arXiv:2602.02249v1 [cs.NI] 2 Feb 2026

评估用于附近智能设备间自组织通信的声学数据传输方案

Florentin Putz，PHILIPPFORTMANN，JANFRANK，CHRISTOPHHAUGWITZ，Mario Kupnik，以及Matthias Hollick，达姆施塔特工业大学，德国

声学数据传输借助智能手机和物联网设备中无处不在的扬声器与麦克风，为蓝牙和近场通信提供了一种极具吸引力的替代方案。然而，该领域的大多数研究依赖于仿真或有限的设备端测试，因此所提出方案在真实环境中的可靠性难以评估。我们系统回顾了31项面向商用设备的声学通信研究，发现没有任何一项提供可获取的源代码。在联系作者并重新实现三种有前景的方案后，我们搭建了一个包含八种代表性声学通信系统的测试平台。基于在真实室内环境和消声室中进行的超过11000次智能手机传输，我们提供了一套系统且可重复的方法，用于评估这些方案在真实环境条件下的可靠性与泛化能力。我们的结果表明，许多现有方案在实际使用中面临挑战，主要原因是室内严重的多径传播，以及不同设备型号之间音频特性的差异。为支持未来研究并促进更稳健的评估，我们发布了这些重新实现，以及首个真实环境声学传输综合数据集。总体而言，我们的发现凸显了严格设备端测试的重要性，并强调需要稳健的设计策略，以弥合仿真结果与可靠物联网部署之间的差距。

CCS 概念：• **网络** → **移动自组网络**；短距离网络；网络可靠性；**网络性能评估**；• **以人为中心的计算** → **泛在与移动设备**；**智能手机**；• **硬件** →信号处理系统。

附加关键词和短语：声学通信；超声波通信；近超声；基于音频的通信；声波数据传输；可复现性；概念性复现；泛化能力；数据集；开放数据；开源；音频录音；物联网；普适计算；实验测试平台；系统性文献研究；重新实现；性能评估；设备间通信；智能手机间通信；短距离；自组；互操作性；自发；Nearby；邻近性

ACM参考文献格式：
Florentin Putz、Philipp Fortmann、Jan Frank、Christoph Haugwitz、Mario Kupnik 和 Matthias Hollick。2026。评估附近智能设备间自组织通信的声学数据传输方案。*ACM*物联网汇刊 7，1，第8篇文章（2026年2月），31 页。https://doi.org/10.1145/3779439

1 引言

智能手机和物联网 (IoT) 设备可以通过内置扬声器和麦克风利用声波进行通信 [39]。鉴于智能手机、智能手表、智能音箱和家庭自动化系统等设备普遍配备音频输入输出，声学数据传输有望成为蓝牙和近场通信 (NFC) 的实用替代方案。在过去十年中，人们越来越关注将声学通信用作无线短距离自组信道，以支持普适计算和移动计算的应用场景，例如移动支付 [68]，上下文感知计算 [40, 41]，或基于位置的服务 [45, 71]。其受限于位置的特性也很适合普适计算场景，在这些场景中，用户会自然而然地

作者联系信息：Florentin Putz，fputz@seemoo.de；PhilippFortmann，pfortmann@seemoo.de；Jan Frank，jfrank@seemoo.de；Christoph Haugwitz，christoph.haugwitz@tu-darmstadt.de；Mario Kupnik，mario.kupnik@tu-darmstadt. de；Matthias Hollick，mhollick@seemoo.de，达姆施塔特工业大学，德国达姆施塔特。

©2026 版权所有者/作者所有。出版权已授权给 ACM。本文为作者版本，仅供个人使用。禁止再分发。最终版本记录发表于 *ACM* 物联网汇刊，https://doi.org/10.1145/3779439。

让设备彼此靠近，以便它们能够“交谈” [11], ，这与用户的直觉行为相符 [6, 25, 26]。配备扬声器和麦克风的低成本物联网硬件广泛普及，甚至促成了商业应用，例如智能音箱制造商 Sonos 利用声学通信实现空间感知和设备配对 [27]。

尽管提出声学数据传输系统的研究日益增多 [4, 13, 14, 33,38, 70], ，但许多系统仅通过仿真或在非常有限的真实环境条件下进行测试。然而，真实设备测试至关重要，因为它可以考量硬件缺陷、用户操作、真实的信号传播以及背景噪声等因素——这些因素众所周知，很难精确建模。此外，声学信道与无线电信道有很大差异：声波的传播速度比电磁波慢 87,000 多倍，因此会产生数十毫秒量级的时延扩展 [33, 59]。声学信号还会受到变化极大的环境噪声影响；而且由于大多数智能手机和物联网设备的采样率较低，带宽最多仅为 22 千赫，因此高吞吐量传输颇具挑战。现有评估通常聚焦于单个设备或场景，因此难以了解某种方案能否推广到多样化的真实环境中。

与此同时，学术研究与商业实践之间存在明显差距。智能设备上的声学数据传输商业产品通常只能达到 10–100 比特/秒 [9, 13, 27] ，偶尔可达 200 比特/秒 [62], ，并优先考虑在嘈杂环境中的稳健性和可靠性。相比之下，研究人员通常报告高得多的吞吐量，可达 500–10,000 比特/秒[4, 14, 70] ，甚至更高 [69]。这种差异引发了一个根本问题：这些系统能否在真实场景中保持可靠性？还是说，硬件局限、用户行为或环境干扰等额外约束最终会迫使实用系统以较低的数据速率运行？

解决这些问题的一大障碍是，该领域的出版物往往不提供源代码，因而限制了独立验证，也妨碍了后续工作。本文通过考察具有代表性的声学数据传输方案在各种实际场景中的泛化能力，初步着手整合现有研究成果。具体而言，我们探讨以下研究问题：

研究问题1： “获取或重新实现该领域提出的系统有多困难？”
研究问题2： “这些方案在自组织智能手机通信用例中的泛化能力如何？”**研究问题3：** “哪些实际挑战会影响智能设备之间的近距离声学数据传输？”

1.1 贡献

本文首次对其他研究人员提出的**声学数据通信方案的泛化能力进行独立评估，**并在真实场景中对其进行比较。我们的贡献如下：

- **系统性文献研究**。 我们开展了一项系统综述，并识别出31篇提出面向商用现成（COTS）设备的声学数据传输系统的出版物（第4节）。遗憾的是，这些研究均未提供实现源代码。
- **代表性方案的获取与重新实现**。我们联系了作者以获取软件实现。仅有三位作者提供了可运行的代码，一位提供了无法运行的实现，其余作者要么未回复，要么已无法访问。为了更好地理解声学通信研究的泛化能力，我们还重新实现了三种有前景的方案（第5节）。
- **泛化能力评估**。我们总共评估了八种声学通信方案的实际可靠性 [4, 13, 14, 33, 38, 70], ，其中包括广受欢迎的

开源项目 ggwave [12] (第 6 节)。我们的测试平台重点考察了多种智能手机型号，将其作为物联网设备中一种实用且易于获取的子集；不过，相关见解也适用于任何具备相当音频输入和输出能力的设备。我们通过在各种物联网应用场景中评估这些方案，并分析设备距离、型号、用户操作和背景噪声等因素如何影响通信可靠性（以传输错误率衡量），扩展了先前的研究。

- **测量数据集**。我们公开了重新实现的方案和分析代码，并同时发布该领域首个数据集——包含 11900 段声学数据传输的真实录音，以支持透明、可复现的研究 [51]。

1.2 范围

本研究从研究人员或开发者希望将这项技术集成到智能设备中，以便在近距离交换数据的角度，分析声学数据传输方案，为物联网研究社区提供支持。从这一角度来看，声学通信的采用应像蓝牙或近场通信一样无缝，无需掌握底层无线标准的专业知识。因此，我们在多种智能手机型号上评估代表性方案，因为智能手机是最常见、用途最广泛的物联网设备形态。我们不专注于物理层性能或仿真，而是将完整的通信系统视为黑盒，包括作者原始描述中的所有组件（如同步和纠错），以确保分析其在真实设备上的实际适用性。

2 系统模型

本研究评估声学数据传输方案：通过声波将数据从一台设备广播到 Nearby 设备的软件实现。从系统开发者的角度来看，每种方案都相当于一个黑盒，可在一台设备上发送数据，并在另一台设备上接收数据。在评估中，我们将此功能建模为一个包含两个函数的接口，如图 1所示：

(1) 一个**发射器**，wav =TX(b_{TX})，用于将数据位转换为 WAV 文件，以便通过扬声器播放。(2) 一个**接收器**，b_{RX} = RX(wav)，用于从包含传输过程麦克风录音的 WAV 文件中提取数据。

该接口将调制器和解调器的数字信号处理与音频播放和录音。尽管有些方案会在内部处理播放和录音，但为了清晰起见，我们将二者分开处理。在实践中，开发者可能不会存储中间 WAV 文件，而是直接从缓冲区流式传输采样点。尽管如此，我们的接口从根本上反映了现实世界的使用方式，并有助于在统一的播放和录音条件下公平比较各个方案。

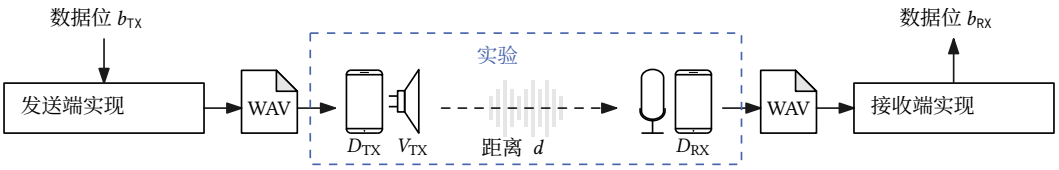


图 1. **声学数据传输模型**。 在我们的实验中，我们在设备 D_{TX}（音量为 V_{TX}）上播放 TX 实现生成的 WAV 文件，并在设备 D_{RX} 上录制。随后，使用接收端实现对生成的 WAV 文件进行解码。

2.1 场景

我们关注附近智能设备之间的数据交换场景：在这种场景中，可以修改软件（例如在智能手机上安装应用），但不能更改硬件或固件。我们重点关注无需与通信伙伴进行特殊安排即可实现通信的广播方案。我们根据通信距离这一影响可用性的关键因素，将声学通信的常见用例（第 3 节）分为三类。不同距离下的实际可用性差异显著：近至设备必须相互接触（如近场通信），远至可以将设备举在距离通信伙伴数米之外。本文始终聚焦于这些具体的通信场景，它们涵盖了当前大多数用例：

