利用基于超表面的物理深度伪造欺骗无线感知

Johannes Kortz*、Vignesh Natarajkumar*†、Christof Paar* 和 Paul Staat‡
*马克斯·普朗克安全与隐私研究所， {*johannes.kortz, christof.paar*}*@mpi-sp.org*
‡波鸿鲁尔大学，*paul.staat@rub.de*
†印度理工学院马德拉斯分校，*vigneshn@smail.iitm.ac.in*

arXiv:2611.32386v1 [cs.CR] 24 Nov 2026

摘要——**无线通信网络正从单纯的信息传输扩展到感知：在一体化感知与通信范式中，现有基础设施和信号被重复利用，以获取设备周围环境的物理信息，旨在为医疗保健和自动驾驶车辆等高风险应用提供信息。从安全角度看，无线感知天然信任对无线电环境的物理观察。然而，与无线通信一样，开放的无线电传播特性使攻击者能够篡改受害者的信号。具体而言，攻击者可以利用可编程超表面操纵无线信道，从而误导感知系统。然而，以往基于超表面的操纵仅针对方差或多普勒频移等单个信道特征。因此，这些方法需要了解受害者感知系统，还会产生容易被检测到的不可信信道。**

在本文中，我们展示了超表面能够对无线感知进行物理上可信的操纵，且不依赖于受害者感知系统所使用的特征。我们的攻击在物理层重现真实的预录事件，从而合成无线信道的物理深度伪造。为此，我们开发了一种攻击框架，可在时间、频率和天线维度上控制受害者信道的幅度与相位。这里的关键挑战在于，建立超表面与受害者信道之间交互的深度学习模型。我们证明，攻击者可以在多种设置下解决这一挑战，例如利用 Wi-Fi 波束成形反馈信息或复制受害者环境。我们既在 MIMO 信道上开展受控实验室研究，也使用现成 Wi-Fi 感知对该攻击进行了验证。我们的结果从根本上对无线感知的可信度提出了挑战。

1. 引言

蜂窝网络和 Wi-Fi 等无处不在的无线网络连接着家庭、办公室、工业环境和公共空间中的数十亿台设备。迄今为止，这些网络仅用于传输数字数据。然而，在一体化感知与通信（ISAC）这一根本性范式转变中，已部署的无线通信基础设施和信号将被重新用于无线感知，以获取周围物理环境的信息。标准化机构正在积极研究 ISAC [1], [2], [3] 和设备制造商 [4], [5],，首批现场测试已在中国 [6], 、欧盟 [7], 和美国 [8]报告。

在 Wi-Fi 中，IEEE802.11bf 已将感知标准化为一项原生能力 [9]。从技术上讲，它建立在接收器为每个接收数据包估计的信道状态信息（CSI）之上，该信息刻画了无线电波与物理环境的交互方式。如今，基于 Wi-Fi 的感知已经能够推断个体的活动 [10],位置 [11],生命体征 [12], 和身份 [13] 。未来，无线感知有望为安全与安保关键型应用提供信息，例如自主无人机和工厂机器人的碰撞规避 [2], [14]。

从安全角度来看，无线感知建立在这样一个假设之上：对无线电环境的物理观察能够代表物理现实。如果这一假设被违反，伪造的生命体征可能导致错误的医疗治疗，而欺骗性运动可能在入侵检测中触发误报。这是一个现实的担忧，因为无线感知面临着与无线通信相同的安全挑战：无线电传播是一种开放介质，因此可能受到对手控制。因此，无线感知容易遭受操纵。具体而言，已有研究表明，超表面——数字可配置无线电反射器——能够让对手实时操纵无线信道并误导感知 [15], [16], [17], [18]。例如，IRShield 随机切换超表面配置，以混淆物理事件 [15], ；而 RISiren 切换配置，合成能够欺骗感知事件的多普勒频移 [16]。二者都依赖于在不同配置之间快速切换，这会产生自然传播中不存在的突发信道不连续性。

更根本地说，现有的欺骗攻击针对感知系统从信道中提取的个别特征，例如多普勒频移。这会带来两个后果。首先，这类操纵假设受害者感知系统处于白盒设置，因为要操纵正确的特征，就必须知道目标系统依赖哪种特征。

其次，孤立地改变某项特征，会使其余信道特征与伪造事件不一致，因此合成信道并不对应于所感知事件在物理上合理的发生情况。

在本研究中，我们展示了人体运动和生命体征能够以逼真的方式进行欺骗，且不受受害者感知系统的影响，从而从根本上挑战

无线感知的可信度。我们将这种操纵称为物理深度伪造，并提出了一种用于合成此类伪造的新型攻击框架。我们不操纵单个特征，而是利用超表面重现预录真实事件的信道，使感知系统可能提取的所有特征在构造上都逼真且一致。这带来了两项挑战。首先，攻击者必须找到能够产生目标信道的超表面配置，而穷举搜索在计算上难以实现。我们将这一问题形式化，并通过训练一个描述超表面对信道影响的深度学习模型，再利用梯度下降在模型中搜索配置来解决这一问题。其次，该模型必须捕捉超表面对受害者信道的影响，但攻击者无法接触受害传感器本身。因此，我们考虑了五种攻击者设置，它们在获取信道知识的方式上各不相同，范围从直接观察受害者信道到完全无法访问信道。

在一项针对160MHz 多输入多输出（MIMO）信道的实验室研究中，我们的框架能够跨时间、频率和天线复现真实物理事件的信道。我们还发现，攻击难度取决于受害者的测量噪声，以及联合合成的信道系数数量。随后，我们在现成的 Wi-Fi 设备上演示了该攻击。在此，攻击者被动利用波束成形反馈信息（BFI）以明文形式传输这一点，欺骗了三种已发表的、基于 BFI 的呼吸估计器，表明这类感知系统并不安全。除了这种直接观察，我们还研究了信道知识受限的攻击者：他们通过受害者装置的物理复制体，或借助窃听天线来校准超表面。仅使用真实录制数据训练的分类器，仍会在58% 的情况下将由窃听天线生成的深度伪造归类为预期活动（随机猜测概率：20%）。最后，我们展示了超表面无需了解感知系统，便能克隆物理物体的反射。总而言之，我们的贡献如下：

- 我们提出针对无线感知的物理深度伪造，即能够紧密追踪另一场景物理状况的信道测量。由于物理深度伪造复现的是信道本身，因此不受受害者感知处理方式的影响。我们还表明，先前基于超表面的攻击很容易从原始信道数据中被检测出来。
- 我们提出利用超表面进行对抗性任意信道合成，以欺骗无线感知，并将其形式化为寻找超表面配置的问题，再通过深度学习求解。我们表明，超表面攻击者能够在时间、频谱和空间域中，准确重现真实事件的幅度和相位信道剖面。
- 我们表明，即使无法直接观测受害者信道，攻击者也能够校准超表面的影响。
- 我们证明，基于BFI的感知容易受到物理操纵攻击，因为它会向攻击者泄露感知数据。我们在现成的 Wi-Fi 设备上演示了这种攻击。

2. 预备知识

2.1. 无线感知

无线电信号从发射器传输到接收器时，会沿着多条路径在环境中传播：包括直达路径，以及墙壁、物体和人体产生的各种反射。这些路径的叠加形成信道响应，其中编码了传播环境中的物理变化，例如人员移动或呼吸。无线感知 利用这一点，将信道响应视为对物理环境的测量。这正是雷达的工作原理 [19], ，也可以通过 Wi-Fi [10] 或超宽带（UWB） [20]等常规通信信号实现。通信接收器已经会针对每个接收的数据包估计信道，例如Wi-Fi 中的 CSI，以补偿传播效应；而商用芯片组会将这些估计值提供给软件 [21], [22], [23]。Wi-Fi 设备还会显式交换信道估计以进行波束成形。接收器测量信道，并将其作为所谓的 BFI 报告给发射器。由于 BFI 以未加密的形式传输 [24],，任何处于通信范围内的设备都能窃听这些反馈，并将其用于感知 [25]。

基于此类信道估计，无线感知可以提取有关物理环境和个体的信息，例如其位置 [11], 生命体征 [12], 身份 [13], 以及活动 [26]。这已促成了老年护理中的跌倒检测 [27], 独自留在车内儿童的检测 [28], 以及物理入侵检测 [29]等应用。

无线感知目前正从研究领域走向标准化和产品化。在通感一体化的框架下，物理感知将成为一种原生网络功能，从底层构建并内置于网络中，而不再是对通信信号的机会性复用 [30], [31]。即将推出的 6G [14]和 Wi-Fi [9] 标准将使基站、路由器和智能手机能够分析其周围的物理环境，国际电信联盟也将感知技术列为 2030 [3]之后的关键技术趋势。作为全球蜂窝网络标准背后的标准化组织，3GPP 已研究了 32 个感知用例，其中包括自主无人机和工厂机器人碰撞规避等安全关键型应用 [2]。目前，欧盟 [7],中国 [6], 和美国 [8],已在开展现场测试，基于 Wi-Fi感知的家庭安防产品也已实现商业化 [32]。

2.2. 超表面技术

2014 年前后，可编程超表面从超材料物理学中兴起，用于定制无线电传播环境 [33]。在无线通信研究中，它们通常被称为可重构智能表面

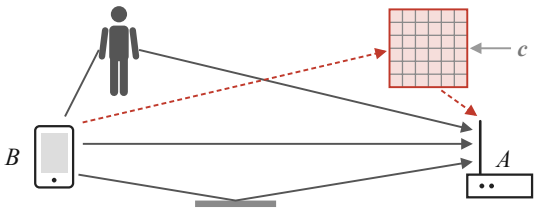


图1：系统模型。发射器 B 经由多条传播路径到达受害传感器 A ，其中包括良性反射和恶意反射。

（RISs），并被视为未来 6G网络的一种潜在构成要素 [34]。如今，超表面已发展为商业产品 [35], [36]。通过电子方式控制其对入射波的反射方式，超表面使传播介质本身成为可调资源 [37],例如，可用于重定向无线电波、改善覆盖范围或消除盲区 [38], [39], [40]。

典型的超表面是由 L 个相同的单元反射器组成的平面阵列，印制在低成本印刷电路板（PCB） [41]上。每个元件的反射相位通过数字逻辑信号独立调谐，例如来自微控制器的信号。在常见的1比特控制情况下，每个元件在两种反射相位状态 $0°$与 $180°$之间切换，对应的反射系数分别为 $+1$ 与-1。因此，超表面配置是一个二进制向量$\mathbf{c} \in \{0,1\}^L$，并且每一种表面配置都会在物理上表现为一种不同的反射模式。

何种配置能够产生期望行为（例如最大化发射器与接收器之间的信号功率），取决于周围的无线电环境：反射分量必须与直接传播路径相干叠加。由于二进制超表面可提供 2^L 种配置，实际控制依赖于借助信道测量 [33], [42], [43],[44], [45]的反馈进行迭代优化。

3. 针对无线感知的物理深度伪造攻击

本节介绍系统模型和攻击者模型，并定义攻击目标。随后，我们将物理深度伪造作为一种独特的攻击类型加以介绍，这类攻击旨在误导无线感知系统，对手会对受害者环境进行物理操控。

3.1. 系统与攻击者模型

系统模型 我们考虑这样一种双基地无线感知场景：两台无线设备部署在室内环境中，例如住宅或办公室。设备 A 是受害传感器，用于接收设备 B发射的感知信号。A根据 B的信号估计无线信道 $\mathbf{H}$ ，并据此推断物理事件。根据感知波形和硬件能力， A 可能会在时间、频率和空间域上观测信道。$\mathbf{H}$ 是所有传播路径的叠加，这些路径从 B 传播至 A。$\mathbf{H} = \mathbf{H}_{\text{static}} + \mathbf{H}_{\text{event}}$。直达路径和墙面反射记为 $\mathbf{H}_{\text{static}}$。 A 用于推断物理事件的传播路径（例如，移动的人产生的反射）构成分量$\mathbf{H}_{\text{event}}$。

