加密流量侧信道泄漏的不可避免性

刘光杰[1],程 光[2*],刘伟伟[3]

(1. 南京信息工程大学电子与信息工程学院,江苏南京 210044;2. 东南大学网络空间安全学院,江苏南京 211189;
3. 南京理工大学自动化学院,江苏南京 210094)

摘 要: TLS 1.3与QUIC的普及使载荷内容不可见,流量分析转而依赖侧信道特征,但侧信道泄漏在加密通信中为何不可避免,长期缺乏严谨论证。本文从信息论与系统设计出发,构建形式化模型$\Sigma=(\Gamma,\Omega)$,其中加密通信模型$\Gamma=(A,\Pi,\Phi,N)$描述"应用生成、协议封装、加密变换、网络传输"的因果链,观察模型Ω刻画外部观测能力。该框架将完整通信过程抽象为因果可测的马尔可夫链$X\to\Xi_A\to\Xi_P\to\Xi_C\to\Xi_N\to Y$,使语义变量到可观测特征之间的互信息严格可定义。基于复合信道结构、数据处理不等式与有界Lipschitz统计量的稳定传递性,提出并证明"侧信道存在性定理":对于可辨识的语义对,在系统满足映射非退化性(度量期望有界$E\left[d\left(z_P,z_N\right)|X\right]\leqslant C$)、协议层统计可辨识性(期望差$\geqslant\bar{\Delta}$)、统计量Lipschitz连续性、观测非退化性(保留比例$\rho>0$)以及可辨识性传递条件($C<\bar{\Delta}/2L_\varphi$)的前提下,观测特征与语义变量的互信息$I(X;Y)$必然严格为正且存在显式下界。推论表明,在效率优先的多元语义系统中,只要存在至少一对应用在统计上可区分,侧信道泄漏就不可避免。3个关键因素共同决定泄漏边界:映射非退化常数C受效率约束限制,反映带宽、时延等实用性要求;语义可辨识性$\bar{\Delta}$源于应用多样性,体现不同应用在统计特征上的固有差异;观测非退化性ρ由分析者能力决定。本文进一步通过全变差与Chernoff信息的下界链条,建立了从信息论下界到分类准确率的量化联系,揭示了多次观测使识别错误率指数衰减的必然性。理论分析表明,降低泄漏面临三种困境:增大度量偏差需牺牲效率,减小语义可辨识性将破坏应用功能,而观测非退化性由分析者控制。因此,侧信道并非协议实现的偶发瑕疵,而是满足实用性约束的网络通信系统的内在属性,正确的工程目标是在给定效率约束下最小化泄漏的约束优化问题。本文首次为加密流量侧信道建立严格的信息论基础,为攻击可达性提供可检验的预测,为防御机制提供可量化的性能基准,并为效率-隐私权衡的工程决策提供数学依据。

关键词: 侧信道;信息论;加密流量分析;存在性定理;效率-隐私权衡

基金项目: 国家自然科学基金(No.U24366001,No.U22B2025)

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2026)02-0837-14

电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20251034

The Inevitability of Side-Channel Leakage in Encrypted Traffic

LIU Guangjie[1], CHENG Guang[2*], LIU Weiwei[3]

(1. *School of Electronic and Information Engineering*, *Nanjing University of Information Science and Technology*, *Nanjing*, *Jiangsu* 210044, *China*; 2. *School of Cyber Science and Engineering*, *Southeast University*, *Nanjing*, *Jiangsu* 211189, *China*;
3. *School of Automation*, *Nanjing University of Science and Technology*, *Nanjing*, *Jiangsu* 210094, *China*)

Abstract: The widespread adoption of TLS 1.3 and QUIC renders payload content invisible, shifting traffic analysis toward side-channel features. However, rigorous justification for "why side-channel leakage is inevitable in encrypted communications" has long been lacking. This paper establishes a strict foundation from information theory and system design by constructing a formal model $\Sigma=(\Gamma,\Omega)$, where the encrypted communication system $\Gamma=(A,\Pi,\Phi,N)$ describes the causal chain of "application generation-protocol encapsulation-encryption transformation-network transmission", and the observation model Ω characterizes external observation capabilities. This framework abstracts the complete communication process as a causally measurable Markov chain $X\to\Xi_A\to\Xi_P\to\Xi_C\to\Xi_N\to Y$, enabling the mutual information between semantic variables and observable features to be rigorously defined. Based on the composite channel structure, data processing inequality, and stable propagation of bounded Lipschitz statistics, we propose and prove the "Side-Channel Existence Theorem": for distinguishable semantic pairs, under the conditions that the system satisfies mapping non-degeneracy (bounded metric expectation $E\left[d\left(z_P,z_N\right)|X\right]\leqslant C$), protocol-layer statistical distinguishability (expectation difference $\geqslant\bar{\Delta}$), Lipschitz continuity of statistics, observation non-degeneracy (preservation ratio $\rho>0$), and the distinguishability propagation condi-

收稿日期:2025-12-30;录用日期:2026-02-24;责任编委:秦湛

*通信作者:程光

tion ($C<\bar{\Delta}/2L_{\varphi}$), the mutual information $I(X;Y)$ between observed features and semantic variables is necessarily strictly positive with an explicit lower bound. The corollary demonstrates that in efficiency-prioritized multi-semantic systems, side-channel leakage is inevitable as long as at least one pair of applications is statistically distinguishable. Three key factors jointly determine the leakage boundary: the mapping non-degeneracy constant C is constrained by efficiency requirements, reflecting practical demands such as bandwidth and latency; semantic distinguishability $\bar{\Delta}$ stems from application diversity, embodying inherent differences in statistical characteristics across applications; and observation non-degeneracy ρ is determined by analyst capabilities. This paper further establishes a quantitative connection from information-theoretic lower bounds to classification accuracy through the chain of total variation and Chernoff information bounds, revealing the inevitability that multiple observations cause recognition error rates to decay exponentially. Theoretical analysis shows that reducing leakage faces a trilemma: increasing metric deviation requires sacrificing efficiency, reducing semantic distinguishability disrupts application functionality, while observation non-degeneracy is controlled by analysts. Therefore, side channels are not incidental flaws in protocol implementations but inherent properties of network communication systems subject to practicality constraints, and the correct engineering objective is a constrained optimization problem that minimizes leakage under given efficiency constraints. This paper establishes, for the first time, a rigorous information-theoretic foundation for encrypted traffic side channels, providing verifiable predictions for attack feasibility, quantifiable performance benchmarks for defense mechanisms, and mathematical basis for engineering decisions on efficiency-privacy tradeoffs.

Keywords: side channel; information theory; encrypted traffic analysis; existence theorem; efficiency-privacy tradeoff
Foundation Item(s): National Natural Science Foundation of China (No.U24366001, No.U22B2025)

0 引言

随着TLS 1.3、QUIC等加密协议的广泛部署,现代网络通信中的载荷内容已得到强有力的密码学保护。然而,加密流量分析通过观测加密流量的元数据特征(包长度、时序、方向等)推断敏感信息(如用户访问的网站、使用的应用、传输的内容类型)仍然取得了令人瞩目的成功。在封闭环境下,网站指纹识别准确率可达91%~95%[1],应用识别准确率也超过90%[2]。针对加密匿名流量的识别,在真实校园网关环境测试中,中等基准率场景下精度可达96%,在1 000∶1的极低基准率下仍能保持93%以上[3]。这一现象引发了一个根本性的理论问题:为什么即使使用如AES-256-GCM等计算上安全的加密算法,流量分析也仍然有效?

从密码学角度看,加密算法保证了载荷内容的机密性,在没有密钥的情况下,攻击者无法从密文恢复明文。但加密协议必然无法隐藏通信的元数据:源/目的地址、包长度、时间戳等信息是网络路由与传输控制的必需。这些元数据构成了侧信道(side channel)的基础。侧信道这一概念最早出现于密码学硬件实现领域。Kocher[4]提出的时序攻击通过测量加密操作的执行时间恢复密钥,Kocher等人[5]的差分功耗分析利用加密设备的功耗变化提取密钥信息。这些攻击的共同特征是利用密码算法实现过程中的物理副产物(时间、功耗、电磁辐射等),而非直接破解算法本身。随后,侧信道概念扩展至网络通信领域。Hintz[6]的工作展示了如何通过分析加密网页的流量模式识别用户访问的网站,推动了网络流量侧信道的深入研究。网络流量侧信道与硬件侧信道在本质上一致,即加密保护了载荷内容,但通信过程本身的可观测特征(包大小、时序、方向等)成为新的信息泄漏源。本文聚焦于网络流量侧信道,但继承了硬件侧信道研究的核心洞察:侧信道并非加密算法的缺陷,而是实现与部署过程中效率优先导致的必然副产品。

网络流量侧信道泄漏显示了应用层行为模式在传输过程中的“指纹传递”。具体而言,应用对象的大小差异映射为加密记录序列的长度模式,HTTP请求-响应的时序逻辑体现为包到达间隔分布,客户端服务器交互的方向性与突发特征反映协议语义,协议栈的逐层封装放大微小差异(TCP分段边界、拥塞控制窗口调整)。这些机制使得不同应用的流量在统计上可区分。视频流呈现周期性突发、即时通信表现为小包双向交互,网页浏览显示短时高密度资源加载。加密操作虽然改变了载荷内容,但无法消除这些源于应用逻辑的统计指纹。

然而,仅仅指出“指纹传递”并不能回答侧信道为何不可避免。更深层的问题是,在给定的系统约束下(计算资源有限、通信效率要求、协议兼容性需求),加密系统设计为何必然导致侧信道的存在?这种存在性是否有可量化的理论下界?效率与隐私之间的权衡关系能否以严格的数学函数刻画?