(1) 近距离 ($d \leq 10\text{cm}$)：在短距离内传输公钥或小型文件（≈4096 bit），类似于近场通信。示例包括移动支付 [8, 63], 安全设备配对[31, 64, 70], 和访问控制 [8, 45]。**(2) 中距离** ($10\text{cm} < d \leq 1\,\text{m}$)：在中等距离内传输公钥指纹或哈希值（≈128bit）。示例包括物联网设备配置 [27, 58], 用于隔离设备的侧信道 [54], 以及个人电子设备和助听器的控制信道 [43]。**(3) 远距离** ($d > 1\text{m}$)：在较远距离内传输短标识符（≈16 bit）。示例包括在公共标牌附近或商店内提供位置特定信息 [10], 室内物联网通信 [18, 23], 以及增强媒体体验，例如社交电视 [36] 和通过电视与广播中的隐藏音频数据实现的第二屏幕服务 [28]。

3 相关工作

本节将简要回顾声学数据传输方面的既有研究，然后在第 4 节介绍我们的系统性文献研究。

尽管智能手机和物联网设备的音频硬件主要用于媒体播放和语音通信，但也可以改用于感测 [5, 66, 72]，甚至用于空中声学数据传输。后一种情况下，扬声器通过频率或振幅变化在声波中编码信息，这与电磁通信类似 [39], ，从而无需专用射频（RF）硬件，即可让包括智能手机和语音控制家电在内的各类设备交换数据 [11]。该通信信道受到典型 0–22 千赫音频范围的限制（这是由于常见采样率为 44.1 千赫），并且存在明显的频率选择性 [21], 以及高音量下的非线性（见第 7.3 节）。数据可以通过可听声音或近超声频率传输；与 Wi-Fi 或蓝牙不同，这种方式无需网络设置或配对，因此可以在数米范围内进行自组织广播。由于声学通信完全由软件定义，因此易于部署在现有设备上，并可根据各种通信需求进行定制 [38]。

普适计算和移动计算领域的研究探讨了利用声学数据传输实现上下文感知计算 [40, 41] 以及近距离设备之间的位置受限通信 [1, 67]。有限的传输范围也引起了人们对安全设备交互的兴趣 [23, 52], ，包括设备配对 [15, 16, 50, 53, 70] 以及保护植入式医疗设备 [17]。研究人员评估了声学传输的用户体验 [38, 39, 42, 52], ，因为数据交换的可听度是一项新特征，使其区别于以往的无线技术。各类商业应用利用声学数据传输实现移动支付 [8, 63],、上下文感知应用 [37],、基于位置的服务 [9, 60], 、数据传输 [62], 和设备配对 [31, 64]。尽管这类技术具有应用潜力，但此前尚无研究尝试在真实设备上对这些系统进行系统评估。缺少已发布的源代码也阻碍了对先前研究结果的复现。

4 文献研究

我们开展了系统性的文献调查，收集所有提出新型空中声学软件定义通信系统的论文，范围如下。

4.1 范围

为确保我们的研究与普适计算及日常消费者互动相一致，我们将文献调研重点放在描述适用于商用现成设备（如智能手机）的空中声学设备间通信实际实现的论文上。我们排除了与其他类型声学通信相关的研究，例如水下或工业用途的研究，以及不适用于通用数据传输系统的研究。

4.2 方法

为找出与声学数据传输方案相关的文献，我们首先收集了大量与该主题相关的论文。随后，我们进一步筛选，仅保留包含适用于目标使用场景的研究实现的论文。我们的方法以 PRISMA 为指导 [48], 。PRISMA 是一种用于以结构化方式评估大量医学研究的框架。PRISMA 提供了一个包含三个步骤的流程图：

(1) **识别：**确定数据来源和检索策略。(2) **筛选：**根据预期范围筛选研究。(3) **纳入：**选择要纳入的论文。

4.2.1 识别。本次文献研究以 Web of Science（WoS）作为信息来源，[1] 这是一个全面收录期刊和会议论文集的数据库。我们选择 WoS，是因为其目录内容丰富、定期更新，且具备先进的筛选功能，支持使用嵌套逻辑运算符；这些功能对于将结果缩小到我们的特定范围至关重要。我们的检索策略兼顾了覆盖面和可操作性。2023年5月18日，我们设计并执行了两项互斥的检索式[2]，共得到 631 篇论文。这些检索式针对 2000 年至 2022年发表在声学、计算机科学和工程学领域的论文，重点检索标题中的相关关键词。我们通过屏蔽与水下通信和超声成像相关的特定关键词，排除了航海或医学领域的论文。第二项检索式扩大了检索范围，纳入了更多关键词组合。详情和关键词请参见检索式网址[2]。

4.2.2 筛选。 我们对所有已识别的论文进行处理，以判断其与本研究范围的相关性。在第一轮中，我们审阅了全部631篇论文的标题和摘要，剔除了422篇明显超出范围的论文，例如聚焦于基于语音的通信、音频跟踪或经由金属屏障通信的论文。在第二轮中，我们审阅了其余209篇论文的全文，并根据其对普适计算用例的适用性，以及是否描述了面向商用现成设备的可测试实现，对其进行分类。我们排除了依赖专用硬件的论文，例如用于多输入多输出（MIMO）通信的麦克风或扬声器阵列，以及将数据嵌入音乐中的论文，除非该方案可用作通用通信系统。

4.2.3 纳入标准。 我们的文献研究总共找到了 31 篇论文，它们与本研究的范围直接相关，并描述了一种可测试的通信方案（如表 1所示）。

[1] https://clarivate.com/products/web-of-science/, 最后访问：2025年 03 203- - 检索式 1: https://www.webofscience.com/wos/woscc/summary/4c89e96d-5de5-4543-9167-23d73086d8e6-8a0aeac3/ relevance/1; 检索式 2: https://www.webofscience.com/wos/woscc/summary/392d145e-a4ac-4aff-8995-ad2103dd116b- 898dca59/relevance/1。

表 1。**所调查的声学通信方案，**按吞吐量排序。右侧表格承接左侧。出版物中未提供的数据以“?”标出。

Name	调制	Frequencies [kHz]	Inaudible	Throughput [bps]
Moriyama 等人 [44]	幅移键控	~19	✓	5
Kim 等人 [30]	16 阶频移键控	18 – 20	✓	5
Ka 等人 [28]	啁啾	18.5 – 19.5	✓	15
Lee2015 [33]	啁啾	19.5 – 22	✓	16
Lee2020 [34]	啁啾	18.5 – 19.5	✓	16
Hornych 等人 [20]	幅移键控、频移键控	18 – 20	✓	33
TagScreen [36]	啁啾	18 – 20	✓	36
Haus 等人 [18]	正交频分复用 + 幅移键控	22 – 24	✓	64
Miegel 等人 [43]	开关键控，频移键控	16 – 20	✓	84
Rustle [24]	自定义[1]	0 – 15	✗	90
Nearby [13]	直接序列扩频 + 多频移键控	18.5 – 20	✓	95
Bang 等人 [3]	FSK	18 – 22	✓	170
Sadeq等人 [54]	FSK	18 – 19	✓	440
HRCSS [4]	啁啾	18 – 22	✓	500
Dolphin [65, 73]	正交频分复用 + 幅移键控	8 – 20		500

Name	调制	Frequencies [kHz]	Inaudible	Throughput [bps]
Gonçalves等人 [14]	正交频分复用 + 相移键控	0 – 4	✗	600
PriWhisper [70, 71]	MFSK	9 – 17	✗	1000
DigitalVoices [38]	幅移键控、（多进制）频移键控、扩频	1 – 10	✗	1280
A-NFC [35]	QPSK	16–21	✓	2200
Dhwani [45]	正交频分复用 +相移键控	6 – 7	✗	2400
Hush [47]	正交频分复用 +正交振幅调制	17.5 – 21	✓	4900
He 等人 [19]	调幅、频移键控	0 – 22	✗	6000
Yamamoto 等人 [69]	16-正交振幅调制	?	?	32000
Arentz等人 [1]	ASK	20–23	✓	?
Bits-over-Sound [29]	差分二进制相移键控	?	✗	?
Jeon等人 [22]	PSK	20	✓	?
禁止进入 [32]	啁啾	?	?	?
Sasano 等人 [56]	正交频分复用 +相移键控	6.4 – 8	✗	?
SilentInformer [57]	FSK	17 – 19	✓	?

[1] Rustle 的定制方案称为“隐蔽调制”。

5 获取声学数据传输实现

为了独立评估现有的声学通信系统，并将其嵌入物联网测试平台，我们首先需要获得可正常运行的实现。我们采用了三种不同的方法来获取这些实现（相关挑战的讨论请参见第7.1节 ）。

5.1 公开可用的实现

所调查的31篇论文中，没有一篇提及公开可用的实现。

5.2 作者请求

我们通过电子邮件联系了每位作者，如图2所示。只有六位作者回复，其中仅三份回复提供了可运行的实现：

- **数字语音**由 Lopes等人 [38] 于2001年提出，是有记录以来最早的空中声学通信方案之一。该方案使用八种载波频率进行二进制幅移键控（ASK）。作者提供了其 Java 源代码。
- **HRCSS**由 Cai等人 [4], 提出，是一种基于啁啾的通信方案，可同时使用多个啁啾载波。提供给我们的 MATLAB源代码缺少其双向探测功能，因此我们仅评估数据传输。
- **Gonçalves等人** [14] 提出了一种方案，在每个子载波上使用正交频分复用（OFDM）和相移键控（PSK），并加入一段短啁啾序列以实现同步。作者向我们提供了 MATLAB源代码。

第四份回复（Hush [47]）无法在当前版本的 Android 上使用。[3]

[3]Hush 的作者 [47] 分享了他们的 Android 实现，但我们遇到了严重的运行时错误，导致信号传输和录音受阻。这些问题可能是由安卓生态系统的变化以及代码不完整所致。

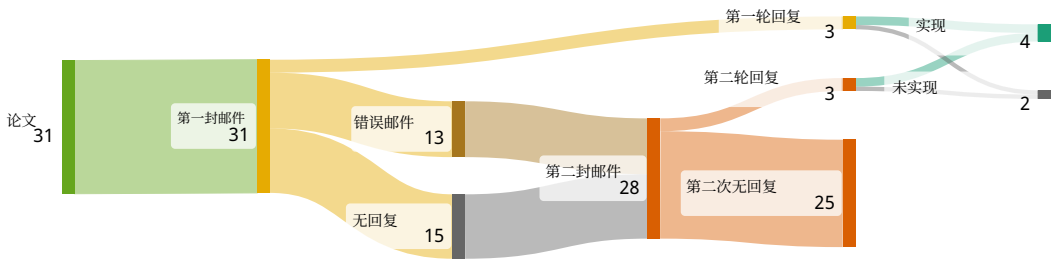


图 2. **作者的回复**我们就声学数据传输相关的 31 篇出版物联系作者，请他们提供实现。我们联系了第一作者和所有共同作者。对于 13 篇论文，提供的电子邮件地址已过时，因此我们手动搜索了最新的联系方式。如果三周后仍未收到回复，我们会再次联系。在收到的四份实现中，由于第四份存在运行时错误，我们评估了其中三份。其他回复的作者表示，他们已无法访问这些实现，原因包括主要作者离开了大学或硬件故障等。

