracy of hidden service de-anonymization under different training strategy

模型	TCP	桶分割	RemoCellEst
DF[34]	0.796 2	0.750 9	**0.948 5** (+0.152 3)
TF[18]	0.780 1	0.167 5	**0.937 4** (+0.157 3)
NetAugment[21]	0.815 5	0.793 7	**0.946 6** (+0.131 1)
WFCAT[8]	0.724 7	0.704 9	**0.913 5** (+0.188 8)
Wang-Mix[13]	0.814 5	0.312 8	**0.937 5** (+0.123 0)

注:加粗代表溯源效果最优的流量层,下划线代表溯源效果次优的流量层,括号内为RemoCellEst相较于效果次好的流量层的性能提升。

与上述现象形成对比的是,RemoCellEst在所有对比方法上均给出了最优或接近最优的溯源效果,并且相对于同一方法下的次优输入表现出稳定提升。这一结果说明,在跨位点结构差异显著的远程场景中,决定性能上限的关键因素首先是输入表示的一致性与可比性。当重建表示能够更好地对齐Tor的协议语义并抑制累计估计误差时,无论采用端到端学习(DF)、对比学习训练策略(TF)、数据增强(NetAugment、WFCAT)还是二者叠加(Wang-Mix),模型都能恢复到较高的识别水平。总而言之,在隐藏服务溯源场景中,基于协议驱动的流量重建方法更能针对位点迁移的主要挑战,成为提升隐藏服务溯源性能的关键方法。

6.3 时间戳近似对隐藏服务溯源准确率的影响分析

尽管RemoCellEst在重建流程中通过事件边界约束与结构化切分降低了时间戳近似误差的传播范围,但在网络级可观测条件下,线性插值仍难以刻画中继内部排队与调度引起的细粒度时序变化。为避免时间近似成为影响结论的潜在因素,有必要在隐藏服务溯源场景中系统评估该时间处理对不同训练范式的实际影响。基于这一前提,本文按不同流量结构(TCP/桶分割/RemoCellEst)与输入特征(结构主导/时间主导/时间与方向联合/时间鲁棒表示学习)开展实验对比,分别评估以方向与结构为主的模型DF[34]、仅以时间序列为输入的对照模型、时间与方向联合表示的Tik-Tok[41],以及基于时间信息的鲁棒表示学习方法RF[19],以量化时间信息在隐藏服务溯源中的重要性,评估时间近似对模型判别能力的敏感性。值得注意的是,桶分割方法将TLS记录长度量化并展开为Cell序列,但仅提供TLS记录的时间戳,未定义记录内各Cell的时间分配规则。为在时间相关模型的对照实验中统一输入格式,本文采用一种最小假设的时间赋值策略:将该TLS记录的时间戳赋给由该记录展

开得到的所有 Cell。该处理仅用于对照评估，其不刻画记录内 Cell 的细粒度时序差异。

表 6 的实验结果表明，仅使用时间戳序列的对照模型整体识别能力有限，表明在远程隐藏服务溯源设置下，时间信息难以单独构成稳定的判别线索。在 RemoCellEst 重建的输入表示下，结构主导的指纹模型在不显式引入时间戳信息时反而能够取得更优或更稳定的性能。这一现象说明，在远程跨位点场景中，RemoCellEst 带来的主要收益来源于对报文序列结构的一致性恢复，而线性插值时间戳所提供的细粒度时序信息并未形成稳定的额外增益。造成这一现象的一个关键原因在于隐藏服务溯源与传统网站指纹攻击的可观测流量形态存在差异：网站指纹攻击通常能够覆盖连接建立与页面加载初期的密集交互阶段，时间特征更具一致性与区分度；而隐藏服务一旦建立长期电路与连接后，观测流量往往夹杂大量维持连接的间歇性报文，其时间间隔更长且更不稳定，使跨位点时间信息更稀疏，从而削弱以时间为核心的表示（如仅时间戳序列或基于时间的鲁棒表示学习）的有效性。其结果也表明纯时间序列与 RF 在 3 种输入层上整体准确率显著低于结构主导模型，而引入方向联合后（Tik-Tok）性能明显改善但仍未超过结构主导的 DF。尽管如此，RemoCellEst 在上述不同流量结构与不同输入特征设定下均取得最高准确率，并相对同一模型下效果第二好的输入层保持稳定提升，表明本文方法的优势主要来自对跨位点流量报文结构的一致性重建，这种重建过程带来了更具普适性的改进。