攻击者模型 攻击者可以物理访问受害者环境，并能放置超表面以影响$\mathbf{H}$ between A and B之间的信道。攻击者可以检查受害者环境，因此了解感知硬件和受害者装置的几何结构，包括 $A$and B的位置。此外，攻击者可以窃听受害者环境中传输的所有无线消息，以观察物理层信号，但无法读取加密数据。攻击者无法访问 A: 他们既不能读取或修改 A's的信道估计，也不知道 A如何推断物理事件，例如感知处理、特征和推理函数。

超表面 攻击者的超表面由 L个单元组成，并为受害者信道增加第三个分量，

$$\mathbf{H} = \mathbf{H}_{\text{static}} + \mathbf{H}_{\text{event}} + \mathbf{H}_{\text{meta}}(\mathbf{c}), \tag{1}$$

其中，$\mathbf{c} \in \{0,1\}^L$ 是超表面配置。$\mathbf{H}_{\text{meta}}$是由 B 到超表面单元、再到 A的 L 条拼接信道叠加而成的，

$$\mathbf{H}_{\text{meta}}(\mathbf{c}) = \sum_{l=1}^{L} r(c_l)\,\mathbf{H}_l, \tag{2}$$

其中，$\mathbf{H}_l$ 是从 B 经由单元 l 到 A 的信道，而$r(c_l)$ 是单元 l 处于状态 c_l时的反射系数，即， $r(0) = +1$ 和 $r(1) = -1$，参见第 2.2 节。攻击者以至少与 A 对信道进行采样一样快的速度随时间切换 $\mathbf{c}$ ，从而部分控制 A 观测到的 $\mathbf{H}$。攻击者无法移除静态路径，即，这些路径会相对于 $\mathbf{H}_{\text{static}}$塑造信道。

攻击者目标 通过重新配置超表面，攻击者旨在使 A推断出一个实际上并未发生的目标事件，例如，让人误以为有人穿过空房间。

攻击者设置 为合成目标信道，攻击者需要了解配置$\mathbf{c}$ 如何影响 A处的信道。由于攻击者无法访问 A，因此无法直接观测这种影响。因此，我们区分五种攻击者设置，以细化上述攻击者模型；这些设置的区别在于攻击者如何获取信道知识，如表 1 所示。各设置按信道知识从多到少排列，我们的评估也遵循这一顺序。

3.2 物理深度伪造

我们现在通过抽象一个通用感知模型，来定义针对无线感知的物理深度伪造；随后，我们将以基于超表面的无线感知操纵来实例化该模型。

表 1：攻击者设置及其获取信道知识的相应能力。

类别	设置	攻击者控制	信道知识来源	Sec.
观察	**S1** 不诚实伙伴	控制 B	B 可直接测量互易信道	5
	S2BFI窃听者	被动窃听	通过未加密的波束成形信息直接访问	6
估计	**S3**物理复制体	离线重建装置	在受害者装置的离线副本上测量	7.1
	S4 对称代理	窃听天线 E	通过 A 与 E之间的几何对称性推断	7.2
克隆	**S5** 环境克隆	初始环境访问	无（超表面模拟合法环境）	8

3.2.1. 通用感知模型。我们将第3.1节中的设置抽象为一系列映射，这些映射将物理ˆ事件 X 映射为推断事件 X。尽管该模型旨在用于无线感知，但它同样适用于任意感知系统，例如摄像机、激光雷达或雷达，因此我们可以独立于感知方式来定义攻击：

$$X \xrightarrow{\pi} Y \xrightarrow{\varepsilon} \mathbf{M} \xrightarrow{\Phi} f \xrightarrow{g} \hat{X}, \tag{3}$$

其中，物理变换 π 将事件 X（例如，某人行走）映射为传感器处的信号 Y（例如，通过从 B 到 A的无线电传播）。测量函数 ε 将Y 转换为测量值$\mathbf{M}$（例如， A的信道估计）。处理过程 Φ 提取特征 f（例如，多普勒频移），推理函数 g（例如，活动分类器）据此推断事件 X。

测量是由 T 快照组成的序列 $\mathbf{M} = (\mathbf{M}_1, \ldots\ldots, \mathbf{M}_T)$，例如，来自连续数据包的信道估计，其中每个 $\mathbf{M}_i \in \mathbb{C}^d$ 捕获传感器的 d 自由度（DoF）。

3.2.2. 攻击目标。令 X 表示实际发生的物理事件˜，令 $X \neq X$ 表示攻击者试图使系统推断出的目标事件。对于事件 X的测量，我们记为 $\mathbf{M}(X) = \varepsilon(\pi(X))$。攻击者无法在物理上呈现 X。相反，他们通过配置 $\mathbf{c}$操纵环境，例如使用第3.1节中的超表面配置，从而合成测量值 $\mathbf{M}(\mathbf{c}) = \varepsilon(\pi(X; \mathbf{c}))$。我们假设攻击期间环境保持静态，即不会发生真实事件，因此合成测量值仅取决于 $\mathbf{c}$。我们区分攻击成功率的三个等级，其强度依次递增：

$$\text{(inference level)} \quad g(\Phi(\mathbf{M}(\mathbf{c}))) = g(\Phi(\mathbf{M}(\tilde{X}))), \tag{4}$$

$$\text{(feature level)} \quad \Phi(\mathbf{M}(\mathbf{c})) \approx \Phi(\mathbf{M}(\tilde{X})), \tag{5}$$

$$\text{(measurement level)} \quad D\big(\mathbf{M}(\mathbf{c}), \mathbf{M}(\tilde{X})\big) \leq \delta_\varepsilon, \tag{6}$$

其中 $D(\cdot, \cdot)$ 是距离度量， δ_ε 是由测量函数的质量决定的容差。

> **定义*1***（物理深度伪造）。物理深度伪造是一种测量 $\mathbf{M}(\mathbf{c})$，由攻击者通过对环境进行物理操纵而诱发，旨在使其与˜目标事件的测量 X相匹配。如果$D\big(\mathbf{M}(\mathbf{c}),\mathbf{M}(\tilde{X})\big) \leq \delta_\varepsilon$，则该攻击在测量层面上成功。

我们沿用近期关于模仿人体运动的人形机器人研究中的术语 [46]，并将制造物理深度伪造的尝试称为物理深度伪造攻击。与媒体深度伪造类似，物理深度伪造通过深度学习合成，并冒充真实事件。从构造上说，它独立于 Φ 和 g，且任何仅基于$\mathbf{M}$运行的下游机制都无法检测到它。因此，它构成了攻击目标，并符合我们的攻击者模型；在该模型中，攻击者并不掌握 Φ 和 g的相关知识。不同于物理对抗样例 [47], [48], ——后者会使测量结果接近实际事件的测量结果，并利用$\mathbf{M}$的弱点——物理深度伪造会复现目标事件的测量结果，因此即使是理想的推理函数也会受到欺骗。

3.2.3. 实现 g Phy理深度伪造p造与超表面。在第3.1节的设置中，测量$\mathbf{M}$ 是 A的信道估计$\mathbf{H} = (\mathbf{H}_1, \ldots, \mathbf{H}_T)$，其中$\mathbf{H}_i \in \mathbb{C}^{s\times k}$对应 s 条空间链路和 k 个子载波，*i.e.*， $d = s \cdot k$。超表面配置$\mathbf{c}$通过在式（1）的信道中添加$\mathbf{H}_{\text{meta}}(\mathbf{c})$ 来影响传播 π。为简单起见，我们假设攻击期间环境静态，$\mathbf{H}_{\text{event}} = \mathbf{0}$，且物理深度伪造需要

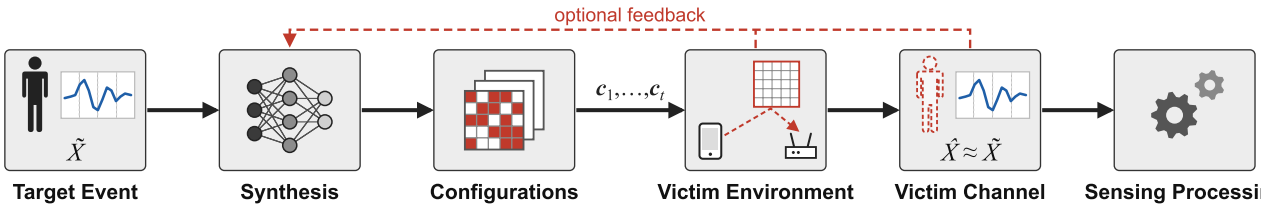
$$\mathbf{H}_{\text{static}} + \mathbf{H}_{\text{meta}}(\mathbf{c}) \approx \mathbf{H}_{\text{static}} + \mathbf{H}_{\text{event}}(\tilde{X}), \tag{7}$$

即，超表面路径取代真实的˜目标事件 X 发生时会产生的路径。

为合成目标信道，攻击者需要一个模型 ˆ$\mathbf{H}(\mathbf{c})$，该模型能够针对给定的超表面配置 $\mathbf{c}$，预测 A 处的信道。我们考虑了几种获取该模型的攻击者设置，如表1所总结。

4. 使用超表面合成物理深度伪造

物理深度伪造要求攻击者使 A's˜信道符合目标信道 $\mathbf{H}_{\text{static}} + \mathbf{H}_{\text{event}}(X)$。因此，攻击者必须找到一系列超表面配置 $(\mathbf{c}_i)_{i=1}^T$，以重现目标信道的录制内容。由于可能的配置数量为 2^L，穷举搜索并不可行。此外，每种配置对 A's 信道的影响取决于受害者环境，而攻击者无法通过 A 观察该环境。最后，这一系列配置必须在时间、频率和天线维度上共同重现目标。下文将讨论攻击者如何应对这些挑战。



˜图2：物理深度伪造攻击流程。攻击者根据目标事件的录制内容 X合成一系列超表面配置 $\mathbf{c}_1$，……，$\mathbf{c}_T$。在受害者环境中重放这些配置后，它们会塑造受害者信道，使得ˆ˜感知处理推断出 $X \approx X$。是否获取受害者环境的反馈均可。

4.1. 前向代理模型

为了搜索配置，攻击者需要预测给定配置下 A 处的信道。式（2）中的理想化物理模型需要每个单元的信道$\mathbf{H}_l$，而攻击者无法测量这些信道；该模型还忽略了相邻单元之间的耦合。因此，我们利用神经网络从信道测量中学习这一映射，构建前向代理模型 $\hat{\mathbf{H}}(\mathbf{c})$。

代理模型将配置$\mathbf{c}$ 映射为时间 A处的信道$\hat{\mathbf{H}}(\mathbf{c}) \in \mathbb{C}^{s\times k}$。它将两个并行分支相加：一个线性分支，用于模拟式（2）；另一个卷积分支，按照第4.3节的做法，将$\mathbf{c}$ 视为 16×16 孔径图像，并捕捉相邻单元之间的耦合。我们使用已知配置的信道测量来训练代理模型。

根据攻击者能够观察到的每种配置的信道知识量，我们考虑了不同的攻击者设置，如表 1所示。当受害者信道可观测时，攻击者使用通过 B （S1）或通过窃听波束成形信息（S2）获取的受害者信道本身来训练代理模型。当无法获取受害者信道时，攻击者会使用另一个无线信道（信道副本或辅助信道）进行训练，然后用该代理模型估计受害者环境（S3、S4）。在只能初次访问受害者环境一次、且无法访问受害者信道的情况下（S5），攻击者会分析环境一部分区域的反射，并使用在自身设置上训练的超表面对其进行克隆。我们整理了一个包含已知配置的数据集，其中涵盖随机图案、它们的位翻转邻居、稀疏和结构化图案，以及波束转向配置，以便数据集能够体现超表面每个单元的影响。如果只能进行少量测量，我们会按照 [49] 微调在另一环境中预训练的代理模型（见第5.2节）。

如果代理模型在留出配置上的平均距离 $D\big(\mathbf{H}(\mathbf{c}), \hat{\mathbf{H}}(\mathbf{c})\big)$ 保持在由系统噪声给出的容差 δ_ε 范围内，我们就认为其足够准确。我们将在第5.2节进一步讨论所需的训练工作量，模型架构和训练细节见附录A。

4.2. 配置搜索

我们通过对前向信道进行迭代的基于梯度的搜索，获得目标信道$\mathbf{c}$ $\mathbf{H}$ $\in\mathbb{C}^{s\times k}$ 的配置。