近年来,尽管加密流量分析的攻击方法不断创新并取得显著性能提升[7-8],但对侧信道存在的理论必

然性仍缺乏系统论证。Li 等人[9]用互信息 $I(X;Y)$ 量化网站指纹泄漏,在针对 Tor 网络的实验中发现,即使 Tor 加密了通信内容并通过三层中继隐藏端点身份,元数据特征仍会泄漏网站信息:单个特征的信息泄漏 $I(F;W)$ 最高达 3.45 bit,综合特征可达约 6.6 bit。这种非零互信息源于 Tor 为保证低延迟和低开销而未对流量元数据进行混淆或填充。Cai 等人[10]证明任何达到 ε-安全性的防御在 n 个网站的封闭世界中都必然会产生可计算的带宽开销下界,揭示了防御的理论成本,但这一下界仅针对特定场景,缺乏通用性。差分隐私框架为流量分析防御提供了可量化的隐私保证[11-12],但其设计理念是在隐私与效用之间寻求平衡,而非探究效率约束下信息泄漏的根本原因。

本文从信息论与系统设计出发,给出侧信道存在性的严格基础推理。首先构建形式化模型 $\Sigma=(\Gamma,\Omega)$,其中加密通信系统 $\Gamma=(A,\Pi,\Phi,N)$ 刻画应用生成、协议封装、加密变换与网络传输过程,观察模型 Ω 限定外部观测的层与位置,并把“生成—封装—加密—传输—观测”形式化为一条因果可测的复合信道 $X\to\Xi_A\to\Xi_P\to\Xi_C\to\Xi_N\to Y$。在此框架下,我们把效率优先设计的约束外化为映射非退化性:存在度量 d 与常数 $C<\infty$,使得协议层到网络层的轨迹映射在度量意义下保持有界偏差 $E\left[d\left(z_P,z_N\right)|X\right]\leqslant C$。该度量捕获点过程的长度、时序、方向等多维度联合统计结构,而有界偏差 C 的存在性源于带宽、时延、吞吐量等实用性要求。结合有界 Lipschitz 统计量的稳定传递性,我们建立了从协议层可辨识性到观测层信息泄漏存在性的严格推导链条。

本文的核心结论是“侧信道存在性定理”:对于可辨识的语义对,在满足系统映射保持非退化性(度量期望界 $E\left[d\left(z_P,z_N\right)|X\right]\leqslant C$)、协议层存在统计可辨识性(期望差 $\geqslant\bar{\Delta}$)、统计量满足 Lipschitz 连续性、观察模型保持非退化性(保留比例 $\rho>0$),以及可辨识性传递条件($C<\bar{\Delta}/2L_\varphi$)下,观测特征与语义变量之间的互信息满足显式下界:

$$I(X;Y)\geqslant\frac{1}{2\ln 2}\left(\frac{\rho\left[\bar{\Delta}-2L_\varphi C\right]}{2}\right)^2>0 \tag{1}$$

这一下界揭示了侧信道泄漏的不可避免性:映射非退化常数 C 受效率约束限制、语义可辨识性 $\bar{\Delta}$ 源于应用多样性、观测非退化性 ρ 由分析者能力决定,三者共同构成不可逾越的泄漏边界。推论表明,在效率优先的多元语义系统中,只要存在至少一对应用在统计上可区分,侧信道泄漏量就必然为正。

因此,侧信道并非某一协议实现的偶发瑕疵,而是满足实用性约束的网络通信系统的内在属性。正确的工程目标不是追求不可达的零泄漏,而是在给定效率约束下最小化泄漏的约束优化问题。

1　相关理论工作

网络中侧信道分析的理论基础可追溯至信息论与匿名性度量的早期研究。Chaum[13]提出的混合网络为匿名通信奠定基础,但如何量化匿名性长期缺乏严格的数学框架。Díaz 等人[14]首次提出用 Shannon 熵 $H(X)=-\Sigma p_i\log p_i$ 度量匿名性,引入归一化熵 $d=H(X)/H_M$ 区分匿名集大小与概率分布的影响。这一工作的关键结论是概率分布比集合大小更重要,10 人均匀分布($d=1$)的匿名性强于 100 人但某人被识别概率 90% 的情况($d\approx 0.47$)。后续工作引入更丰富的熵度量工具:Deng 等人[15]将 Rényi 熵 $H_\alpha(X)=\frac{1}{1-\alpha}\log\Sigma p_i^\alpha$ 应用于匿名性度量,不同 α 值捕获分布的不同方面($\alpha\to 1$ 时退化为 Shannon 熵)。Serjantov 等人[16]将互信息 $I(X;Y)$ 引入匿名协议分析。Chatzikokolakis 等人[17]进一步将匿名协议解释为噪声信道,用信道容量 $C=\max I(X;Y)$ 刻画泄漏上界,并引入相对熵 $D(P\parallel Q)=\Sigma p_i\log(p_i/q_i)$ 度量攻击前后分布的变化。这些早期工作奠定了用信息论量化隐私泄漏的理论框架,但主要关注匿名通信协议(如混合网络、洋葱路由),未系统分析加密协议本身与侧信道的关系。

统计泄漏攻击的理论研究揭示了长期观察的威胁。Kedogan 等人[18]证明:在开放环境中,只要用户有习惯性通信模式,长期观察必然使匿名性退化。数学上,这等价于条件熵 $H(X|Y_1,Y_2,\cdots,Y_t)$ 随观察次数 t 递减至接近 0,揭示了匿名性随时间指数衰减的规律。Danezis 等人[19]提出统计披露攻击,通过流量分析识别匿名系统中的发送者-接收者对。这些研究揭示了时间相关性的影响,但未涉及加密协议设计与侧信道的因果关系,即为什么协议设计的基本约束会导致侧信道的存在。

差分隐私框架为隐私保护提供了可量化保证。Dwork 等人[11]提出差分隐私:机制 M 满足 (ε,δ)-DP,若对相邻数据集 D,D' 和任意输出集 S,$\Pr[M(D)\in S]\leqslant e^\varepsilon\Pr\left[M\left(D'\right)\in S\right]+\delta$,这一框架的优势在于提供精确的隐私参数和可组合性:k 次 (ε,δ)-DP 操作在基础组合下提供 $(k\varepsilon,k\delta)$-DP 保证。Vuvuzela 和 Stadium 首次将差分隐私应用于大规模消息系统元数据保护[20-21]。NetShaper 首次为网络侧信道防御建立了形式化的差分隐私框架[12]。该工作的关键理论贡献是将抽象的 ε 隐私参数与具体的系统性能指标(带宽、延迟)通过

可计算的数学关系连接起来。具体而言，给定隐私预算ε和网络条件，NetShaper可计算出实现该隐私保证所需的最小带宽开销和延迟增量。这一框架将“隐私vs性能”的定性权衡转化为可优化的数学问题。然而，该方法仍假设“一定程度的泄漏可接受”（通过$\varepsilon>0$编码），而非从严格数学角度证明侧信道作为功能实现副产品的不可避免性。

网站指纹攻击与可证明防御的研究为理解侧信道提供了实证基础与理论尝试。早期网站指纹攻击从Hintz对加密网页的流量分析[6]到Panchenko等人的支持向量机方法[22]，逐步证明了侧信道的有效性。Dyer等人[23]论证了多数“高效”流量整形方案因可观测侧信道仍然失效，并提出BuFLO基线防御，采用固定速率、定长包与最小持续时间策略。Cai等人[10]建立了与攻击无关的评测框架与带宽下界，据此提出Tamaraw防御：对上下行分别定速并按块填充，给出攻击准确率的可证明上界。Wang等人[24]提出Walkie-Talkie防御，采用半双工突发整形与成对混淆机制，实测平均带宽开销约31%，时延增加约34%。Huang等人[25]提出不对称防御（STAP）将攻击准确率降至48.3%，且带宽开销仅18%。Wright等人[26]将流量整形形式化为凸优化问题：给定源分布X和目标分布Y，寻找变形矩阵$\boldsymbol{A}\geqslant 0$使得$\boldsymbol{A}X=Y$（其中$\sum_i A_{ij}=1$确保每列为概率分布），同时最小化期望的带宽开销$\sum_{i,j}x_j a_{ij}\left|s_i-s_j\right|$，其中$s_i,s_j$为包大小。这些工作揭示了防御的理论成本，但仅针对网站指纹这一特定场景，未建立通用的效率-泄漏关系，也未解释为什么在没有防御的情况下侧信道必然存在。

互信息泄漏的量化为侧信道分析提供了信息论工具。Li等人[9]在ACM CCS首次系统应用互信息$I(X;Y)$量化网站指纹泄漏。在针对Tor网络的100个网站封闭环境实验中，他们发现即使Tor使用加密和混淆技术，流量特征仍会泄漏大量网站信息。单个特征的信息泄漏量$I(F;W)$最高可达3.45 bit（来自四舍五入的出站数据包计数特征），54.55%的特征泄漏少于1 bit，而综合多个特征后总信息泄漏可达约6.6 bit，接近理论上限$\log_2 100\approx 6.64$ bit。Cherubin[27]提出基于Bayes错误下界的(ξ,Φ)-privacy度量，从信息论角度界定防御的安全性。这些工作揭示了加密后仍存在信息泄漏的现象，未从系统设计角度解释其必然性，即为什么在满足实用性条件的加密系统中，$I(X;Y)$必然大于0？

与此类直接以互信息或Bayes错误下界度量泄漏不同，Fu等人[28-30]及后续工作将加密流量特征建模为随机变量或信号，在频域谱特征、长度模式或流交互图空间中进行差分熵、信息损失、KL散度与可分性/鲁棒性分析，侧重于评估给定特征与检测方法在恶意流量检测、隧道流识别等具体任务中的有效性。而本文则从复合信道—互信息下界的系统级视角出发，在不预先固定特征形式的前提下讨论加密流量侧信道泄漏是否不可避免及其理论下界。