6.4 基于传输数据量差异的对比实验

为进一步分析传输数据规模对 RemoCellEst 效果的影响，本实验在 RePHST 数据集上，从真实 Cell 数量角度对流量样本进行分组：将每条流的 Cell 总数不超过 1 000 记为 Low 组（共计 43 个类别），超过 1 000 记为 High 组（共计 15 个类别）。前文已表明桶分割方法的累计估计误差随传输数据量增大而放大，因此可以预期其在 High 组中对远程指纹攻击的影响更为显著。基于这一划分，实验分别在 Low/High 两组上，用隐藏服务代理侧流量训练指纹模型，在隐藏服务侧进行测试，以比较真实 Cell 序列、桶分割重建序列以及 RemoCellEst 重建序列在跨位点指纹攻击中的识别性能。

表 6 时间戳近似对隐藏服务溯源准确率的影响

Table 6 The Impact of Timestamp Approximation on the accuracy of hidden service de-anonymization layer

模型	TCP	桶分割	RemoCellEst
DF[34]	0.817 8	0.750 9	**0.948 5**(+0.130 7)
时间戳序列	0.387 8	0.644 5	**0.728 9**(+0.084 4)
Tik-Tok[41]	0.676 6	0.803 2	**0.940 9**(+0.137 7)
RF[19]	0.504 1	0.529 7	**0.641 4**(+0.111 7)

注：加粗代表溯源效果最优的流量层，下划线代表溯源效果次优的流量层，括号内为 RemoCellEst 相较于效果次好的流量层的性能提升。

从表 7 可以看出，真实 Cell 序列给出的攻击性能可视为上界：High 组和 Low 组的准确率分别为 0.950 3 和 0.967 1。RemoCellEst 在 High 组上几乎与上界完全重合，说明在大流量场景下，缓冲仿真得到的重建序列已足以支持与真实 Cell 几乎等价的远程指纹攻击能力；在 Low 组上虽略有下降（准确率从 0.967 1 降至 0.920 7），但整体仍保持较高的识别精度。RemoCellEst 在 Low 组上性能下降的原因可能在于：在 Cell 层，连接和流的起点可以通过 ConnID/StreamID 等字段显式标识，指纹模型能够聚焦于真实会话对应的 Cell 序列；相反，在 TCP 层无法直接定位某一隐藏服务访问在电路中的精确起始位置，特别是隐藏服务在初始化阶段已与入口节点建立并长期保持连接，期间存在一定比例的背景数据包和控制流量。当传输数据量较少时，数据传输在整体 TCP/TLS 序列中只占很小一段，这部分“有效信号”更容易被初始化与保持连接的背景流量稀释。RemoCellEst 虽然可以将 TCP 流量映射到近似的 Cell 序列，但仍难以完全剔除这类背景片段，导致在 Low 组中信号噪声比偏低，从而使远程指纹模型的识别精度略有下降。

表 7 RePHST 数据集下，不同传输数据量对隐藏服务溯源能力对比

Table 7 Comparison of hidden service de-anonymization capabilities under different data transmission volumes in the PHST dataset

流量层	High				Low			
	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall	F1
Cell (upper bound)	0.950 3	0.958 2	0.950 3	0.950 3	0.967 1	0.968 2	0.967 1	0.966 6
RemoCellEst	**0.950 3**	**0.958 0**	**0.950 3**	**0.950 6**	**0.920 7**	**0.914 5**	**0.920 7**	**0.907 6**
Bucket-based	0.791 0	0.856 9	0.791 0	0.771 0	0.717 1	0.741 6	0.717 1	0.706 6

