

网络加密流量侧信道攻击研究综述

李玎^{1,2}, 祝跃飞^{1,2}, 芦斌^{1,2}, 林伟^{1,2}

(1. 信息工程大学, 河南 郑州 450001;
2. 数学工程与先进计算国家重点实验室, 河南 郑州 450001)

摘要: 网络加密流量侧信道攻击通过分析、提取网络应用通信过程中泄露的数据包长度、时间等侧信道信息, 能够识别用户的身份和行为, 甚至还原用户输入的原始数据。基于信息论建立了网络加密流量侧信道攻击模型, 使用统一的模型框架分析了代表性的指纹攻击、击键攻击和语音攻击的方法和效果, 讨论了基于隐藏数据包长度和时间信息的防御方法, 结合技术发展前沿对未来可能的研究方向进行了展望。

关键词: 网络应用; 加密流量; 侧信道攻击; 信息增益

中图分类号: TP311

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2021050

Survey of side channel attack on encrypted network traffic

LI Ding^{1,2}, ZHU Yuefei^{1,2}, LU Bin^{1,2}, LIN Wei^{1,2}

1. Information Engineering University, Zhengzhou 450001, China
2. State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou 450001, China

Abstract: By analyzing and extracting information such as packet size and timing leaked during Web application communication, side channel attack on encrypted network traffic is able to recognize users' identity and behavior and even restore the original data entered by users. A model of side channel attack on encrypted network traffic according to information theory was developed. Based on the unified model, the methods and results of representative attacks such as fingerprinting attacks, keystroke attacks and speech attacks were analyzed in detail. Furthermore, defense methods of hiding packet size and timing information were discussed. At last, possible research directions were prospected with the frontiers of technology development.

Keywords: web application, encrypted traffic, side channel attack, information gain

1 引言

随着互联网技术的发展, 应用程序逐渐摆脱

主机端的限制, 通过网络向用户提供服务。与传统的主机应用相比, 网络应用 (web application) 通常以浏览器作为载体, 可以在不需要安装或升

收稿日期: 2020-07-03; 修回日期: 2020-09-24

通信作者: 李玎, ld258166011@outlook.com

基金项目: 国家重点研发计划 (2019QY1302)

Foundation Item: The National Key R&D Program of China (2019QY1302)

论文引用格式: 李玎, 祝跃飞, 芦斌, 等. 网络加密流量侧信道攻击研究综述[J]. 网络与信息安全学报, 2021, 7(4): 114-130.

LI D, ZHU Y F, LU B, et al. Survey of side channel attack on encrypted network traffic[J]. Chinese Journal of Network and Information Security, 2021, 7(4): 114-130.

级的前提下,为用户提供多样化的便捷服务,如信息索引^[1]、协同编辑^[2]、音视频通话^[3-4]等。由于网络应用分为浏览器和服务端两个部分,用户的敏感数据不可避免地会在网络中传输,这给了攻击者窃取用户隐私信息的可乘之机。为了保护用户数据在网络上进行安全传输,开发者通常采用安全外壳(SSH, secure shell)协议、传输层安全(TLS, transport layer security)协议等网络安全协议加密应用层数据,从而保证用户数据的机密性。在此基础上,用户还可以使用虚拟专用网络、洋葱路由等匿名网络隧道加密技术,进一步保证数据的安全性和身份的隐秘性。由于加密后的网络应用数据隐藏了原始的明文特征,传统基于特征匹配的深度包检测等流量检测方法不再适用。因此,研究者开始将目光转向加密流量侧信道攻击方法。

侧信道攻击又称旁路攻击,其概念最早来源于密码学,是指绕过传统的暴力破解或密码分析方法,利用密码系统实现中泄露的硬件执行时间差异、功率消耗、电磁泄露等物理信息攻击密码系统^[5-6]。这些物理信息往往包含与密钥、明文相关的额外信息,攻击者利用这些额外信息,通过统计学分析方法,可以推断出部分原始数据或密钥。这些泄露的物理信息并不来源于密码系统算法逻辑中的缺陷,因此很难被设计者发现。网络加密流量侧信道攻击借鉴了传统侧信道攻击的思想,绕过攻击数据加密算法本身,利用网络加密流量可见的数据包长度、到达时间、传输方向和各层协议元数据,以及数据流的统计信息,试图推断出用户敏感信息。近年来,研究者提出了多种网络加密流量侧信道攻击方法,该领域成为学术界研究的热点。例如,指纹攻击^[7-8]能够判断用户正在浏览的敏感网站或使用的手机应用,击键攻击^[9-10]能够推测用户键盘输入的口令或搜索的查询,语音攻击^[11-12]能够识别用户使用网络电话交谈的内容或说话人的身份等。

网络加密流量容易遭受侧信道攻击的原因主要有3个。①网络应用一般分为客户端和服务端,因而应用程序的部分数据流和控制流必须通过网络进行传输。其中,数据流包含用户个性化的请求数据和服务器返回的相关应答数据,由此

可以推断用户身份或隐私信息;控制流一般包含网络应用特有的通信特征(如定期查询、更新同步等),由此可以识别具体的网络应用。②侧信道攻击者利用互联网服务提供商、自治系统、无线接入点等天然的中间人位置,可以在用户无感的情况下被动窃听流经的网络流量。③虽然加密后的数据包隐藏了应用层数据载荷,但加密流量不可能完全消除原始数据的所有特征,如网络资源大小在数据包长度和数量上的体现、用户行为所引发的网络请求及应答的时间信息和数据方向等。

当前基于网络加密流量的侧信道攻击目标类型繁多,攻击和评价方法各异,在学术界缺乏统一的模型框架对现有攻击方法进行系统总结。针对该问题,本文建立了网络加密流量侧信道攻击模型,通过理论分析和实例验证的方式,详细分析了几种有代表性的网络加密流量侧信道攻击方法的原理及效果,讨论了针对性的对抗防御方法,对未来技术发展方向进行了展望。本文的贡献主要有以下几点:

- 1) 基于信息论构建了网络加密流量侧信道攻击模型,为相关问题研究提供理论帮助;
- 2) 使用统一的模型框架对多种类型的加密流量侧信道攻击进行分析归纳,方便后续的研究应用;
- 3) 讨论网络加密流量侧信道攻击的限制和防御方法,提出未来的改进方向。

2 网络加密流量侧信道攻击模型

网络流量由网络数据流组成,一个网络数据流通常由包含相同五元组(源IP、源端口、目的IP、目的端口和协议号)的连续数据包组成。网络流量是网络终端之间数据通信的载体,即抽象的信息传输信道。对于未加密的网络流量,由于原始应用数据作为数据包载荷在网络中传递,攻击者可以直接攻击信息传输的主信道,即流量载荷数据。而对于网络加密流量,由于应用数据被加密隐藏在数据包载荷中,攻击者仅能获得网络流量侧信道中包含的信息,即流量中连续数据包的长度、到达时间和数据包各层协议头部的元数据,包括IP地址、端口号、TLS协议中的服务器名称等。利用基本的数据包特征,可以进一步得

到其他统计特征,如累计上行数据包数、平均数据包时间间隔、每秒下行字节数等,从而建立侧信道攻击模型。

网络加密流量侧信道攻击的目标是通过观测分析网络流量确定产生这些流量的用户行为。对于攻击者而言,用户行为可以描述为一个不确定性集合 $\mathcal{X}$,该集合可以是用户访问的网站、使用的应用程序、键盘输入或语音通话的内容等用户敏感数据。根据香农的信息熵理论^[13],如果集合的规模 $|\mathcal{X}|$ 可以缩小到原来的 1/2,就意味着攻击者获得了 1 bit 的信息量。因此,侧信道攻击的过程就是攻击目标集合规模减小的过程,对于攻击者也是用户行为不确定性降低的过程^[14]。

假设随机变量 X 表示用户行为, X 的取值 $x \in \mathcal{X}$ 为用户行为集合的元素,且 $\mathcal{X}$ 中的元素 x 具有先验概率 $p(x)$,则对于攻击者, X 的初始不确定性可以表示为 X 的(香农)信息熵,即

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \lg p(x) \quad (1)$$

一般情况下,侧信道攻击者并不具备用户行为的先验概率,即 $p(x_i) = \frac{1}{|\mathcal{X}|}$, $x_i \in \mathcal{X}$ 。此时,

X 的初始不确定性最大,其信息熵转化为最大熵,即 $H(X) = \lg |\mathcal{X}|$ 。以击键攻击为例,假设用户输入字符集是小写字母,则一次击键 X 的初始不确定性 $H(X) = \lg 26 \approx 4.7$ bit。

设随机变量 Y 表示用户行为产生的网络加密流量, Y 的取值 $y \in \mathcal{Y}$ 为加密流量中一组数据包的侧信道信息,包括数据包的长度、到达时间和元数据,即 $y = (p_1, \dots, p_n)$, $p_i = \langle \text{length}, \text{time}, \text{matedata} \rangle$ 。在网络应用和环境确定的情况下,用户行为 X 和其产生的网络加密流量 Y 是相关的。此时,对于观测到 Y 的攻击者, X 的剩余不确定性为 X 条件熵 $H(X|Y)$,其值为 $Y = y$ 的条件下 X 的信息熵 $H(X|Y = y)$ 关于先验概率 $p(y)$ 的加权平均数,即

$$\begin{aligned} H(X|Y) &= \sum_{y \in \mathcal{Y}} p(y) H(X|Y = y) = \\ &= -\sum_{y \in \mathcal{Y}} p(y) \sum_{x \in \mathcal{X}} p(x|y) \lg p(x|y) \end{aligned} \quad (2)$$