代理模型上进行迭代式基于梯度的搜索，获得目标信道 $\mathbf{c}$ 的超表面配置 $\mathbf{H}$，其中利用距离度量 $D\big(\mathbf{H}, \hat{\mathbf{H}}(\mathbf{c})\big)$ 的梯度，将其用作损失函数。由于代理模型具有可微性，并且运行在图形处理器（GPU）上，这种搜索策略成本低且可并行化。该方法的另一项优势是，搜索能够生成训练数据集中不存在的超表面配置。我们将 $\mathbf{c}$ 的二进制状态视为可训练参数，并冻结代理模型的权重。我们为每个元素优化一个连续逻辑值，并在每次代理模型评估之前将逻辑值进行阈值化，以得到二进制配置，从而确保每个被评估的配置都是可实现的。梯度通过直通估计器穿过阈值。我们通过单元素翻转和成对翻转来细化最佳配置，从而降低误差。搜索参数和细化细节见附录A。

如第3.1节所述，我们可以在时间、频率和空间维度上设计信道响应，以攻击各种受害传感器。对于信道估计$\mathbf{H}_i \in \mathbb{C}^{s\times k}$，其中 $1 \le i \le T$，我们可以聚焦于空间链路的子集 $\mathcal{S} \subseteq \{1,\ldots,s\}$ 以及子载波的子集 $\mathcal{K} \subseteq \{1,\ldots,k\}$。随后，逆向设计和距离计算仅针对信道估计 $\mathbf{H}'_i \in \mathbb{C}^{|\mathcal{S}|\times|\mathcal{K}|}$进行。这样便可以在特定频段内或跨空间链路设计欺骗攻击。第5.3节描述了特定帧中合成质量与自由度数量之间的关系。逆向设计针对每个时间步独立执行，生成一个配置序列 $(\mathbf{c}_i)_{i=1}^{T}$，并以录制内容的原始采样率在超表面上重放。不同于在固定配置之间切换，每个配置都会在其对应的时间步复现目标信道，因此观测到的信道会遵循目标事件的自然变化，而不是在固定状态之间跳跃。在时间上耦合各帧会将关于目标时间变化的假设纳入攻击之中，从而限制可合成活动的范围。

4.3. 实验设置

我们介绍了各项实验中共用的硬件。各项实验的具体设置将在相应章节中介绍。

超表面 攻击者的超表面采用 Heinrichs等人 [50]公开的设计。它是一块 FR4印刷电路板，$L = 256$ 个单元按 16×16 网格排列，尺寸为

36.0 厘米 $\times$ 24.7厘米。每个单元都采用二进制相位控制，并针对5GHz Wi-Fi 频段进行了优化。我们通过 USB 串行通信对配置$\mathbf{c}$ 进行编程，从而将每个单元 l 设为状态 $c_l = 0$ 或 $c_l = 1$，即，反射相位为 $0°$ 或 $180°$ （$r(c_l) = \pm 1$，见式 (2)）。更多技术细节见 [50]。

实验室设置 我们的实验室设置（图 3）测量一个 4×4 MIMO 信道，即，$s = 16$ 空间链路。是德科技 P9372A 矢量网络分析仪（VNA）与射频（RF）开关配合使用，记录每条链路的复数频率响应。对于每条链路，我们在以5.57GHz为中心的160MHz 频段内测量129 个频率点，中频带宽为500kHz。测量 MIMO 信道的单个帧大约需要0.13s。超表面放置在距离发射阵列0.2m 处，并将攻击者控制的路径加入信道。电动平台改变传播条件。直线轴以10cm为步长，将发射阵列与接收阵列之间的距离从1 m 调整至1.5m；旋转台将超表面相对于发射阵列的角度设置为 $0°$、$30°$ 或 $45°$。为表征不同环境，我们针对轴向距离和超表面角度的每种组合，测量150,800种已知的超表面配置。

Wi-Fi 设备 对于商用硬件上的实验，我们使用 ESP32-C5 设备和 Netgear Nighthawk RAX50 路由器，它们运行一个信道带宽为 20 兆赫的 IEEE 802.11ac 网络。所有网络均配置为 Wi-Fi 信道 112，与超表面的有效带宽相匹配。在设置 S3 和 S4 中，ESP32-C5 设备同时充当发射器和响应器，并从接收器中提取 CSI。在设置 S2 中，一个 ESP32-C5 充当合法参与者，因此参与波束成形协议，此时可以以明文形式窃听 BFI。

指标 对于目标帧 $\mathbf{H}_i$ 和合成信道 $\hat{}\mathbf{H}_i$，其中 $\mathbf{H}_i$，$\hat{}\mathbf{H}_i \in \mathbb{C}^{s\times k}$，我们使用归一化均方误差（NMSE）作为距离度量：

$$\mathrm{NMSE}_i(\hat{\mathbf{H}}_i, \mathbf{H}_i) = \frac{\left\|\hat{\mathbf{H}}_i - \mathbf{H}_i\right\|_F^2}{\left\|\mathbf{H}_i\right\|_F^2}, \tag{8}$$

其中，$\|\cdot\|_F$ 表示弗罗贝尼乌斯范数。因此，误差通过相应目标帧的能量进行归一化，使该指标与绝对信道增益无关。为总结一组 N 帧上的结果，我们将每帧的数值转换为分贝并取平均值：

$$\mathrm{NMSE} = \frac{1}{N}\sum_{i=1}^{N} 10\log_{10}\mathrm{NMSE}_i(\hat{\mathbf{H}}_i, \mathbf{H}_i). \tag{9}$$

在对数域中取平均相当于
逐帧数值的几何平均值，对单个异常帧不太敏感。为了说明这些数值的意义，我们使用相同的指标报告噪声底，即静态环境下重复测量之间的归一化均方误差，该指标反映了系统的测量噪声。随后，我们将容差水平 δ_ε 设为所得分布的第99百分位数。

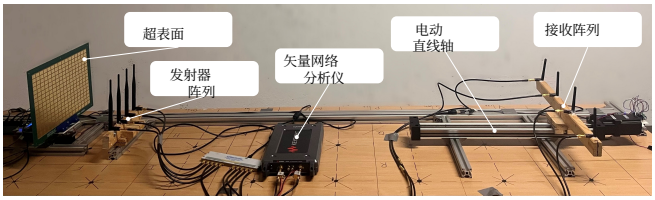


图3：用于测量 4×4MIMO 信道的基于矢量网络分析仪的实验室设置，其中包括对抗性超表面的传播路径。

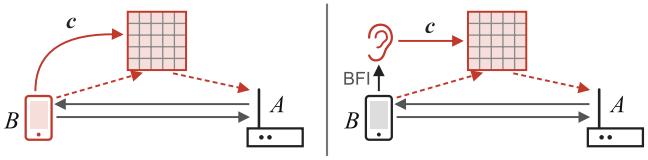


图4：基于观测到的信道知识的攻击者设置。左侧（S1）：攻击者控制 B，其观测到与 A 之间的互易信道，并驱动超表面配置$\mathbf{c}$。右侧（S2）：B 是合法设备，会发射未加密的波束成形信息，被动攻击者可监听该信息以推导$\mathbf{c}$。

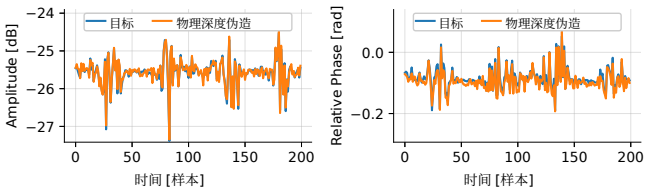
5. 攻击可行性

我们首先在设置 S1 中评估物理深度伪造，在该设置中，攻击者控制发射器 B（见图4）。由于信道具有互易性，B 测量到 A 的信道，因此可以观察每种超表面配置对受害者信道的影响。只要受害者使用其无法控制的设备所发出的信号进行感知，就会出现这种设置，例如，公共 Wi-Fi 网络中的路由器使用未知客户端设备发送的数据包进行感知。任何加入该网络的设备都可以充当 B。我们基于 VNA 的实验室设置模拟了这种场景。

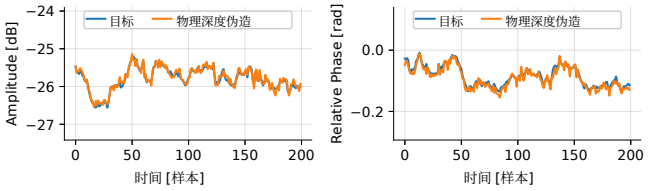
5.1. 使用物理深度伪造进行任意信道合成

在我们的第一个实验中，我们使用实验室感知装置录制了三种目标事件（一个人行走、拍手和坐下），每种事件都重复多次。然后，我们将对抗性超表面置于环境中，使用已知配置的信道测量训练前向代理模型，并为每段录制内容合成一个配置序列。我们在超表面上重放该序列，并将生成的信道与目标进行比较。在此实验中，每种配置合成一个自由度即，一个时间样本中、某一频率下单个空间链路的复信道系数。图 5 展示了随时间变化的幅度和相位。重放的深度伪造信道与原始录制内容高度吻合。深度伪造信道的归一化均方误差介于 -37.3分贝和 -36.3分贝之间，接近于 -41.1分贝，这也是我们装置的可重复性极限。

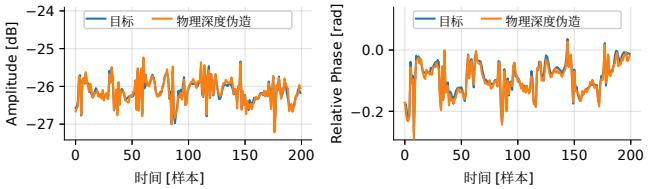
感知系统会跨频率和空间观测信道，这使得超表面合成更加困难：



(a) 行走

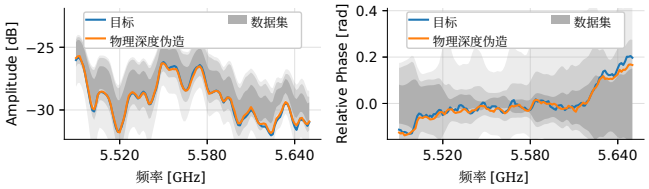


(b) 拍手

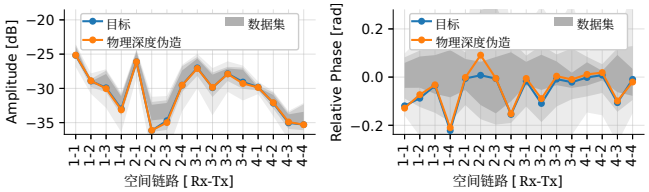


(c) 坐下

图 5：时域物理深度伪造示例，为多种目标感知事件合成信道幅度和相位曲线。



(a) 多个频段

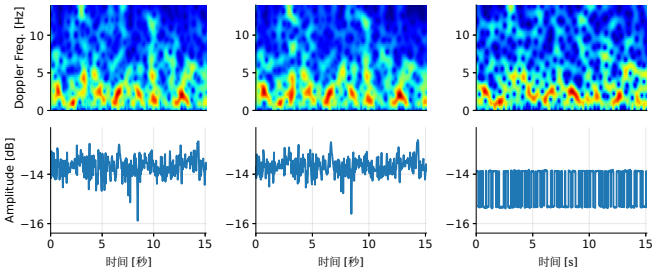


(b) 多条空间链路

图 6：拍手事件某一时刻的频域和空域物理深度伪造示例。

在时域中，每个样本都有自己的配置；而在跨频率和空间的情况下，单一配置必须同时匹配多个自由度。图 6 展示了拍手事件中一个时间样本的合成结果。在这里，合成在单条空间链路的整个 160 MHz 频段上达到 -34.1dB 的归一化均方误差，而在单个频率下覆盖全部 16 条空间链路时达到 -35.2dB 。

以往的超表面攻击操纵的是感知系统提取的特征，而不是原始信道，因此会暴露在检查原始测量值的防御者面前。



(a) 目标　(b) 物理层深度伪造　(c) RISiren [16]