注：加粗代表溯源效果最优的流量层，下划线代表溯源效果次优的流量层。

相比之下，桶分割方法在两组中的性能均明显落后：High 组准确率仅为 0.791 0，Low 组进一步降至 0.717 1，且精度也显著低于 RemoCellEst。结合前文关于累计估计误差随数据量增大而放大的分析，表明桶分割在大流量条件下误差累积导致攻击能力严重衰减，而 RemoCellEst 能在不同流量规模下稳定逼近

Cell上界,显著提升了跨位点远程指纹攻击的可用性。

6.5 跨位点流量之间跳数对远程指纹模型性能影响分析

为验证远程指纹攻击中指纹模型在位点迁移条件下普遍存在性能衰减现象,本文联合RePHST与ReSurface两套数据集,构造覆盖不同迁移距离,并在相同的指纹识别模型与相同的训练/测试流程下开展对比实验。具体而言,RePHST包含隐藏服务代理侧(Proxy HS)与隐藏服务侧(HS)的流量,对应跨6跳的典型远程迁移;ReSurface包含Client-Entry、Entry-Middle、Middle-Exit 3个位点流量,可构造跨1跳与跨2跳的迁移组合。考虑到实际远程溯源中敌手仅能获得网络层观测,本文分别采用桶分割与RemoCell-Est将上述网络层流量重建为Cell层序列,并在不同迁移距离下评估模型准确率变化,以揭示位点迁移导致的报文结构差异如何引发性能衰减,并为后续分析协议驱动重建在缓解远程迁移退化中的作用提供依据。

从表8可以看出,当训练流量与未知流量的观测位点相距较近(1~2跳)时,指纹模型整体仍能维持较高识别精度(约0.84~0.93),说明在近距离迁移条件下,跨位点差异虽已出现,但其带来的结构性失真尚未充分累积到主导模型判别的程度。在该区间内,RemoCellEst在多数配置上能够带来一定提升,表明更精细的Cell层重建有助于缓解由中继重封装与记录边界变化引起的分布偏移。尤其在未知流量向数据发起端(更接近出口)迁移的配置中,性能改善更为明显,反映出口侧具备更强的信元分簇能力,从而放大跨位点流量结构差异,从而一致性更高的Cell层输入来维持可分性。

表8　远程指纹攻击中,跨位点流量在相距不同中继跳数情况下,RemoCellEst和桶分割方法的准确率对比

Table 8　The comparison of the accuracy between the RemoCellEst and bucket-based divided methods for cross-position traffic for remote fingerprinting attacks under different numbers of relays

间隔跳数	训练流量位置	测试流量位置	桶分割	RemoCellEst
1	Client-Entry	Entry-Middle	**0.872 2**	0.843 9(–0.028 3)
	Entry-Middle	Client-Entry	**0.911 7**	0.911 3(–0.000 4)
	Entry-Middle	Middle-Exit	0.932 5	**0.944 9**(+0.012 4)
	Middle-Exit	Entry-Middle	0.889 6	**0.939 7**(+0.049 7)
2	Client-Entry	Middle-Exit	0.842 4	**0.846 7**(+0.004 3)
	Middle-Exit	Client-Entry	0.889 6	**0.893 0**(+0.093 4)
6	Proxy HS	HS	0.750 9	**0.948 5**(+0.197 6)
	HS	Proxy HS	0.724 7	**0.912 7**(+0.188 0)