根据以上定义,网络加密流量泄露的用户行为信息量即为 X 减少的不确定性。对于攻击者,

该值是观测到 Y 获得的信息增益,可以表示为 X 和 Y 的互信息,即

$$I(X;Y) = H(X) - H(X|Y) \quad (3)$$

为了进一步刻画攻击者获得的信息增益对于原始不确定性的比重,即相对信息增益,通常利用归一化的方法得到标准化互信息(NMI, normalized mutual information),即

$$\text{NMI}(X;Y) = \frac{I(X;Y)}{H(X)} \quad (4)$$

在上述击键攻击的例子中,如果攻击者通过监听用户一次击键产生的网络加密流量 Y 可以将 $\mathcal{X}$ 的范围减小到包含 8 个字母的集合,则 X 剩余的不确定性 $H(X|Y) = \lg 8 = 3$ bit,攻击者获得的信息增益 $I(X;Y) = 4.7 - 3 = 1.7$ bit,相对信息增益 $\text{NMI}(X;Y) = 0.36$ 。

考虑极端情况下,当 X 和 Y 相互独立时,即用户行为产生的网络流量不包含用户行为的任何信息,有 $H(X|Y) = H(X)$,此时攻击者的信息增益 $I(X;Y) = 0$,相对信息增益 $\text{NMI}(X;Y) = 0$,这是网络安全传输最理想的目标。当用户行为 X 可以由产生的网络流量 Y 唯一确定时,有 $H(X|Y) = 0$,此时攻击者的信息增益 $I(X;Y) = H(X)$,相对信息增益 $\text{NMI}(X;Y) = 1$,即攻击者可以推断出原始的用户行为,这个理论的最大值等价于未加密的网络流量。

在衡量侧信道攻击效果时,除了使用攻击者获得的相对信息增益 $\text{NMI}(X;Y)$,通常还会借助识别准确率 Acc 来对比攻击方法的优劣。直观地,相对信息增益和识别准确率之间呈正相关。当相对信息增益接近于 1 时,攻击目标范围被压缩到很小的集合,此时攻击可以获得较高的识别准确率。但不同的是,识别准确率依赖于攻击中使用的分类特征,导致一些针对分类模型的对抗防御方法可以有效降低模型的识别准确率,但侧信道泄露的信息量却未必减少。事实上,识别准确率是对剩余不确定性集合 $\mathcal{X}$ 中元素一次性猜测的成功率,对于攻击者没有试错的机会,一旦没有猜中就记作一次识别失败。由此可以得出,当攻击者不具备先验知识,即 $p(x_i) = \frac{1}{|\mathcal{X}|}$, $x_i \in \mathcal{X}$ 时,识别准确率和相对信息增益之间的关系为

$$\text{Acc} = 2^{-(1-\text{NMI}(X,Y))\text{lb}|X|} \quad (5)$$

仍以击键攻击为例, $|X|=26$ 时识别准确率与相对信息增益之间的关系如图 1 所示。当攻击者获得的相对信息增益为 0.5 时, 识别的准确率仅为 0.196; 而当相对信息增益为 0.9 时, 识别的准确率可以达到 0.722。

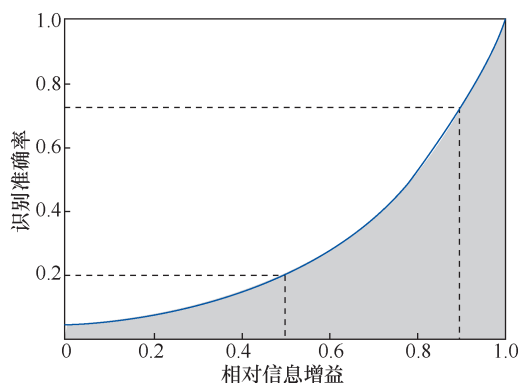


图1 识别准确率与相对信息增益的关系 ($|X|=26$)
Figure 1 Relationship between recognition accuracy and relative information gain ($|X|=26$)

在网络加密流量侧信道攻击方法中, 攻击者攻击目标的范围 X 一般是明确的, 如用户使用浏览器可能访问的网站集、使用网络电话可能交谈的语句集、使用键盘可能输入的字符集等, 该集合的大小 $|X|$ 等于攻击者采取暴力搜索时需要猜测的次数。侧信道攻击的关键在于提取加密流量中与用户行为相关的特征, 即确定用户行为和加密流量之间的概率关系 $p(x|y)$, 从而带来额外的信息增益, 有利于减少攻击者猜测的次数, 提高攻击识别准确率。例如, 指纹攻击通过提取加密流量的统计特征建立和用户访问网站的概率关系, 击键攻击建立数据包长度和时间模式、按键之间的概率关系, 语音攻击利用语音波形的编码特征建立数据包长度模式和语音音素之间的概率关系等。

3 网络加密流量侧信道攻击研究进展

第2节从概念上阐述了网络加密流量侧信道攻击的基本原理, 即攻击者从加密流量侧信道泄露的信息中获得的用户敏感数据的信息增益。根据不同的攻击目标集合 X , 侧信道攻击从网络加密流量集合 Y 中获取信息泄露的渠道和方式有所不同。本节针对当前取得良好效果的几种网络加密流量侧信道攻击技术进行详细分析。

3.1 指纹攻击

指纹攻击是一种通过被动监听网络加密流量, 分析推断用户访问的网站或应用的攻击方法。在计算机领域中, 传统的指纹算法是指数字认证中的哈希算法, 目的是将任意长度的数据用一个较小的定长比特串表示。在不考虑哈希碰撞的情况下, 该比特串可以唯一地标识原始数据。随着越来越多的网络应用使用加密技术保护自身网络数据, 攻击者无法直接从网络流量中获取原始数据的相关信息。因此, 研究者借鉴传统指纹的思想, 试图从网络加密流量侧信道中提取出能够唯一标识网络应用的指纹信息。

根据攻击目标类型, 指纹攻击可分为网站指纹攻击和手机应用指纹攻击, 二者的攻击场景如图 2 所示。指纹攻击的目标是从预先设定的对象集合中正确判别用户访问的对象, 其攻击方法通常包含 4 个步骤。① 攻击者设定攻击的目标范围, 即识别对象集合 X , 如敏感网站集合、监控手机应用集合等。② 攻击者从观测流量 $y \in Y$ 中手动或自动化地提取流量特征, 如前 n 个上行数据包的长度、加载资源的数据包持续时间、匿名通道中信元的序列和数量等, 建立能够区分不同识别对象的指纹信息。③ 攻击者构造或收集带标记的流量样本数据集, 包含 X 中所有元素产生的流量样本。④ 攻击者使用机器学习算法训练一个分类模型, 以识别用户访问的对象 $x \in X$ 。

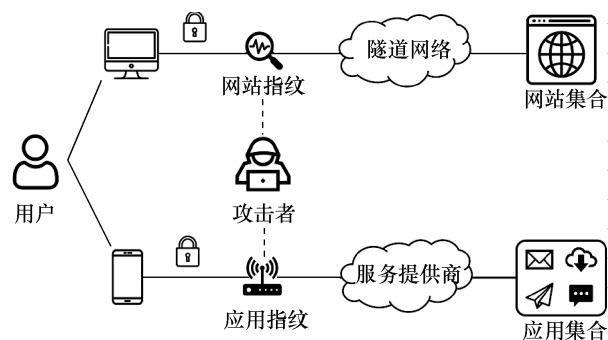


图2 网站指纹攻击和手机应用指纹攻击场景
Figure 2 Scenarios of website fingerprinting attack and mobile App fingerprinting attack

3.1.1 网站指纹攻击

随着计算能力和网络性能的提升, 目前绝大多数网站支持通过安全套接字 (SSL, secure sockets layer) 协议、TLS 等加密方式保护用户隐

私信息。根据统计显示,当前互联网上有超过 80% 的网站通过 HTTPS 加密^[15]。传统的基于内容模式匹配的方法已经无法直接从网络流量中获取用户信息,因此大量的研究者开始将目光转向侧信道攻击,尝试从网络加密流量中提取网站特有的指纹信息。

网站指纹攻击的原理主要依赖于用户访问网站时浏览器的行为。当用户在浏览器中访问一个网站的 Web 页面时,浏览器首先向服务器请求该网站主页的 HTML 页面文件,然后检查页面文件中引用的相应资源文件。当本地没有缓存的资源文件时,浏览器按照顺序向服务器请求对应的资源文件。根据 HTTP Archive 网站统计显示,当前访问一个网站平均需要加载超过 70 个资源文件,包括字体、图像、JavaScript 脚本、CSS 样式等^[16]。不同网站的资源文件在类型、数量、大小和加载顺序上一般存在明显差异,并且 SSL/TLS 加密不会隐藏数据包长度和时间信息,使攻击者可以通过分析网络加密流量中隐含的网站资源序列化特征,识别用户浏览的网站。

为了检测网站指纹攻击的鲁棒性,研究者通常将攻击场景分为封闭世界和开放世界。在封闭世界中,用户访问的网站集合 $\mathcal{X}$ 的规模通常较小,攻击者可以构建所有网站的指纹。因此,封闭世界检验的是模型基本的分类识别能力,一般通过模型识别的准确率 Acc 来衡量攻击的效果。而在模拟真实攻击场景的开放世界中,用户访问的网站集合 $\mathcal{X}$ 的规模通常较大,而攻击范围是 $\mathcal{X}$ 中一个较小的子集 $\mathcal{S}$,即攻击者无法构建攻击目标 $\mathcal{S}$ 之外未知网站的指纹。因此,开放世界可以检验模型的鲁棒性,一般通过模型识别的真阳率 TPR 和假阳率 FPR 来衡量攻击的效果。