匿名系统的形式化安全分析尝试在更严格的框架下证明隐私保证。Camenisch等人[31]在通用可组合（UC）框架下形式化洋葱路由，提供可组合的安全定义。Feigenbaum等人[32]用概率I/O自动机建模洋葱路由，针对主动时序攻击提供形式化的匿名性分析[32]。Danezis等人[33]的Sphinx模型为混合网络提供紧凑消息格式，在随机预言模型下证明不可关联性与路径长度隐藏。这些形式化方法在特定威胁模型下提供了严格的安全证明，但主要针对特定匿名协议的安全性分析，未从信息论角度建立加密通信中侧信道泄漏不可避免性的通用理论框架。

尽管上述研究在各自领域取得重要进展，但对侧信道存在性的理论基础仍存在明显空白。（1）缺乏形式化因果框架。从应用语义到可观测特征的信息传递过程如何严格建模？加密、封装、传输的逐层作用如何在数学上刻画？（2）效率约束与泄漏的内在联系不明。系统设计中的效率优先如何必然导致侧信道？在满足何种可检验条件时泄漏不可避免？（3）缺乏可计算的泄漏边界。给定系统参数，侧信道泄漏的下界是什么？如何从理论下界预测实际攻击性能？本文通过构建形式化模型，证明存在性定理、建立显式下界，系统填补这些理论空白，为理解侧信道的根本原因提供严格的数学基础。

2 侧信道分析的形式化模型

2.1 基本定义与建模约定

本文的核心做法是把生成、封装、加密、传输、观测的全过程抽象为一条由可测映射与随机信道组成的复合通道，使语义变量到可观测特征之间的互信息严格可定义并可被数据处理不等式刻画。下面首先介绍建模的基本定义和约定。

时间与随机性。时间建模为连续非负轴$T=R_{\geqslant 0}$。所有随机变量（语义X、应用侧随机性U_A、协议侧U_Π、加密侧U_Φ、网络侧U_N、观测侧U_Θ）定义在公共概率空间(Ω_0,F_0,P)上，该空间的具体构造不影响后续分析，用于保证联合分布与条件期望的良定义性。

点过程与序列表示。点过程Ξ_A,Ξ_P,Ξ_C,Ξ_N（为直观起见，后文简称为“消息序列、明文包序列、密文包序列、到达包序列”）及观测特征Y的样本路径取值于相应的可测空间。

统一轨迹空间与度量。为避免跨层统计量定义域不一致的问题,引入统一的标记点过程轨迹空间Z(例如用有限标记计数测度空间,或将窗口内轨迹嵌入Skorokhod空间)。设e_P: range$(\Xi_P)\to Z$、e_N: range$(\Xi_N)\to Z$为逐层到Z的可测嵌入。下文所有窗口级统计量均视为φ: $Z\to[-M,M]$的可测函数。度量d定义在Z上,既可同时度量时序与标记(长度、方向)差异,又与后文的Lipschitz条件相容。为避免不必要的技术负担,假定X、Y、range(Ξ_A)、range(Ξ_P)、range(Ξ_C)、range(Ξ_N)以及Z均为标准Borel空间。

为明确区分系统设计与外部观测,我们给出如下侧信道分析模型定义。

定义1　侧信道分析模型。侧信道分析模型记为$\Sigma=(\Gamma,\Omega)$。其中,加密通信模型记为$\Gamma=(A,\Pi,\Phi,N)$,由4个因果算子组成:应用生成A、协议封装Π、加密变换Φ、网络传输N;Ω为观察模型,刻画侧信道分析者对若干网络层与位置的被动监听与特征提取能力。

在该定义下,系统Γ仅决定如何产生与传送密文,观察模型Ω仅决定分析者能看见什么。泄漏量$L(\Gamma,\Omega)$是两者复合作用的结果。这样的分离使后文的存在性定理只需最弱的系统与观察假设即可成立。

2.2　加密通信模型

加密通信模型$\Gamma=(A,\Pi,\Phi,N)$由4个因果算子组成,分别作用于时间轴上的消息点过程与包序列。我们以因果可测映射刻画这些变换,仅要求其保持时间因果性与可测性,而不对概率分布的具体形式(如独立性、平稳性)作强假设。

(1)应用层(算子A)

令Ξ为应用语义空间(如网站ID、应用类别、视频内容等),令$X\in\Xi$为语义变量。应用在时间轴上产生消息序列:

$$\Xi_A=\left\{\left(\tau_k,m_k\right)\right\}_{k\geqslant 1},\ \tau_k\in T,m_k\in M \tag{2}$$

它由X与外生噪声共同决定,可写为因果可测映射$\Xi_A=G_A\left(X,U_A\right)$。这里$M$为消息集合,$U_A$为应用侧随机性(用户行为、业务逻辑抖动等)。我们不要求平稳或独立,仅要求G_A因果、可测。

(2)协议层(算子Π)

协议栈把消息序列Ξ_A映射为分段与封装后的明文包序列:

$$\Xi_P=\left\{\left(t_i,\ l_i,\ \mathrm{dir}_i,\ h_i,\ b_i\right)\right\}_{i\geqslant 1} \tag{3}$$

其中,$t_i\in T$为发送时刻,$l_i=\left|h_i\right|+\left|b_i\right|$为长度,$\mathrm{dir}_i\in\{\uparrow,\downarrow\}$为方向,$h_i$为各层头部拼接,$b_i$为明文负载。用因果映射表示为$\Xi_P=\Pi\left(\Xi_A,U_\Pi\right)$,$U_\Pi$表示协议侧随机性(Nagle、分段边界、拥塞控制自适应等)。Π保持因果性:t_i仅依赖于$\left\{\left(\tau_k,m_k\right):\tau_k\leqslant t_i\right\}$与过去的协议状态。

(3)加密层(算子Φ)

加密把明文包序列Ξ_P映射为密文包序列:

$$\Xi_C=\left\{\left(t_i',\ l_i',\ \mathrm{dir}_i,\ h_i',\ c_i\right)\right\}_{i\geqslant 1}=\Phi\left(\Xi_P,U_\Phi\right) \tag{4}$$

其中,h_i'为可见或半可见的头部字段,c_i为密文负载。我们只用到两条通用性质:(a)语义独立性:在理想密码假设下,给定输入长度与公开参数,c_i对明文内容条件独立;(b)长度变换的确定性:存在确定函数g_{len}使$l_i'=g_{\mathrm{len}}\left(l_i;\ \theta_\Phi\right)$,其中$\theta_\Phi$含块大小、记录开销、可选填充等。现实协议(TLS 1.3、QUIC)满足这一直观性质:内容被置乱,但长度与时序只受小幅对齐与记录化影响。时间映射保因果:$t_i'\geqslant t_i$,且$t_i'-t_i$由实现与调度决定。

(4)网络层(算子N)

网络把发送侧密文序列Ξ_C转为观察路径上的到达包序列:

$$\Xi_N=\left\{\left(\tilde{t}_j,\tilde{l}_j,\mathrm{dir}_j,\tilde{h}_j,\tilde{c}_j\right)\right\}_{j\geqslant 1}=N\left(\Xi_C,U_N\right) \tag{5}$$

其中,U_N刻画排队、路由、丢包、重传、乱序、成帧及多路复用等不确定性。N被视为随机时变队列网络的因果信道,不需作独立或马尔可夫假设。

对Γ的一个基本假设是其在协议层产生统计上可区分的流量模式。为此,我们给出如下语义可辨识性的定义。

定义2　语义可辨识性。给定窗口$T>0$,称系统Γ在协议层诱导的Z表示上具有$\bar{\Delta}$-可辨识性,若存在语义对$x\neq x'\in\Xi$与有界可测统计量φ: $Z\to[-M,M]$,则使得:

$$\left|E\left[\varphi\left(e_P\left(\Xi_P\big|_{[0,T]}\right)\right)\middle|X=x\right]-E\left[\varphi\left(e_P\left(\Xi_P\big|_{[0,T]}\right)\right)\middle|X=x'\right]\right|\geqslant\bar{\Delta} \tag{6}$$

称系统Γ在协议层上可辨识,若存在$\bar{\Delta}>0$,则使其具有$\bar{\Delta}$-可辨识性。

这一定义刻画了系统的固有属性:不同语义在经过应用生成与协议封装后,在统一轨迹空间Z的某个统计量的条件期望上产生足够大的差异。$\bar{\Delta}$的存在排除了“期望差任意小”的退化情况,确保可辨识性在统计意义上是可操作的。常见的统计量φ包括窗口总字节数、上下行包数比、包间隔的加权平均等。不同应用语义往往在这些统计量上呈现差异:视频流的高字节数、即时通信的平衡上下行比、网页浏览的短时高密度等模式,即使在协议封装后仍保持统计可分性。

现实的加密通信系统设计遵循效率优先原则：在保证密码学安全的前提下，最大化资源利用效率以承载上层业务需求。业务对多个性能维度有可量化要求。

带宽开销：协议引入的额外字节数应控制在可接受范围（如 TLS 1.3 约 5%）。

端到端时延：加密、成帧、调度引入的延迟应低于应用感知阈值（如网页 < 200 ms，VoIP< 150 ms）。

吞吐量：不应因填充或人工延迟导致有效吞吐率显著下降。

协议兼容性：需与现有网络基础设施（MTU、拥塞控制、NAT 穿透等）兼容。

效率优先设计的直接后果是从明文协议层 Ξ_P 到到达包序列 Ξ_N 的映射必然保持非退化性。

定义 3　映射非退化性。在固定窗口 $[0,T]$ 上，称加密通信系统 Γ 具有窗口级映射非退化性，若存在常数 $C_T<\infty$，使得对所有 $x\in\Xi$ 有

$$E\left[d\left(e_P\left(\Xi_P\big|_{[0,T]}\right),e_N\left(\Xi_N\big|_{[0,T]}\right)\right)\big|X=x\right]\leqslant C_T \tag{7}$$