注:加粗代表溯源效果最优的流量层。

当训练与测试位点间的中继数量增加至6跳(Proxy HS与HS之间)时,模型性能出现显著退化:在桶分割的粗粒度重建方法作为模型输入下,准确率下降至0.750 9与0.724 7,明显低于1~2跳区间。其结果明显表明随着中继跳数增加,Cell分片所产生的缓冲残留会在逐跳解密与重加密过程中产生较大差异,使跨位点的结构差异逐渐累积形成较大的系统性偏差。相比之下,RemoCellEst通过对缓冲残留进行动态追踪并在重建过程中显式管理残留状态,使重建的Cell序列在远距离跨位点条件下保持更一致的结构,从而在6跳溯源设置中将准确率恢复至0.948 5与0.912 7。上述结果表明,RemoCellEst的流量重建方法可抑制跨位点流量缓冲残留带来的噪声,消除传统重建方法带来的累计估计误差,其构建的指纹模型在远距离跨位点条件下的识别能力可以得到显著恢复并趋于稳定。

6.6 消融实验

之前的分析表明,突发分割中提出的“分割策略”可降低TCP头和TLS头部载荷过多导致的过计;在缓冲模拟中提出的“段末残留处理”可保留缓冲残留,缓解由于缓冲残留中存在的Cell分片带来的漏计,二者联用将获得最小的跨位置流量差异性并带来最佳识别效果;相反,简单分割与残留丢弃可能放大缓冲区残留引起的计数偏差,削弱远程指纹攻击能力。

为了评估RemoCellEst方法在突发分割中提出的“分割策略”以及在缓冲模拟中提出的“段末残留处理”对Cell重建精度以及远程指纹模型性能的影响,本文控制变量设置了2组实验。在第1组中,本文消除了分割策略,利用无方向信息的突发分段来分割双向流量;在第2组中,本文在对分段突发估计Cell数量后,直接将剩余的缓冲残留丢弃。本文在RePHST数据集上,对2组实验进行比较。

表9中消融结果表明，2项设计均对识别性能有实质贡献，且“段末残留处理”的作用更为关键。完整的RemoCellEst达到0.948 5的准确率；去除突发分割策略后，其准确率下降了约0.3%；再进一步去除段末残留处理方法，相较于完整版本，性能下降约0.7%。实验结果表明，2个策略在重建Cell序列时确实存在其重要性，共同提升了重建Cell层流量在远程指纹攻击中的能力。

表9 RemoCellEst中分割策略与段末残留处理对隐藏服务溯源的影响

Table 9 The impact of segmentation strategy and End-of-segment residual processing in RemoCellEst for hidden service de-anonymization

	分割策略	段末残留处理	Accuracy	Precision	Recall	F1
RemoCellEst	ü	ü	**0.948 5**	**0.941 8**	**0.948 5**	**0.942 5**
RemoCellEst(wo 分割策略)	û	ü	0.945 5	0.938 8	0.945 5	0.939 9
RemoCellEst(wo 段末残留)	ü	û	0.940 5	0.934 9	0.940 5	0.933 9

注：加粗代表溯源效果最优的流量层。

6.7 基于单个位点流量指纹攻击对比实验

针对目前流量指纹攻击主要存在网站指纹攻击及隐藏服务溯源这2种攻击场景，本文开展单个位点流量指纹攻击实验，排除跨位置流量差异，以评估RemoCellEst在同一观测点条件下的有效性。

具体而言，在隐藏服务溯源场景中，本文使用RePHST数据集中隐藏服务侧流量用于模型训练和测试。实验设置训练集与测试集的比例为1∶1，即每个受监测的隐藏服务类别均抽取100个双向流量样本用于模型训练，剩余100个样本用于测试。值得注意的是，在真实的隐藏服务溯源场景中，直接获取带标注的隐藏服务侧流量来训练指纹模型是一种理想场景。因此，该实验只为在同位置、充足标注条件下对RemoCellEst的能力进行衡量，用于验证该方法有效性与比较桶分割重建方法的相对优劣。

在网站指纹攻击场景中，本文利用Wang等人[28]提供的公开数据集进行实验。该数据集覆盖明网与隐藏服务500余个监控站点，提供TCP/TLS/Cell多层流量。已有结论表明：监控站点数增大将降低网站指纹模型的识别性能，而训练样本数增多可提升性能。为与隐藏服务溯源场景的训练规模对齐，并构造更严苛与可比的条件，针对每个监控站点统一采样100条双向流量样本训练、100条双向流量样本用于指纹模型测试。