(1) 基于简单指纹匹配的攻击方法

早期网站指纹攻击主要针对网站自身 SSL/TLS 加密的流量,由于传输层安全协议只加密应用层数据,传输数据的流量特征得以保留。因此,攻击者通过监听网络流量可以直接统计出网站请求资源的大小、数量等网站指纹。Cheng 等^[17]最早发现不同网站请求的 HTML 文件大小和其他资源大小存在特异性,构建了包含 HTML 和其他所有资源大小的二维网站指纹,并通过指

纹库匹配的方式识别未知指纹。Hintz 等^[18]通过对 SSL 代理软件 SafeWeb 的研究,提出了一种比 Cheng 等更加精细化的网站指纹攻击方法,将包括 HTML 在内的所有请求文件的大小序列作为网站指纹信息。Sun 等^[19]也在 Cheng 研究的基础上,在开放世界中测试了指纹攻击方法的鲁棒性。文献[19]首先基于网页请求的资源数量和各个资源的大小构建网页指纹,然后利用 Jaccard 系数衡量两个网页指纹的相似度,并根据一个预设的相似度阈值对未知网页进行预测。Fegghi 等^[20]注意到之前基于请求资源大小特征指纹攻击方法虽然识别准确率高,但通过简单的数据包填充防御方法就可以使攻击失效。为此,文献[20]提出了一种仅利用上行数据包时间信息的网站指纹攻击方法,使用动态时间规整(DTW, dynamic time warping)算法计算两个数据包时间序列的最佳规整路径,并通过定义的 F 距离衡量不同网站的相似度。

(2) 基于数据流统计特征的攻击方法

随着隧道加密技术的发展,很多用户选择使用 SSH、VPN 等单跳网络代理来隐藏自身网络活动。由于隧道隐藏了网络连接的状态及目的 IP 地址等协议头部字段,侧信道攻击者无法直接区分请求不同资源的 TCP 连接。同时,普及的 HTTP/1.1 支持流水线技术,使网站请求资源时流量的上下行边界特征变得模糊。2005 年, Bissias 等^[21]首次提出针对 SSH 隧道加密保护下的网站指纹攻击方法,不再使用网站单个资源的大小特征,而是基于服务器应答数据包的长度序列和时间间隔序列的分布特征建立网站指纹。预测未知网站时,简单地使用信号处理中的互相关函数量化两个序列的相关度,从而对指纹库中的网站进行排序。Liberatore 等^[22]随后提出一种改进的 SSH 隧道网站指纹攻击方法,将访问网站时具有相同方向和长度的数据包统计值作为指纹特征向量,并分别使用 Jaccard 系数相似性度量和朴素贝叶斯(NB, naive Bayes)监督分类模型进行预测,前者只比较数据包出现的种类,后者以每种数据包的数量作为输入。Herrmann 等^[23]在其研究基础上,对特征提取和模型构建进行了改进,提出了一种基于多项式朴素贝叶斯(MNB, multinomial

naive Bayes)模型的攻击方法,并使用 TF-IDF (term frequency-inverse document frequency)变换和余弦归一化(cosine normalization)强化特征向量的表示能力。顾晓丹等^[24]利用访问网站时上下行流量的特征差异,分别构建 SSH 匿名流量的上下行指纹,然后通过 NB 算法对网站指纹进行相似度匹配。

为了进一步满足人们对匿名访问的需求,以 Tor 匿名网络为代表的多跳隧道网络开始兴起。除了保障用户链路的匿名性, Tor 还采用了多种抵抗传统指纹攻击的方法,如采用固定长度为 512 byte 的信元作为数据传输单元,隐藏原始数据的长度信息;使用令牌桶机制限制读取数据和发送的速率,混淆数据请求的时序特征。为了实现针对 Tor 网络的网站指纹攻击,研究者相继提出了多种基于精细化关联特征的识别方法。2009 年, Hermann 等^[23]首次尝试基于数据包长度特征对 Tor 网络进行网站指纹攻击,但收效甚微。Panchenko 等^[25]在其基础上设计了一组精细化的流量统计特征,并使用支持向量机(SVM, support vector machine)分类器代替此前的 NB 分类模型。Cai 等^[26]注意到网络数据包重传、丢包和乱序的特点与文本中插入、删除和换位的操作非常类似,因此将带方向数据包长度序列转化为字符串,使用能够容忍字符串间细微差异的 Damerau-Levenshtein 编辑距离作为比较数据流之间相似性的核函数,并据此构建了 DLSVM 分类模型。Wang 等^[27]对 Cai 等采用的 DL 编辑距离进行了改进,提出了一种新的基于最佳字符串对齐距离(OSAD, optimal string alignment distance)的 SVM 核函数。随后, Wang 和 Cai 等^[28]合作提出了一种适应开放世界场景的网站指纹攻击方法,使用带有权重调节的改进 kNN 分类模型有效训练了超过 3 000 个混合类型的流量特征。Panchenko 等^[29]利用可视化方法设计了基于累加长度(CUMUL)的指纹特征向量,并将前 100 个值与 4 个传统的流量统计值合并,共同作为 SVM 分类模型的输入。为了改进 Wang 等^[28]特征集合过于庞大导致的训练困难, Hayes 等^[30]提出了一种基于随机森林(DF, random forest)的特征选择方法,将决策树中最重要的叶子节点作为新的指纹特征训练一个

kNN 分类模型。

(3) 基于深度学习特征自动化提取的攻击方法

随着深度学习在计算机视觉和自然语言处理等领域不断取得成功,研究者尝试将深度学习的思想应用在网站指纹攻击方法中。通过训练神经网络模型从原始数据中自动提取隐藏的关联特征,可以代替传统方法中人工设计微调特征的繁复过程。2016 年, Abe 等^[31]最早尝试使用栈式去噪自编码器(SDAE, stacked denoising autoencoders)模型自动提取 Tor 网络流量的指纹特征。在构造神经网络输入时,直接使用前 5 000 个信元的方向组成特征向量,上下行分别记作 1 和-1,长度不足的填充 0。这种模型输入构造方法目前在学术界被广泛认可,并在后续的研究中一直被沿用。为了弥补 Abe 等训练模型时数据量较小的问题, Rimmer 等^[32]构建了大规模的流量数据集,并细致比较了 SDAE、卷积神经网络(CNN, convolutional neural network)和长短期记忆(LSTM, long short-term memory)网络 3 种深度学习模型提取网站指纹特征的效果。文献[32]还通过复现传统的 kNN^[28]、SVM^[29]和 RF^[30]方法,证明了深度学习方法的检测效率更高,所提取特征抵抗概念漂移的能力更强。为了进一步提高深度学习模型识别的准确率, Sirinam 等^[33]基于 Rimmer 等使用的 3 层 CNN 模型,训练了一个包含 8 个卷积层、2 个全连接层的 CNN 模型,同时进行了防止过拟合和精细化参数调优。随后, Bhat 等^[34]在 Sirinam 研究的基础上,为了适应小规模训练数据,使用数据包间隔时间代替原有只包含方向的特征向量,然后以 ResNet-18 作为基础模型,使用具有长时间理解能力的扩展因果卷积(dilated causal convolutions)代替原有卷积操作。Rahman 等^[35]提出一种利用数据包时间特征强化攻击效果的方法,对方向特征向量中的每个值乘以数据包到达时间得到新的神经网络输入向量。马陈城等^[36]为了更好地提取 Tor 流量的上层特征,提出将 Tor 网络突发流量(即方向相同的信元序列)作为网站指纹深度分析的特征。

近年来,学术界对网站指纹攻击的研究数量众多,一方面是基于加强网络监管、打击网络犯罪的现实需求,另一方面得益于相关研究成果的

公开性。研究者通常会公开模型源码和实验数据集,在保证实验结果可复现的同时,方便后续研究者进一步研究。表1从现有网站指纹攻击方法中选取了不同技术发展阶段中有代表性的方法,并从多个角度进行了分析对比。

3.1.2 手机应用指纹攻击

针对手机应用指纹的侧信道攻击方法随着智能手机的技术发展而出现。早期的手机应用指纹分析主要集中在明文的HTTP,通过提取流量中HTTP头部中的请求URL、Host等特征字段构建应用指纹,从而识别用户使用的应用^[37]。这类方法属于深度包检测技术,优点是简单高效、检测率高,但无法识别HTTPS加密流量。为了满足人们对隐私保护的需求,主流的手机操作系统强制采用加密方式建立连接,如iOS 9开始为应用默认开启应用传输安全(ATs, App transport secu-

rity)功能,Android 9开始默认禁止应用发起未加密的连接。根据统计资料显示,截至2019年,全球超过80%的安卓应用默认使用加密的网络连接通信^[38]。

对于攻击者而言,手机应用的指纹攻击通常比网站指纹更加复杂,这主要基于两个原因。①网站指纹攻击很大程度上利用了不同网站加载资源数量、大小的差异性,而手机应用请求的资源数量一般较少,且大部分在安装时预置在本地。②手机应用通信的加密流量中不仅包含应用资源的请求数据,而用包含很多公共服务数据(如位置同步、广告信息等),使手机应用可能同时访问多种类型的服务,或多个应用同时访问同一个服务器的资源,这进一步减弱了应用之间的差异性。考虑到手机应用产生流量的复杂性,研究者们在提取流量指纹特征时,通常会先将流量分解为