5.3 重新实现声学数据传输方案

我们重新实现了三篇高引用论文中的方案，它们分别代表了三种目标使用场景，各自具有不同的吞吐量和最大传输距离：

- **Lee 等人** [33]声称自己率先创建了一个在实用声学通信中使用啁啾信号的原型，目标是远距离用例。在他们的系统中，每个通信帧都以一个较长的啁啾前导码开头，后接 16 个独立的啁啾信号。

- **Nearby** 由 Getreuer 等人提出， [13] 并已集成到谷歌的多款产品中，包括 Google Nearby。它针对中距离应用场景，采用多频移键控 (MFSK)、直接序列扩频 (DSSS) 和单边带调制 (SSB) 的组合。

- **PriWhisper**由 Zhang 等人提出 [70], ，是一种面向近距离应用场景的声学通信系统，与近场通信类似。该系统采用多频移键控，并加入友好干扰以实现物理层安全。不过，在我们的研究中，我们仅关注 PriWhisper 的通信能力，并未实现其安全功能。

我们通过仿真和实际实验验证了这些重新实现，确保其表现与原始评估结果相近。Lee 等人提出的系统 [33]有详尽的说明，使我们能够按照其论文中的描述实现所有方面。实现另外两个系统时，我们需要补充未记录的细节（见附录 B）。我们将 MATLAB 重新实现作为复现包的一部分发布 [51]。

5.4 总结

我们最终获得了六种代表最先进水平的声学通信系统学术实现。此外，我们还从开源项目 ggwave 中纳入了两种非学术实现进行比较，[4] 分别采用其可听（**ggwave**$_a$）和不可听（**ggwave**$_i$）配置。因此，共有八种方案进入评估（表 2）。

该项目自 2017 年以来便未更新，作者在通信中也承认，他们知道该项目缺乏维护。尽管如此，我们仍然赞扬 Novak 等人 [47] 公开了他们的源代码。[4]GitHub 上星标数最多的“声波数据传输”项目：https://github.com/topics/data-over-sound （截至 2025年3月3日）。

表 2。 **评估的声学通信方案**，按吞吐量排序。第一列注明实现的来源：P——论文中引用；A——联系作者获得；R——由我们自行重新实现；N——用于比较的非学术解决方案。在最后三列中，我们根据各方案标明的吞吐量和最大距离，对其先验适用于 第 2.1 节所定义的三类用例的程度进行分类。

Via	Name	调制	频率	不可听	吞吐量[1]	近	中	远
R	Lee等人 [33]	Chirp BOK	19.5–22 千赫	✓	15 比特/秒	✗	✗	✓
A	数字语音 [38]	MFSK 直接序列扩频	1–3 千赫	✗	80 比特/秒	✗	✓	–[2]
R	Nearby [13]	+多频移键控	18.5–20 千赫	✓	84 比特/秒	✗	✓	✓
N	ggwave$_a$ [12]	MFSK	1.8–6.3 千赫	✗	268 比特/秒	✗	✓	–[2]
N	ggwave$_i$ [12]	MFSK	15–19.5 千赫	✓	268 比特/秒	✗	✓	–[2]
A	HRCSS [4]	OCSS 正交频分复用 + 正交相移键控	18–22 千赫	✓	500 比特/秒	✓	✓	✓
A	Gonçalves等人 [14]		0.1–4 千赫	✗	600 比特/秒	✓	✗	✗
R	PriWhisper [70]	MFSK	9–17 千赫	✗	729 比特/秒	✓	✗	✗

[1] 我们列出净数据速率，因为从用户的角度来看，这是经过纠错后可用的数据速率。[2] 未指定最大距离，因此无法进行先验评估。

6 评估

本节将系统评估声学数据传输在各种使用场景中的泛化能力和实际适用性。例如，室内物联网通信 [18] 需要在数米的距离上实现可靠传输，而移动支付系统 [8, 63]只需在几厘米的距离上即可有效运行。当集成到智能手机应用中（例如谷歌的 Nearby 系统 [13]）时，声学数据传输必须适应具有不同音频硬件和频率响应的各种智能手机型号。此外，传输会在不同的环境条件下进行，并且必须能够抵御各种背景噪声。我们的评估将系统地考察这些因素。

6.1 方法

我们开展了四项实验，分别改变了距离（第 6.3 节）、设备型号（第 6.4 节）、噪声干扰（第 6.5 节）和传输环境（第 6.6 节）。根据我们的系统模型（图 1），我们将每种实现视为黑盒，并按照作者的设计及其指定参数使用。[5] 我们的目标是从链路层（第二层）的角度，将每种方案作为完整的通信系统进行评估，包括成帧、同步和纠错。

我们随机选择了要传输的数据 b_{tx} 。发射器将这些数据转换为音频，并可自由使用 0 至 24 千赫的基带。每个 WAV 文件在从一部智能手机传输到另一部之前，都会归一化至特定的最大振幅。接收到的音频被录制为 WAV 文件，并由接收器实现进行处理；该实现会解调信号并输出数据 b_{rx}。随后，我们验证了 b_{rx} 是否与 b_{tx}匹配。借助这一流程，我们能够将每种方案视为黑盒，评估其在设备之间传输数据的能力。每种方案至少重复测量 20 次 [22], ，而 Lee等人提出的方案 [33] 由于传输时长较短，重复测量了 100 次。若出现测量误差，我们还会进行额外测量。实验设置如图 3 所示，并

[5]大多数原始论文只给出一组物理层参数。有两个例外：（1）PriWhisper [70], 改变了 M-FSK 方案中的频率数量——为了符合其宣称的吞吐量，我们采用了最高数量；（2）DigitalVoices [38], 改变物理层参数仅用于评估声音悦耳度，这不在我们的研究范围内。

详细说明见附录A，以提高透明度，其中包括完整的硬件和软件配置。

6.1.1 指标。为评估可靠性，我们引入总错误率（TER），这是一种链路层指标，综合了比特错误率（BER）和分组错误率（PER）的要素，便于直接比较不同方案。补充材料中还提供了底层原始 BER 数据。在我们的系统模型（图 1）中，TER 通过比较纠错后接收的数据位错误率来衡量最终错误率，b_{rx} 与 b_{tx}进行比较。如果接收器返回错误（例如，由同步失败导致）而非数据，我们便将该次传输的 TER 视为 100%。[6] 这种分类方式将此类传输视同每一位都被错误解码的传输，因为从用户角度来看，两种情况都意味着传输完全无法使用。以往关于声学数据传输的研究通常只报告 BER [14, 22, 57]或 PER， [13, 70], 但两者各有缺陷。BER 忽略无法解码的传输，因此即使大多数传输都无法解码，BER 仍可能为 0%；而 PER 则将单个位错误与一千个位错误等同看待。结合这两种指标后，TER 能更好地反映现实中的可靠性，同时考虑单个位错误和完整的传输失败。[7]

6.1.2 参数。我们进行了四项实验，系统地调整以下参数，以确定它们对可靠性的影响。这些参数是以往使用真实设备开展的研究中最常评估的参数 [13, 14, 33, 70]。

- **距离** (第 6.3 节)。传输距离 d 会显著影响方案是否适用于我们指定的用例类别。例如，无法在数米距离内可靠传输的方案不适用于远距离用例。我们将默认距离设为 $d = 50$ 厘米，这对于中距离用例而言较为典型。
- **设备型号** (第 6.4 节)。发射器 *D*TX 和接收器 *D*RX 会影响传输可靠性，因为智能设备中的音频硬件并非专为数据传输而设计。每台设备的频率响应各不相同，尤其是在近超声频率下 [21]。我们选择 Google Pixel 4a 智能手机作为默认发射器 *D*TX，因为它在音量设置[8] 为 $V\mathrm{TX} = 19/25$ 时失真相对较低。通过实验，我们发现这一设置在较高音量与因放大器和扬声器的非线性导致的低失真之间取得了良好平衡。默认接收器 *D*RX 为 Samsung Galaxy S20 Ultra 智能手机，因其自噪声较低而被选用。
- **噪声干扰**（第 6.5 节）。实际传输通常发生在不理想的条件下，例如环境噪声很大，或用户操作智能手机时产生的干扰。默认情况下，我们的测试是在没有外部噪声的安静环境中进行的。
- 环境（第 6.6 节）。环境会影响声学信道的多径剖面。默认情况下，我们在典型的室内环境中评估这些方案，使用安静的办公室，并在走廊进行长距离测量。由于这些环境条件

[6]我们的测量结果表明，当接收消息有很大一部分缺失时，错误率可能超过 50%。为正确惩罚完全失败的情况——即未收到任何消息时——我们将传输错误率（TER）设为 100%。[7]总之，除非接收器返回错误而非数据（此时 TER 为 100%），否则 TER 等于比特错误率（BER）。举例说明（假设每帧包含 100 个数据位）：（1）正确接收 99 个数据位，TER 为 1%（与 BER 相同）。（2）完全无法解码，TER 为 100%（此时 BER 未定义）。（3）所有数据位均接收错误，TER 也为 100%（与 BER 相同）。（4）在情况（1）和情况（2）中各进行 50 次传输，平均 TER 为 50.5%；相比之下，BER 会产生误导性地偏低，为 1%，而分组错误率（PER）则会产生误导性地偏高，为 100%。[8]在智能手机上，音量通常只能设置为少数几个离散音量索引之一，最大索引因设备而异。

无法完全复现，我们还在消声室（图 3）中进行了测试。消声室没有混响或环境噪声，可提供受控且可复现的测试环境。最后，我们还在一间配有机械通风的大型报告厅进行了测试。办公室、报告厅和消声室的背景噪声平均值分别为 53.60 dB声压级、60.01 dB声压级和 27.65 dB声压级（均为未加权均方根值）。

6.2 初步测试

6.2.1 设置。在开展大规模基于智能手机的测试之前，我们在近乎理想的条件下进行了初步评估，以确认每种实现均可正常运行。我们使用高质量音频设备，通过 Neumann KH 80 DSP 录音室扬声器传输每种方案，并用 Earthworks M23R 参考麦克风进行录制。扬声器和麦克风在办公室内同轴放置，距离为 50 厘米（详情请参见 附录 A.4.1）。

6.2.2 结果。Lee等人、DigitalVoices、Nearby、PriWhisper 的方案以及 ggwave 的两个变体均实现了 0% 的完美 TER（图 4）。Gonçalves等人的方案总体准确，但错误率为 0.8%，这可能是因为该方案是针对更近的距离设计的。然而，HRCSS 实现的 TER 极高，尽管使用了高质量设备，这一结果仍出乎意料。遗憾的是，我们不得不认定该实现不适用于实际测试，因此将其排除在后续测试之外（讨论见 第 7.4.3 节）。