从表10中可以看出，RemoCellEst在单点流量指纹攻击下依旧具备比Bucket-divided更好的识别能力，即使在大类别、样本受限的设置中也显著优于Bucket-divided。结果表明，RemoCellEst可以作为更有效的信元层流量重建方法，在网站指纹攻击中可以显著提升识别精度与鲁棒性。本文提出的RemoCellEst可以在单个位点的流量指纹攻击中提供一种新的数据特征提取视角，能在不控制Tor中继的情况下开展细粒度的流量指纹攻击。

表10 单个位点流量指纹攻击下RemoCellEst与桶分割方法性能对比

Table 10 The performance comparison between RemoCellEst and bucket-based divided methods under single-position traffic fingerprinting attacks

攻击场景	分类数	Bucket-divided		RemoCellEst	
		Accuracy	Precision	Accuracy	Precision
隐藏服务溯源	58	0.928 1	0.934 0	**0.936 0**	**0.934 4**
网站指纹识别(明网服务)	501	0.568 8	0.686 1	**0.727 7**	**0.728 7**
网站指纹识别(隐藏服务)	514	0.492 4	0.482 2	**0.603 2**	**0.581 9**

注：加粗代表溯源效果最优的流量层。

7 总结与展望

针对远程隐藏服务溯源中由于训练位置和预测位置不同，流量在中继转发过程中由于缓存、加解密和重封装带来的报文序列的结构性变化，从而引起指纹模型性能衰减的问题，本文提出了一种上下文感知的信元序列重建方法。通过系统刻画Tor传续协议，指出引起跨位点报文序列结构性变化的主要因素是报文序列中信元分片的位置和分布差异，从而在不同中继解密与重封装过程中在不同报文序列位置上形成缓冲残留。这种机理导致的非线性动态变化使得“上下文无关”的信元重建方法会产生显著的累计估计误差。基于此机理，本文提出的上下文感知的远程信元序列重建方法RemoCellEst通过显式模拟中继缓冲区行为与缓冲残留补偿机制，能够动态追踪残留字节并抑制误差累积，从而还原结构一致的跨位点信元序列。在基于真实Tor网络采集的RePHST与ReSurface数据集的实验表明，RemoCellEst能够适配多种指纹

模型,并在不同传输数据量与不同位点间距设置下表现出更稳定的位点迁移性能。在隐藏服务溯源的典型长跳远程场景中,RemoCellEst的准确率最高可达94.85%,显著优于桶分割等现有信元重建方案。本文将远程指纹攻击的迁移退化问题归因到可解释、可建模的报文序列结构性差异,并给出在网络级敌手能力下实现可部署的数据特征重构方法,为后续Tor加密流量分析、隐藏服务溯源、跨位点一致性建模及流量关联[5-6]、垫脚石检测[42]等相关研究提供可解释框架。

本文关注和解决由Tor协议封装演化引起的跨位点报文结构噪声与系统偏差,其目标是在多跳远程迁移条件下恢复输入表示的一致性,从而避免指纹模型因结构差异而发生性能衰减。在已有的研究中,网站指纹攻击还会面临其他重要因素,例如多标签页/多客户端访问导致的并发流与多路复用[43-45],以及填充、变形、调度扰动等对抗性防御[46-48]。在隐藏服务溯源场景下,也会存在如多客户端访问、服务侧防御机制等复杂场景。基于本文所建立的机制分析与重建框架,未来工作将面向更复杂的真实条件开展扩展研究:一方面,在RemoCellEst提供的更一致Cell序列重建基础上,评估并发流与多路复用对跨位点重建一致性与溯源性能的影响机理,并探索与之匹配的鲁棒模型或在线适配策略;另一方面,研究隐藏服务侧可能采用的典型对抗性流量防御对缓冲残留带来的跨位点差异与累计估计误差的相关性,并在噪声、并发与防御同时存在的条件下评估远程溯源能力的可用边界与潜在适配方法,从而推动远程隐藏服务溯源技术在更贴近真实网络环境的条件下,实现更可靠的稳健性验证。

参考文献

[1] SOCRadar Cyber Intelligence Inc. SOCRadar 2024 annual dark web report - SOCRadar[EB/OL]. (2025-02-27)[2026-01-25]. https://socradar.io/resources/report/socradar-2024-annual-dark-web-report/.