表1 网站指纹攻击方法比较
Table 1 Comparison of website fingerprinting attack methods

作者	隧道类型	模型/方法	流量特征			攻击场景	网站数量	结果 ^[*]
			长度	时间	方向			
Cheng ^[17]	SSL	match	✓		✓	封闭	489	96.7%
Sun ^[19]	SSL	Jaccard	✓		✓	开放	2 191 / 100 000	75% / 1.5%
Fegghi ^[20]	TLS	DTW+kNN		✓	✓	封闭	100	95%
Bissias ^[21]	SSH	correlation	✓	✓	✓	封闭	100	23%
Liberatore ^[22]	SSH	Jaccard, NB	✓		✓	封闭	1 000	75%
Hermann ^[23]	SSH	MNB	✓		✓	封闭	775	96.7%
Cai ^[26]	Tor	DLSVM	✓		✓	封闭	100	87.3%
Wang ^[28]	Tor	kNN	✓		✓	封闭	100	91%
						开放	100 / 5 000	85% / 0.6%
Panchenko ^[29]	Tor	SVM	✓		✓	封闭	100	91.4%
						开放	100 / 9 000	96.6 / 9.6%
Hayes ^[30]	Tor	RF+kNN	✓	✓	✓	开放	30/100 000	85% / 0.02%
Abe ^[31]	Tor	SDAE			✓	封闭	100	88%
						开放	100 / 7 000	86% / 2%
Rimmer ^[32]	Tor	SDAE, CNN, LSTM			✓	封闭	900	94.3%
						开放	200 / 400 000	71.3% / 3.4%
Sirinam ^[33]	Tor	CNN			✓	封闭	95	98.3%
						开放	95 / 20 000	95.7 / 0.7%
Bhat ^[34]	Tor	CNN		✓	✓	封闭	100	97.8%
						开放	100 / 10 000	89.2% / 1.1%

注: “✓”表示攻击利用了该流量特征。[*]对于开放世界场景,结果为准确率;对于封闭世界场景,结果为真阳率和假阳率。

较小的基本单元,一般使用突发(burst)和数据流(flow)的概念进行描述。其中,burst描述了短时间内出现的大量数据包,通常随着应用启动、后台同步或用户操作等过程而产生。使用burst描述突发流量的难点在于确定burst阈值(即时间范围),该值作为模型的超参数,一般根据经验选取^[39-40]。flow描述了一个burst中具有相同五元组的数据包序列,并且可以根据选取的特征进一步细分为上行序列、下行序列和双向序列。flow区别于传统TCP/IP数据流的概念,其起止时间取决于burst的时间范围。

(1) 基于burst统计特征的攻击方法

早期的研究者注意到不同手机应用产生的加密流量在burst的统计特征上存在差异,因此,通过提取burst统计特征可以建立起识别不同手机应用的指纹。Stöber等^[41]最早提出一种基于kNN和支持向量机(SVM, support vector machines)分类模型的识别方法。由于手机应用在无用户交互的情况下会在后台定期产生更新或同步请求的加密流量,文献[41]设计了针对后台burst流量的23个数据流统计特征作为应用指纹,并使用信息增益量化了每个特征的重要程度,发现数据包长度的绝对中位差(MAD, median absolute deviation)可以获得最高16.6%的相对信息增益。Wang等^[42]通过研究应用中用户行为所产生的加密流量,提出了基于数据包长度和时间模式的手机应用指纹攻击方法。与Stöber等的方法不同,文献[42]去除了非用户行为产生的后台流量,将每个用户行为产生的流量描述为burst,然后从双向数据包的长度和时间信息中提取了20个流量统计特征,用来构建RF分类模型。Alan等^[43]通过对大量安卓应用流量数据的研究,发现仅通过应用启动时的前64个数据包的长度模式就能够较好区分不同的应用。文献[43]借鉴了Herrmann等^[23]的网站指纹攻击的方法,对特征向量进行TF-IDF变换和归一化处理,然后通过MNB算法构建一个分类模型。

(2) 基于flow相似性比对的攻击方法

基于burst统计特征的指纹攻击方法虽然能够识别应用类型,但难以进一步识别用户的特定行为。为了提取更加精细化的手机应用指纹,研

究者相继提出了多种基于flow相似性比对的攻击方法。Conti等^[39]提出一种识别用户特定行为的指纹攻击方法。数据预处理阶段利用目的IP地址过滤属于目标应用的流量,并截取前4.5 s的burst限制flow的时间范围。模型以数据包长度的DTW最佳规整路径作为聚类算法的距离测度,使用无监督的层次聚类(HC, hierarchical clustering)算法将flow聚合成不同的簇。模型训练阶段以不同簇的大小作为分类特征,衡量用户行为与各个簇的相关度,并为每个应用训练一个RF模型。Saltaformaggio等^[44]改进了Conti等通过目的地址过滤应用流量的方法。文献[44]首先将flow划分为5 ms的用户行为测量单元,针对测量单元设计了26个统计特征,然后利用k-means算法对测量单元进行无监督聚类,将聚类结果作为用户行为的特征向量训练一个SVM分类模型。Taylor等^[40]同样在Conti研究的基础上,提出了一种扩展性更好的手机应用指纹攻击方法。文献[40]将burst的时间范围从4.5 s缩短到1 s以适应网络性能的提升,并在训练阶段使用了6种不同要素组合的分类模型,即分类方式为多分类或二分类,分类算法为SVM或RF,分类特征为flow的数据包长度序列或40个数据流统计特征。模型预测阶段,为了防止坏结果的产生,采用概率阈值的方式得到预测的结果。6个月后,Taylor等在不同安卓设备上再次进行了实验^[45],验证了模型的泛化能力。朱迪等^[46]在Taylor研究的基础上,提出了一种使用密度聚类算法过滤干扰样本,然后使用RF分类模型识别手机应用加密流量的攻击方法。

表2从多个角度对比了当前手机应用指纹攻击中有代表性的方法。这些方法普遍利用加密流量的数据流统计特征,通过机器学习算法训练分类模型,可以准确区分不同的手机应用,甚至可以识别用户的特定行为。因此,手机应用指纹攻击方法可以作为网络监管的有效手段。

3.2 击键攻击

击键攻击又称击键记录侧信道攻击,是一种利用侧信道泄露的信息跟踪键盘或鼠标活动,试图还原用户输入或选择内容的攻击方法。随后研究人员发现,通过监听该装置每次击键发出的电磁波峰值,可以在很远的地方恢复出高达75%的

表 2 手机应用指纹攻击方法比较
Table 2 Comparison of mobile app fingerprinting attack methods

作者	攻击目标		模型/方法	流量特征		准确率
	类型	数量		载体	数量	
Stöber ^[41]	应用	14	kNN, SVM	burst	23	90%
Wang ^[42]	应用	13	RF	burst	20	93%
Alan ^[43]	应用	1 595	MNB	packet	64	88%
Conti ^[39]	特定行为	50	DTW+HC+RF	flow	小于 600	95%
Saltaformaggio ^[44]	特定行为	35	k-means+SVM	flow	26	78%
Taylor ^[40]	应用	110	RF, SVM	flow	40	99%
Taylor ^[45]	应用	65	RF	flow	54	96%

明文。此后，研究者开始从多个侧信道展开对击键攻击的研究，如用户手指或手部动作、击键引起的物理震动、电磁或声音信号、CPU 负载峰值等。随着网络技术的发展，应用程序逐渐摆脱主机端限制，开始基于网络传输为用户提供高交互、低时延的服务。同时，计算能力的提升使网络应用普遍采用 SSH、TLS 等加密方式保护用户隐私信息。因此，针对网络加密流量的击键攻击方法通过分析数据包长度和时间等侧信道信息，试图推测用户原始的击键内容。

与指纹攻击相比，击键攻击一般更加困难，因为攻击目标集合 $\mathcal{X}$ 的规模通常非常大。以键盘输入为例，假设随机变量 X 表示一个由字母和空格组成的长度为 n 的字符串，则理论上攻击目标集合的规模 $|\mathcal{X}| = 27^n$ 。如果攻击者没有用户输入字典等先验知识，则对于攻击者的初始不确定性取得最大熵，即 $H(X) = \lg 27^n \approx 4.75n \text{ bit}$ 。在现实中，用户输入的内容一般符合某些自然语言特征，如名字、词汇等，使 $\mathcal{X}$ 中一些字符串不可能出现，即 $p(x) \rightarrow 0$ 。此时，根据式(1)计算的信息熵会大幅度减小，对于攻击者实际的初始不确定性可以下降到可猜测的范围。

(1) 基于用户击键时间模式的攻击方法

早期研究者提出的击键攻击方法主要基于网络流量泄露的用户击键模式 (typing pattern)，即击键时的时间间隔特征。击键模式包含人的生物学行为特征，因此还被广泛用作身份验证的生物测定方法^[47-50]。Dhakal 等^[51]通过收集分析超过 16

万人的键盘记录，发现用户的击键时间间隔 (IKI, inter-key interval) 具有明显生物特征。例如，快速打字者的平均 IKI 为 122 ms，标准差为 12 ms，而慢速打字者的平均 IKI 为 481 ms，标准差为 123 ms。此外，对于快速打字者，不同手部组合下的双字符对 (bigram) IKI 有稳定的分布差异。例如，双手交替字母对的平均 IKI 为 109 ms，同一手指字母对的平均 IKI 为 145 ms。因此，在稳定的击键模式下，攻击者可以利用侧信道泄露的 IKI 信息获得关于原始输入的信息增益。Song 等^[9]提出一种利用 SSH 连接数据包泄露的击键时间信息推测用户口令的攻击方法。由于 SSH 连接属于实时交互的网络应用，用户的每次击键会立即产生一次数据交换。文献[9]首先利用 SSH 登录时特殊的数据包长度确定用户输入口令的起始位置 (即口令长度)，然后通过分析数据包的时间间隔信息推测可能的双字符对。为了研究用户的击键模式，文献[9]收集了字母和数字双字符对的 IKI 值，并通过高斯分布参数估计发现攻击者可以从每次击键泄露的时间信息中获得约 1 bit 的信息增益，大大缩减了口令暴力搜索的范围。文献[9]进一步设计了基于隐马尔可夫模型 (HMM, hidden Markov model) 和 n -Viterbi 算法的分类预测模型，对于给定的 IKI 序列，模型通过预测输出 n 个可能的字符序列。