6.3 不同距离的影响

6.3.1 实验设置。 接下来，我们在 5 cm 至 40 m 的距离范围内评估了每种方案。测试分别在安静的室内环境（代表日常使用场景）和混响与噪声极少的消声室中进行。在这两种环境中，我们测量了最远 5 m 的距离；对于超过 5 m（最远 40 m）的距离，则使用办公室旁边的走廊进行测试。

6.3.2 结果。 在消声室(图 5a)中，大多数方案在短距离下都实现了较低的错误率。Lee等人的方案在所有测试距离下均保持零错误率。两种 ggwave 方案几乎完美运行，仅在 2m 和 5m 处略有例外。Nearby 表现也十分出色，只有在 50 cm 处出现一个小幅异常值。Gonçalves 等人的方案在距离小于 10 cm 时将错误率维持在 0.5% 以下，但在较远距离时升至约 13%。DigitalVoices 和 PriWhisper 在近距离时的错误率低于 1%，但超过 20 cm 后性能明显下降。

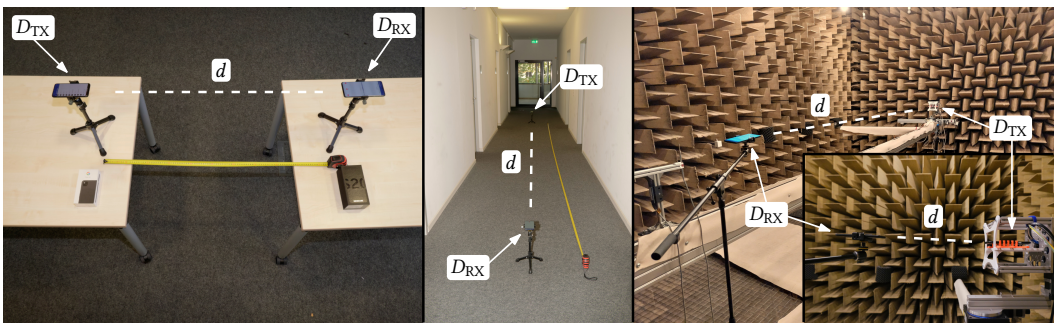


图 3. **实验设置**。 左图展示了我们的办公室设置，其中设备 D_{TX} 向距离为 d 的设备D_{RX} 传输。中图展示了我们在办公室走廊中的设置，用于测试 5 m 至 40 m 的较远距离。右图展示了我们在消声室中两个不同距离下的设置。

在办公室环境(图 5b)中，由于多径效应和环境噪声更强，错误率显著上升。DigitalVoices、PriWhisper、Nearby 和 Gonçalves 等人的方案都表现出明显的距离依赖性，距离越远，性能越差。不过，在不超过 10 cm 的短距离下，除 DigitalVoices 外，所有方案的 TER 均低于 0.5%。尽管 DigitalVoices 在近距离下的表现不如消声室中，但令人意外的是，它在 50 – 500 cm 范围内有所改善，这表明 DigitalVoices 接收器尚未得到充分优化。我们调查了 DigitalVoices 表现差异的原因，发现对录音进行后处理和放大，能显著降低其在消声环境中的错误率。不过，为了按照相关出版物中的描述比较各方案，我们没有在图表中纳入后处理结果。Nearby 在 1 m 以内表现良好，但在更远距离时变得不可靠，其中在 10 m 处出现一个异常值，这可能是由于测试位置过渡到了走廊。这种非单调的距离表现与 Getreuer 等人对 Nearby 的原始评估结果一致 [13,图 13],；该评估将这种变化归因于复杂的多径效应。Lee等人的方案在所有距离下都表现出色，仅在 5 m 处出现一个轻微异常值。不可听 ggwave 变体在最远 20 m 的距离内仍然有效，而可听变体只能在 1 m 以内工作，之后便无法解码。

6.4 不同设备型号的影响

6.4.1 设置。 本实验评估了声学数据传输方案在典型室内环境中、在多种智能手机上的表现。我们选取了五款智能手机，它们在发布日期、价格和制造商方面覆盖广泛（表 3）。我们使用21种不同的发射器-接收器组合对每种方案进行了测试。我们的设置是将一部智能手机作为发射器，沿轴向放置在距离两到三台同时工作的接收器50厘米处，以提高测量过程的效率。由于我们有两部Pixel 4a智能手机，我们还评估了发射器与接收器型号相同时的性能。

6.4.2 结果。 如图 6所示，Lee等人的方案和两种ggwave变体在所有设备组合中始终表现出极低的错误率。DigitalVoices的性能有所不同：最佳结果出现在Pixel 6 Pro向Nexus 6P传输时（传输错误率 $\approx 12\%$），而 Pixel 4a 和 Pixel 6 Pro 作为接收器时大多失败。Nearby 总体表现良好，但在使用 Samsung S20 作为发射器时以及 Pixel 4a 的某些传输中除外。PriWhisper 在某些组合中有效，尤其是以 Pixel 6 Pro 作为发射器时（传输错误率低至 3%），但在其他情况下因同步问题而失败。其整体性能较差可归因于 50cm 的设备距离，这凸显了其在预期

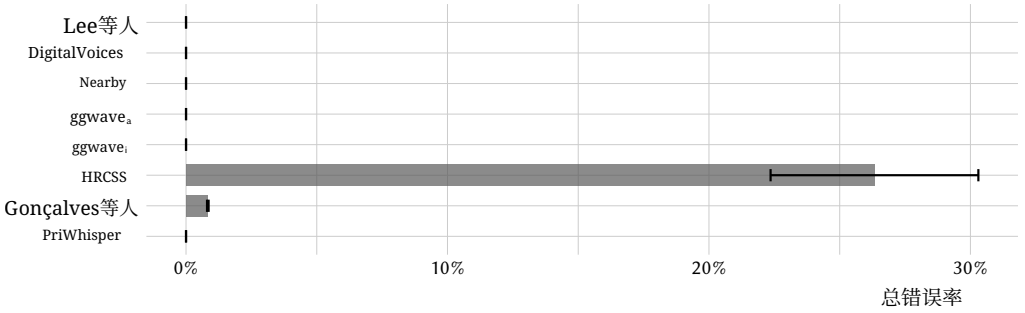
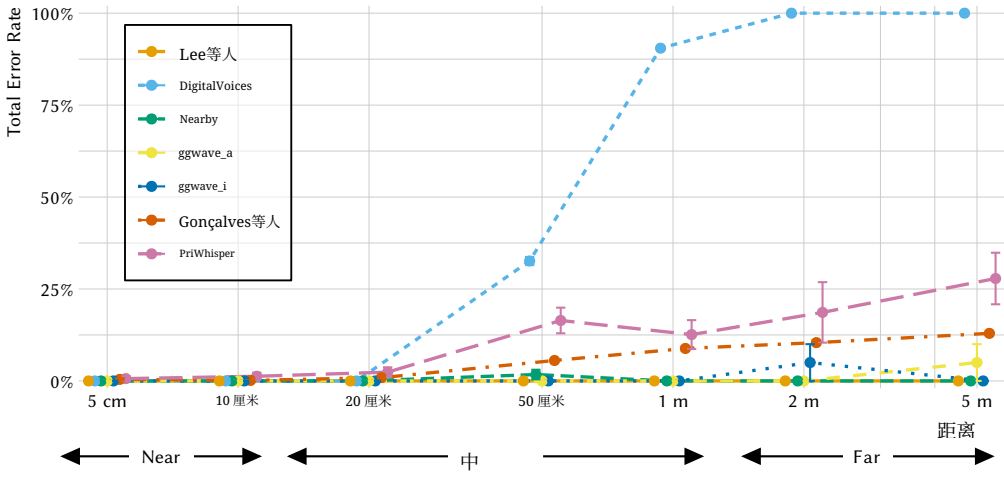
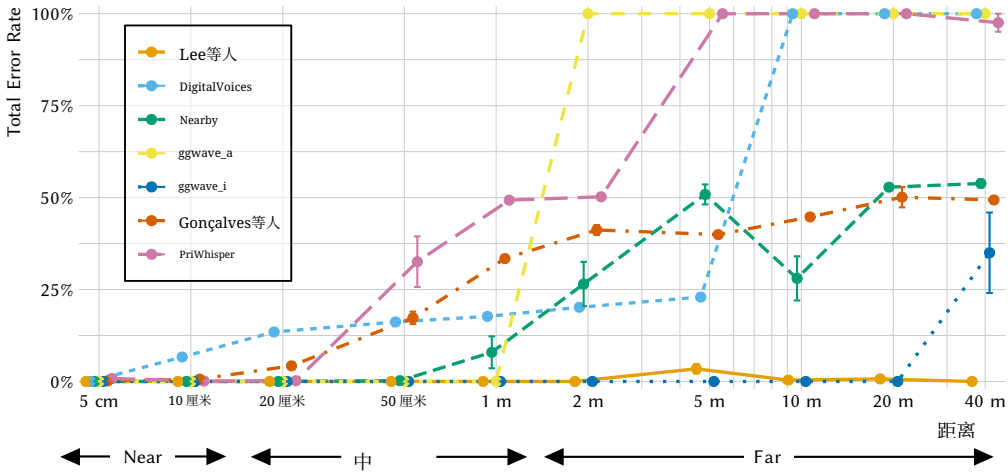


图4。 **最佳硬件场景下的初步测试**。 柱形越短越好。方案按吞吐量排序（见 表 2）。误差条表示标准误差。参数： $d = 50$厘米； D_{TX} =扬声器 Neumann KH 80 DSP； D_{RX} = 麦克风 Earthworks M23R； $V_{TX} = -40$分贝 接口增益；安静的办公环境；N=240 次测量。



(a) 消声室。



(b) 办公室环境。

图 5. **不同距离下声学传输的可靠性**。传输错误率（TER）越低越好。消声室中的测量距离最大为 5 m。在办公室环境中，我们最多测量到 40 m，但 10–40 m 距离的测量是在走廊中进行的。距离轴采用对数刻度，但我们仅在标示的离散距离处进行测量。误差条表示标准误差。为避免重叠，不同方案的数据点向两侧错开。参数：D_{TX} = Pixel 4a 智能手机； D_{RX} = Samsung S20 Ultra 智能手机； V_{TX} = 音量指数 19/25；消声室（图 5a）和安静的办公环境（图 5b）；N=3740次测量。

数厘米的使用范围之外的局限性。在这种情况下，不同设备型号对可靠性的影响差异显著，从令人满意到完全无效不等。Gonçalves 等人的方案性能更为稳定，错误率从 5%（最佳情况：OPPO Reno 6 到 Pixel 4a）到 32%（最差情况：Pixel 4a 到 Nexus 6P）。