图 7：我们实施的攻击 (b) 与 RISiren 特征操纵 (c) 的比较。RISiren 对多普勒频谱进行欺骗，而物理深度伪造则跟踪原始信道测量。

例如，RISiren [16] 通过在两种超表面配置之间切换来伪造多普勒频谱。图 7 将 RISiren 与针对同一目标事件的物理深度伪造进行了比较。在多普勒域中，这种切换复现了预期的特征，因此 RISiren 会误导基于多普勒的分类器。然而，在原始信道中，这种切换会留下一个自然传播中不会出现的矩形模式。相比之下，物理深度伪造在两个域中都跟随目标。

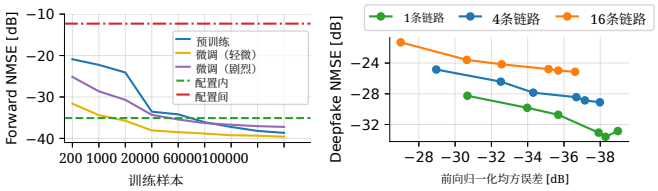
> 超表面能够在时间、频率和空间维度上重现真实事件的原始信道。与特征级操纵不同，物理深度伪造会在原始信道层面匹配目标，因此也能匹配由其导出的每一项特征。

5.2. 代理模型训练开销

攻击的关键步骤之一是将前向代理模型拟合到受害者环境。图 8a 将这一开销表示为达到给定预测误差所需的训练样本数量。上界是配置间离散度，即两个不同配置的信道之间的平均归一化均方误差；达到这一误差时，代理模型的预测效果不比任意另一配置的信道更好。下界是配置内离散度，即同一配置重复测量之间的平均归一化均方误差。这种离散度反映了系统噪声，也标示出攻击者力求达到的误差水平。

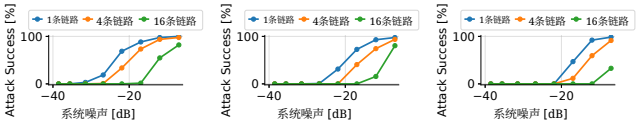
从头开始训练时，代理模型需要30，000 个样本才能达到噪声底，其中误差下降最快的阶段位于1，000 到10，000 个样本之间。微调可以降低训练负担，因此该方法能够很好地扩展到新环境。在一次轻微的物理变化中，我们沿直线轴移动接收天线；经过135，720 个样本预训练的代理模型只需不到1，000 个样本便能达到下界。在一次剧烈的物理变化中，我们移动发射天线，并移动和旋转接收阵列；即使经过微调，该代理模型所需的样本仍少于从头开始训练的代理模型。

由于配置搜索基于代理模型进行，深度伪造质量随代理模型准确度变化：图-



(a) (b)

图 8：深度学习代理模型的性能评估。(a) 代理模型预测信道时的误差与所用训练样本数量之间的关系。(b) 对于拍手活动，在合成不同自由度数量时，前向误差与合成质量之间的关系。



(a) 20 兆赫 (b) 80 兆赫 (c) 160 兆赫

图9：在三个带宽下，1、4 和 16 条空间链路的攻击成功率与受害者系统噪声之间的关系。

ure 8b表明，预测误差较低的代理模型会产生归一化均方误差较低的物理深度伪造。对于给定的代理模型，深度伪造误差随联合合成的自由度数量增加而增大。这种增长反映了配置搜索的误差，因为必须用单一配置匹配更多信道系数。

5.3. 攻击成功率与合成难度

我们评估攻击在受害者测量质量方面的成功情况，并将其量化为系统噪声，即同一信道重复测量之间的归一化均方误差。为模拟噪声更大的受害者感知，我们向所有测量添加复高斯噪声，并将系统噪声从可重复性极限提高至 −7分贝。噪声会影响攻击者的训练数据和目标，也会影响受害者观测到的合成信道。

如果受害者无法将物理深度伪造与真实信道的重复测量区分开来，即其相对于目标的 NMSE 低于容差 δ_ε，则物理深度伪造攻击成功。对于每个噪声水平，我们将 δ_ε 设置为系统噪声的第99百分位数，这对应于假阳性率为 1% 的检测器。图9报告了攻击成功的目标样本比例，带宽分别为 20MHz、80MHz 和 160MHz，空间链路数分别为 1、4 和 16。合成与评估均限制在相应的带宽和空间链路范围内。我们发现，受害者的测量质量是影响攻击成功率的关键因素，因为攻击成功率会随着系统噪声的增加而提高。低噪声受害者几乎不容忍与目标存在任何偏差。随着噪声增大，受害者自身的测量结果变得更加不可重复，容差 δ_ε 增大，攻击成功率趋近于 100%。攻击成功率还会进一步

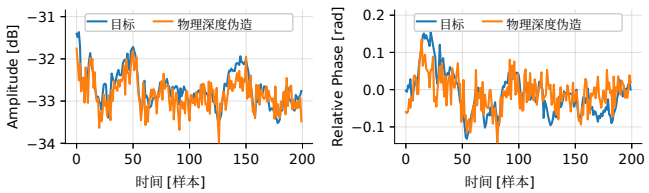


图10：在160MHz下，跨越全部16条空间链路的拍手时域物理深度伪造，其归一化均方误差为−26.7dB。它未达到成功判据，但再现了目标事件的大致结构。

随着联合合成的自由度数量增加而降低。更多的空间链路和更宽的频段都会使成功率曲线向更高噪声水平移动。

> 受害者的测量质量越高，同时合成的自由度越多，攻击难度就越大。

图10展示了一种未达到成功判据、但仍再现了目标事件结构的合成结果。该判据要求物理深度伪造与对真实信道的重复测量无法区分。这是受害者可能提出的最严格判据，因为这样的信道能够挫败任何后续处理。然而，物理事件从不会完全重复，同一事件的两份录制数据之间的差异远远超出系统噪声。因此，我们预计，许多未达到我们判据的合成信道仍会使受害者推断出目标事件。因此，报告的比率只是实际攻击成功率的下界。

6. 基于BFI的感知中的物理深度伪造

Wi-Fi 设备会将信道估计作为波束成形信息报告，以实现波束成形。越来越多的研究将波束成形信息转化为一种感知方式，用于推断例如人体活动 [25],身份 [13], 和生命体征 [51], [52], [53]。本节将展示，基于BFI的感知容易受到物理深度伪造的攻击。该攻击对应图4中的设置 S2：被动攻击者窃听受害者环境中报告的波束成形信息。由于标准规定以明文形式传输波束成形信息， [24], 攻击者观察到的反馈与受害者感知系统收到的反馈相同。通过这些反馈，攻击者观察目标感知事件、了解超表面效应，并合成目标事件的波束成形信息。攻击者不伪造反馈帧，而是利用超表面改变设备 B测量到的信道，使合法报告的波束成形信息承载物理深度伪造。

合成波束成形反馈感知并不直接作用于$\mathbf{H}$ ，而是作用于由其推导出的角度。关键在于，攻击者能够准确观测受害者感知系统所使用的波束成形信息（BFI）。因此，攻击者可以直接在反馈域中学习其超表面的作用。我们训练一个前向代理模型，用于预测给定配置下报告的反馈。第4.2节的配置搜索

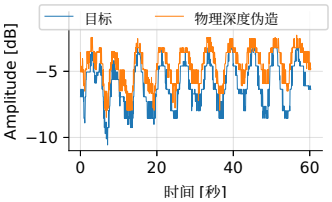


图11：一种用于基于波束成形信息的呼吸频率估计的物理深度伪造。

系统	频率 [次/分钟]	
	目标	深度伪造
LeakyBeam [51]	10.5	9.8
Kanda 等人 [52]	10.0	10.0
BFMSense [53]	10.2	10.2

表2：基于波束成形信息的呼吸频率估计（实际呼吸频率为10.0bpm）。

随后保持不变，其目标是将待欺骗事件的记录波束成形信息（BFI）流作为目标。

突破基于波束成形信息的呼吸感知 我们针对生命体征感知这一经过充分研究的波束成形信息应用评估了攻击。攻击者记录一个以10.0bpm呼吸的人的波束成形信息，合成一组能够重现该反馈的配置序列，并在受害者环境中重放。图11展示了目标信号以及重放的深度伪造内容；后者紧密跟踪由呼吸引起的周期性信号变化。

我们使用三种基于 BFI 的呼吸感知系统评估深度伪造攻击：LeakyBeam[51],Kanda 等人 [52],以及 BFMSense [53]。表2报告了每种估计器针对攻击者的真实目标录音和重放的深度伪造内容所返回的速率，并与10.0bpm的真实速率进行比较。对于所有三种估计器，深度伪造得到的速率与真实值相差不超过0.2bpm ，且与根据真实录音得到的估计值非常接近，尽管攻击者既不了解这些方法，也没有针对其中任何一种方法进行优化。由于深度伪造复现了 BFI 本身，这些方法从中提取的所有特征也都得以复现。

基于 BFI 的感知并不安全：感知系统所使用的明文反馈，也让被动攻击者能够使合法设备报告一场从未发生过的事件的反馈。

7. 受限信道知识下的合成

在设置 S1 和 S2 中，攻击者直接测量超表面对受害者信道的影响。本节讨论图12所示的两种设置，在这两种设置中，攻击者无法直接观测信道，但仍能合成用于物理深度伪造的超表面配置。在设置 S3 中，攻击者构建受害者感知系统及其几何结构的物理复制体，在该复制体上进行训练，然后将超表面移至受害者环境中。在设置 S4 中，攻击者使用放置在受害者环境内的窃听天线。我们的评估表明，即使在受限信道知识的情况下，物理深度伪造仍然可行。

我们在同一受害者感知系统上评估这两种设置：由两个商用现成

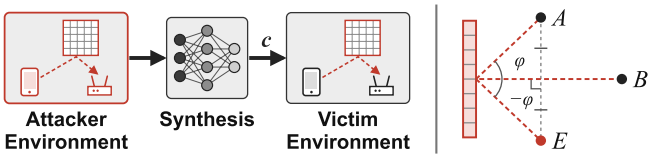


图12：信道知识受限时的攻击者设置。左图（S3）：攻击者在物理复制体中合成配置c，并在受害者环境中重放这些配置。右图（S4）：窃听天线 E 关于穿过 B与超表面的对称轴镜像受害者天线 A 。

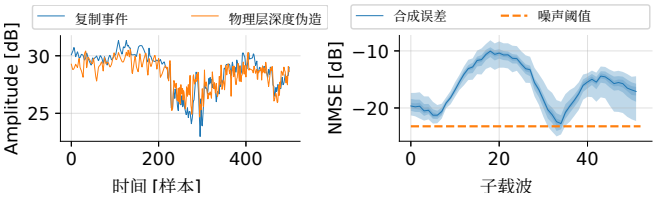
ESP32-C5设备之间构成的单条Wi-Fi链路。受害者在5 GHz 频段、带宽为20兆赫的信道112上提取信道状态信息。按照传统的信道状态信息预处理方法 [10]，我们去除由采样时间和载波频率偏移引起的每个数据包的线性相位。下文中，我们首先使用物理复制体和对称代理评估合成质量，然后在两种设置下攻击活动识别系统。

7.1. 物理复制体

在 S3 中，攻击者检查受害者环境，并将感知系统重建为物理复制体。由于 Wi-Fi感知会在有人存在的空间中复用商品化基础设施，访客可以记录接入点的型号及其安装位置，购买相同的设备，并以相同的几何结构重建设置。在复制体中，攻击者记录自身的感知事件，然后安装其超表面、训练替代模型，并合成这些事件的物理深度伪造，全程无需访问受害者环境。在攻击过程中，攻击者将同一个超表面移入受害者环境，将其安装在相对于感知设备与复制体中相同的位置，并重放合成配置。如果复制体足够相似，那么其中能够合成的任何事件都可以在受害者环境中被欺骗。在相同假设下，该复制体也可以是射线追踪模拟。