(2) 基于 AJAX 交互数据长度特征的攻击方法

虽然 SSH 加密流量泄露的击键时间信息可以有效减少暴力搜索猜测的次数，但仅靠少量的时间信息还不足以推断用户输入的内容，攻击者

试图从网络应用加密流量中获取更多关于用户输入的信息。为此,研究者开始关注基于 AJAX (asynchronous Javascript and XML) 技术的实时交互网络应用,如网络会诊、在线地图、增量搜索等。使用 AJAX 技术的网络应用支持通过异步更新方式对用户击键(如鼠标点击、键盘输入等)进行实时应答,而不需要重新加载新的页面。以增量搜索为例,典型的基于 AJAX 技术的增量搜索应用如图 3 所示。当用户在增量搜索框中输入字符时,后台 JavaScript 程序会发送包含查询字符串的 HTTP 请求,接收服务器返回的查询建议列表,并将列表显示在增量搜索下拉框中。在用户搜索的过程中,不同请求和应答的大小特征被保留在数据包长度模式中,攻击者可以通过分析加密流量推断用户原始输入,收集用户的政治倾向、医疗状况、身份标志等敏感信息。

采用 AJAX 技术的网络应用服务器会对用户击键做出快速响应,网络侧信道攻击者通过分析击键序列产生的网络流量特征可以推断用户击键的内容。Chen 等^[10]最早提出一种基于模糊集缩减(ASR, ambiguity set reduction)的击键攻击方法。文献[10]首先对交互式的网络应用进行状态化描述,即某一时刻应用的浏览器端和服务器端处在某一特定的状态,并将用户击键集合定义为模糊集,包含鼠标点击的表项或键盘输入的字符。当

用户击键时,应用会发生状态转移,同时产生一个带方向的数据包长度序列,该序列被称为网络流(web flow)向量。攻击者通过匹配事先穷举的网络流向量可以缩小应用状态的搜索范围,从而推断用户可能的输入。Sharma 等^[52]在 Chen 研究的基础上,针对谷歌搜索引擎提出了一种类似的击键攻击方法,通过分析 HTTP 请求和应答中的搜索字段,验证了利用数据包特殊长度推断搜索字段的可行性。然而,谷歌搜索在 2012 年率先改变了推荐策略。当用户输入搜索查询时,服务器会根据当前时事热点、地理位置和用户的搜索历史生成查询建议列表,因此相同的查询可能返回不同的结果。为此,Schaub 等^[53]提出了一种基于随机化算法的数据包长度模式匹配算法,并且只将下行数据包长度作为特征向量。对于相同的搜索查询,文献[53]假设服务器返回的数据包长度符合高斯分布,并通过容忍参数 ϵ 确定长度区间 $[x-\epsilon, x+\epsilon]$ 。攻击者通过监听数据包长度序列,计算各个长度属于某个已知前缀字符串的概率乘积,作为该查询字符串的可能性预测值。

(3) 基于多维度流量特征的攻击方法

随着网络应用普遍采用随机化策略,传统基于通信数据长度特征的攻击方法不再被使用,研究者开始利用多维度的流量特征实现用户击键攻击。Oh 等^[54]提出了一种针对 Tor 匿名网络保护下

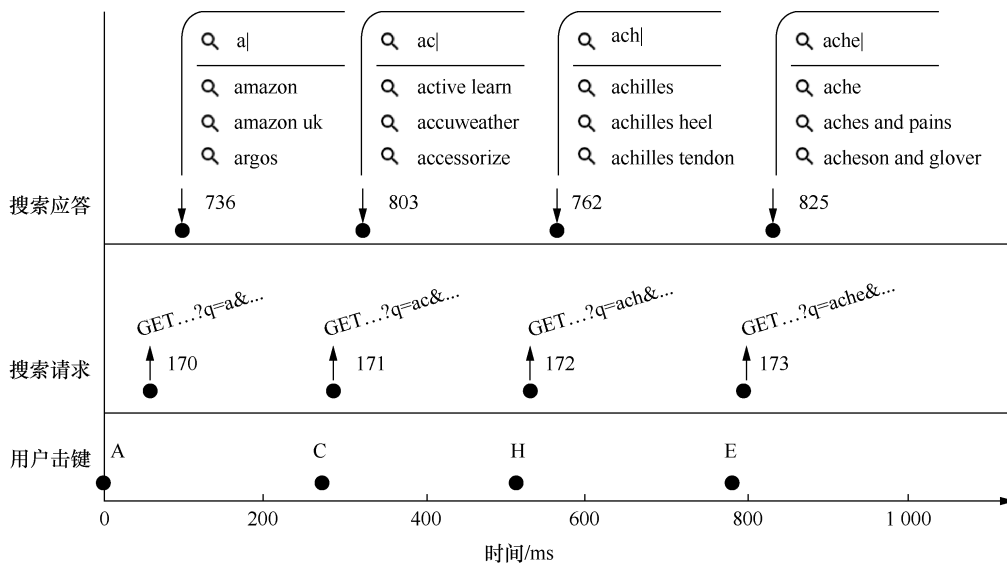


图3 基于 AJAX 技术的增量搜索应用

Figure 3 Incremental search application based on AJAX technologies

的搜索查询攻击方法。服务器返回的查询建议列表和结果使攻击者通过下行流量特征能够区分不同的查询关键字,因此文献[54]借鉴指纹攻击的思想,构造了以下行流量为主的 247 个统计和序列特征作为关键字指纹,然后利用分类模型实现对用户输入关键字的识别。文献[55]使用深度学习方法重复了上述实验,并选择下行数据包长度序列作为神经网络输入特征。Monaco 等^[56]考虑到 Oh 等的方法只能识别特殊的关键字,并且增量搜索服务器返回的结果会随时间和位置等因素发生变化,因此提出一种基于上行 AJAX 请求数据包特征的击键攻击方法。该方法首先利用接收数据包长度增量的确定有穷自动机(DFA, deterministic finite automaton)检测用户击键数据包,然后利用搜索请求 URL 中百分号编码空格字符的特殊长度划分英文单词,同时通过匹配 HAPCK 静态霍夫曼编码后单词长度的增长模式裁剪字典规模,最后使用识别数据包时间间隔序列的 RNN 模型预测可能的单词,并借助语言模型对预测单词进行集束搜索(beam search),得到候选的查询字符串。文献[57]分析了 5 个知名网站的增量搜索 AJAX 模型,发现采用 keydown 事件触发请求的模型可以在加密流量中保留原始的击键信息,而采用 keyup 或计时器超时事件会合并紧邻的搜索请求,导致加密流量损失被合并的击键信息。

表 3 对当前加密流量击键攻击方法进行了横向对比。由于击键攻击的目标范围非常大,现有攻击方法一般针对目标集合的子集,并且对攻击场景进行了适当的假设。虽然击键攻击结果无法达到指纹攻击的准确率,但攻击者通过网络加密

流量泄露的侧信道信息,可以有效减小猜测用户击键输入的范围。

3.3 语音攻击

网络技术的发展使人们逐渐从传统电话转向网络通话服务,即 VoIP (voice over Internet protocol)。语音攻击主要针对基于网络的在线语音通话技术,是一种通过分析用户语音通信的网络加密流量,试图还原语音通话内容或鉴别用户身份的攻击方法。由于网络传输的流量具有开放性,网络通话技术通常在传输语音数据前,通过加密技术保证数据的机密性。语音流量加密虽然有助于保护用户隐私,但同时可能被不法分子用来隐匿犯罪活动。因此,针对网络加密流量的语音攻击技术对于网络监管具有重要意义。

语音攻击不同于传统的自动语音识别^[58-59],原始语音经过编码压缩和数据加密处理后,攻击者无法直接从流量数据中还原语音信息。因此,语音攻击试图从网络流量的数据包长度模式和时间间隔等侧信道信息中建立与编码前语音的关系,其攻击方法通常包含 4 个步骤。① 攻击者设定攻击的目标范围,即识别对象集合 $\mathcal{X}$,如语句集合、语言集合、用户身份集合等。② 攻击者从观测流量 $y \in \mathcal{Y}$ 中提取与识别对象相关的流量特征,如语音帧的时间间隔、数据包长度模式、滤波器参数阈值等,建立能够刻画识别对象的特征判别模型。③ 攻击者构造或收集带标记的流量样本数据集,包含 $\mathcal{X}$ 中所有元素产生的流量样本。④ 攻击者使用机器学习算法训练一个分类模型,并利用训练好的模型识别攻击目标 $x \in \mathcal{X}$ 。

在传统 VoIP 服务的通信协议中,语音数据和控制信息通常是在单独的通道中传输。其中,控

表 3 针对加密流量的击键攻击方法比较
Table 3 Comparison of keylogging attack methods against encrypted traffic