表 3. **设备型号实验中的智能手机**，位于第 6.4 节。音量列显示了每台设备选定的音量索引和最大音量索引。我们将每台设备的音量设为其最大音量的约 75%；通过实验，我们发现这一设置能在保持较低的放大器和扬声器非线性失真时，实现较高的音量，因而是一个不错的折中方案。

设备名称	产地	发布日期	价格	音量
Google Pixel 4a	USA	2020年10月	349€	19/25
Google Pixel 6 Pro	USA	2021年10月	900€	19/25
华为 Nexus 6P	中国	2015年9月	700€	11/15
OPPO Reno 6	中国	2021年9月	499€	12/16
Samsung Galaxy S20 Ultra	韩国	2020年2月	1300欧元	11/15

总体而言，以 Pixel 6 Pro 作为发射器时，错误率始终最低。Samsung Galaxy S20 Ultra 作为接收器时整体表现良好，但搭配 DigitalVoices 和 PriWhisper 的某些发射器时除外。这凸显了设备特性的巨大差异，使某些方案难以推广。

6.5 不同类型噪声和设备操作的影响

6.5.1 设置。 前述测试是在相对安静的环境中进行的，但实际场景往往存在不同类型的干扰，这些干扰可能因环境及周围状况，或因人为操作设备而降低声学传输的可靠性。本实验旨在评估不同类型的噪声和人为操作对接收器的影响。我们将 Pixel 4a 用作发射器，将 Samsung Galaxy S20 Ultra 用作接收器，两者在安静的室内环境中沿轴线对齐，间距为 50cm。我们针对以下七种旨在模拟现实干扰的噪声和人为操作分别进行了实验：

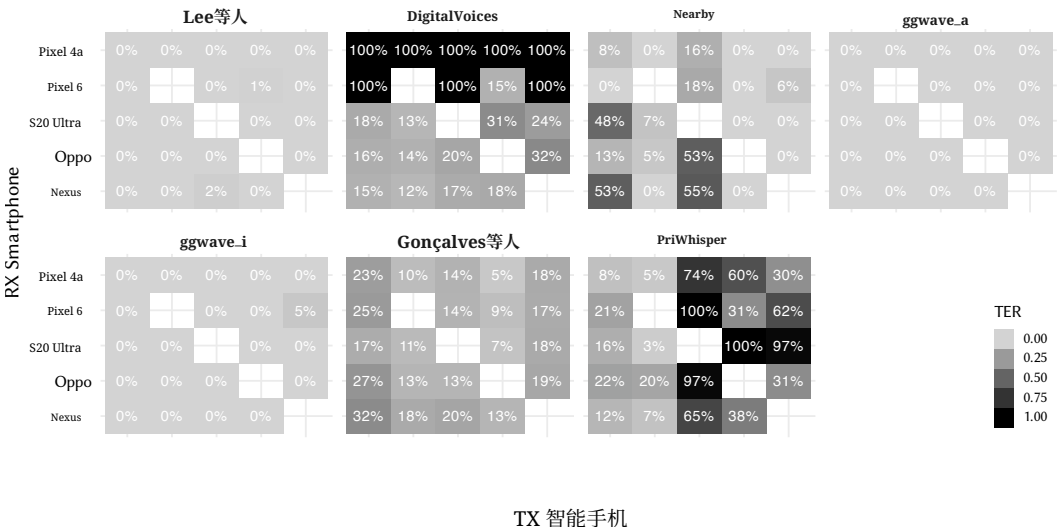


图6. **不同设备组合下声学传输的可靠性**。 传输错误率越低越好。方案按吞吐量排序。热图展示了每种方案在 21 种不同发射器-接收器组合下的可靠性。参数： d = 50厘米； $D_{TX}/D_{RX}/V_{TX}$ = 参见表 3；办公室环境；N=4620次测量。

- **环境音**。 通过距离接收器 1m 的 Neumann扬声器播放来自咖啡馆、火车站和市场的环境现场录音。三种环境音均分别进行了实验（详细设置信息见 A.4.2 节）。
- **圆周运动**。 在桌面上以圆周运动方式移动接收器，使距离变化 ±5cm。
- **拍手**。 每隔 2 秒在接收器附近拍手。
- **口袋**。 每隔 2 秒将接收器放入裤子口袋并取出，摆放时使发射器的扬声器朝向口袋中的接收器。
- **上下移动**。 大约每隔 2 秒将接收器抬起并稳稳地放在桌面上。
- **相反方向**。 将智能手机摆放成麦克风和扬声器朝向相反方向。这种摆放方式模拟两个人手持手机、彼此面对的情形。
- **并排放置**。 智能手机彼此相邻放置，长边相贴。
- **手持**。 发射端智能手机由一人手持，因而引入了典型手持使用时的自然移动。

6.5.2 结果。 传输过程中将智能手机放入口袋，会导致所有方案的性能显著下降（图 7）。不过，Nearby、不可听 ggwave 变体和 Lee等人的方案能够抵御所有其他受测噪声类型。可听 ggwave 变体会受到环境噪声影响，导致其完全无法使用，这可能是因为环境噪声干扰了传输带宽。拍手会严重影响 DigitalVoices、PriWhisper 和 Gonçalves 等人的方案，使它们基本失效。无论是绕圈移动，还是反复抬起并放下，桌面相关运动与噪声的相互作用对性能的影响都相对较小。值得注意的是，DigitalVoices在拍手声和放入口袋产生的噪声下表现明显不佳（传输错误率 $> 50\%$），但应对桌面相关运动时表现较好（传输错误率 $\approx 6\%$）。相比相向方向，并排方向对可靠性的损害更大，不过并非所有方案都会受到影响。

6.6 不同环境的影响

6.6.1 设置。为确定环境及其多径剖面的影响，我们测量了 Pixel 4a（发射器）与 Samsung Galaxy S20 Ultra（接收器）之间的传输。两台设备沿轴线对齐，相距 50 cm。评估在多个地点进行：小型和大型办公室、会议室以及讲堂。在讲堂中，我们评估了静态设置和动态设置两种情况；在动态设置中，人们在房间内轻声走动，同时确保设备之间的视线不受阻挡，从而动态改变多径环境（见附录 A.4.3）。

6.6.2 结果。如图 8所示，Lee等人的方案、Nearby 和两个 ggwave 变体在所有测试环境中均实现了完美的可靠性。其余方案在办公室中的表现最差，但在会议室和讲堂中有所改善。有趣的是，附近轻微的走动似乎不会对可靠性产生负面影响。

6.7 各用例总结

我们按方案和用例汇总了此前的测量结果，如图 9所示。该图展示了各方案在其预期用例（参见表 2）中的表现。由于这些用例对应典型的室内环境，因此我们排除了在消声室和使用高品质录音室设备进行的测量，因为这些情况并不代表典型的使用场景。我们将在第 7.2 节中更详细地讨论这些结果。

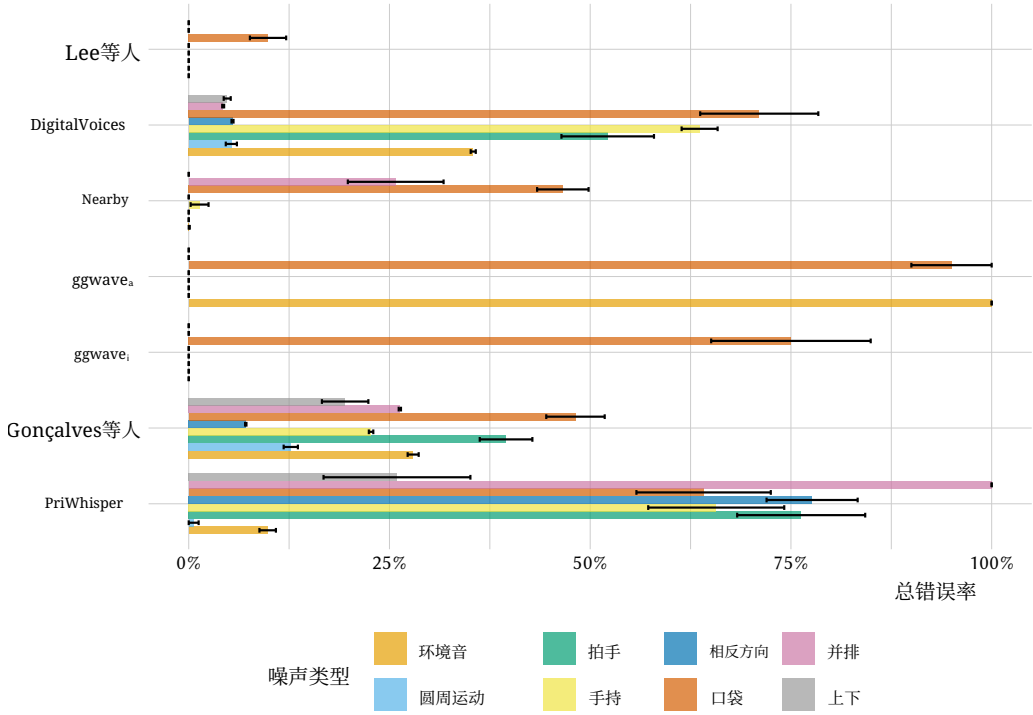


图 7。 **不同类型噪声干扰传输时的可靠性**。 柱形越小越好。方案按吞吐量排序。对于环境噪声，我们使用不同的现场录音分别进行了三次实验。误差条表示标准误差。参数： d = 50厘米； D_{TX} = Pixel 4a 智能手机； D_{RX} = Samsung S20 Ultra 智能手机； V_{TX} = 音量指数 19/25；包含不同类型受控噪声的办公室环境（参见 第 6.5 节）；N=2200次测量。

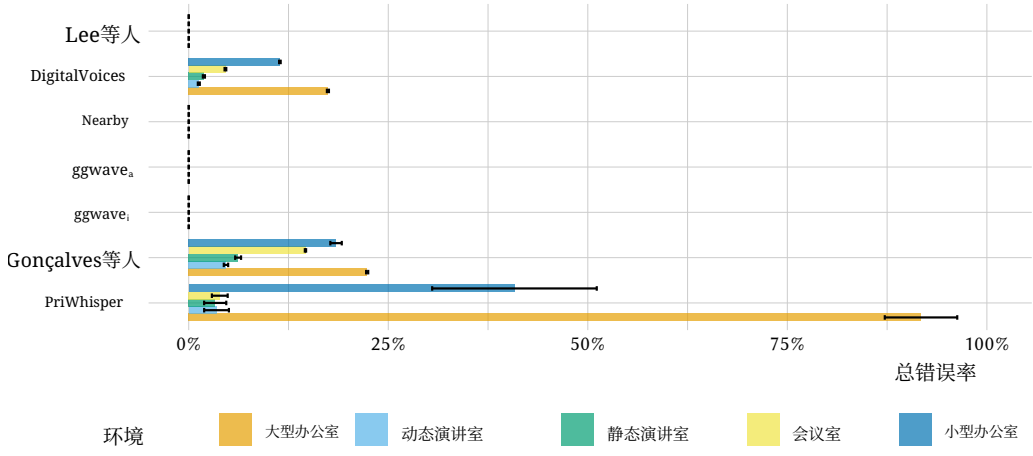


图 8。**不同传输环境中的可靠性**。 柱形越小越好。方案按吞吐量排序。误差条表示标准误差。参数： d = 50厘米； D_{TX} = Pixel 4a 智能手机； D_{RX} = Samsung S20 Ultra 智能手机； V_{TX} = 音量指数 19/25；N=1100次测量。