直观上，这要求加密与传输映射不会把明文包序列的统计结构过度扭曲，即点过程的标记（长度、方向）与时序的联合分布在映射前后保持“接近”。

非退化性直接源于效率约束，如过度填充使度量在长度分量上发散（带宽开销超标），过度延迟或抖动使度量在时序分量上发散（端到端时延超标，破坏实时性），完全打乱包序列的标记与时序结构使度量发散（协议语义破坏，业务不可用）。“窗口级”指度量作用于有限时间窗口 $[0,T]$（如 10 s 或一次会话）内的点过程片段，允许单包扰动（如 TCP 重传、乱序）但约束累积偏差。度量 d 的具体形式可以多样，只要能捕获点过程的长度、时序、方向的联合统计结构。本文结论不依赖特定的度量选择，后续定理的成立仅依赖抽象的非退化性，即存在某个合适的度量 d 使 $E[d]\leqslant C_T<\infty$。在典型网络场景中，一个常用的选择是把 d 定义为“长度偏差+包数偏差+时延偏差”的加权和，或者由机器学习和深度学习刻画的隐式的关于流特征的某种高维空间的距离。此时，C_T 可以被理解为在该窗口下由带宽、包数开销和端到端时延抖动共同限定的最大可接受偏差。

上述语义可辨识性和映射非退化性共同构成了加密通信系统的内在属性。它们刻画了“在满足业务可用性前提下，系统必然保留的统计变异性”。下一节将讨论侧信道分析者如何通过观测模型 Ω 捕获这些属性，从而最终导出侧信道泄漏的存在性。

2.3　观察模型

观察模型 Ω 刻画侧信道分析者对若干网络层与位置的被动接入能力，以及对到达包序列的特征抽取过程。观察者虽然无法直接访问明文协议层 Ξ_P，但可以通过观测到达包序列 Ξ_N 并提取特征来推断语义信息。本节定义观察模型的构成、观测特征的生成，以及保证观测通道不退化为常数的非退化性条件。

定义 4　观察模型。侧信道分析模型 $\Omega=(L_{\mathrm{acc}},O_{\mathrm{acc}},\Theta)$ 由 3 个组件构成。

L_{acc}：可访问层集合（如 IP 层、UDP/TCP 层、QUIC/TLS 记录层等）。

O_{acc}：可观测位置集合（链路、交换机端口、主机网络接口等）。

Θ：因果可测的特征抽取算子。

给定到达包序列 Ξ_N，观测者通过 Θ 得到特征序列 $Y=\Theta(\Xi_N;L_{\mathrm{acc}},O_{\mathrm{acc}},U_\Theta)$，其中 U_Θ 表示观测侧不确定性（时间戳精度、采样策略、计数粒度等）。

特征抽取算子 Θ 通常包括以下操作：提取包长度序列、计算包到达时间间隔、统计上下行包数比、构造突发模式描述符、计算窗口级统计量（总字节数、包数、速率等）等。Θ 的具体形式取决于观测者的技术能力与分析目标，但必须保持因果性：时刻 t 的观测特征仅依赖于 t 时刻及之前的到达包。

把系统模型 Γ 与观测模型 Ω 首尾相接，得到从语义到观测特征的完整因果链，如图 1 所示。

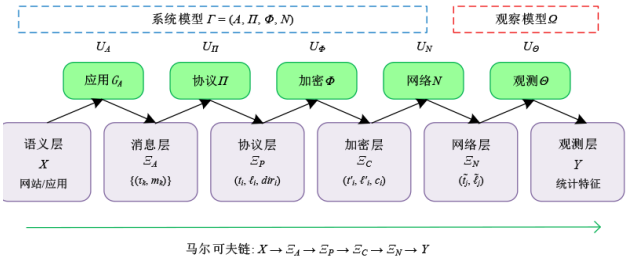

图 1　从语义到观测特征的完整因果链

Figure 1　Complete causal chain from semantics to observational features

$$X\xrightarrow{G_A}\Xi_A\xrightarrow{\Pi}\Xi_P\xrightarrow{\Phi}\Xi_C\xrightarrow{N}\Xi_N\xrightarrow{\Theta}Y \tag{8}$$

其中：X 为敏感语义变量；Ξ_A,Ξ_P,Ξ_C,Ξ_N 分别为消息序列、明文包序列、密文包序列、到达包序列（均为点过程）；Y 为观测特征序列。这条因果链诱导两个关键性质。

命题 1　观测信道的存在性与可测性。在上述各空间为标准 Borel 空间且 N、Θ 为随机核（或可测映射与随机核的复合）的条件下，存在从 X 到 Y 的随机核 $K_\Sigma(dy|x)$，使得 $Y|X=x\sim K_\Sigma(\cdot|x)$，故 $I(X;Y)$ 良定义。该结论源于随机核在复合下的闭包性与正则条件概率

的存在性。

证明　在标准 Borel 空间上,可测映射与随机核的复合仍为随机核,且正则条件概率存在[34]。由 G_A, Π, Φ, Θ 为因果可测映射,N 为因果随机信道,复合得到 $K_\Sigma(dy|x)$。

命题 2　数据处理结构。复合链诱导马尔可夫关系:

$$X \to \Xi_A \to \Xi_P \to \Xi_C \to \Xi_N \to Y \tag{9}$$

从而对任意中间层变量 Z 有 $I(X;\ Y) \leqslant I(X;\ Z)$。特别地,令 $Y_{\text{raw}} = \Xi_P$ 或 $Y_{\text{raw}} = \Xi_C$,则 $I(X;\ Y) \leqslant I\left(X;\ Y_{\text{raw}}\right)$。

证明　由因果性与可测性,链上每一箭头均为随机核的复合,满足马尔可夫性。由数据处理不等式[35],对于马尔可夫链 $X \to \Xi_A \to \Xi_P \to \Xi_C \to \Xi_N \to Y$,有 $I(X;\ Y) \leqslant I(X;\ Z)$ 对任意中间变量 Z 成立。

观察模型 Ω 的一个关键属性是其能否保持到达包序列的统计变异性。若观测映射 Θ 把所有输入都压缩为常数(如只记录“存在流量”而丢弃所有长度、时序、方向信息),则无论系统 Γ 产生多么丰富的统计模式,观测特征 Y 都无法反映语义差异。为排除这种退化情况,我们给出如下定义。

定义 5　非退化观测。给定窗口 T 与定义 2 中的统计量 φ,称观测模型 Ω 在该统计量下非退化,若存在常数 $\rho \in (0, 1]$ 与有界可测映射 $\psi: Y \to [-1, 1]$(仅依赖于 T 与 φ,与具体的语义对及其先验无关),则使得对任意语义对 $x \neq x' \in \Xi$ 与任意 $\delta > 0$,有

$$\left| E\left[\varphi\left(e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=x\right] - E\left[\varphi\left(e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=x'\right] \right| \geqslant \delta$$
$$\Rightarrow \left| E[\psi(Y)|X=x] - E\left[\psi(Y)|X=x'\right] \right| \geqslant \rho\delta \tag{10}$$

直观上,这要求观测映射至少保留统计量 φ 的条件期望差异的一个正比例,即使有信息损失($\rho < 1$),也不会完全抹平统计差异($\rho > 0$)。常数 ρ 与映射 ψ 仅取决于观测模型 Ω、窗口 T 与统计量 φ 的结构,而不依赖于被比较的具体语义对或其先验分布,体现了观测模型的固有性质。

这一定义排除了极端退化的观测模型,例如只记录“是否有流量”的二值指示器、只统计会话总时长而丢弃所有细粒度信息的聚合器、对所有流量返回固定特征向量的常数映射等。现实的侧信道分析模型通常满足非退化性,因为观测者希望最大化可提取的信息量,会保留包长度、时序、方向等关键维度。

至此,我们已经构建了完整的侧信道分析模型:系统 Γ 在协议层产生可辨识的统计模式,经过满足映射非退化性的加密与传输,到达观测路径;非退化的观测模型 Ω 捕获这些模式并提取特征 Y。下一节将基于这一框架,证明在系统满足窗口级映射非退化性(定义 3)的前提下,观测特征对语义的互信息 $I(X;\ Y)$ 必然严格大于 0,从而侧信道泄漏不可避免。

3　侧信道泄漏存在性定理

本节基于上一节的建模框架,给出侧信道泄漏不可避免性的严格陈述与证明。首先针对二元语义对建立期望差的稳定传递链并给出互信息的显式下界,然后推广到多元语义空间的一般情况。整个论证仅依赖因果可测的复合信道结构、数据处理不等式以及有界度量的稳定传递,不依赖具体协议细节或特定的度量选择。

3.1　技术性加强假设

设 X 为应用语义变量,Y 为观察者在模型 $\Sigma = (\Gamma, \Omega)$ 下得到的可观测特征。上一节已给出 $X \to \Xi_A \to \Xi_P \to \Xi_C \to \Xi_N \to Y$ 的马尔可夫链与对应随机核 $K_\Sigma(dy|x)$(命题 1)。

定义 2 给出了语义可辨识性的基本概念:在协议层诱导的统一轨迹空间 Z 上存在统计量 φ,使不同语义的条件期望差 $\geqslant \bar{\Delta}$。为保证这一可辨识性能够稳定传递到观测层,需要对统计量 φ 附加 Lipschitz 连续性约束。

假设 1　Lipschitz 稳健性。定义 2 中的 φ 关于 Z 上的度量 d 为 L_φ-Lipschitz 连续:

$$\left|\varphi(z) - \varphi(z')\right| \leqslant L_\varphi d(z, z'),\ \forall z, z' \in Z \tag{11}$$

这一假设保证统计量在轨迹扰动下的稳定性,即轨迹的小变化只导致统计量的小变化。常见的有界 Lipschitz 统计量包括窗口总字节数的截断版、上下行包数比(天然有界)、包到达间隔的饱和版等。实际上,绝大多数用于流量分析的统计特征都可以通过适当的截断或归一化处理,转化为 Lipschitz 连续函数。