作者	攻击目标		模型/方法	流量特征			信息增益	准确率
	类型	数量		方向	长度	时间		
Song ^[9]	输入口令	36 ⁸	HMM+Viterbi	双向	✓	✓	5.7 bit	
Chen ^[10]	点击表项	2 670	ASR	双向	✓		10 bit	
Schaub ^[53]	查询关键字	10 / 6 812	stochastic algo	下行	✓			34%
Oh ^[54]	查询关键字	100 / 40 000	SVM, kNN, RF	双向	✓			48.2%
Monaco ^[56]	查询字符串	4 000	DFA+IC+RNN	上行	✓	✓		15.8%

注:“✓”表示攻击利用了该流量特征。

制信息一般通过应用层协议传输,如标准的公开协议会话发起协议(SIP, session initiation protocol)、可扩展通信和表示协议(XMPP, extensible messaging and presence protocol),以及特殊应用的控制协议如Skype等。而语音数据是通过基于UDP的应用层协议实时传输协议(RTP, real-time transport protocol)进行传输。语音攻击主要针对数据通道,利用数据通道加密流量泄露的侧信道信息对原始语音内容进行攻击。

在VoIP语音数据被加密封装之前,都会通过专门针对人声的语音编码器进行编码压缩处理。不同编码方式之间虽然有细微差异,但一般基于基本的码激励线性预测(CELP, code-excited linear prediction)编码模型,如图4所示。CELP以8 kHz至32 kHz的频率对语音采样,以20 ms的时间间隔为单位将样本划分为帧。编码器首先计算每个帧的线性预测系数(LPC, linear prediction coefficients),并将帧分别与固定码本和自适应码本中的编码进行匹配,找到最能代表原始语音内容的编码;然后将编码索引和线性预测系数及增益一起作为输出,组成RTP的载荷数据。由于语音通信中一方大部分时间是静默状态,CELP通常采用可变码率(VBR, variable bit rate)来自适应地选择对应比特率的码本,即讲话时选择高比特率的码本以表达复杂的帧。而在高比特率的码本中,编码索引的位数更多。因此,VBR模式下VoIP加密数据包的长度能够体现通话信息的特征,语音攻击正是利用了该侧信道泄露的信息。

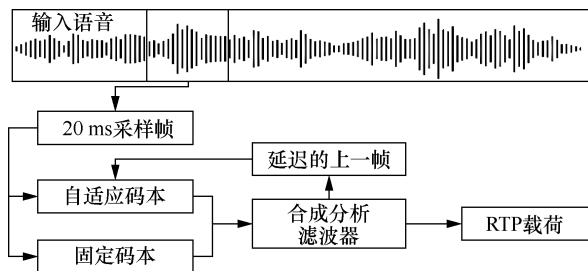


图4 CELP 语音编码器
Figure 4 CELP speech encoder

(1) 基于编码音素长度模式比对的攻击方法

传统的语音识别技术首先将语音帧标记为状态,然后将状态组合成发音的基本单元——音素。根据CMU发音词典^[60],所有的英文单词都可以

由39个基本音素组合而成。研究者借鉴语音识别的思想,以音素为单位划分语音加密流量,然后建立不同音素与数据包长度模式的关系。Wright等^[61]提出一种基于 n -gram模型的VoIP讲话者语言识别方法,通过衡量同一种语言的讲话者之间和不同语言的讲话者之间的概率分布差异,定义了语言之间概率分布的可区分性。文献[11]以音素为基本单元得到对应的VoIP数据包长度模式,通过训练Profile HMM模型来识别加密流量中符合特定模式的语句。White等^[62]在Wright研究的基础上,实现了对VoIP加密流量中任意音素的建模和识别。文献[62]首先利用最大熵(max-entropy)模型识别音素边界数据包,将加密数据包按照音素分组,然后使用最大熵和Profile HMM模型识别数据包分组具体属于哪个音素,并利用音素约束模型将音素序列划分成英文单词,最后通过比较语音编辑距离识别音素分组,并在trigram语言模型的纠正下得到原始语音内容。Dupasquier等^[63]研究了针对Skype加密流量的语音攻击方法,发现Skype语音加密流量的独立音素可以良好地被区分,进而通过音素的组合可以识别出完整的语句。为了屏蔽噪声和不同讲话者带来的影响,文献[63]使用卡尔曼滤波器来提取语音信号的核,并利用DTW算法较好地解决了两个存在时间差异的相同语句的相似性度量问题。

(2) 基于说话人语音模式特征的攻击方法

除了识别还原语音内容,用户身份鉴别即说话人识别是语音攻击的另一个重要领域。Khan等^[64]提出一种利用VoIP语音加密流量识别说话人身份的方法,为每个说话人建立一个基于数据包长度的嵌套二分集合(END, ensemble of nested dichotomie)模型,然后利用训练好的模型对数据包长度模式进行识别。在另一项识别用户身份的工作中,Backes等^[12]发现VoIP语音加密流量中的停顿模式能够反映出讲话者的身份特征,利用语音端点检测(VAD, voice activity detection)滤波器将加密流量按照数据包长度划分为讲话状态和静默状态,然后利用“静默-讲话-静默”三元组的相对频率构建说话人的特征模型,通过L1最近邻分类器对说话人进行检测。

表 4 分析对比了当前针对网络加密流量的语音攻击方法。现有方法主要针对开源的语音编码系统 Speex, 实验结果中对语言、用户身份和特定语句的准确率可以达到 50% 以上。虽然对任意语句内容的识别准确率相对较低, 但任何侧信道泄露都会给攻击者带来一定的信息增益, 有助于识别用户的敏感信息。

4 网络加密流量侧信道攻击防御方法

由于网络加密流量侧信道攻击能够给用户行为和隐私信息带来巨大威胁, 相应的防御方法应运而生。攻击者通过监听网络加密流量实现侧信道攻击, 本质上是利用了加密流量泄露的数据包长度、时间、方向等基本特征。这些特征经过分析组合可以形成高维的序列特征和统计特征, 作为机器学习等模型的输入。因此, 网络加密流量侧信道攻击防御方法的基本思想是隐藏数据包的长度和时间信息。

为了隐藏数据包长度信息, 研究者提出了多种基于数据长度混淆的防御方法。Sun 等^[19]针对网站指纹攻击提出了 3 种基本的数据包填充方式, 即线性填充 (固定长度的整数倍)、指数填充 (2 的幂次) 和随机填充, 并且提出数据包变形的思想, 即使用 HTTP 流水线技术和比特范围请求随机化数据包的长度和数量。随后, Liberatore 等^[22]在其基础上补充了二值填充和 MTU 填充方法。Wright 等^[11, 61]针对其实现的 VoIP 语音攻击, 采用线性填充方式验证了防御方法的有效性, 随后又提出一种流量变形防御方法^[65], 通过动态填充特殊长度的数据包, 使之符合其他类型目标的特征。Luo 等^[66]借鉴长度混淆的思想, 提出了一种基于

数据包变形的防御方法 HTTPoS, 在客户端通过 HTTP Range 头部字段实现对数据的切分, 动态改变数据包的长度和数量。Panchenko 等^[25]针对网站指纹攻击提出了一种引入背景噪声流量的诱骗防御方法, 在用户加载网站时, 随机加载另一个网站, 使其混淆真实网站的流量特征。Wang 等^[67]借鉴 Decoy 诱骗的思想, 提出了一种名为 Walkie-Talkie 的防御方法。该方法在浏览器半双工模式下学习网站特有的数据包突发指纹, 并在用户访问网站时按照其他网站的指纹塑造数据包序列。Monaco 等^[56]在其实现的增量搜索击键攻击中测试了数据包填充防御方法, 结果以 50% 的概率填充 1 个字节就可以有效降低攻击成功率。此外, Schwarz 等^[68]提出的伪造用户击键方法可以有效防御击键攻击。

为了进一步隐藏数据包的时间信息, 现有防御方法一般采取调整数据发送时间和插入伪造数据的做法。Song 等^[9]提出一种隐藏 SSH 加密流量中击键时间信息的防御方法。该方法在 SSH 客户端设置一个固定周期为 50 ms 的发送计时器: 若周期内没有击键, 则客户端发送一个伪造数据包; 若周期内有多次击键, 则客户端合并数据后发送一个数据包。Dyer 等^[69]注意到针对网站指纹攻击的防御方法都无法隐藏加密流量的持续时间、数据总量和突发数量等“粗糙”特征, 理论上只有发送固定长度和时间间隔的数据包, 并且持续固定的时间, 才能完全抵抗侧信道攻击。为此, 文献^[69]提出了一种名为 BuFLO 的防御方法, 使通信双方不论实际数据量大小, 都以恒定的时间间隔发送固定长度的数据包, 并且持续一定的最短时间。该方法虽然从理论上实现了防御侧信道攻

表 4 针对加密流量的语音攻击方法比较
Table 4 Comparison of attack methods against encrypted voice traffic

作者	攻击目标	编码器	模型/方法	数据集	准确率
Wright ^[61]	语言	Speex	n-gram	OGI 电话录音	66%
Wright ^[11]	特定语句	Speex	Profile HMM	TIMIT 音素语音	51%
White ^[62]	语句	Speex	max-entropy+HMM	TIMIT 音素语音	10%
Dupasquier ^[63]	特定语句	SILK	Kalman-filter+DTW	合成语音	60%
Khan ^[64]	用户身份	Speex	ENDs	CSLU 讲话者语音	73%
Backes ^[12]	用户身份	Speex	VAD+3-tuples	公开政治家演讲	48%