6.7.1 异常值。 对于不可听 ggwave 变体，在中距离应用场景中有一次传输失败（d = 50厘米， D_{TX} = Nexus 6P， D_{RX} = Pixel 6 Pro）；在远距离用例中，有几次传输失败（d = 40米， D_{TX} = Pixel 4a， D_{RX} = Samsung S20 Ultra）。如图 5b所示，Lee等人提出的方案中的异常值大多出现在距离为 5 米时。Nearby 在中距离应用场景中表现良好（传输错误率中位数 = 0%），但某些特定设备组合（例如， D_{TX} = Samsung S20 Ultra， D_{RX} = Nexus 6P）存在少数异常。

7 讨论

7.1 获取软件实现颇具挑战

为回答我们的第一个研究问题（**RQ1**），即研究人员如何获取可用于评估的声学数据传输系统的可运行实现，我们开展了系统性文献研究，并记录了尝试不同获取方法的经历（第 5 节）。在调查中确定的 31 篇出版物中，没有一篇提及公开可用的实现，只有四位作者应请求提供了他们的实现。缺乏易于获取的研究产物，严重阻碍了声学数据传输领域研究的复现，进而减缓了科学进步。若无法轻松访问现有研究以进行验证并在其基础上继续开展工作，要证明研究成果优于最先进水平就会困难得多。

我们还尝试根据论文中的描述重新实现这些方案，这一过程既耗时，也很少有其他研究人员会去做。在此过程中，我们发现原始出版物中存在几处不准确和遗漏之处，往往需要根据专业判断作出推测来解决（附录 B）。这一挑战并非我们所在领域独有，在计算机科学的其他领域也有相关记载 [7]。此类重新实现具有

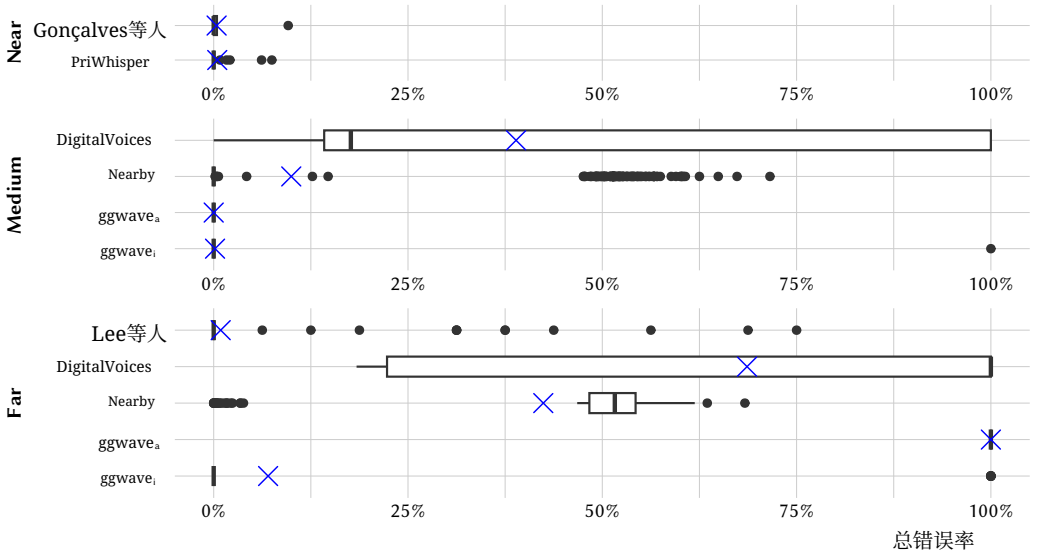


图 9. **安静的室内环境中的测量结果汇总** (第 6.3 节、第 6.4 节和第 6.6 节)，针对三类用例（定义见第 2.1 节，对应关系见表 2）。传输错误率越低越好。每类方案按吞吐量排序。箱线图显示中位数（黑色横线）、平均值（蓝色叉号）和异常值（黑色圆点），箱线两端分别标示第 25 和第 75 百分位数。参数： d = 可变（根据第 2.1 节中的场景）； $D_{TX}/D_{RX}/V_{TX}$ = 可变，依据表 3；安静的办公环境；N=3300次测量。

表4。 **各方案对不同用例类别的适用性（研究问题2）**。方案按吞吐量排序。若某方案的平均 TER 约为 0%，则视为适用（✓）；否则视为不适用（✗）。如果仅出现少量异常值，且 TER 中位数仍为 0%，该方案仍可能适用，但需要进行额外修改（例如纠错）。高亮条目表示相较于既有认知有所改变的新发现（表 2）。第一列说明我们如何获得该实现：P——论文中引用；A——联系作者获得；R——由我们自行重新实现；N——用于比较的非学术解决方案。

Via	Name	频率	带宽	不可听	吞吐量[1]	近	中	远
R	Lee等人 [33]	19.5–22 千赫	2.5 千赫	✓	15 比特/秒	✗	✗	✓
A	数字语音 [38]	1–3 千赫	2 千赫	✗	80 比特/秒	✗	✗	✗
R	Nearby [13]	18.5–20 千赫	1.5 千赫	✓	84 比特/秒	✗	✓[2]	✗
N	ggwave$_a$ [12]	1.8–6.3 千赫	4.5 千赫	✗	268 比特/秒	✗	✓	✗
N	ggwave$_i$ [12]	15–19.5 千赫	4.5 千赫	✓	268 比特/秒	✗	✓	✓
A	HRCSS [4]	18–22 千赫	4 千赫	✓	500 比特/秒	✗	✗	✗
A	Gonçalves等人 [14]	0.1–4 千赫	3.9 千赫	✗	600 比特/秒	✓[2]	✗	✗
R	PriWhisper [70]	9–17 千赫	8 千赫	✗	729 比特/秒	✓[2]	✗	✗

[1] 我们列出净数据速率，因为从用户的角度来看，这是经过纠错后可用的数据速率。[2] 需要进一步纠错。

与原始版本相比，出现错误或偏差的风险很高。考虑到投入了大量时间，这令人沮丧。为提高研究的可重复性和透明度，我们鼓励其他研究人员遵循我们的建议，参见第7.5节，并发布其软件制品。

7.2 面向智能手机自组通信用例的泛化能力

为评估先前研究中提出的声学数据传输方案的实用性（**研究问题2**），我们评估了这些方案对第 2.1 节中定义的三类用例的泛化能力。对于每一类用例，我们首先查阅了各方案的原始出版物，根据其中规定的吞吐量和最大距离，确定方案的先验适用性。例如，Lee等人的方案吞吐量仅为 15 比特/秒，因此从设计上讲，并不适合在近距离传输公钥或小文件，因为耗时会过长。在许多情况下，我们的实验发现这些方案的实际适用性有限；鉴于其原始出版物中的说法，这一结果令人意外。表4展示了我们根据评估对各方案所作的事后评估。具体而言，表中所有突出显示的结果都提供了新的见解；此前现有文献中的测试范围有限，尚未发现这些结果。这种差异表明，有必要在不同距离和设备上开展广泛测试，以证明方案在实际中的适用性。

对于远距离（最长可达 40 米），只有不可听 ggwave 变体（有效距离最长为 20 米）和 Lee等人的方案能够保持可靠性，其中后者的可靠传输距离最长可达 40 米。其他所有方案在这些距离下均无法正常工作。在中距离下，两个 ggwave 变体均有效，其中不可听变体尤为稳健。Nearby 方案的表现也不错，但难以适用于所有测试过的设备组合。其他方案不适用于此用例，原因要么是错误率高，要么是吞吐量不足。对于近距离用例，只有 PriWhisper 和 Gonçalves等人的方案在保持低错误率的同时具备足够的吞吐量。不过，这两种方案仍出现了一些错误数据包，表明需要改进或增加纠错技术。

我们分析的方案均无法适用于所有用例（见表4），因为不同应用对可靠性、吞吐量、延迟和可听度之间的权衡各有侧重。有些方案具有灵活性；例如，ggwave 可配置为以牺牲可靠性为代价提高吞吐量，也可将频段移至近超声频段以实现不可听。

对于安全配对或移动支付等以数据完整性为首要要求的应用，较低的 TER 是最关键的要求，而高吞吐量则居于次要地位。Lee等人的方案和 PriWhisper 很适合这一上下文，不过 PriWhisper 可能需要额外的纠错，才能确保在具有挑战性的环境中保持性能。相比之下，多媒体应用和内容分发（例如第二屏服务、基于位置的信息）等用例通常优先考虑高吞吐量和不可听性，以确保流畅的用户体验，并且通常能够容忍轻微且可纠正的错误。不可听 ggwave 变体是一个候选方案，但其数据速率无法超过 268 比特/秒。值得注意的是，针对更高数据速率设计的方案通常难以在多样化的室内环境中可靠运行。

鉴于高吞吐量通常会降低可靠性，许多实际应用的一种稳健设计模式是将声学信道用作低带宽触发：一条简短而可靠的声学消息（例如唯一标识符或哈希）可以发起通过 Wi-Fi 或蜂窝网络等高吞吐量信道建立的连接，以传输大量数据。

7.3 声学数据传输的实际挑战

我们的测量表明，在商用智能手机和物联网设备上进行声学数据传输所面临的挑战，很少出现在仿真或受控实验室环境中（**RQ3**）。由于音频硬件和传播信道在不同设备和环境之间差异很大，并不存在单一的传输规范。因此，我们并不规定一种新的调制解调器架构，而是总结了任何未来方案为实现实际适用性所应遵循的设计要求 。

7.3.1 严重的多径传播。 由于多径传播发生在相对较慢的343m/s声速下，混响会产生较长的时延扩展。由于此时延扩展取决于环境和几何结构，在一种场景中表现良好的方案在另一种场景中可能表现不同。我们观察到，即使在同一环境中，设备的位置和朝向也会显著影响可靠性，从而导致错误率出现相当大的波动。这些发现支持了Getreuer 等人的观察，即声学传输的可靠性对环境条件敏感 [13, 第8A节]。**设计要求：** 应能在长达数十毫秒的时延扩展 [33, 59] 以及由此产生的码间干扰下正确工作（例如，使用保护间隔）。