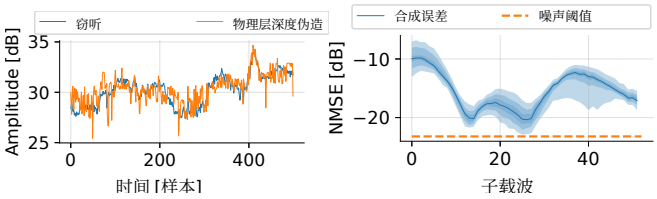
我们在另一间房间中搭建了这样一个复制品，使用与受害者装置相同的设备和天线类型，并保持设备和超表面的几何结构相同。我们仅使用来自复制品的测量数据训练代理模型，然后将超表面移至受害者装置。为了量化相似度，我们对两个装置应用相同的一系列随机超表面配置，并计算由此产生的信道响应之间的皮尔逊相关系数。复制品与受害者装置之间的相关系数达到 0.73，而受害者装置自身重复测量之间的相关系数为 0.96。

图 13a 展示了受害者接收器在单个子载波上挥手的物理深度伪造，以及其合成目标和复制品中的攻击者的录音。物理深度伪造可在两套设备之间传输，并在受害者环境中达到测量级误差。



（a）挥手示例　（b）每个子载波的归一化均方误差

图 13：使用受害者装置的物理复制品实施攻击的结果。（a）成功进行挥手物理深度伪造的示例。（b）合成误差随频率变化。



（a）挥手示例　（b）每个子载波的归一化均方误差

图 14：使用对称窃听天线进行攻击的结果。（a）挥手的物理深度伪造示例。（b）每个子载波的合成误差。

图 13b 展示了全部 150 条录制数据中，物理深度伪造与目标事件之间的逐子载波归一化均方误差。合成误差在整个频段内有所变化。表现最佳的子载波可达到测量级合成质量，而完整 20 兆赫 频段的平均值为 -14.7分贝。

7.2. 对称窃听天线

攻击者在 S4 中并不复制受害者装置，而是将一根窃听天线（代理）与超表面一同放置在受害者环境内，位于受害者天线的镜像位置，如图 12 所示。这在公共空间中是可行的，因为任何人都可以进入这些空间。放置代理所需的物理访问权限与放置超表面相同，因此属于我们的威胁模型范围。在代理处，攻击者记录受害者环境中的目标事件，并测量超表面的影响。然而，代理处的信道不同于受害者观测到的信道。

为了将配置从代理传输至受害者天线，攻击者利用几何结构的对称性。攻击者将超表面限制为波束转向配置，每种配置都会将反射波导向一个方位角和仰角。对于每个时间样本，攻击者选择能够在代理处最佳合成目标的波束，并以反转后的方位角重放该波束，从而在受害者天线处产生相应的效果。

图 14a 显示，在受害者天线处测得的挥手物理深度伪造与代理处录制的目标高度一致。图 14b 中每个子载波的误差与复制品的误差遵循相同模式，在整个频段上的平均归一化均方误差为 -14.5dB 。

7.3. 攻击活动识别

我们使用同一个分类器，对物理复制体（S3）和对称代理（S4）进行端到端评估，合成覆盖整个20兆赫频段。分类器可识别五种活动（拍手、跳跃、坐下、行走和挥手），每种活动都在受害者接收器处录制了 30 次。我们将所有录制数据重采样到相同长度，以确保持续时间不会成为判断线索；随后移除静态路径，并对所得的多普勒频谱图进行分类我们使用小型一维卷积网络进行分类。采用五折交叉验证，每种活动的 20 个录制样本用于训练，4 个用于验证，6 个用于测试。表 3 报告了各折结果的均值和标准差。

分类器在真实录制数据上的准确率达到 92。0% ，这是两种设置的参考值。对于每次攻击，我们对三类输入进行分类。目标是攻击者捕获的、用于复现事件的真实录制数据。物理深度伪造 是在超表面重放合成配置时，于受害者接收器处测得的数据。打乱 深度伪造会在时间上随机打乱样本，从而破坏运动模式，但保留幅度统计。当准确率为 19。3% 和 20。7% 时，该基线与 20% 的随机水平一致。

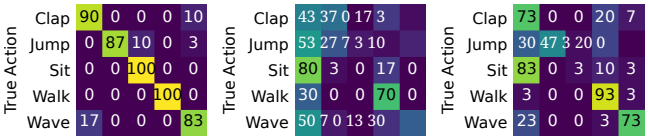
这两种设置的区别在于攻击者如何获取目标。在代理设置中，攻击者录制与受害者相同的事件，目标的准确率达到90。7%，与受害者自己的录制数据相当，因为动作的多普勒特征在代理处得以保留。物理深度伪造达到58。0%，是打乱基线的三倍。在复制设置中，攻击者依赖自己对活动的重复操作，因此目标的准确率仅为46。7%。物理深度伪造保留了其中大部分准确率，达到34。0%。

图15显示，在这两种设置中，被误分类的深度伪造主要集中在拍手和行走上。我们将此归因于合成误差，该误差因数据包而异，因此在多普勒频谱图中表现为宽带噪声。这种噪声会掩盖坐下等缓慢活动；这些活动在较高的多普勒频率下能量很少，并会使分类器倾向于判定为行走，这一点在打乱基线中也有所观察。

> 即使攻击者无法观测受害者信道，物理深度伪造仍可通过使用受害者装置的物理复制体或窃听天线来欺骗活动识别。

8. 环境克隆

我们的最终评估针对表 1 中的设置S5，在该设置下，攻击者一次性获得对受害者环境的初始访问权限。在这次访问期间，攻击者分析环境某一部分的反射行为，例如，



(a) 真实 (b) 复制品（S3） (c) 代理（S4）

图15：针对（a）真实测试录音以及来自（b）物理复制体和（c）对称代理的物理深度伪造，活动分类器的混淆矩阵。

表 3：受害者分类器对真实录音以及复制品（S3）和代理（S4）攻击中的目标、物理深度伪造和时间打乱深度伪造的准确率（五折的均值 ± 标准差）。

场景	输入	准确率 [%]
受害者	真实	92.0 ± 5.6
复制品（S3）	目标	46.7 ± 14.3
	物理深度伪造	34.0 ± 13.4
	打乱	20.7 ± 4.3
窃听（S4）	目标	90.7 ± 2.8
	物理深度伪造	58.0 ± 9.6
	打乱	19.3 ± 2.8

一个物体，但无法获知感知系统的任何信息。因此，S5 不同于第3.1节中攻击者知晓受害者装置几何结构的假设。攻击者不再以感知系统为目标，而是克隆环境的这一部分，如图16所示。在自己的实验室中，攻击者合成一种超表面配置，其响应$\mathbf{X}$ 近似于所分析的反射 $\tilde{}$特征$\mathbf{X}$。返回实施攻击时，攻击者将配置好的超表面放置为目标的替代物，或移动它以模仿目标的运动。

我们使用一根朝向物体的发射天线和一个由七根相邻天线组成的接收阵列来评估这项攻击。由于 S5 中的攻击者不知道感知系统位于何处，因此我们要求克隆体能够同时在全部七根天线上重现物体反射。这样的配置可近似呈现阵列所覆盖角度范围内物体的散射场，而七根天线则代表可能的接收器位置。

我们使用第4.3节中的矢量网络分析仪（VNA），在 5.56 GHz 下测量信道，采用 312.5kHz 的子载波间隔和 20MHz 的带宽，以匹配标准 Wi-Fi。作为目标，我们使用图17所示的三个物体：两张贴有金属胶带、分别呈字母 X（目标1）和 Y（目标2）形状的纸板片，以及一个表面具有不规则结构的全金属物体，即复杂反射体（目标3）。目标1和目标2是一对高度相似的物体。对于每个目标，我们都会记录一次静态参考测量 $\tilde{\mathbf{X}} \in \mathbb{C}^{7\times 57}$。

与之前的设置一样，我们使用超表面对信道影响的替代模型来合成深度伪造

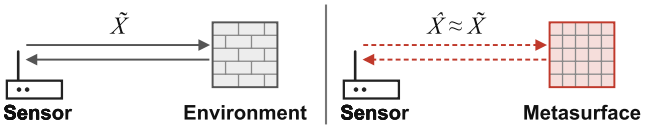


图16：环境克隆（S5）。左：传感器观测到环境一部分的反射，即$\tilde{}$目标事件$\mathbf{X}$。右：攻击者用能够模拟其反射的超表面替换这部分，因此传感器推断出$\hat{}$ $\tilde{}\mathbf{X} \approx \mathbf{X}$。



图17：克隆实验中使用的物体。从左到右依次为目标1–3：贴有金属胶带、形状分别为字母X和Y的纸板片，以及一个表面带有不规则表面结构的全金属物体。

信道。不同于设置 S1–S4 中超表面位于接收器附近的情况，此时超表面占据了其旨在替代的物体所在的位置。对于每个参考 $\tilde{}\mathbf{X}$，我们随后使用替代模型寻找一个超表面 $\hat{}$配置，使其响应$\mathbf{X}$ 与参考响应近似。

$\tilde{}$**结果**图18对比了接收天线和频率维度上的参考信号 $\mathbf{X}$ 与$\hat{}$合成深度伪造信号$\mathbf{X}$ 。对于全部三个目标，深度伪造信号的幅度和相位都与参考信号高度吻合。在所有天线和子载波上，目标1、2和3的深度伪造信号归一化均方误差分别达到-17.2dB、 -19.1dB和 -15.1dB 。这远低于两个不同目标之间最小的归一化均方误差，即目标1和2之间的 -11.2 dB 。因此，每个深度伪造信号与其目标的接近程度，至少比与最相似的其他物体还要高3.9dB 。

位移鲁棒性 先前的实验在单一固定位置评估克隆效果。在实际情况下，攻击者无法将克隆物体放置在合成时使用的确切位置，而感知系统可能会在一系列位置上观测该物体。为测试合成配置在发生此类位移时是否仍然有效，我们使用直线执行器移动物体和超表面，使其经过 89 个位置，相邻位置间隔为 0.5 厘米 。超表面保留在第一个位置合成的配置。在每个位置，我们测量所得响应，并计算其相对于对应目标参考的归一化均方误差。

图19展示了结果，同时给出了随机配置的归一化均方误差以及作为参考的目标间和目标内归一化均方误差均值。在整个范围内，每个深度伪造的结果都低于随机配置和目标间归一化均方误差，即，在任意位置，克隆都比任意两个目标彼此之间更接近其自身目标。所有曲线都以约2.7厘米的周期振荡，即半个波长；与此同时，参考功率也随

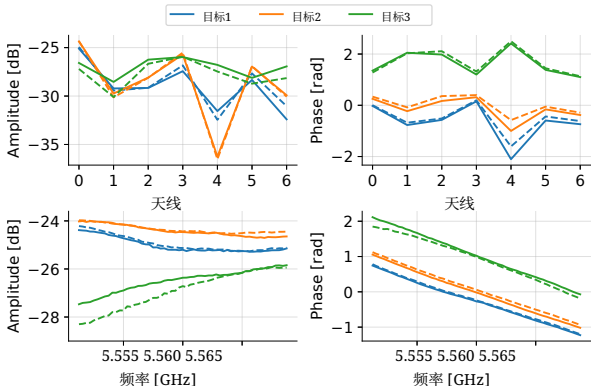


图18：目标1–3的物体克隆结果，比较每个目标的ˆ测量参考信号$\mathbf{X}$（实线）与ˆ合成深度伪造信号$\mathbf{X}$（虚线）。各列分别显示幅度（左）和相位（右）；各行分别显示5.56GHz 时七个接收天线上的变化（上）以及天线0处随频率的变化（下）。

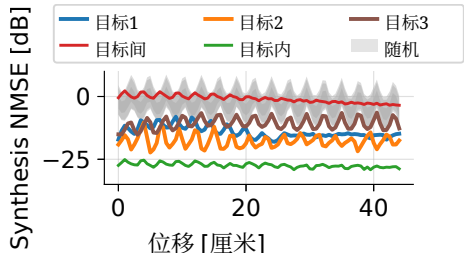


图19：超表面沿直线执行器移动时环境克隆的合成误差。

直接路径与反射路径之间的干涉而振荡，而绝对合成误差保持不变。合成曲线和目标内曲线在整个范围内没有呈现系统性趋势，而目标间归一化均方误差则随位移减小。我们认为，后者是由于物体远离阵列时，静态路径相对于物体反射的主导作用逐渐增强，使各目标彼此越来越相似，而与克隆无关。总体而言，在所有位置上，深度伪造仍然与其目标保持关联。