击，但带来了巨大的带宽消耗和网络时延，难以在实际中应用。随后，Cai 等^[70]通过分析指纹攻击利用的流量特征以及防御方法需要的带宽下界，提出了一种较好平衡带宽消耗和安全性的 Tamaraw 防御方法。Juarez 等^[71]为了降低 BuFLO 以恒定速率发送数据包的方式带来的网络开销，提出了一种基于自适应填充算法的防御方法 WTF-PAD。该方法根据数据包到达时间间隔将当前网络状态分为突发和空闲状态，并且只在空闲状态下插入伪造突发状态的数据包，实现隐藏真实的突发状态。

表 5 分析对比了当前针对网络侧信道攻击有代表性的防御方法。现有方法虽然可以有效地防范特定的侧信道攻击，但往往存在以下问题。① 简单的防御方法如数据包填充、定时发送等虽然适用性强，但会带来严重的带宽消耗和时间开销，给实际应用部署造成困难。② 复杂的防御方法如 HTTPOS、WTF-PAD 等虽然效果明显，但一般针对特殊的攻击方法，对于未知侧信道攻击的防御效果具有不确定性。Li 等^[72]通过量化网站指纹攻击中各个特征带来的信息增益，揭示了识别准确率与侧信道泄露信息量之间的关系，表明不能单从降低识别准确率来衡量网站指纹攻击防御方法的有效性。

5 结束语

网络加密流量侧信道攻击和防御方法在过去 20 年经历了螺旋式的发展，其内生动力一方面源于用户保护自身隐私的需求，另一方面来自政府、安全部门对打击网络犯罪的需要。当前网络加密流量侧信道攻击研究有了较好的成果，并且一些成果具有良好的实际应用价值，成为加强网络监管、防范网络威胁的重要手段。然而，现有研究中还有一些不足，仍然有很大的发展空间。结合近年来网络侧信道攻击方法的技术发展前沿理论，本文提出以下几个值得深入研究的方向。

(1) 指纹攻击中机器学习模型的可解释性。近年来以网站指纹为代表的指纹攻击方法充分利用神经网络的特征学习能力，突破了 Tor 匿名网络的侧信道安全防御措施。虽然监督式的机器学习模型具有强大的预测能力，但模型分类提取的特征一般缺乏可解释性，因此无法证明分类特征在实际应用中的有效性。如何准确地度量分类特征在侧信道泄露的信息量，是未来值得研究的重要方向。

(2) 增量搜索击键攻击的可扩展性。虽然现有研究提出了多种针对搜索引擎的击键攻击方法，但目前的方法与具体应用高度相关，没有从

表 5 加密流量侧信道攻击的防御方法比较
Table 5 Comparison of defense methods against side channel attack on encrypted traffic

作者	防御目标			防御方法	隐藏信息			开销 ^[*]	防御率 ^[*]
	指纹	击键	语音		长度	时间	数量		
Song ^[9]		√		send timer		√		~500%	100%
Sun ^[19]	√			exp padding	√			~100%	64%
Liberatore ^[22]	√			MTU padding	√			145.3%	88.7%
Wright ^[65]	√		√	morphing	√			35.6% / 15.4%	38.9% / 23.9%
Panchenko ^[25]	√			Decoy	√		√	85%	94.4%
Luo ^[66]	√	√		HTTPOS	√		√	~50%	98% / 100%
Dyer ^[69]	√			BuFLO	√	√	√	129.2%	85.5%
Juarez ^[71]	√			WTF-PAD	√	√	√	80%	78%
Wang ^[67]	√			Walkie-Talkie	√		√	31%	70.5%
Monaco ^[56]		√		rand padding	√			0.3%	47%

注：[*] 开销表示额外的带宽消耗；防御率表示防御后攻击识别准确率降低的比例，即 $1 - \text{Acc}_{\text{Defended}} / \text{Acc}$ 。

系统角度提出增量搜索模型与侧信道信息泄露之间的量化关系。此外,当前研究中假设的击键攻击对象都是英文输入,缺乏对中文输入及相关生物特征的深入研究。相比指纹攻击能够监管用户访问敏感网站的行为,击键攻击可以进一步实现对用户输入内容的分析检测,因此具有更深远的研究价值和实际意义。

(3) 面向新型编码加密语音的攻击方法。现有 VoIP 语音的侧信道攻击方法主要针对 Speex 语音编码器,并且假设每个数据包包含单个帧。目前新型语音编码器 Opus 由 IETF 工作组设计,融合了 Skype 的 SILK 和 Xiph.Org 的 CELT 技术,前者采用与 CELP 类似的线性预测编码,用于 8 kHz 以下的语音频率范围,后者基于改进型离散余弦变换,用于 8 kHz 以上的音乐频率范围。如何设计面向新型编码加密语音的攻击方法,成为网络侧信道语音攻击的一个研究趋势。

参考文献:

- [1] JANSEN B J, RIEH S Y. The seventeen theoretical constructs of information searching and information retrieval[J]. *Journal of the American Society for Information Science and Technology*, 2010, 61(8): 1517-1534.
- [2] JUNG Y, LIM Y, KIM M. Possibilities and limitations of online document tools for design collaboration: the case of google docs[C]//*Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2017: 1096-1108.
- [3] SULAIMAN N, CARRASCO R A, CHESTER E G. Performance evaluation of voice call over an IP based network[C]//*Proceedings of the 41st Annual Conference on Information Sciences and Systems*. 2007: 392-395.
- [4] JANA S, PANDE A, CHAN A, et al. Mobile video chat: issues and challenges[J]. *IEEE Communications Magazine*, 2013, 51(6).
- [5] BRUMLEY B B, TUVERI N. Remote timing attacks are still practical[C]//*16th European Symposium on Research in Computer Security*. 2011: 355-371.
- [6] YAROM Y, FALKNER K. FLUSH+RELOAD: a high resolution, low noise, L3 cache side-channel attack[C]//*Proceedings of the 23rd USENIX Security Symposium*. USENIX Association. 2014: 719-732.
- [7] WANG T, GOLDBERG I. On realistically attacking tor with website fingerprinting[J]. *Privacy Enhancing Technologies*, 2016, 2016(4): 21-36.
- [8] AL-NAAMI K, CHANDRA S, MUSTAFA A M, et al. Adaptive encrypted traffic fingerprinting with bi-directional dependence[C]//*Proceedings of the 32nd Annual Conference on Computer Security Applications*. 2016: 177-188.
- [9] SONG D X, WAGNER D A, TIAN X. Timing analysis of keystrokes and timing attacks on SSH[C]//*Proceedings of 10th USENIX Security Symposium*. 2001.
- [10] CHEN S, WANG R, WANG X, et al. Side-channel leaks in web applications: a reality today, a challenge tomorrow[C]//*IEEE Symposium on Security and Privacy*. 2010: 191-206.
- [11] WRIGHT C V, BALLARD L, COULL S E, et al. Spot me if you can: uncovering spoken phrases in encrypted VoIP conversations[C]//*IEEE Symposium on Security and Privacy*. 2008: 35-49.
- [12] BACKES M, DOYCHEV G, DÜRMUTH M, et al. Speaker recognition in encrypted voice streams[C]//*15th European Symposium on Research in Computer Security*. 2010: 508-523.
- [13] SHANNON C E. A mathematical theory of communication[J]. *The Bell System Technical Journal*, 1948, 27(3): 379-423.
- [14] BACKES M, DOYCHEV G, KÖPF B. Preventing side-channel leaks in web traffic: a formal approach[C]//*20th Annual Network and Distributed System Security Symposium*. 2013.
- [15] Let's Encrypt. Percentage of Web pages loaded by firefox using HTTPS[EB].
- [16] HTTP Archive. Page weight[EB].
- [17] CHENG H, AVNUR R. Traffic analysis of ssl encrypted web browsing[R]. 1998.
- [18] HINTZ A. Fingerprinting websites using traffic analysis[C]//*Privacy Enhancing Technologies*. 2002: 171-178.
- [19] SUN Q, SIMON D R, WANG Y, et al. Statistical identification of encrypted web browsing traffic[C]//*IEEE Symposium on Security and Privacy*. 2002: 19-30.
- [20] FEGHNI S, LEITH D J. A Web traffic analysis attack using only timing information[J]. *IEEE Transactions Information Forensics and Security*, 2016, 11(8): 1747-1759.
- [21] BISSIAS G D, LIBERATORE M, JENSEN D D, et al. Privacy vulnerabilities in encrypted HTTP streams[C]//*Privacy Enhancing Technologies*. 2005: 1-11.
- [22] LIBERATORE M, LEVINE B N. Inferring the source of encrypted HTTP connections[C]//*Proceedings of the 13th ACM Conference on Computer and Communications Security*. 2006: 255-263.
- [23] HERRMANN D, WENDOLSKY R, FEDERRATH H. Website fingerprinting: attacking popular privacy enhancing technologies with the multinomial naïve-Bayes classifier[C]//*Proceedings of the first ACM Cloud Computing Security Workshop*. 2009: 31-42.
- [24] 顾晓丹, 杨明, 罗军舟, 等. 针对 SSH 匿名流量的网站指纹攻击方法[J]. *计算机学报*, 2015, 38(4): 135-147.
- [25] GU X, YANG M, LUO J, et al. Website fingerprinting attack based on hyperlink relations[J]. *Chinese Journal of Computers*, 2015, 38(4): 135-147.
- [26] 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*. 2011: 103-114.
- [27] CAI X, ZHANG X C, JOSHI B, et al. Touching from a distance: website fingerprinting attacks and defenses[C]//*Proceedings of the 2012 ACM Conference on Computer and Communications Security*. 2012: 605-616.