7.3.2 设备异构性。 扬声器和麦克风响应、放大器增益以及非线性各不相同：测试显示，在相同音量设置下，不同手机之间的声压级差异最高可达 30 dB。以往研究往往只考察特定设备型号，因此实际泛化能力有限。大多数声学链路都是单向的，因而无法采用基于反馈的增益控制。开发者只能根据每台设备的最大输出调节音量；我们的测试表明，大多数设备在音量设置超过约 75% 时会出现非线性失真，但具体阈值因设备而异。**设计要求：** 在未进行闭环校准的情况下（例如采用归一化和鲁棒调制），至少承受 30 dB 的信噪比波动和中等程度的非线性失真。

7.3.3 带宽受限。 物联网音频硬件通常采用采样率为 44.1 千赫或 48 千赫的模数转换器和数模转换器，因此实际带宽被限制在 22 千赫以下。当传输必须不可听时，可用带宽会缩减至 4 千赫以下，而且音频硬件在此频段具有很强的频率选择性 [21]。**设计要求：** 选择近超声波频段以实现不可听性时，应做好应对强频率选择性和低吞吐量的准备。

7.3.4 环境噪声。 与射频通信不同，声学频段会受到各种常见噪声源的影响，例如交通噪声、机械噪声、人类语音以及设备操作产生的噪声，这些噪声占据了可用频谱的很大一部分。突发噪声和平稳噪声都很常见。**设计要求：** 在持续噪声和脉冲噪声（例如，采用强纠错和时间交织）存在时，仍须保持链路完整性。

7.3.5 吞吐量与可靠性。 我们调查的部分出版物在 表 1 中追求非常高的吞吐量，但我们的评估表明，提高吞吐量通常会以更高的错误率为代价。商用系统通常以 ≤200比特/秒 [9, 13, 27, 62], 为目标，这可能是出于实际应用中的可靠性考虑。 **设计要求：** 除非可行链路自适应（例如 HRCSS式探测），否则应优先考虑稳健性并采取积极的错误控制；而在单向物联网场景中，这通常是不可能的。

7.4 与原始评估的比较

我们测试的若干方案也曾由其原作者进行评估。尽管我们的实验设置经过有意调整，以评估泛化能力，但我们将结果进行比较，以确定原始发现是否在概念上 可复现 [61]。总体而言，考虑到设备特性、环境以及我们的重新实现存在差异而产生偏差是意料之中的，我们的结果与原始评估一致。唯一的例外是 HRCSS [4], ，其提供的实现即使在最佳条件下也产生了高得不切实际的错误率。DigitalVoices 和 ggwave 此前没有可靠性评估。

7.4.1 PriWhisper。 Zhang等人 [70] 在室内使用 Samsung Galaxy S3 和 Google Nexus S 进行了基础可靠性评估，报告的分组错误率为 0.5%。在我们的近距离测试中，我们观察到了相似的比特错误率（约为 0.5%），不过我们的分组错误率略高。尽管存在这些差异，但我们的结果仍具有可比性，因为我们的重新实现可能不像原始实现那样经过优化，并且我们使用了不同的智能手机型号。

*7.4.2 Gonçalves*等人。 Gonçalves等人 [14] 在两台便携式计算机上对其方案进行了室内评估，报告的比特错误率在 10cm 距离处约为 3%，在 100cm 距离处约为 25%。我们的实验显示出相似的趋势。尽管他们的噪声测试（使用来自附近扬声器的背景音乐）缺少详细参数，但我们在 50cm 处进行的噪声实验证实了类似的可靠性下降。

7.4.3 HRCSS。Cai等人的HRCSS [4] 最初看起来很有前景，但我们使用高质量设备进行初步测试时，提供的MATLAB实现未能通过测试。为严谨地说明为何将其排除，我们详细分析了所提供的MATLAB源代码。我们发现，较高的错误率源于同步失败。具体而言，该实现中的前导码检测算法与HRCSS论文 [4]所述的方法不一致。具体来说，代码中的错误导致接收器错误地计算并选取信号峰值，因此无法正确识别传输的起始位置。虽然我们没有尝试修复代码，但通过手动向解码器提供正确的前导码起始索引，我们验证了同步问题是根本原因；此举显著改善了传输错误率。

我们注意到，HRCSS 偶尔会取得出色的结果，但其性能在不同距离、音量和设备上的表现极不稳定。论文提到，他们在评估中使用了 Java 实现，但我们只收到 MATLAB 代码，因此无法测试 Java 版本。尤其遗憾的是，他们的研究独特之处在于使用真实设备，将其方案与现有方案（Lee等人 [33]提出的方案）进行了基准比较。

*7.4.4 Lee*等人 Lee等人 [33] 在室内远距离场景下，使用 Android 设备评估了他们的方案；在 10 m 和 25m 的距离上，数据包成功率均超过 97%。我们的实验也证实了这一高成功率（在 10 m、20 m 和 40m 的距离上均超过 98%）。此外，我们在 21 种设备组合上进行了兼容性测试：使用智能手机作为发射器，在 50cm 距离下测试，结果表明该方案具有出色的稳健性和通用性。

7.4.5 Nearby。 Getreuer 等人 [13] 使用 Nexus 6 发射器和 Nexus 5/5X 接收器测试了 Nearby，报告在最远 2 米处数据包成功率超过 94%，不过在更大距离上错误率有所上升。我们的测试证实，1 米以内错误率较低，而在 1 米和 2 米处错误较高，这可能是由于我们的重新实现、设备型号和数据包大小存在差异。另一个差异是，Getreuer 等人采用了一种纠错形式，即将每次传输重复三次，并利用冗余来纠正比特错误，但我们没有评估这种方法，因为它实际上会将吞吐量降至三分之一。不过，使用这种纠错形式时，分组错误率会有所改善。

7.5 迈向更严谨且可复现的评估

我们的评估表明，许多现有的声学数据传输系统在实际场景中难以有效工作。这凸显出一项关键需求：需要更严谨且可复现的评估方法，以妥善评估新方案如何泛化到真实环境条件。由于环境变化和文档不完整，复现先前工作面临挑战，这在历史上限制了比较分析，大多数研究都是孤立地评估新方案。

为此，我们提出了一套指导方针，以实现更严谨、可复现的评估。我们在 第 6 节 中详细介绍的黑盒分析框架，是这些原则的实用实现。我们的方法有一项重要优势，即具有通用性；由于它与系统无关，因此可应用于任何单向声学数据传输系统，从而确保其适用于评估未来的方案。我们鼓励未来的出版物采用这一框架。

7.5.1 优先进行真实环境评估，而非仿真。 仿真有助于初步设计，但最终评估必须在真实设备上进行，且测试场景应能代表目标使用场景。我们在 第 6 节 提出的方法为如何系统地测试关键实际挑战的影响提供了范例：

- **距离**。 在与预期应用相关的各种距离下进行测试。
- **设备异构性**。 使用多种发射器和接收器设备（例如，不同型号的智能手机或物联网硬件）。
- **环境噪声**。 评估系统在环境噪声（使用现场录音或白噪声）以及间歇性突发噪声（例如，设备操作产生的噪声）下的性能。
- **多径效应**。在不同大小和声学特性的环境中进行评估。

7.5.2 聚焦全面且相关的指标。传输错误率（TER）可以说是最关键的指标，因为它直接反映了系统的可靠性，这对于希望将声学数据传输方案集成到其物联网或智能手机项目中的开发者而言尤为重要。可听度和延迟等其他因素无需通过实验确定，因为它们可直接从系统设计中得知。在时间要求严格的应用中，编码和解码期间额外产生的处理延迟也可能很重要。

7.5.3 通过开放科学实践确保可重复性。为便于验证并在既有工作基础上进一步推进，我们倡导采用以下最佳实践 [2, 7, 46, 55, 61]：

- **发布软件制品**。 发布完整的实现（发射器和接收器代码），并尽可能附上清晰的运行说明和分析脚本。我们研究的论文中没有一篇

提供了代码，迫使人们进行容易出错且耗时的重新实现。我们在复现包中提供了对 Nearby [13], PriWhisper [70], 和 Lee等人的方案 [33] 的重新实现 [51]。

- **描述实验设置**。 明确说明所有系统参数和实验条件（例如发射/接收设备、音量、增益、有效载荷长度、距离、环境、背景噪声），并附上设置图片。准确定义评估指标。我们的设置详见附录A。
- **发布数据制品**。 分享原始测量数据（例如音频录音），以支持评估的复现和再现 [49]。这种透明度有助于独立验证，并使其他研究人员能够改进这项工作（例如设计更好的解码器）。我们发布了来自八种方案的11900条录制音频传输数据集，以及一个包含每条录音原始比特错误率结果的CSV文件 [51]。

7.5.4 建立标准化测试平台。 为了实现最大程度的可复现性，评估可以在标准化测试平台中进行，例如消声室。为了模拟现实条件，可以引入播放环境现场录音或白噪声的受控噪声源。

7.6 局限性

7.6.1 实现差异性。 由于声学数据传输领域缺乏研究透明度，且针对所提出方案没有公开可用的实现，我们不得不自行重新实现一些方案。这一过程十分繁琐，也容易出错。我们对 Nearby [13], PriWhisper [70],和 Lee等人提出的方案的重新实现 [33]可能与原始实现存在差异，因为相关出版物没有提供完整的实现细节。在某些情况下，我们不得不依靠合理推测，并在 附录 B 中进行了说明。此外，对于 Nearby 和 PriWhisper 的重新实现，为满足其在单次传输中的预期使用场景需求，我们采用了比原始出版物所报告的更长的消息长度。我们通过仿真和实际实验验证了重新实现，结果表明，Nearby 和 PriWhisper 能够处理这些更长的消息长度，同时保持较高的可靠性，这说明在这些更具挑战性的条件下，它们在概念上仍可复现。这一积极结果表明，我们的重新实现并未降低原始实现的性能。在实验过程中，我们的所有重新实现通常都能在其预期使用场景中保持稳健可靠。

7.6.2 环境差异性。 我们在安静的室内房间和一条走廊进行了实验，并意识到不同的环境条件可能会影响结果。为确保准确性，对于结果不一致或受到意外噪声干扰影响的实验，我们都进行了重复测试。不过，有一项测量结果格外突出：在设备型号实验中，Nearby 的表现明显差于距离实验，尽管两项实验使用了相同的设备和传输距离。这并非测量误差；唯一的区别是办公室不同。这印证了 Nearby 作者的观察：声学数据传输的可靠性对环境条件十分敏感 [13, 第 8A 节]。因此，鉴于办公室等多径环境本质上具有差异性，且无法由他人精确复现（参见 第 7.3.1 节），我们还在消声室中进行了测试；消声室是一种可控且可复现的测试环境。