[2] DeepStrike. Dark web daily activity: What really happens in 2025[EB/OL]. (2025-11-18)[2026-01-25]. https://deepstrike.io/blog/dark-web-daily-activity-2025.

[3] Qin Yi, Zheng Tianming, Wu Yue, et al. Tracing Tor hidden service through protocol characteristics[C]//2022 International Conference on Computer Communications and Networks. Piscataway: IEEE, 2022: 1-9.

[4] Chen Muqian, Wang Xuebin, Shi Jinqiao, et al. Napping guard: Deanonymizing Tor hidden service in a stealthy way[C]//2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications. Piscataway: IEEE, 2020: 699-706.

[5] Nasr M, Bahramali A, Houmansadr A. DeepCorr: Strong flow correlation attacks on Tor using deep learning[C]// Proceedings of 2018 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2018: 1962-1976.

[6] Lopes D, Dong Jindong, Medeiros P, et al. Flow correlation attacks on Tor onion service sessions with sliding subset sum[C]//Proceedings of the 31st Annual Network and Distributed System Security Symposium. The Internet Society, 2024.

[7] Zhou Qiang, Wang Liangmin, Zhu Huijuan, et al. WF-transformer: Learning temporal features for accurate anonymous traffic identification by using transformer networks[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 30-43.

[8] Gong Jiajun, Cai Wei, Liang Siyuan, et al. WFCAT: Augmenting website fingerprinting with channel-wise attention on timing features[J]. IEEE Transactions on Dependable and Secure Computing, 2026, 23(1): 149-163.

[9] Li J, Wang D, Liu Y, et al. Cross-environmental website fingerprinting[C]// IEEE Conference on Computer Communications. 2025: 1-10.

[10] Deng Xinhao, Li Qi, Xu Ke. Robust and reliable early-stage website fingerprinting attacks via spatial-temporal distribution analysis[C]//Proceedings of 2024 on ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2024: 1997-2011.

[11] Zhao Xiyuan, Deng Xinhao, Li Qi, et al. Towards fine-grained webpage fingerprinting at scale[C]//Proceedings of 2024 on ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2024: 423-436.

[12] Shen Meng, Liu Yiting, Zhu Liehuang, et al. Fine-grained webpage fingerprinting using only packet length information of encrypted traffic[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 2046-2059.

[13] Wang Meiqi, Chen Muqian, Li Zeyu, et al. Deanonymize Tor hidden services using remote website fingerprinting[C]//2023 IEEE 22nd International Conference on Trust, Security and Privacy in Computing and Communications. Piscataway: IEEE, 2023: 998-1005.

[14] Zhao Can, Zhang Qingfeng, Qin Yefeng, et al. Quadruplet fingerprinting: Onion website fingerprinting through quadruplet network[C]//2025 28th International Confer-

ence on Computer Supported Cooperative Work in Design. Piscataway: IEEE, 2025: 2563-2568.

[15] Li Zeyu, Wang Yipeng, Wang Xuebin, et al. Proxied traffic fingerprinting for hidden service de-anonymization with burst reshaping[J]. IEEE Transactions on Information Forensics and Security, 2025, 20: 7663-7678.

[16] Khajehpour A, Zandi F, Malekghaini N, et al. Deep inside tor: Exploring website fingerprinting attacks on Tor traffic in realistic settings[C]//2022 12th International Conference on Computer and Knowledge Engineering. Piscataway: IEEE, 2022: 148-156.