3.2　二元语义对的侧信道泄漏定理

首先针对二元语义对,建立严格的泄漏下界。记泄漏量为 $L(\Gamma, \Omega) = I(X;\ Y)$。

定理 1　二元语义侧信道泄漏定理。在命题 1 条件下,固定窗口 $T > 0$。设存在可辨识语义对 $x \neq x' \in \Xi$,满足先验正质量条件 $P(X=x) > 0$、$P(X=x') > 0$。若系统满足以下条件:

(i)映射非退化性(定义函语义对上的实例化):存在度量 d 与常数 $C < \infty$,使得对该语义对有

$$\max\left\{ \begin{array}{l} E\left[d\left(e_P\left(\Xi_P|_{[0,T]}\right), e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=x\right], \\ E\left[d\left(e_P\left(\Xi_P|_{[0,T]}\right), e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=x'\right] \end{array} \right\} \leqslant C \tag{12}$$

(ii)语义可辨识性(定义 2):存在 $\bar{\Delta} > 0$ 与有界可

测统计量 $\varphi: Z\to[-M,M]$ 使得：

$$\left|E\left[\varphi\left(e_P\left(\Xi_P|_{[0,T]}\right)\right)|X=x\right]-E\left[\varphi\left(e_P\left(\Xi_P|_{[0,T]}\right)\right)|X=x'\right]\right|\geqslant\bar{\Delta} \tag{13}$$

（iii）Lipschitz 稳健性（假设 1）：上述统计量 φ 关于 Z 上的度量 d 为 L_φ-Lipschitz 连续：

$$\left|\varphi(z)-\varphi(z')\right|\leqslant L_\varphi d(z,z'),\ \forall z,z'\in Z \tag{14}$$

（iv）非退化观测（定义 5 针对统计量 φ）：存在常数 $\rho\in(0,1]$ 与有界可测 $\psi: Y\to[-1,1]$（仅依赖 T 与 φ，与语义对无关），使对任意语义对 $a\neq b\in X$，有

$$\left|E\left[\varphi\left(e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=a\right]-E\left[\varphi\left(e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=b\right]\right|\geqslant\delta$$
$$\Rightarrow|E[\psi(Y)|X=a]-E[\psi(Y)|X=b]|\geqslant\rho\delta \tag{15}$$

（v）可辨识性传递条件：度量偏差界 C、Lipschitz 常数 L_φ 与可辨识性裕量 $\bar{\Delta}$ 满足：

$$C<\frac{\bar{\Delta}}{2L_\varphi} \tag{16}$$

则观测特征对语义的互信息满足 $I(X;\ Y)>0$，且有显式下界：

$$I(X;\ Y)\geqslant\frac{2}{\ln 2}P(X=x)P(X=x')\left(\frac{\rho\left[\bar{\Delta}-2L_\varphi C\right]}{2}\right)^2 \tag{17}$$

特别地，若限制到二元等先验子问题（即 $X\in\{x,x'\}$ 且 $P(X=x)=P(X=x')=1/2$），则下界化为

$$I(X;\ Y)\geqslant\frac{1}{2\ln 2}\left(\frac{\rho\left[\bar{\Delta}-2L_\varphi C\right]}{2}\right)^2 \tag{18}$$

定理聚焦于二元语义对 (x,x') 的分析，这是证明技术的本质要求：基于 Pinsker 不等式与全变差的论证天然是成对的。条件（i）仅要求该语义对满足映射非退化性，在效率优先的实际系统中，若映射对所有语义都保持非退化性，则对任意语义对自然成立。条件（v）刻画了侧信道存在性的边界：度量偏差 C 不能过大，否则协议层的可辨识性在传递到网络层时会被扰动完全抹平。在网站指纹的典型场景中，可以把 x 与 x' 理解为“访问网站 A”和“访问网站 B”两类语义，其中 $\bar{\Delta}$ 可选取为二者在固定时间窗口内总字节数、上下行包数比、突发持续时间/载荷等聚合统计的期望差异；C 则对应于在同一窗口下由于加密封装、TCP 重传和随机抖动引入的统一轨迹空间 Z 中长度与到达时间的平均偏差；而 ρ 反映观测链路在采样粒度、时间戳精度等限制下仍能保留的这些差异的比例。

证明需要以下引理。

引理 1　期望差与全变差的关系。设 $f: S\to[-M,M]$ 为有界可测函数，P,Q 为 S 上的两个概率测度。若：

$$\left|E_P[f]-E_Q[f]\right|\geqslant\delta \tag{19}$$

则全变差距离满足：

$$\mathrm{TV}(P,Q)\geqslant\frac{\delta}{2M} \tag{20}$$

证明　由全变差的对偶表示如下：

$$\mathrm{TV}(P,Q)=\frac{1}{2}\sup_{\|g\|_\infty\leqslant 1}\left|E_P[g]-E_Q[g]\right| \tag{21}$$

对有界函数 $f: S\to[-M,M]$，归一化 $g:=f/M$ 得 $\|g\|_\infty=1$，因此有

$$\left|E_P[f]-E_Q[f]\right|=M\left|E_P[g]-E_Q[g]\right|\leqslant 2M\cdot\mathrm{TV}(P,Q) \tag{22}$$

若 $\left|E_P[f]-E_Q[f]\right|\geqslant\delta$，则 $\mathrm{TV}(P,Q)\geqslant\frac{\delta}{2M}$。

定理 1 的证明。证明沿 $\bar{\Delta}\to(L_\varphi)\to d\to(C_T)\to\rho\to I(X;\ Y)$ 展开：以 Lipschitz 性将协议层差异界定到轨迹空间，再由 C_T 吸收扰动、以 ρ 刻画可观测比例，最终拼接不等式得到互信息下界。证明分 4 步：在协议层 Ξ_P 建立期望差，传递到网络层 Ξ_N，再传递到观测层 Y，最后转化为互信息。整个推导在统一轨迹空间 Z 上进行，依赖条件（i）~（v）构成的可辨识性传递链。

步骤 1：协议层的期望差可由条件（ii）直接给出。

由条件（ii）语义可辨识性，存在有界统计量 $\varphi: Z\to[-M,M]$，满足：

$$\left|E\left[\varphi\left(e_P\left(\Xi_P|_{[0,T]}\right)\right)|X=x\right]-E\left[\varphi\left(e_P\left(\Xi_P|_{[0,T]}\right)\right)|X=x'\right]\right|\geqslant\bar{\Delta} \tag{23}$$

步骤 2：从协议层到网络层的期望差传递。

引入简化记号：$z_P:=e_P\left(\Xi_P|_{[0,T]}\right), z_N:=e_N\left(\Xi_N|_{[0,T]}\right)$ 为统一空间 Z 中的轨迹表示。由条件（i）映射非退化性，对语义对 $x\neq x'\in X$ 有

$$E\left[d(z_P,z_N)|X=x\right]\leqslant C,\ E\left[d(z_P,z_N)|X=x'\right]\leqslant C \tag{24}$$

由条件（iii）Lipschitz 稳健性，对任意轨迹 $z_P,z_N\in Z$ 有

$$\left|\varphi(z_P)-\varphi(z_N)\right|\leqslant L_\varphi\cdot d(z_P,z_N) \tag{25}$$

取关于 $X=x$ 的条件期望：

$$\begin{aligned}&\left|E\left[\varphi(z_P)|X=x\right]-E\left[\varphi(z_N)|X=x\right]\right|\\&=\left|E\left[\varphi(z_P)-\varphi(z_N)|X=x\right]\right|\\&\leqslant E\left[\left|\varphi(z_P)-\varphi(z_N)\right||X=x\right]\ (\text{Jensen 不等式})\\&\leqslant E\left[L_\varphi\cdot d(z_P,z_N)|X=x\right]\ (\text{Lipschitz 性})\\&=L_\varphi\cdot E\left[d(z_P,z_N)|X=x\right]\\&\leqslant L_\varphi\cdot C\ (\text{条件i})\end{aligned} \tag{26}$$

对 $X=x'$ 同理，有

$$\left|E\left[\varphi\left(z_P\right)|X=x'\right]-E\left[\varphi\left(z_N\right)|X=x'\right]\right|\leqslant L_\varphi\cdot C \quad (27)$$

应用三角不等式:

$$\begin{aligned}&\left|E\left[\varphi\left(z_N\right)|X=x\right]-E\left[\varphi\left(z_N\right)|X=x'\right]\right|\\&\geqslant\left|E\left[\varphi\left(z_P\right)|X=x\right]-E\left[\varphi\left(z_P\right)|X=x'\right]\right|\\&\quad-\left|E\left[\varphi\left(z_P\right)|X=x\right]-E\left[\varphi\left(z_N\right)|X=x\right]\right|\\&\quad-\left|E\left[\varphi\left(z_P\right)|X=x'\right]-E\left[\varphi\left(z_N\right)|X=x'\right]\right|\\&\geqslant\bar{\Delta}-L_\varphi C-L_\varphi C\\&=\bar{\Delta}-2L_\varphi C=:\delta_N\end{aligned} \quad (28)$$

由条件(v),$C<\dfrac{\bar{\Delta}}{2L_\varphi}$,从而 $\delta_N=\bar{\Delta}-2L_\varphi C>0$。

步骤3:从网络层到观测层的期望差传递。

由条件(iv)非退化观测,应用于上一步得到的期望差 $\delta_N>0$ 与语义对(x,x'),存在有界观测统计量 ψ:$Y\to[-1,1]$,使得:

$$\left|E[\psi(Y)|X=x]-E\left[\psi(Y)|X=x'\right]\right|\geqslant\rho\delta_N=\rho\left(\bar{\Delta}-2L_\varphi C\right) \quad (29)$$

步骤4:从期望差到互信息(含先验权重)。

由引理1,应用于观测层条件分布与统计量 ψ:$Y\to[-1,1]$(即 $M=1$),得:

$$\mathrm{TV}\left(P_{Y|X=x},P_{Y|X=x'}\right)\geqslant\frac{\rho\delta_N}{2}=\frac{\rho\left(\bar{\Delta}-2L_\varphi C\right)}{2} \quad (30)$$