> 超表面无需了解感知系统，即可克隆物理对象的反射；即使发生位移，该克隆体仍然比其他任何物体更接近其目标。

9. 讨论与未来工作

9.1. 实验设置与结果

为记录训练样本，攻击者让超表面在其配置空间中循环，这会扰动受害者信道，可能引起注意。然而，随机配置看起来更像信道噪声，而非合理的物理事件。攻击者还可以采样一个有序配置序列，其中只有少数元素不同，从而使信道变化保持微小且渐进。在 S3 和 S5 中，攻击者完全在受害者环境之外进行采样。

攻击者将超表面放置在感知设备附近，因为超表面路径会随着其与 A 和 B 之间距离的增加而衰减。更大的孔径可以补偿距离，而攻击者只需诱导出与人体反射体所造成量级相当的信道变化。

9.2. 局限性

我们的攻击假设，在获取信道知识与发起攻击之间，环境保持静态。环境发生变化时，例如设备或人员移动，合成的超表面配置就会过时，因此攻击者必须更新其信道知识。然而，在设置 S1 和 S2 中，攻击者会持续观察受害者信道，并可随时更新其信道知识。此外，攻击者需要物理访问权限，才能将超表面放置在受害者环境中。该攻击还受限于超表面的工作频段，无法在超出这些频段的频率上合成信道。此外，我们仅合成发生在 A 和 B 之间视距（LoS）之外的事件，这与反射事件的感知相匹配。不过，只要超表面能够诱导出所需幅度的信道变化，发生在视距（LoS）内的事件仍然可以被合成。最后，物理深度伪造会重现目标事件的录制内容，而攻击者必须事先获取该录制内容。我们认为，目标测量的精确复制品是该攻击最强的变体，但攻击成功率并不严格要求达到这一程度，因为无线感知必须具有内在的领域不变性，这是由于无线电传播会在不同环境中以不同方式编码所感知的事件。

9.3. 对策

对策可以阻止攻击者获取表 1 中列出的信道知识，使其无法确定合适的超表面配置。通过确保感知合作方诚实，可以缓解 S1。通过在信道探测期间对波束成形信息进行加密或随机预编码 [51] 等保护措施，可以阻止 S2。在 S3 和 S4 中，攻击者会将从复制品或窃听天线获得的信道知识转移到受害者信道。机密传感器配置（例如随机波束成形）可以阻止攻击者复制受害者装置。在 S5 中，攻击者会创建受害者环境的电磁克隆体，因此完全不需要信道知识。这些对策针对特定设置，无法保证能够防范通过其他途径获取信道知识的攻击者。无论设置如何，都有两种通用对策。首先，隐藏物理环境可防止攻击者放置超表面，从而形成一个物理上受限且已知无害的无线信道。其次，受害者可以通过在更多自由度上进行感知，并对各链路进行一致性检查、降低系统噪声，以及将跳频范围扩展到超表面带宽之外，来增加合成难度。

通过跨链路一致性检查、降低系统噪声，以及在超表面的带宽范围之外进行跳频。

9.4. 未来工作

我们独立地为每个时间瞬间合成超表面配置。未来的工作可以合成完整的配置序列，以确保时间平滑性。此外，还可以通过在连续配置之间进行渐进式衰落来对配置序列进行上采样。更进一步地，我们使用单个超表面一次性合成所有自由度。由于随着联合合成的自由度数量增加，合成质量会下降，因此可以由多个超表面分担负载，例如，为每个频率范围或空间链路配备一个专用超表面。最后，物理深度伪造还可以扩展到存在超表面的其他感知领域，例如声学 [54] 和光学 [55] 感知。

10. 相关工作

超表面安全应用 安全领域的文献研究超表面在良性和恶意目的下的应用。超表面已被用于在特征层面操纵无线感知，既包括良性 [15],[56] 应用，也包括恶意 [16], [57] 应用。近期的分析研究表明，超表面可以欺骗单基地雷达的到达角估计 [58] ，以及车载通感一体化网络中的多普勒和离开角估计 [59]。在无线通信领域，超表面已被用于防御或实施干扰攻击 [60], [61], [62], 以及阻挠或协助窃听 [63], [64], [65]。超表面还可以随机化无线信道 [66], ，例如支持对发射机的身份认证 [67] 及其位置验证 [68]。

无线感知的操纵对无线感知的物理操纵最初源于国防部门，旨在欺骗敌对雷达，例如，通过诱饵 [69], 信号重放 [70], 和恶意反射 [71]。如今，无线感知已应用于许多民用场景，因此操纵也成为一种民用威胁。在汽车雷达中，对手可以隐藏真实物体或注入虚假物体，方法包括使用工程反射器 [72], [73]或再辐射信号 [74], [75]。同样，也可以利用特制物体和智能表面操纵对人体和室内环境的感知，以伪造生命体征 [76],绕过入侵检测 [57], [77],模拟跌倒 [16], 并欺骗室内定位 [78]。物理物体还可以欺骗自动驾驶车辆中的相机-激光雷达感知 [79], [80],、X射线成像 [81], 和卫星雷达 [82]。

防御传感器欺骗 传感器通常容易受到操纵后的输入影响 [83], [84], [85], [86],，因为它们默认所捕获的物理信号是真实的。对策通常依赖异常检测，无论是在无线感知 [87], [88] 领域 [89]还是其他领域。然而，这类防御预设攻击会产生可检测的异常 [90], [91]。因此，它们会增加攻击难度，但面对操纵方式与预期物理观察 [92]保持一致的对手时，先天上必然失效。

11. 结论

在本研究中，我们提出了针对无线感知系统的物理深度伪造攻击。我们对该攻击进行了形式化描述，并利用超表面操纵物理感知环境，实现了这一攻击。这类新型攻击能够准确重现真实物理事件的无线信道。从安全角度来看，我们的攻击相当于传统消息级重放攻击在信道层面的对应形式。我们展示了，无论无线感知系统使用何种信号特征进行推断，该攻击都能欺骗它们。我们的研究揭示了无线感知易受物理操纵的脆弱性，因为感知本质上信任其对无线电环境的物理观测，而该环境可能受对手控制。鉴于全球正在努力推进通感一体化，我们认为，安全关键型应用中的无线感知需要采取额外的安全措施。

致谢

我们感谢 NovoFlect GmbH、Simon Tewes、Markus Heinrichs 和 Rainer Kronberger 提供超表面原型。此外，我们感谢 Christian Zenger 作为第一作者的博士导师给予指导，并感谢 Ilyes Lajili 和 Ali Kalife 提供技术支持。本研究部分经费由德国研究基金会（DFG）通过德国卓越战略项目 EXC 2092 CaSa（项目编号：390781972）资助。

参考文献

[1] ETSI, 《ETSI发布首份6G通感一体化用例报告》，2025年，访问日期：2026年11月17日。 [在线]。可访问：https://www.etsi.org/newsroom/press-releases/2520-etsi-publishes-first-report-on-isac-use-cases-for-6g[2] 3GPP, 《一体化感知与通信可行性研究（第19版）》，第三代合作伙伴计划（3GPP），技术报告TR 22.837，2024年，版本19.4.0。[3] ITU-R, 《面向2030年及以后的地面国际移动通信系统未来技术趋势》，国际电信联盟无线电通信部门，瑞士日内瓦，报告ITU-R M.2516-0，2022年11月。[4] Qualcomm, 《无线感知与通信相结合：在6G时代可以期待哪些全新效率与体验？ | Qualcomm》，2024年，访问日期：2026年11月17日。 [在线]。可访问：https://www.qualcomm.com/news/onq/2024/09/when-wireless-sensing-meets-communications-what-new-efficiencies-and-experiences-can-you-expect-in-the-6g-era[5] Huawei, 《一体化感知与通信（ISAC）——从概念到实践》，2022年，访问日期：2026年11月17日。 [在线]。可访问：https://www.huawei.com/en/huaweitech/future-technologies/integrated-sensing-communication-concept-practice[6] C. Meng、C. He、Y. Jiang、X. Fan、D. Wang、L. Wang、J. Ji和Y. Zhang, 《大海捞针：从真实世界海量噪声中追踪5G-A基站数据里的无人机》，载于第*24*届年度移动系统、应用与服务国际会议论文集。英国剑桥：ACM，2026年6月，第14–27页。

[7] 爱立信，《爱立信与Erillisverkot为芬兰当局演示5G通感一体化无人机检测》，新闻稿，2026年9月，访问日期：2026年11月17日。 [在线]。可访问：https://www.ericsson.com/en/news/2026/9/ericsson-erillisverkot-demo-5g-isac-drone-detection[8] AT&T和爱立信，《未来展翅：AT&T与爱立信在AT&T体育场外演示无人机检测》，新闻稿，2026年7月，访问日期：2026年11月17日。 [在线]。可访问：https://about.att.com/story/2026/att-ericsson-drone-detection.html[9] IEEE，《信息技术标准——系统间电信与信息交换——局域网和城域网——特定要求——第11部分：无线局域网介质访问控制（MAC）和物理层（PHY）规范——修正案4：无线局域网感知增强》，IEEE标准协会，标准IEEE Std 802.11bf-2025（对IEEE Std 802.11-2024的修正），2025年。[10] Y. Ma、G. Zhou和S. Wang，《利用信道状态信息进行WiFi感知：综述》，*ACM*计算调查，第52卷，第3期，第1–36页，2019年。[11] Y. Zhu、Z. Xiao、Y. Chen、Z. Li、M. Liu、B. Y. Zhao和H. Zheng，《连Alexa也如此？商用WiFi设备何以变成对抗性运动传感器》，载于第*27*届网络与分布式系统安全研讨会，NDSS '20系列。美国加利福尼亚州圣迭戈：互联网协会，2020年。[12] J. Liu、Y. Wang、Y. Chen、J. Yang、X. Chen和J. Cheng，《利用现成WiFi设备追踪睡眠期间的生命体征》，载于第*16*届*ACM*移动自组网与计算国际研讨会论文集，MobiHoc '15系列。中国杭州：ACM，2015年，第267–276页。[13] J. Todt、F. Morsbach和T. Strufe，《BFId：利用波束成形反馈信息的身份推断攻击》，载于*2025*年*ACM SIGSAC*计算机与通信安全会议论文集。中国台北：ACM，2025年11月，第2399–2413页。[14] F. Dong、F. Liu、Y. Cui、W. Wang、K. Han和Z. Wang，《6G感知网络中的感知即服务：通感一体化资源分配的统一框架》，*IEEE*无线通信汇刊，第22卷，第5期，第3522–3536页，2023年5月。[15] P. Staat、S. Mulzer、S. Roth、V. Moonsamy、M. Heinrichs、R. Kron-berger、A. Sezgin和C. Paar，《IRShield：对抗性物理层无线感知的对策》，载于*2022*年*IEEE*安全与隐私研讨会（*SP*）。美国加利福尼亚州旧金山：IEEE，2022年5月，第1705–1721页。[16] C. Jiang、J. Yang、X. Li、Q. Li、X. Zhang和J. Ren，《RISiren：通过超表面实施无线感知系统攻击》，载于*2024*年*ACM SIGSAC*计算机与通信安全会议论文集。美国犹他州盐湖城：ACM，2024年12月，第3332–3345页。[17] J. Yang、C. Jiang、J. Zhang、X. Li、B. Yang、S. Zhang、X. Meng、Z. Wang、D. Fang和Z. Tang，《Metaegis：通过超表面混淆防御物理层无线感知泄漏》，*ACM*交互式、移动、可穿戴与普适技术论文集，第10卷，第1期，2026年3月。[18] D. Zivanovic、J.-C. Liao、Z. Shaikhanov、H.-T. Chen、C.-C. Chang、S. Addamane、D. M. Mittleman和E. W. Knightly，《Metaheart：利用动态超表面欺骗振动生物特征》，载于*2025*年*IEEE*通信与网络安全会议（*CNS*），2025年，第1–9页。[19] M. I. Skolnik，雷达系统导论，第3版，麦格劳-希尔电气工程丛书。美国马萨诸塞州波士顿、伊利诺伊州伯尔里奇、爱荷华州迪比克：麦格劳-希尔，2001年。[20] Q. Huang、L. Qu、B. Wu和G. Fang，《基于压缩感知的超宽带穿墙成像》，*IEEE*地球科学与遥感汇刊，第48卷，第3期，第1408–1415页，2010年。[21] D. Halperin、W. Hu、A. Sheth和D. Wetherall，《工具发布：利用信道状态信息采集802.11n跟踪数据》，*ACM SIGCOMM*计算机通信评论，第41卷，53，2011 1,第。.