[17] Yuqi Qing, Yin Qilei, Deng Xinhao, et al. Training robust classifiers for classifying encrypted traffic under dynamic network conditions[C]//Proceedings of 2025 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2025: 3564-3578.

[18] Sirinam P, Mathews N, Rahman M S, et al. Triplet fingerprinting: More practical and portable website fingerprinting with n-shot learning[C]//Proceedings of 2019 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2019: 1131-1148.

[19] Shen Meng, Ji Kexin, Gao Zhenbo, et al. Subverting website fingerprinting defenses with robust traffic representation[C]//32nd USENIX Security Symposium. USENIX Association, 2023: 607-624.

[20] Mitseva A, Panchenko A. Stop, don't click here anymore: Boosting website fingerprinting by considering sets of subpages[C]//33rd USENIX Security Symposium. USENIX Association, 2024: 4139-4156.

[21] Bahramali A, Bozorgi A, Houmansadr A. Realistic website fingerprinting by augmenting network traces[C]//Proceedings of 2023 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2023: 1035-1049.

[22] Jansen R, Wails R, Johnson A. Repositioning real-world website fingerprinting on tor[C]//Proceedings of the 23rd Workshop on Privacy in the Electronic Society. New York: ACM, 2024: 124-140.

[23] Fu Chuanpu, Li Qi, Bertino E, et al. Training with only 1.0‰ samples: Malicious traffic detection via cross-modality feature fusion[C]//Proceedings of 2025 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2025: 3930-3944.

[24] Liu Zixuan, Zhao Yi, Liu Zhuotao, et al. A hard-label black-box evasion attack against ML-based malicious traffic detection systems[C]//Proceedings of Network and Distributed System Security Symposium. The Internet Society, 2026.

[25] Shen Meng, Wu Jinhe, Ai Junyu, et al. Swallow: A transfer-robust website fingerprinting attack via consistent feature learning[C]//Proceedings of 2025 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2025: 1574-1588.

[26] Wang Tao, Goldberg I. Improved website fingerprinting on Tor[C]//Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society. New York: ACM, 2013: 201-212.

[27] Panchenko A, Lanze F, Pennekamp J, et al. Website fingerprinting at internet scale[C]//Proceedings of the 23rd Annual Network and Distributed System Security Symposium. The Internet Society, 2016.

[28] Wang Meiqi, Li Yanzeng, Wang Xuebin, et al. 2ch-TCN: A website fingerprinting attack over Tor using 2-channel temporal convolutional networks[C]//Proceedings of 2020 IEEE Symposium on Computers and Communications. Piscataway: IEEE, 2020: 1-7.

[29] Fu Chuanpu, Li Qi, Shen Meng, et al. Detecting tunneled flooding traffic via deep semantic analysis of packet length patterns[C]//Proceedings of 2024 on ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2024: 3659-3673.

[30] Zhou Guangmeng, Guo Xiongwen, Liu Zhuotao, et al. TrafficFormer: An efficient pre-trained model for traffic data[C]//2025 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2025: 1844-1860.

[31] Wang Tao, Cai Xiang, Nithyanand R, et al. Effective attacks and provable defenses for website fingerprinting[C]// Proceedings of the 23rd USENIX Conference on Security Symposium. USENIX Association, 2014: 143-157.

[32] Abe K, Goto S. Fingerprinting attack on Tor anonymity using deep learning[C]//Proceedings of the APAN. Semantic Scholar, 2016: 15-20.

[33] Rimmer V, Preuveneers D, Juárez M, et al. Automated website fingerprinting through deep learning[C]//25th Annual Network and Distributed System Security Symposium. The Internet Society, 2018.

[34] Sirinam P, Imani M, Juarez M, et al. Deep fingerprinting: Undermining website fingerprinting defenses with deep learning[C]//Proceedings of 2018 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2018: 1928-1943.

[35] Kwon A, AlSabah M, Lazar D, et al. Circuit fingerprint-

ing attacks: Passive deanonymization of Tor hidden services[C]//Proceedings of the 24th USENIX Security Symposium. USENIX Association, 2015: 287-302.