对于一般先验分布,利用互信息与条件分布全变差的标准关系(参见文献[35]),有

$$I(X;Y)\geqslant\frac{2}{\ln 2}P(X=x)P\left(X=x'\right)\mathrm{TV}^2\left(P_{Y|X=x},P_{Y|X=x'}\right) \quad (31)$$

代入全变差下界:

$$I(X;Y)\geqslant\frac{2}{\ln 2}P(X=x)P\left(X=x'\right)\left(\frac{\rho\left(\bar{\Delta}-2L_\varphi C\right)}{2}\right)^2 \quad (32)$$

由条件(v)与先验正质量条件,右端严格为正,从而 $I(X;Y)>0$。

在二元等先验子问题$\left(P(X=x)=P\left(X=x'\right)=1/2\right)$下,上式化为

$$I(X;Y)\geqslant\frac{2}{\ln 2}\cdot\frac{1}{4}\left(\frac{\rho\left(\bar{\Delta}-2L_\varphi C\right)}{2}\right)^2=\frac{1}{2\ln 2}\left(\frac{\rho\left(\bar{\Delta}-2L_\varphi C\right)}{2}\right)^2 \quad (33)$$

证毕。

互信息下界与先验有关:$I(X;Y)$ 的下界按可辨识语义对的先验质量 $P(X=x)P(X=x')$ 加权。在未知先验时,可用保守下界 $p_{\min}^2$ 替代,其中 $p_{\min}$ 为支撑集上的最小先验质量。证明的核心洞察是:条件(i)~(v)构成一条期望差的稳定传递链,即可辨识性在每一步都以可控损失 $L_\varphi C$ 传递,只要条件(v)保证总损失不超过初始裕量 $\bar{\Delta}$ 的一半,最终在观测层就可保持为正。

3.3　多元语义空间的泄漏不可避免性

基于二元定理,我们现在推广到一般多元语义空间的情况,建立侧信道泄漏的不可避免性。

推论1　多元语义侧信道存在性。设语义空间 Ξ 非平凡($|\Xi|\geqslant 2$ 且先验支撑至少包含两个元素)。在命题1条件下,固定窗口 $T>0$。若加密通信系统 Γ 与观测模型 Ω 满足以下条件。

(i)效率优先设计:存在度量 d 与常数 $C<\infty$,使得对所有 $x\in\Xi$ 有

$$E\left[d\left(e_P\left(\Xi_P|_{[0,T]}\right),e_N\left(\Xi_N|_{[0,T]}\right)\right)|X=x\right]\leqslant C \quad (34)$$

(ii)语义多样性:存在至少一对可辨识语义 $x\neq x'\in\Xi$ 与有界 Lipschitz 统计量 $\varphi:Z\to[-M,M]$(Lipschitz 常数为 L_φ),使得:

$$\left|E\left[\varphi\left(e_P\left(\Xi_P|_{[0,T]}\right)\right)|X=x\right]-E\left[\varphi\left(e_P\left(\Xi_P|_{[0,T]}\right)\right)|X=x'\right]\right|\geqslant\bar{\Delta}>0 \quad (35)$$

且 $P(X=x)>0$、$P(X=x')>0$

(iii)观测者理性:观测模型 Ω 针对该统计量 φ 满足非退化性(定义5),存在 $\rho\in(0,1]$。

(iv)可辨识性传递条件:$C<\dfrac{\bar{\Delta}}{2L_\varphi}$。

则观测特征对语义的互信息满足 $I(X;Y)>0$。

证明　由条件(i),效率优先设计对所有语义成立,特别地对语义对(x,x')成立。结合条件(ii)(iii)(iv),该语义对满足定理1的所有条件。因此

$$I(X;Y)\geqslant\frac{2}{\ln 2}P(X=x)P\left(X=x'\right)\left(\frac{\rho\left[\bar{\Delta}-2L_\varphi C\right]}{2}\right)^2>0 \quad (36)$$

证毕。

推论1表明:在效率优先的多元语义系统中,只要存在至少一对应用在统计上可区分,侧信道泄漏就不可避免。这一结论的普遍性源于以下三个方面。

(1)效率优先是系统级约束(条件i)。现实系统必须满足带宽、时延等性能要求,故对所有语义都有 $C<\infty$。

(2)语义多样性是应用的必然结果(条件ii)。不同应用类型,如视频流(包大且密集)、网页浏览(包小且稀疏)、即时通信(双向对称)、文件传输(单向集中)在包大小分布、时序模式、上下行比等统计特征上必然存在差异,这源于应用逻辑本身而与加密无关。

(3)观测者理性是分析者目标(条件iii)。侧信道

分析者以最大化信息提取为目标，会保留关键统计特征，故$\rho>0$。

条件(ii)仅要求“存在一对可辨识”，而非“所有语义两两可辨识”，这是极弱的假设。在包含$n\geqslant 2$个应用的现实系统中，几乎不可能让所有应用产生统计不可区分的流量。这需要彻底改变应用的工作方式，违背应用设计的初衷。

因此，在“效率优先的可用系统+非平凡的应用场景+理性的观测者”这一普遍情形下，侧信道泄漏$I(X;\ Y)>0$不可避免。

4 理论分析与讨论

本节解释存在性定理的操作含义，讨论从信息论下界到实际攻击性能的转化，以及定理揭示的效率—隐私权衡。

4.1 从信息论下界到攻击可达性的操作解释

定理1与推论1断言在效率优先的系统中$I(X;\ Y)>0$，并给出显式下界。为了将这一信息论陈述转化为对实际攻击性能的可操作预测，我们建立从全变差到分类准确率的精确联系。

二元情形的准确率下界。对于二元等先验问题$\left(P(X=x)=P(X=x')=1/2\right)$，最优贝叶斯分类器的错误率与准确率满足精确关系：

$$\begin{aligned}P_e^*&=\frac{1-\mathrm{TV}\left(P_{Y|X=x},P_{Y|X=x'}\right)}{2}\\ \mathrm{Acc}^*&=\frac{1+\mathrm{TV}\left(P_{Y|X=x},P_{Y|X=x'}\right)}{2}\end{aligned}\tag{37}$$

由定理1步骤4，我们已得到全变差下界：

$$\mathrm{TV}\left(P_{Y|X=x},P_{Y|X=x'}\right)\geqslant\frac{\rho\left(\bar{\Delta}-2L_\varphi C\right)}{2}\tag{38}$$

其中，C为映射非退化常数(定理条件(i))。

代入上式，得到最优准确率的下界：

$$\mathrm{Acc}^*\geqslant\min\left\{1,\ \frac{1}{2}+\frac{1}{4}\rho\left(\bar{\Delta}-2L_\varphi C\right)\right\}\tag{39}$$

因统计量$\varphi:Z\to[-M,M]$有界，故$\bar{\Delta}\leqslant 2M$，结合非退化观测定义中$\rho\in(0,1]$与$\psi:Y\to[-1,1]$的设定，上式右端自然不越界。

这是最优贝叶斯分类器准确率的下界，从而使分类器达到该水平。例如，若$\rho=0.8$、$\bar{\Delta}=1.0$、$L_\varphi C=0.2$，则

$$\mathrm{Acc}^*\geqslant\frac{1}{2}+\frac{1}{4}\times 0.8\times(1.0-0.4)=0.62\tag{40}$$

即最优准确率至少为62%。

对于一般先验或多类情形($M>2$)，可在最难区分的二元子问题上应用该界(下界可能较松)，或用“一对多”的并合策略得到保守下界。

Fano不等式给出的错误率下界。作为补充，我们也可用Fano不等式刻画信息论意义下的错误率下界。设语义空间Ξ包含$M\geqslant 2$个元素，P_e^*为最小贝叶斯错误概率。经典Fano不等式给出：

$$H(X|Y)\leqslant H_2\left(P_e^*\right)+P_e^*\log_2(M-1)\tag{41}$$

其中，$H_2(p)=-p\log_2 p-(1-p)\log_2(1-p)$为二元熵(以比特为单位)，$H(X|Y)=H(X)-I(X;\ Y)$为后验熵。整理得：

$$P_e^*\geqslant\frac{H(X)-I(X;\ Y)-1}{\log_2(M-1)}\tag{42}$$

对于二元等先验情形，$H(X)=1$，上式等价于$I(X;\ Y)\geqslant 1-H_2(P_e^*)$，从而$P_e^*\geqslant H_2^{-1}(1-I(X;\ Y))$。

例如，若$I(X;\ Y)=0.1$比特，则$P_e^*\geqslant H_2^{-1}(0.9)\approx 0.317$，即最优错误率至少为31.7%，准确率至多约68.3%。需注意Fano不等式给出的是错误率的下界而非上界，它阐明了即使最优分类器也无法突破信息论限制，但不能直接预测实际攻击的可达性能。

多次观测的累积效应。在实际侧信道分析中，攻击者往往可以观测多次会话$Y^{(1)},Y^{(2)},\cdots,Y^{(n)}$(同一用户访问同一网站的不同会话，或拼接同一流的多个时间窗口)。假设给定语义X的条件下，各次观测独立同分布，则联合互信息满足可加性：

$$I(X;\ Y^{(1:n)})=\sum_{i=1}^{n}I\left(X;\ Y^{(i)}\right)=n\cdot I(X;\ Y)\tag{43}$$

此时，错误率以Chernoff信息为指数衰减。对于二元等先验问题，最优贝叶斯错误率的大偏差渐近行为由精确定理刻画：

$$-\lim_{n\to\infty}\frac{1}{n}\log P_e^*(n)=C\left(P_{Y|X=x},P_{Y|X=x'}\right)\tag{44}$$