[22] Y. Xie、Z. Li 和 M. Li，《使用商用WiFi进行精确功率延迟剖析》，载于第*21*届移动计算与网络国际年会论文集，MobiCom '15 系列。法国巴黎：ACM，2015年，第53–64页。[23] F. Gringoli、M. Schulz、J. Link 和 M. Hollick，《释放您的CSI：现代Wi-Fi芯片组的信道状态信息提取平台》，载于第*13*届无线网络测试平台、实验评估与特征分析国际研讨会论文集，WiNTECH '19 系列，2019年，第21–28页。[24] 《IEEE信息技术标准——电信与系统间信息交换——局域网和城域网——特定要求——第11部分：无线局域网媒体访问控制（MAC）和物理层（PHY）规范》。 [在线]。网址：https://ieeexplore.ieee.org/docu ment/7786995/

[25] K. Bao, H. Li, Y. Zhang, Z. Meneghello, and F. Restuccia, "Densesense: 重新思考基于 MU-MIMO Wi-Fi 波束成形反馈的无线感知", *Computer Networks*，第 258，111020, 卷，第 页。在线。可访问：2025 [ ] https://www.sciencedirect.com/science/ar ticle/pii/S1389128624008521

[26] S.Yousefi,H.Narui, S.Dayal,S.Ermon,andS. Valaee,"一项关于使用 WiFi信道状态信息进行行为识别的调查"，*IEEE Communications Magazine*，第 55 卷，第 10 期，第 98–104 页，2017年。

[27] Y. Wang, K. Wu, and L. M. Ni, "Wifall: 基于无线网络的无设备跌倒检测，" *IEEE*移动计算汇刊，第*16*卷，第2期，第581–594页，2016年。

[28] Q. Xu, B. Wang, F. Zhang, D. S. Regani, F. Wang, and K. R. Liu, "智能汽车中的无线人工智能：汽车能有多智能？"*IEEE Access*，第8卷，第55 091–55112页，2020年。

[29] J. Lv、D. Man、W. Yang、X. Du 和 M. Yu，"利用 PHY 层信息实现基于 WLAN 的鲁棒室内入侵检测"，*IEEE Access*，第 6 卷，第 30 117–30127 页，2017年。

[30] F. Liu、Y. Cui、C. Masouros、J. Xu、T. X. Han、Y. C. Eldar 和 S. Buzzi，"通感一体化：迈向面向 6G 及以后的双功能无线网络"，*IEEE*通信领域精选期刊，第 40 卷，第 6 期，第 1728–1767 页，2022 年 6 月。[31] M. Z. Chowdhury、Md. Shahjalal、S. Ahmed 和 Y. M. Jang，"6G无线通信系统：应用、需求、技术、挑战与研究方向"，*IEEE*通信学会开放期刊，第 1 卷，第 957–975 页，2020 年。[32] Comcast，"Comcast推出Xfinity Shield，重新定义智能家庭防护"，新闻稿，2026 年，访问日期：2026年11月17日。[在线]。网址：https://corporate.comcast.com/press/releases/comcast-introduces-intelligent-home-protection-a-new-category-for-the-connected-home-with-xfinity-shield[33] N. Kaina、M. Dupr´e、G. Lerosey 和 M. Fink，"利用二进制可调超表面塑造混响介质中的复杂微波场"，《科学报告》，第 4 卷，第 1 期，6693 页，2014 年。[34] H. Tataria、M. Shafi、A. F. Molisch、M. Dohler、H. Sj¨oland 和 F. Tufvesson，"6G无线系统：愿景、需求、挑战、见解与机遇"，《*IEEE*会刊》，第 109 卷，第 7 期，第 1166–1199 页，2021 年 7 月。[35] Greenerwave，"Greenerwave"，2025年，访问日期：2026年11月17日。[在线]。网址：https://www.greenerwave.com/[36] NovoFlect，"NovoFlect — 可重构智能表面"，2025年，访问日期：2026年11月17日。[在线]。网址：https://novoflect.de/[37] W. Jiang、B. Han、M. A. Habibi 和 H. D. Schotten，"迈向6G之路：综合综述"，*IEEE*通信学会开放期刊，第 2 卷，第 334–366 页，2021 年。[38] C. Liaskos、S. Nie、A. Tsioliaridou、A. Pitsillides、S. Ioannidis 和 I. Akyildiz，"面向下一代无线系统、借助可编程室内无线环境实现高容量与安全性的新型通信范式"，《自组网》，第 87 卷，第 1–16 页，2019 年 5 月。

[39] Y. Liu、X. Liu、X. Mu、T. Hou、J. Xu、M. D. Renzo 和 N. Al-Dhahir，“可重构智能表面：原理与机遇”，*IEEE Communications Surveys & Tutorials*，第 23 卷，第 3 期，第 1546–1577 页，2021 年。[40] E. Basar、M. Di Renzo、J. D. Rosny、M. Debbah、M. Alouini 和 R. Zhang，“通过可重构智能表面实现无线通信”，*IEEE Access*，第 7 卷，第 116 753–116 773 页，2019 年。[41] B. Rana、S.-S. Cho 和 I.-P. Hong，“可重构智能表面硬件综述”，*IEEE Access*，第 11 卷，第 29 614–29 634 页，2023 年。[42] B. Feng、J. Gao、Y. Wu、W. Zhang、X.-G. Xia 和 C. Xiao，“可重构智能表面辅助网络中的优化技术”，*IEEE Wireless Communications*，第 28 卷，第 93 页，2021 年，第 6 期，第 87 页。

[43] Y. 邹, Cheng龙, D. Niyato, 和 W. 崔, “面向基于智能反射面辅助无线网络的鲁棒波束成形优化”，*IEEE*认知通信与网络汇刊，第 8 卷第 2 期，第 856–870 页，2021 年。

[44] S. Tewes、M. Heinrichs、R. Kronberger 和 A. Sezgin，《基于IRS的呼吸追踪与共址商用WiFi收发器》，*IEEE*《物联网期刊》，第 10 卷第 8 期，第 6870–6886 页，2022 年。[45] X. Pei、H. Yin、L. Tan、L. Cao、Z. Li、K. Wang、K. Zhang 和 E. Bj¨ornson，《RIS 辅助无线通信：原型开发、自适应波束成形与室内外现场试验》，*IEEE*通信汇刊，第 69 卷，2021 年，第 8640、8627 页。

[46] T. Kim, J. Choi, H. Noh, H. Junyeup, 和 J. Choi，“I-robot：在人形机器人中识别机器人与人体运动”，载于*Trustworthy AI for Good (AI4GOOD)*研讨会 @ *ICML 2026*，2026年。

[47] A. Kurakin、I. Goodfellow 和 S. Bengio，“物理世界中的对抗样本”，arXiv:1607.02533，2016年。[48] K. Eykholt、I. Evtimov、E. Fernandes、B. Li、A. Rahmati、C. Xiao、A. Prakash、T. Kohno 和 D. Song，“针对深度学习视觉分类的稳健物理世界攻击”，载于*2018 IEEE/CVF*计算机视觉与模式识别会议。美国犹他州盐湖城：IEEE，2018年6月，第1625–1634页。[49] A. Kumar、A. Raghunathan、R. Jones、T. Ma 和 P. Liang，“微调可能扭曲预训练特征，导致分布外表现不佳”，载于国际学习表征会议（*ICLR*），2022年。[50] M. Heinrichs、Aydin Sezgin 和 R. Kronberger，“适用于5 GHz WiFi频段的开源可重配置智能表面”，载于*2023*年国际天线与传播研讨会（*ISAP*）。马来西亚吉隆坡：IEEE，2023年10月。[51] R. Xiao、X. Chen、Y. He、J. Han 和 J. Han，“借我你的波束：WiFi明文波束成形反馈的隐私影响”，载于*2025*年网络与分布式系统安全研讨会论文集。美国加利福尼亚州圣迭戈：互联网协会，2025年。[52] T. Kanda、S. Kondo、H. Shimomura、T. Sato、H. Awano 和 K. Yamamoto，“基于波束成形反馈的呼吸与心率估计：迈向固件无关的WiFi感知”，*IEEE Access*，第13卷，第146 008–146 019页，2025年。[53] E. Yi、D. Wu、J. Xiong、F. Zhang、K. Niu、W. Li 和 D. Zhang，“BFMSense：利用波束成形反馈矩阵进行 WiFi 感知 ”，载于第*21*届*USENIX*网络系统设计与实现研讨会（*NSDI 24*）。美国加利福尼亚州圣克拉拉：USENIX协会，2024年4月，第1697–1712页。[54] Z. Tian、C. Shen、J. Li、E. Reit、Y. Gu、H. Fu、S. A. Cummer 和 T. J. Huang，“可编程声学超表面”，《先进功能材料》，第29卷，第13期，文章编号1808489，2019年3月。[55] A. M. Shaltout、V. M. Shalaev 和 M. L. Brongersma，“利用有源超表面进行时空光调控”，《科学》，第364卷，第6441期，文章编号eaat3100，2019年5月。

[56] X. Wang and C. Caloz，“基于时变超表面的扩频选择性伪装”，*IEEE* 天线与传播学报，第 69 卷，第 1 期，第 286–295 页，2021 年 1 月。[57] Y. Zhou, C. Li, H. Chen, and Q. Zhang，“RIStealth：利用可重构智能表面实施针对基于 WiFi 的入侵检测的实用隐蔽物理层攻击”，载于第 *21* 届 *ACM* 嵌入式网络传感器系统会议论文集，系列：SenSys ’23。土耳其伊斯坦布尔：ACM，2023 年，第 195–208 页。[58] I. Gavras, G. T. F. de Abreu, and G. C. Alexandropoulos，“基于 RIS 的对手感知欺骗：最大化 CRB 的设计与诱骗分析”，arXiv:2601.17320，2026 年 1 月。[59] T. Shui, P.-H. Chou, W. Saad, and M. Chen，“通感一体化（ISAC）中基于 RIS 的欺骗分析与检测”，arXiv:2508.18100，2025 年 8 月。[60] H. Yang, Z. Xiong, J. Zhao, D. Niyato, Q. Wu, H. V. Poor, and M. Tornatore，“智能反射表面辅助抗干扰通信：一种快速强化学习方法”，*IEEE* 无线通信学报，第 20 卷，第 3 期，第 1963–1974 页，2021 年。[61] P. Staat, H. Elders-Boll, M. Heinrichs, C. Zenger, and C. Paar，“镜子，墙上的镜子：基于快速软件控制表面的无线环境重构攻击”，载于*2022* 年 *ACM* 亚洲计算机与通信安全会议论文集，系列：ASIA CCS ’22。日本长崎：ACM，2022 年 5 月，第 208–221 页。[62] P. Mackensen, P. Staat, S. Roth, A. Sezgin, C. Paar, and V. Moonsamy，“利用可重构智能表面进行空间域无线干扰”，载于*2025*年网络与分布式系统安全研讨会论文集。美国加利福尼亚州圣迭戈：互联网协会，2025 年。[63] M. Wei, H. Zhao, V. Galdi, L. Li, and T. J. Cui，“超表面赋能的物理层智能无线攻击”，*Nature Electronics*，第 6 卷，第 8 期，第 610–618 页，2023 年 8 月。[64] X. Li, C. Feng, F. Song, C. Jiang, Y. Zhang, K. Li, X. Zhang, and X. Chen，“Protego：通过可编程超表面保障无线通信安全”，载于第 *28* 届移动计算与网络国际年会论文集，系列：MobiCom ’22。澳大利亚新南威尔士州悉尼：ACM，2022 年 10 月，第 55–68 页。[65] Z. Shaikhanov, F. Hassan, H. Guerboukha, D. Mittleman, and E. Knightly，“超表面中间人攻击：从理论到实验”，载于第 *15* 届 *ACM* 无线与移动网络安全和隐私会议论文集，系列：WiSec ’22。美国德克萨斯州圣安东尼奥：ACM，2022 年 5 月，第 257–267 页。[66] P. Staat, H. Elders-Boll, M. Heinrichs, R. Kronberger, C. Zenger, and C. Paar，“智能反射表面辅助低熵环境无线密钥生成”，载于*2021 IEEE* 第 *32* 届个人、室内与移动无线电通信国际研讨会（*PIMRC*），2021 年 9 月，第 745–751 页。[67] S. Tomasin, H. Zhang, A. Chorti, and H. V. Poor，“部分可控信道上的质询-响应物理层身份认证”，*IEEE Communications Magazine*，第 60 卷，第 12 期，第 138–144 页，2022 年 12 月。[68] Y. Zhang, J. Li, Y. Wang, D. Han, Y. Zhang, A. Shekhawat, and G. Trichopoulos，“RIS-CLA：利用可重构智能表面复兴基于 CSI 的连续位置身份认证”，载于*2026 IEEE*安全与隐私研讨会（*SP*）。美国加利福尼亚州旧金山：IEEE 计算机学会，2026 年 5 月，第 2358–2373 页。[69] H. Griffiths，“诺曼底登陆日欺骗行动 TAXABLE 和 GLIMMER”，*IEEE* 航空航天与电子系统杂志，第 30 卷，第 3 期，第 12–20 页，2015 年 3 月。[70] S. Roome，“数字射频存储器”，*Electronics & Communication Engineering Journal*，第 2 卷，第 4 期，第 147–153 页，1990 年 8 月。[71] V. Kozlov, D. Vovchuk, and P. Ginzburg，“利用时变散射体进行雷达距离欺骗”，*IEEE* 天线与传播学报，第 71 卷，第 5 期，第 4486–4491 页，2023 年 5 月。