- [27] WANG T, GOLDBERG I. Improved website fingerprinting on Tor[C]//Proceedings of the 12th Annual ACM Workshop on Privacy in the Electronic Society. 2013: 201-212.
- [28] WANG T, CAI X, NITHYANAND R, et al. Effective attacks and provable defenses for website fingerprinting[C]//Proceedings of the 23rd USENIX Security Symposium. USENIX Association. 2014: 143-157.
- [29] PANCHENKO A, LANZE F, PENNEKAMP J, et al. Website fingerprinting at internet scale[C]//23rd Annual Network and Distributed System Security Symposium. 2016.
- [30] HAYES J, DANEZIS G. K-fingerprinting: a robust scalable website fingerprinting technique[C]//Proceedings of the 25rd USENIX Security Symposium. 2016: 1187-1203.
- [31] ABE K, GOTO S. Fingerprinting attack on tor anonymity using deep learning[J]. Proceedings of the Asia-Pacific Advanced Network, 2016, 42: 15-20.
- [32] 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.
- [33] SIRINAM P, IMANI M, JUÁREZ M, et al. Deep fingerprinting: undermining website fingerprinting defenses with deep learning[C]//Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. 2018: 1928-1943.
- [34] 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(4): 292-310.
- [35] RAHMAN M S, SIRINAM P, MATTHEWS N, et al. Tik-Tok: the utility of packet timing in website fingerprinting attacks[J]. Proceedings on Privacy Enhancing Technologies, 2020(3).
- [36] 马陈城, 杜学绘, 曹利峰, 等. 基于深度神经网络 burst 特征分析的网站指纹攻击方法[J]. 计算机研究与发展, 2020, 57(4): 746-766.
- MA C D, DU X H, CAO L F, et al. Burst-analysis website fingerprinting attack based on deep neural network[J]. Journal of Computer Research and Development, 2020, 57(4): 746-766.
- [37] DAI S, TONGAONKAR A, WANG X, et al. NetworkProfiler: towards automatic fingerprinting of android APPs[C]//Proceedings of the IEEE INFOCOM 2013. 2013: 809-817.
- [38] Google Security Blog. An update on android TLS adoption[EB].
- [39] CONTI M, MANCINI L V, SPOLAOR R, et al. Analyzing Android encrypted network traffic to identify user actions[J]. IEEE Transactions on Information Forensics and Security, 2016, 11(1): 114-125.
- [40] TAYLOR V F, SPOLAOR R, CONTI M, et al. AppScanner: automatic fingerprinting of smartphone APPs from encrypted network traffic[C]//IEEE European Symposium on Security and Privacy. 2016: 439-454.
- [41] STÖBER T, FRANK M, SCHMITT J B, et al. Who do you sync you are? smartphone fingerprinting via application behaviour[C]//Proceedings of the Sixth ACM Conference on Security and Privacy in Wireless and Mobile Networks. 2013: 7-12.
- [42] WANG Q, YAHYAVI A, KEMME B, et al. I know what you did on your smartphone: inferring App usage over encrypted data traffic[C]//2015 IEEE Conference on Communications and Network Security. 2015: 433-441.
- [43] ALAN H F, KAUR J. Can android applications be identified using only TCP/IP headers of their launch time traffic?[C]//Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks. 2016: 61-66.
- [44] SALTAFORMAGGIO B, CHOI H, JOHNSON K, et al. Eavesdropping on fine-grained user activities within smartphone apps over encrypted network traffic[C]//10th USENIX Workshop on Offensive Technologies. USENIX Association. 2016.
- [45] TAYLOR V F, SPOLAOR R, CONTI M, et al. Robust smartphone app identification via encrypted network traffic analysis[J]. IEEE Transactions on Information Forensics and Security, 2018, 13(1): 63-78.
- [46] 朱迪, 陈丹伟. 基于密度聚类 and 随机森林的移动应用识别技术[J]. 计算机工程与应用, 2020, 56(4): 63-68.
- ZHU D, CHEN D W. Technology of mobile application identification based on density-based clustering and random forest[J]. Computer Engineering and Applications, 2020, 56(4): 63-68.
- [47] JOYCE R, GUPTA G K. Identity authentication based on keystroke latencies[J]. Communication ACM, 1990, 33(2): 168-176.
- [48] MONROSE F, RUBIN A D. Keystroke dynamics as a biometric for authentication[J]. Future Generation Computer Systems, 2000, 16(4): 351-359.
- [49] ARAÚJO L C F, SUCUPIRA L H R, LIZÁRRAGA M G, et al. User authentication through typing biometrics features[J]. IEEE Transactions on Signal Processing, 2005, 53(2): 851-855.
- [50] 王林, 贺冰清. 采用击键特征曲线差异度的用户身份认证方法[J]. 计算机工程与应用, 2018, 54(22): 160-166, 196.
- WANG L, HE B Q. Keystroke dynamic identity authentication algorithm based on diversity degree of key-stroke feature curve[J]. Computer Engineering and Applications, 2018, 54(22): 160-166, 196.
- [51] DHAKAL V, FEIT A M, KRISTENSSON P O, et al. Observations on typing from 136 million keystrokes[C]//Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018: 1-12.
- [52] SHARMA S A, MENEZES B L. Implementing side-channel attacks on suggest boxes in web applications[C]//Proceedings of the First International Conference on Security of Internet of Things. 2012: 57-62.
- [53] SCHAUB A, SCHNEIDER E, HOLLENDER A, et al. Attacking suggest boxes in web applications over https using side-channel stochastic algorithms[C]//Risks and Security of Internet and Systems. 2014: 116-130.
- [54] OH S E, LI S, HOPPER N. Fingerprinting keywords in search queries over tor[J]. Proceedings on Privacy Enhancing Technologies, 2017(4): 251-270.
- [55] OH S E, SUNKAM S, HOPPER N. P1-FP: extraction, classification, and prediction of website fingerprints with deep learning[J]. Proceedings on Privacy Enhancing Technologies, 2019(3): 191-209.
- [56] MONACO J V. What are you searching for? a remote keylogging attack on search engine autocomplete[C]//Proceedings of 28th USENIX Security Symposium. USENIX Association, 2019: 959-976.
- [57] MONACO J V. Feasibility of a keystroke timing attack on search engines with autocomplete[C]//2019 IEEE Security and Privacy

- Workshops. 2019: 212-217.
- [58] 蒋竺芳. 端到端自动语音识别技术研究[D]. 北京: 北京邮电大学, 2019.
- JIANG Z. Research in End-to-end Automatic Speech Recognition Technology[D]. Beijing: Beijing University of Posts, 2019.
- [59] 张瑞珍, 韩跃平, 张晓通. 基于深度 LSTM 的端到端的语音识别[J]. 中北大学学报(自然科学版), 2020, 41(3): 244-248.
- ZHANG R, HAN Y, ZHANG X. End-to-end speech recognition based on depth-gated LSTM[J]. Journal of North University of China (Natural Science Edition), 2020, 41(3): 224-248.
- [60] CMU. The CMU pronouncing dictionary[EB].
- [61] WRIGHT C V, BALLARD L, MONROSE F, et al. Language identification of encrypted voip traffic: alejandra y roberto or alice and bob?[C]//Proceedings of 16th USENIX Security Symposium. USENIX Association. 2007.
- [62] WHITE A M, MATTHEWS A R, SNOW K Z, et al. Phonotactic reconstruction of encrypted VoIP conversations: hookt on fon-iks[C]//IEEE Symposium on Security and Privacy. 2011: 3-18.
- [63] DUPASQUIER B, BURSCHKA S, MCLAUGHLIN K, et al. Analysis of information leakage from encrypted skype conversations[J]. International Journal of Information Security, 2010, 9(5): 313-325.
- [64] KHAN L A, BAIG M S, YOUSSEF A M. Speaker recognition from encrypted VoIP communications[J]. Digital Investigation, 2010, 7(1-2): 65-73.
- [65] WRIGHT C V, COULL S E, MONROSE F. Traffic morphing: an efficient defense against statistical traffic analysis[C]//16th Annual Network and Distributed System Security Symposium. 2009.
- [66] LUO X, ZHOU P, CHAN E W W, et al. HTTPoS: sealing information leaks with browser-side obfuscation of encrypted flows[C]//18th Annual Network and Distributed System Security Symposium. 2011.
- [67] WANG T, GOLDBERG I. Walkie-talkie: an efficient defense against passive website fingerprinting attacks[C]//Proceedings of the 26th USENIX Security Symposium. 2017: 1375-1390.
- [68] SCHWARZ M, LIPP M, GRUSS D, et al. KeyDrown: eliminating software-based keystroke timing side-channel attacks[C]//25th Annual Network and Distributed System Security Symposium. 2018.
- [69] DYER K P, COULL S E, RISTENPART T, et al. Peek-a-boo, I still see you: why efficient traffic analysis countermeasures fail[C]//IEEE Symposium on Security and Privacy. IEEE Computer Society, 2012: 332-346.
- [70] 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. 2014: 227-238.
- [71] JUÁREZ M, IMANI M, PERRY M, et al. Toward an efficient website fingerprinting defense[C]//21st European Symposium on Research in Computer Security. 2016: 27-46.
- [72] 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. 2018: 1977-1992.

[作者简介]



李玎(1992-), 男, 河南郑州人, 信息工程大学博士生, 主要研究方向为网络信息安全、机器学习。



祝跃飞(1962-), 男, 浙江杭州人, 博士, 信息工程大学教授、博士生导师, 主要研究方向为信息安全、公钥密码。



芦斌(1983-), 男, 河南郑州人, 博士, 信息工程大学副教授, 主要研究方向为网络信息安全、机器学习。



林伟(1986-), 男, 湖南常德人, 博士, 信息工程大学讲师, 主要研究方向为软件保护与分析、网络信息安全。