其中，$C(\cdot,\cdot)$为Chernoff信息(此处对数底为e，单位为nats)，log表示自然对数。

为建立从全变差到Chernoff信息的下界链条，引入Bhattacharyya系数$\mathrm{BC}=\int\sqrt{p(y)q(y)}\,\mathrm{d}y$与Bhattacharyya距离$B=-\ln\mathrm{BC}$。已知关系：

$$\mathrm{TV}(P,Q)\leqslant\sqrt{1-\mathrm{e}^{-2B}},\ C(P,Q)\geqslant B\tag{45}$$

从而：

$$C(P,Q)\geqslant B\geqslant-\frac{1}{2}\ln(1-\mathrm{TV}^2(P,Q))>0\tag{46}$$

将我们在定理中得到的$\mathrm{TV}\left(P_{Y|X=x},P_{Y|X=x'}\right)\geqslant\frac{\rho\left(\bar{\Delta}-2L_\varphi C\right)}{2}$代入上式，得到Chernoff信息的显式下界：

$$C(P_{Y|X=x},P_{Y|X=x'})\geqslant-\frac{1}{2}\ln\left(1-\left[\frac{\rho(\bar{\Delta}-2L_\varphi C)}{2}\right]^2\right)>0\tag{47}$$

这保证了错误率以观测次数 n 指数趋于0。

对于多类情形($M>2$),若各类别先验具有正质量且样本在给定 X 条件下独立同分布,则整体贝叶斯错率的误差指数下界由 $\min_{x\neq x'} C\left(P_{Y|X=x}, P_{Y|X=x'}\right)$ 控制。这一结论来自常见的并合与最差对主导论证:总体错误率由最难区分的语义对控制。

这解释了实践中观测到的普遍现象。

(1)长观察窗口提升准确率。在 $\bar{\Delta}(T)$ 的增长不被 $2L_{\varphi}C(T)$ 抵消时,增大时间窗口 T 使 $\delta_N(T)=\bar{\Delta}(T)-2L_{\varphi}C(T)$ 增大,从而全变差下界与准确率下界随 T 增大。

(2)多段会话拼接显著改善识别。拼接 n 个独立会话使互信息线性累积为 $n\cdot I(X;\ Y)$(以比特计),准确率以 Chernoff 指数(以 nats 计)收敛。

(3)指数级收敛到完美识别。在条件独立假设下,错误率以 $\exp(-n\cdot C)$ 指数衰减至0,推论1保证 $I(X;\ Y)>0$ 在效率优先系统中必然成立,从而上述累积效应不可避免,即只要攻击者有足够的观测预算且满足条件独立假设,识别准确率就会趋于完美。这是侧信道泄漏不可消除性的可操作含义。

4.2　效率-隐私权衡的根本性与不可逾越性

定理1的5个条件揭示了降低泄漏 $I(X;\ Y)$ 的唯一途径及其代价。我们分析每个条件的"破坏成本"。

条件(i):映射非退化性 $C<\infty$。这是效率优先设计的直接体现。要破坏该条件(增大 $C\to\infty$),系统必须在以下维度之一付出代价。

长度维度:大量填充使 $E\left[d_{\text{length}}\left(z_P, z_N\right)\right]$ 增大。例如,填充所有包到 MTU(1 500 字节)使小包(如 40 字节 ACK)膨胀数十倍,带宽开销达数量级增长。

时序维度:人工延迟使 $E\left[d_{\text{time}}\left(z_P, z_N\right)\right]$ 增大。例如,引入秒级延迟破坏实时应用(VoIP 要求单向延迟 < 150 ms)。

方向维度:掩护流量改变上下行比,增大 $E\left[d_{\text{direction}}\left(z_P, z_N\right)\right]$。双向掩护使带宽开销翻倍。

定理中的条件(v)要求 $C<\dfrac{\bar{\Delta}}{2L_{\varphi}}$ 才能保证泄漏传递。若要使泄漏下界趋于0,需要 $C\to\dfrac{\bar{\Delta}}{2L_{\varphi}}$,这意味着效率开销趋于某个临界值,该临界值由应用的固有可辨识性 $\bar{\Delta}$ 决定,不可改变。

条件(ii):语义可辨识性 $\bar{\Delta}>0$。这是应用多样性的必然结果。要破坏该条件(使 $\bar{\Delta}\to 0$),需要让所有应用产生统计不可区分的流量。这在实践中几乎不可能。譬如,视频流与网页浏览在带宽需求上相差数量级,即时通信与文件下载在交互模式上本质不同,VoIP 与 HTTP 在时序特征上差异显著。

要完全消除 $\bar{\Delta}$,需要强制所有应用以相同的恒定速率、相同的包大小、相同的双向模式传输,这彻底破坏了应用的功能性,使问题失去意义。

条件(iii)~(iv): Lipschitz 性与观测非退化性。条件(iii)是统计量选择的技术要求,实践中几乎所有有用的统计量(窗口总字节数、包数、上下行比等)都可通过截断或归一化满足 Lipschitz 性。条件(iv)刻画观测者理性,由观测者的技术能力决定而非系统设计者可控。

权衡的不可逾越性。综合上述分析,降低泄漏的代价呈现三难困境:当破坏的条件为增大 C(放松非退化性)时,付出的代价是牺牲效率(带宽 / 延迟);当破坏的条件为减小 $\bar{\Delta}$(均质化应用)时,付出的代价是破坏功能(应用不可用);当破坏的条件为减小 ρ(压缩观测)时,付出的代价是超出控制(由观测者决定)。

在带宽开销要求、端到端延迟要求等固定的业务需求约束下,存在不可逾越的泄漏下界。这不是某一协议实现的缺陷,而是效率优先与语义多样性共同决定的结构性限制。

4.3　防御的理论边界与正确的工程目标

推论1表明零泄漏($I(X;\ Y)=0$)在效率优先的系统中不可达。这一结论对防御机制设计具有重要指导意义。定理1进一步揭示了防御机制的作用机理:定速填充通过增大度量偏差 C 来模糊协议层到网络层的映射,代价是带宽开销;差分隐私机制通过添加噪声降低观测保真度 ρ,代价是效用损失;而完全混淆防御试图减小语义可辨识性 $\bar{\Delta}$,但会破坏应用功能。这些机制的本质是在三种困境构成的约束空间中寻找不同的权衡点。

错误的目标:追求 $I(X;\ Y)=0$。许多防御方案(如 Tor 的流量混淆、VPN 的恒速填充等)隐含地以"消除侧信道"为目标。定理表明这一目标在保持系统可用的前提下不可达。实践中观察到的现象验证了这一结论。

(1)Tamaraw 等强防御在真实 Tor 网络部署中延迟增大 78%,带宽开销达 135%[36],虽能显著降低攻击准确率,但其高开销限制了实际部署。

(2)如 BuFLO 和 CS-BuFLO 等恒速率填充策略可降低识别准确率,但需要 100% 以上的带宽开销[22,37],在实际部署中不可行。

(3)混淆防御如 WTF-PAD 和 FRONT 虽开销较低,但无法抵御最新的深度学习攻击(准确率可达 90% 以上)[36]。

正确的目标：约束优化。定理揭示的正确工程目标是：在给定效率约束与功能要求下，最小化泄漏。形式化为约束优化问题：

$$\min_{\theta\in\Theta} I(X;Y;\boldsymbol{\theta})\ \text{s.t.}\begin{cases}\text{带宽开销}\leqslant\beta_{\max} & (\text{如}10\%)\\ \text{延迟增加}\leqslant\Delta t_{\max} & (\text{如}50\ \text{ms})\\ \text{应用功能完整} & \left(\bar{\Delta}\geqslant\Delta_{\min}\right)\end{cases}\quad(48)$$

其中，$\boldsymbol{\theta}$为填充策略、定时扰动、掩护流量等防御参数的联合向量。

5 结论

本文从信息论与系统设计出发，构建了侧信道分析的形式化模型$\Sigma=(\Gamma,\Omega)$，把生成、封装、加密、传输、观测的全过程抽象为因果可测的马尔可夫链$X\to\Xi_A\to\Xi_P\to\Xi_C\to\Xi_N\to Y$。基于该框架，本文证明了侧信道存在性定理（定理1）：对于可辨识的二元语义对，在满足映射非退化性（$E\left[d\left(z_P,z_N\right)|X\right]\leqslant C$）、语义可辨识性（期望差$\geqslant\bar{\Delta}$）、Lipschitz稳健性（$\left|\varphi(z)-\varphi(z')\right|\leqslant L_\varphi d(z,z')$）、非退化观测（保留比例$\rho>0$）以及可辨识性传递条件（$C<\bar{\Delta}/2L_\varphi$）的前提下，观测特征与语义变量的互信息满足显式下界$I(X;Y)\geqslant\frac{1}{2\ln 2}\left(\frac{\rho\left[\bar{\Delta}-2L_\varphi C\right]}{2}\right)^2>0$。推论1进一步表明：在效率优先的多元语义系统中，只要存在至少一对应用在统计上可区分（源于应用逻辑的固有差异），侧信道泄漏就不可避免。通过全变差-准确率的精确关系与Bhattacharyya-Chernoff下界链条，本文建立了从信息论下界到实际攻击性能的量化联系，揭示了在条件独立假设下多次观测的指数累积效应使识别准确率趋于完美的必然性。

本文分析表明，降低泄漏面临三大困境：增大度量偏差C需牺牲效率，减小语义可辨识性$\bar{\Delta}$将破坏应用功能，而观测非退化性ρ由分析者而非系统设计者控制。因此，正确的工程目标不是追求不可达的零泄漏，而是在给定效率约束与功能要求下最小化泄漏的约束优化问题。值得强调的是，本文给出的存在性定理与显式互信息下界可作为协议演进与防御评估的理论基线：在给定带宽、时延与兼容性等效率约束下，可将不同协议配置或防御策略映射为度量d与非退化常数$C(T)$的变化，从而量化其对泄漏下界的影响，并据此开展面向目标的约束优化与方案比选。