[36] Jansen R, Juárez M, Gálvez R, et al. Inside job: Applying traffic analysis to measure Tor from within[C]//25th Annual Network and Distributed System Security Symposium. The Internet Society, 2018.

[37] Mohd Aminuddin M A I, Zaaba Z F, Samsudin A, et al. The rise of website fingerprinting on Tor: Analysis on techniques and assumptions[J]. Journal of Network and Computer Applications, 2023, 212: 103582.

[38] Cherubin G, Jansen R, Troncoso C. Online website fingerprinting: Evaluating website fingerprinting attacks on Tor in the real world[C]//31st USENIX Security Symposium. USENIX Association, 2022.

[39] The Tor Project. Tor source code: Buffers. c[Z]. 2025.

[40] Bhat S, Lu D, Kwon A, et al. Var-CNN: A data-efficient website fingerprinting attack based on deep learning[J]. Proceedings on Privacy Enhancing Technologies, 2019, 2019(4): 292-310.

[41] Rahman M S, Sirinam P, Mathews N, et al. Tik-Tok: The utility of packet timing in website fingerprinting attacks[J]. Proceedings on Privacy Enhancing Technologies, 2020, 2020(3): 5-24.

[42] Chen Zixuan, Zheng Chao, Li Zhao, et al. Seeing the attack paths: Improved flow correlation scheme in stepping-stone intrusion[C]//2024 27th International Conference on Computer Supported Cooperative Work in Design. Piscataway: IEEE, 2024: 2116-2121.

[43] Yin Haoyu, Liu Yingjian, Guo Zhongwen, et al. From traces to packets: Realistic deep learning based multi-tab website fingerprinting attacks[J]. Tsinghua Science and Technology, 2025, 30(2): 830-850.

[44] Meng Wenwen, Ma Chuan, Ding Ming, et al. Beyond single tabs: A transformative few-shot approach to multi-tab website fingerprinting attacks[C]//Proceedings of ACM on Web Conference 2025. New York: ACM, 2025: 1068-1077.

[45] Gu Xiaodan, Yang Ming, Song Bingchen, et al. A practical multi-tab website fingerprinting attack[J]. Journal of Information Security and Applications, 2023, 79: 103627.

[46] Cui Yuwen, Wang Guangjing, Vu K, et al. A comprehensive survey of website fingerprinting attacks and defenses in tor: Advances and open challenges[PP/OL]. V3. arXiv (2025-11-22)[2025-12-20]. https://arxiv.org/abs/2510.11804.

[47] Xiao Xi, Zhou Xiang, Yang Zhenyu, et al. A comprehensive analysis of website fingerprinting defenses on Tor[J]. Computers & Security, 2024, 136: 103577.

[48] Shen Meng, Ji Kexin, Wu Jinhe, et al. Real-time website fingerprinting defense via traffic cluster anonymization[C]//2024 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2024: 3238-3256.

作者简介

李泽禹　男,1995年5月出生于重庆市。现为北京邮电大学网络空间安全学院博士研究生。主要研究方向为加密流量分析。

E-mail: lzreal@bupt.edu.cn

刘　凯　男,1986年12月出生于山东省济南市。现为北京邮电大学网络空间安全学院博士研究生。主要研究方向为网络性能测量和网络资源管理。

E-mail: liukai@bupt.edu.cn

黄文涛　男,1999年12月出生于山东省德州市。现为北京邮电大学网络空间安全学院博士研究生。主要研究方向为匿名通信、加密流量分析等。

E-mail: hyangwentao@bupt.edu.cn

时金桥　男,1978年1月出生于黑龙江省哈尔滨市。现为北京邮电大学教授、博士生导师。主要研究方向为网络与信息安全,尤其专注于隐私增强技术和数据泄露检测。中国电子学会会员编号:E190158671M。

E-mail: shijinqiao@bupt.edu.cn