[72] X. Chen, Z. Li, B. Chen, Y. Zhu, C. X. Lu, Z. Peng, F. Lin, W. Xu, K. Ren, and C. Qiao, “MetaWave：利用超材料增强标签攻击毫米波感知”，载于第*30*届年度网络与分布式系统安全研讨会，系列：NDSS ’23。美国加利福尼亚州圣迭戈：互联网协会，2023 年。[73] Y. Zhu, C. Miao, H. Xue, Z. Li, Y. Yu, W. Xu, L. Su, and C. Qiao, “TileMask：一种针对自动驾驶中毫米波雷达目标检测的被动反射攻击”，载于*2023*年*ACM SIGSAC*计算机与通信安全会议论文集，系列：CCS ’23。美国纽约州纽约：美国计算机协会，2023 年，第 1317–1331 页。[74] R. Reddy Vennam, I. K. Jain, K. Bansal, J. Orozco, P. Shukla, A. Ranganathan, and D. Bharadia, “mmSpoof：利用反射阵列对汽车毫米波雷达实施韧性欺骗”，载于*2023*年*IEEE*安全与隐私研讨会（*SP*）。美国加利福尼亚州旧金山：IEEE，2023 年 5 月，第 1807–1821 页。[75] R. Komissarov, S. Vaisman, and A. Wool, “针对车载调频连续波雷达的欺骗攻击”，《密码工程杂志》，第 13 卷，第 4 期，第 473–484 页，2023 年。[76] X. Guo, Y. Wei, Y. Ge, Y. Xie, Y. Wang, J. Cheng, and Y. Chen, “通过执行器诱发的模仿攻击毫米波胸部振动感知”，载于*IEEE INFOCOM 2026 - IEEE*计算机通信会议。日本东京：IEEE，2026 年 5 月，第 1–10 页。[77] A. Akinwale, W. Gao, J. Li, A. Li, X. Shang, Y. Zhang, and D. Han, “基于毫米波的物理入侵检测中的实用反射操控”，载于*2025*年*IEEE*通信与网络安全会议（*CNS*）。法国阿维尼翁：IEEE，2025 年 9 月，第 1–9 页。[78] X. Han, J. Huang, M. Li, C. Cai, and T. Zheng, “Loki：通过可微分物体放置对无线室内定位实施物理世界对抗攻击”，*IEEE*《信息取证与安全汇刊》，第 20 卷，第 13 386–13 400 页，2025 年。[79] Y. Cao, N. Wang, C. Xiao, D. Yang, J. Fang, R. Yang, Q. A. Chen, M. Liu, and B. Li, “摄像头和激光雷达均无法察觉：物理世界攻击下自动驾驶多传感器融合感知的安全性”，载于*2021*年*IEEE*安全与隐私研讨会（*SP*）。美国加利福尼亚州旧金山：IEEE，2021 年，第 176–194 页。[80] Y. Zhu, C. Miao, H. Xue, Y. Yu, L. Su, and C. Qiao, “针对自动驾驶中多传感器融合的恶意攻击”，载于第*30*届移动计算与网络国际年会论文集，系列：ACM MobiCom ’24。美国华盛顿州华盛顿：ACM，2024 年，第 436–451 页。[81] A. Liu, J. Guo, J. Wang, S. Liang, R. Tao, W. Zhou, C. Liu, X. Liu, and D. Tao, “X-Adv：针对X射线的物理对抗物体攻击”，载于第*32*届*USENIX*安全研讨会论文集。美国加利福尼亚州阿纳海姆：USENIX协会，2023 年。[82] H. Li, Z. Li, K. Liu, K. Xu, C. Luo, Y. Lv, and Y. Deng, “宽带信息超表面辅助合成孔径雷达目标干扰系统”，《遥感》，第 16 卷，第 9 期，第 1499 页，2024 年 4 月。[83] K. Fu and W. Xu, “信任传感器物理特性的风险”，*Communications of the ACM*，第 61 卷，第 2 期，第 20–23 页，2018 年 1 月。[84] G. Zhang, C. Yan, X. Ji, T. Zhang, T. Zhang, and W. Xu, “DolphinAttack：不可听语音命令”，载于*2017*年*ACM SIGSAC*计算机与通信安全会议论文集。美国得克萨斯州达拉斯：ACM，2017 年 10 月，第 103–117 页。[85] D. F. Kune, J. Backes, S. S. Clark, D. Kramer, M. Reynolds, K. Fu, Y. Kim, and W. Xu, “Ghost Talk：缓解针对模拟传感器的 EMI 信号注入攻击”，载于*2013*年*IEEE*安全与隐私研讨会，美国加利福尼亚州伯克利，2013 年，第 145–159 页。[86] Y. Son, H. Shin, D. Kim, Y. Park, J. Noh, K. Choi, J. Choi, and Y. Kim, “利用陀螺仪传感器上的蓄意声音噪声摇晃无人机”，载于第*24*届*USENIX*安全研讨会论文集，系列：SEC’15。美国：USENIX协会，2015 年，第 881–896 页。[87] R. Geng、D. Zhang、Y. Li、J. Wang、Q. Liang、R. Zhang、X. Xu、Z. Lu、Y. Hu 和 Y. Chen，“mmGuard：针对毫米波雷达感知中物理对抗攻击的对策”，*IEEE*《信息取证与安全汇刊》，第 21 卷，第 5568–5582 页，2026 年。[88] S. Roth 和 A. Sezgin，“感知安全的异常检测”，*IEEE*《物联网期刊》，2026 年。[89] J. Park、R. Ivanov、J. Weimer、M. Pajic 和 I. Lee，“存在瞬态故障时的传感器攻击检测”，载于 *ACM/IEEE*第六届信息物理系统国际会议论文集。美国华盛顿州西雅图：ACM，2015 年 4 月，第 1–10 页。[90] C. Yan、H. Shin、C. Bolton、W. Xu、Y. Kim 和 K. Fu，“SoK：形式化模拟传感器安全的极简方法”，载于 *2020 IEEE*安全与隐私研讨会（*SP*）。美国加利福尼亚州旧金山：IEEE，2020 年 5 月，第 233–248 页。[91] J. Giraldo、D. Urbina、A. Cardenas、J. Valente、M. Faisal、J. Ruths、N. O. Tippenhauer、H. Sandberg 和 R. Candell，“信息物理系统中基于物理的攻击检测综述”，*ACM*计算调查，第 51 卷，第 4 期，第 1–36 页，2019 年 7 月。[92] D. I. Urbina、J. A. Giraldo、A. A. Cardenas、N. O. Tippenhauer、J. Valente、M. Faisal、J. Ruths、R. Candell 和 H. Sandberg，“限制隐蔽攻击对工业控制系统的影响”，载于*2016 ACM SIGSAC*计算机与通信安全会议论文集。奥地利维也纳：ACM，2016 年 10 月，第 1092–1105 页。[93] D. P. Kingma 和 J. Ba，“Adam：一种随机优化方法”，载于 国际学习表征会议（*ICLR*），2015 年。

附录A。前向代理模型与配置搜索

代理模型架构 前向代理模型 $\hat{}\mathbf{H}(\mathbf{c})$（第4节）将 $L = 256$ 二元单元状态映射到信道 $\hat{}\mathbf{H}(\mathbf{c}) \in \mathbb{C}^{s\times k}$。其预测结果是两个并行分支之和。线性分支将单元状态直接映射到信道。卷积分支将配置视为 16×16 孔径图像。 3×3卷积首先将单通道图像映射为256 个特征信道。随后，六个残差块对得到的特征图进行处理，每个残差块都使用两个 3×3 卷积，并结合组归一化和 GELU。每条空间链路都分配有一个经过学习的查询向量，该向量在256 个空间特征向量上进行注意力计算，为每条链路生成一个向量。共享的解码器网络通过 GELU 激活函数，将每个链路向量映射为该$\hat{}$链路在$\mathbf{H}(\mathbf{c})$中的行，即其跨 k 子载波的信道。该架构包含约14.7百万个参数。

代理模型训练 我们对每个 $2 \times s \times k$实值输出进行标准化，并使用 Adam [93] 最小化预测信道与测量信道之间的均方误差。训练采用 3×10^{-4} 的学习率，每当验证损失趋于平稳时，学习率便减半。为微调以适应新环境，我们遵循 [49]：首先拟合一个线性读出层，将现有代理模型的输出映射到新环境，然后以 10^{-5}的学习率对整个架构进行微调。

配置搜索 我们使用基于梯度的配置搜索（第4.2节），进行32 次重启；每次重启都以0.2的学习率运行 Adam1000 次迭代。其中一次重启采用热启动，即从 $[0,1]^{256}$ 中的一个配置开始，我们

先在代理模型的线性分支上优化400 步，然后映射到对数几率；另外31 次重启则从随机对数几率 $\mathbf{z} \sim \mathcal{N}(\mathbf{0},\ \mathbf{I}_{256})$开始。每次迭代中，我们将 Sigmoid 输出二值化，以得到阈值ˆ为0.5的可实现配置，在该配置上计算前向代理模型 $\mathbf{H}(\mathbf{c})$ ，然后使用直通估计器将预测误差反向传播至对数几率。每次重启都会保留过程中遇到的最佳配置。

对于每帧的32 次重启，我们保留误差最低的4 个配置，并分两步对其进行优化。首先，对每个配置执行单元素翻转，依次测试256个元素，并接受任何能降低误差的翻转。我们最多遍历256 个元素4 次；一旦某次遍历没有接受任何位翻转，就提前停止。接下来，我们对4 个配置中最优的那个执行成对翻转，即同时翻转两个元素。我们不测试所有元素对，而是将每对元素的信道变化预测为其各自单次翻转所致变化之和，并仅评估预测误差最低的128 对元素。我们最多重复成对翻转步骤60 次；一旦单元素或成对翻转都无法降低误差，就停止。单元素和成对翻转结合起来，可将深度伪造误差降低零点几分贝。

在单台 NVIDIA H100 GPU 上，我们在一个批次中设计 50 帧，每帧重启 32 次，每帧耗时约3.5s 。攻击者可以在将配置序列重放到硬件上之前，先离线设计该序列。