特别是在网站指纹场景下，本文的理论框架对攻击特征选择具有直接指导意义。根据互信息下界的结构，攻击者应优先选取期望差$\bar{\Delta}$较大的宏观聚合特征（如会话总字节数、上下行包数比、突发载荷等），同时兼顾统计量的Lipschitz稳健性以抵抗网络扰动。应重视时序结构中的辨识性信息，如包间隔分布与突发-静默交替模式。还可利用多会话观测的Chernoff指数效应实现识别准确率的指数级提升。上述启示表明，本文建立的“期望差—Lipschitz稳健性—映射非退化性”分析链条，为流量特征侧信道攻击提供了可解释的理论依据与可操作的特征设计准则。

本文的理论框架在实用化方面仍存在若干开放问题。首先，定理给出的下界依赖于度量d的选择与常数C、L_φ的估计，如何从实测流量数据中识别或验证这些参数，如何针对不同协议族（TLS 1.3、QUIC）与业务类型（视频流、网页浏览）建立参数库，是理论落地的关键步骤。其次，本文的非退化性条件（定义3、定义5）基于期望意义下的度量界，但现实网络存在拥塞突发、路由抖动等瞬态扰动，如何在概率意义上（如高概率界或分位数约束）重新表述条件并推导相应的泄漏下界，将增强结论的鲁棒性。再次，本文聚焦于被动观测场景，但主动探测（如网站指纹攻击中的诱导访问）与自适应攻击（观察者根据中间结果调整策略）可能突破静态下界，如何在博弈论框架下刻画攻防双方的Nash均衡与最优策略，是动态对抗环境下的重要课题。此外，本文基于互信息的信息论度量与差分隐私的(ε,δ)-DP保证之间的精确关系尚未建立，能否证明“满足ε-DP的机制必然导致互信息下降至$f(\varepsilon)$以下”或类似的等价性定理，将为隐私机制设计提供可操作的形式化准则。最后，多任务场景（如先识别应用类别再细分具体网站）的层次化泄漏分析，以及考虑时间相关性与长期观察的动态泄漏累积模型，都是值得深入研究的方向。这些问题的解决将把本文的存在性结论推进为可计算、可验证、可优化的工程实践框架。

参考文献

[1] Wang T, Cai X, Nithyanand R, et al. Effective attacks and provable defenses for website fingerprinting[C]//Proceedings of the 23rd USENIX conference on Security Symposium. New York: ACM, 2014: 143-157.

[2] Shen M, Zhang J P, Zhu L H, et al. Accurate decentralized application identification via encrypted traffic analysis using graph neural networks[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 2367-2380.

[3] Mei H T, Cheng G, Yuan Y L. High precision and efficient anonymous traffic classification in the real-world[J]. IEEE Transactions on Networking, 2025, 33(3): 966-981.

[4] Kocher P C. Timing attacks on implementations of diffie-Hellman, RSA, DSS, and other systems[C]//Advances in Cryptology - CRYPTO'96. Berlin, Heidelberg: Springer,

1996: 104-113.

[5] Kocher P, Jaffe J, Jun B. Differential power analysis[C]// Advances in Cryptology - CRYPTO'99. Berlin, Heidelberg: Springer, 1999: 388-397.

[6] Hintz A. Fingerprinting websites using traffic analysis[M]// Privacy enhancing technologies. Berlin, Heidelberg: Springer, 2003: 171-178.

[7] Lin X J, Xiong G, Gou G P, et al. ET-BERT: A contextualized datagram representation with pre-training transformers for encrypted traffic classification[C]//Proceedings of the ACM Web Conference 2022. New York: ACM, 2022: 633-642.

[8] Shen M, Ye K, Liu X T, et al. Machine learning-powered encrypted network traffic analysis: A comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2023, 25(1): 791-824.

[9] Li S, Guo H, Hopper N. Measuring information leakage in website fingerprinting attacks and defenses[C]//Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2018: 1977-1992.

[10] Cai X, Nithyanand R, Wang T, et al. A systematic approach to developing and evaluating website fingerprinting defenses[C]//Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2014: 227-238.

[11] Dwork C, Kenthapadi K, McSherry F, et al. Our data, ourselves: Privacy via distributed noise generation[C]//Advances in Cryptology - EUROCRYPT 2006. Berlin, Heidelberg: Springer, 2006: 486-503.

[12] Sabzi A, Vora R, Goswami S, et al. NetShaper: A differentially private network side-channel mitigation system[C]// Proceedings of the 33rd USENIX Security Symposium. Berkeley: USENIX Association, 2024: 3385-3402.

[13] Chaum D L. Untraceable electronic mail, return addresses, and digital pseudonyms[J]. Communications of the ACM, 1981, 24(2): 84-90.

[14] Díaz C, Seys S, Claessens J, et al. Towards measuring anonymity[M]//Privacy enhancing technologies. Berlin, Heidelberg: Springer, 2003: 54-68.

[15] Deng Y X, Pang J, Wu P. Measuring anonymity with relative entropy[C]//Formal Aspects in Security and Trust. Berlin, Heidelberg: Springer, 2007: 65-79.

[16] Serjantov A, Danezis G. Towards an information theoretic metric for anonymity[C]//Privacy Enhancing Technologies. Berlin, Heidelberg: Springer, 2003: 41-53.

[17] Chatzikokolakis K, Palamidessi C, Panangaden P. Anonymity protocols as noisy channels[J]. Information and Computation, 2008, 206(2/3/4): 378-401.

[18] Kedogan D, Agrawal D, Penz S. Limits of anonymity in open environments[M]//Information Hiding. Berlin, Heidelberg: Springer, 2002: 53-69.

[19] Danezis G. Statistical disclosure attacks[M]//Security and privacy in the age of uncertainty. Boston, MA: Springer US, 2003: 421-426.

[20] Jelle V D H, Lazar D, Zaharia M, et al. Vuvuzela: Scalable private messaging resistant to traffic analysis[C]// Proceedings of the 25th Symposium on Operating Systems Principles. New York: ACM, 2015: 137-152.

[21] Tyagi N, Gilad Y, Leung D, et al. Stadium: A distributed metadata-private messaging system[C]//Proceedings of the 26th Symposium on Operating Systems Principles. New York: ACM, 2017: 423-440.

[22] Panchenko A, Niessen L, Zinnen A, et al. Website fingerprinting in onion routing based anonymization networks[C]//Proceedings of the 10th Annual ACM Workshop on Privacy in the Electronic Society. New York: ACM, 2011: 103-114.

[23] Dyer K P, Coull S E, Ristenpart T, et al. Peek-a-boo, I still see you: Why efficient traffic analysis countermeasures fail[C]//2012 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2012: 332-346.

[24] Wang T, Goldberg I. Walkie-talkie: An efficient defense against passive website fingerprinting attacks[C]//Proceedings of the 26th USENIX Security Symposium. Berkeley: USENIX Association, 2017: 1375-1390.

[25] Huang J N, Liu W W, Liu G J, et al. STAP: Leveraging state-transition adversarial perturbations for asymmetric website fingerprinting defenses[J]. IEEE Transactions on Network and Service Management, 2025, 22(6): 6200-6214.

[26] Wright C V, Coull S E, Monrose F. Traffic morphing: An efficient defense against statistical traffic analysis[C]// Proceedings of the 16th Network and Distributed Security Symposium. Reston: The Internet Society, 2009: 237-250.

[27] Cherubin G. Bayes, not Naïve: Security bounds on website fingerprinting defenses[J]. Proceedings on Privacy Enhancing Technologies, 2017, 2017(4): 215-231.

[28] Fu C P, Li Q, Shen M, et al. Realtime robust malicious traffic detection via frequency domain analysis[C]//Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM,

2021: 3431-3446.
[29] Fu C P, Li Q, Shen M, et al. Detecting tunneled flooding traffic via deep semantic analysis of packet length patterns[C]//Proceedings of the 2024 on ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2024: 3659-3673.
[30] Fu C P, Li Q, Xu K. Detecting unknown encrypted malicious traffic in real time via flow interaction graph analysis[PP/OL]. V1. arXiv (2023-01-31)[2025-12-29]. https://arXiv.org/abs/2301.13686.
[31] Camenisch J, Lysyanskaya A. A formal treatment of onion routing[C]//Advances in Cryptology - CRYPTO 2005. Berlin, Heidelberg: Springer, 2005: 169-187.
[32] Feigenbaum J, Johnson A, Syverson P. A model of onion routing with provable anonymity[C]//Financial Cryptography and Data Security. Berlin, Heidelberg: Springer, 2007: 57-71.
[33] Danezis G, Goldberg I. Sphinx: A compact and provably secure mix format[C]//2009 30th IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2009: 269-282.
[34] Gray R M. Probability, random processes, and ergodic properties[M]. New York: Springer, 1988.
[35] Cover T M, Thomas J A. Elements of information theory[M]. 2nd ed. Hoboken: J. Wiley, 2006.
[36] Shen M, Ji K X, Wu J H, et al. Real-time website fingerprinting defense via traffic cluster anonymization[C]//2024 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2024: 3238-3256.
[37] Cai X, Nithyanand R, Johnson R. CS-BuFLO: A congestion sensitive website fingerprinting defense[C]//Proceedings of the 13th Workshop on Privacy in the Electronic Society. New York: ACM, 2014: 121-130.

作者简介

刘光杰 男,1980年2月出生于江苏省徐州市。2007年博士毕业于南京理工大学,现为南京信息工程大学电子与信息工程学院教授。主要研究方向为加密网络流量分析、无线隐蔽通信等。
E-mail: gjieliu@njust.edu.cn

程 光 男,1973年2月出生于安徽省黄山市。2003年博士毕业于东南大学,现为东南大学网络空间安全学院教授。主要研究方向为加密网络流量分析、网络安全主动防御等。
E-mail: chengguang@seu.edu.cn

刘伟伟 男,1988年4月出生于江苏省淮安市。2015年博士毕业于南京理工大学,现为南京理工大学自动化学院副教授。主要研究方向为加密网络流量分析、智能体安全等。
E-mail: lwwnjust@njust.edu.cn