7.6.3 突发噪声变异性。由于需要人工操作设备，突发噪声实验面临着环境噪声实验中不存在的若干挑战。这不仅使实验过程耗时更长——560 次传输均需人工操作——还引入了操作上的变异性，影响了实验的可重复性。我们尝试

通过每隔 2 秒触发一次突发噪声来减轻这一问题，但仍存在一些变异性。请注意，对于Gonçalves等人提出的方案，[14] 和 Lee等人提出的方案，[33], 这一设置导致每次传输所受的突发噪声暴露时间更短，因为与其他方案相比，这些方案的传输持续时间较短（Gonçalves等人的方案为 0.4 秒，Lee等人的方案为 1.1 秒，其他方案约为 5 秒；每次传输之间均间隔 2 秒）。因此，这两种方案的结果无法与其他方案的结果直接比较。

8 未来工作

根据我们的评估，我们在第 7.3 节中确定了一系列设计要求，使未来的方案能够满足实际部署的要求。具体而言，对多款智能手机的比较表明，由于音频硬件存在差异，例如麦克风和扬声器的频率响应、放大器非线性以及潜在的芯片上后处理，大多数方案难以适应不同的设备组合。对包括低成本物联网硬件和智能设备在内的各类设备的声学特性开展全面研究，有助于明确方案必须可靠运行的工作范围。这些知识可以简化未来的测试工作，并有助于选择合适的测试设备。

此外，我们发现，以往大多数研究在使用真实设备进行测试时采用了各不相同的实验设置，使得研究结果难以比较。我们的工作强调了提高研究透明度的必要性，尤其是详细说明评估的实施方式。未来的工作应致力于建立并制定声学通信方案的测试标准。我们为多种方案设计统一实验设置的方法可以作为基础（附录A）。然而，测试环境的标准化仍然面临挑战，尤其是在消声室等受控环境之外。

9 结论

我们对声学数据传输方案进行了系统性文献研究，并通过向作者索取实现、亲自重新实现部分方案，以及在真实和受控环境中使用真实设备进行评估，分析了这些方案的泛化能力。研究结果表明，声学数据传输方案往往存在可靠性问题，尤其是在数据速率较高或涉及不同设备型号时。以往研究尚未充分认识或讨论这一问题，且大多依赖理论模拟，或仅进行了有限的真实设备测试。展望未来，我们建议更加重视对实际性能的评估，并提供我们的评估方法，作为未来工作的基础。

研究问题 1. “获取或重新实现该领域提出的系统有多困难？”该领域的可重复性状况不容乐观，因为我们开展的系统性文献研究中，没有一篇出版物提供了可获取的实现。联系作者索取实现的尝试大多未能成功；由于出版物提供的信息不一致，重新实现这些系统既耗时又困难。

研究问题2。“这些方案在自组织智能手机通信使用场景中的泛化能力如何？”我们测试的大多数方案尚无法可靠地用于附近智能手机之间的实际自组织数据传输。尽管许多方案声称吞吐量很高，但在真实条件下却无法维持这一水平。不过，Lee等人的方案 [33] 即使在较远距离下也表现可靠，而非学术方案 ggwave [12] 则适用于近至中距离。

研究问题3。“哪些实际挑战会影响智能设备之间的近距离声学数据传输？”与射频通信相比，智能设备上的声学数据传输还面临其他一些常被忽视的局限性，包括长时延扩展造成的严重码间干扰，以及非高斯噪声和设备多样性带来的额外复杂性。由于仿真无法

准确捕捉这些现实世界中的复杂情况，我们在多种环境与噪声条件下，针对多种设备型号系统地测试了声学方案，以量化这些因素对可靠性和泛化能力的影响。研究结果表明，全面的实际测试有助于厘清特定用例中吞吐量与可靠性之间的权衡，从而为实际应用选择最优系统提供指导。为实现实际应用，未来的方案应能抵御数十毫秒的时延扩展，经受设备异构性导致的大幅音量波动，并处理连续噪声和脉冲噪声。

可用性

为提高研究结果的透明度，我们提供了一个复现包 [51]，其中包含：

- 一份方法透明度与可复现性附录（META），其中详细描述了我们的实验设置，包括发送与接收的具体设置（参见附录A）。
- 我们在 MATLAB 中对三种声学通信方案的重新实现 [13, 33, 70]（参见第 5 节）。
- 包含 11900 个 WAV 文件，记录了八种声学通信方案的传输内容，采集环境包括办公室环境和消声室（见第 6 节）。
- 一个包含原始数据的 CSV 文件，其中包括每种方案录音的比特错误率。
- 完整的 MATLAB 分析代码，可并行解码不同方案的录音；另附生成所有结果图的 R 脚本。

致谢

我们感谢在文献研究后应我们的请求，向我们提供声学数据传输实现的所有作者（第 5 节）。此外，我们感谢匿名审稿人提出的宝贵建议。本研究由黑森州（德国）的 LOEWE计划通过 emergenCITY中心资助 [LOEWE/1/12/519/03/05.001(0016)/72]。

参考文献

[1] Will Archer Arentz 和 Udana Bandara。2011年。面向现代智能手机的近距离超声定向数据传输。载于 第*13*届国际普适计算会议论文集（*UbiComp '11*）。计算机协会，纽约，美国纽约州，481–482。 https://doi.org/10.1145/2030112.2030181[2] Vaibhav Bajpai、Anna Brunstrom、Anja Feldmann、Wolfgang Kellerer、Aiko Pras、Henning Schulzrinne、Georgios Smaragdakis、Matthias Wählisch 和 Klaus Wehrle。2019年。《Dagstuhl 实验性网络研究可复现性入门指南》。*SIGCOMM* 计算机通信评论 49，1（2019 年 2 月），24–30。 https://doi.org/10. 1145/3314212.3314217[3] Green Bang、Myoungbeom Chung 和 Ilju Ko。2016年。基于近场不可听声音的数据通信方法。载于 *2016*年第*18*届国际先进通信技术会议（*ICACT*）。639–642。 https: //doi.org/ 10 1109 2016 7423502 。/ICACT。。Chao Cai、Zhe Chen、Jun Luo、Henglin Pu、Menglan Hu 和 Rong Zheng。2022年。动态信道条件下增强基于啁啾信号的空中声学通信。 *IEEE Transactions on Mobile Computing* 21，9（2022年9月） 2022 3110 3121 10 1109 2021 3051665 ），–。[4] https://doi.org/./TMC..Yongliang Chen、Tao Ni、Weitao Xu 和 Tao Gu。。SwipePass：面向智能手机、基于声学的双因素用户认证。 《*ACM*交互、移动、可穿戴与泛在技术论文集》 6，3（2022年9月），106:1–106:25。 https: //doi.org/ 10 1145 3550292 。/Ming Ki Chong 和 Hans W. Gellersen。2013年。用户群组如何关联无线设备。载于[5] *SIGCHI* 人机交互系统大会论文集（*CHI '13*）。计算机协会，纽约，美国纽约州， 1559 1568 10 1145 2470654 2466207 – 。 https://doi.org/./.Christian Collberg 和 Todd A. Proebsting。。计算机系统研究中的可重复性。[6] 《*ACM* 通讯》59，3（2016年2月），62–69。[7] 2016 https://doi.org/10.1145/2812803

[8] CopSonic Technologies。2025年。CopSonic——超声波认证。https://www.copsonic.com/ [访问日期：2025-03-03]。[9] CUE Audio, Inc. 2025年。Cue——创造难忘体验。https://www.connectwithcue.com/ [访问日期：2025-03-03]。[10] Anind K. Dey。2001年。理解和使用上下文。《Personal and Ubiquitous Computing》5，1（2001年2月），4–7。https://doi.org/ 10 1007 007790170019 。/s Vadim Gerasimov 和 Walter Bender。2000年。会说话的物件：利用声音实现设备间通信及设备与人之间的通信。《IBM Systems Journal》 39, 3 4 2000 530 546 530 546 10 1147 393 0530 。（），– – –。 [11] [12] 2025 [2025 03 03] https://doi.org/ . /sj. . Georgi Gerganov。。ggwave。https://github.com/ggerganov/ggwave 访问日期：- -。[13] Pascal Getreuer、Chet Gnegy、Richard F. Lyon 和 Rif A. Saurous。2018年6月。利用消费级硬件进行超声波通信。《IEEE 多媒体汇刊》20，6（2018年6月），1277–1290。https://doi.org/10.1109/TMM. 2017.2766049[14] Monira A. Gonçalves、Rafael V. Louredo、Michel P. Tcheou 和 Lisandro Lovisolo。2017年。采用正交频分复用和最小二乘信道估计的声学调制解调器。载于《2017年IEEE第9届拉丁美洲通信会议》（LATINCOM）。 1 5 10 1109 2017 8240178 – 。https://doi.org/ . / LATINCOM. . Michael T. Goodrich、Michael Sirivianos、John Solis、Claudio Soriente、Gene Tsudik 和 Ersin Uzun。2009年2月。利用音频进行安全设备配对。《国际安全网络期刊》4，1/2（2009年2月），57–68–57–68。https://doi.org/10.1504/ IJSN.2009.023426[16] Michael T. Goodrich、Michael Sirivianos、John Solis、Gene Tsudik 和 Ersin Uzun。2006年。响亮而清晰：基于音频、可由人验证的认证。载于第26届IEEE分布式计算系统国际会议（ICDCS）。IEEE。https://doi.org/ 10 1109 2006 52 。/icdcs. . Daniel Halperin、Thomas S. Heydt-Benjamin、Benjamin Ransford、Shane S. Clark、Benessa Defend、Will Morgan、Kevin Fu、Tadayoshi Kohno 和 William H. Maisel。2008年。心脏起搏器和植入式心脏复律除颤器：软件无线电攻击与零功耗防御。载于2008年IEEE安全与隐私研讨会（Sp 2008）。129–142。 https://doi.org/ 10 1109 2008 31 。/SP. . Michael Haus、Aaron Yi Ding、Qing Wang、Juhani Toivonen、 Leonardo Tonetto、Sasu Tarkoma 和 Jorg Ott。2019年。利用可见光和超声波增强室内物联网通信。载于ICC 2019——2019年IEEE国际通信会议（ICC）。 1 6 10 1109 2019 8762001 – 。 https://doi.org/ . /ICC. . Yitao He 和 Junyu Bian。。新型超声波广播通信系统。载于2018年第10届国际无线通信与信号处理会议（WCSP）。 1 7 10 1109 2018 8555902 – 。 https://doi.org/ . /WCSP. . Yann Hornych、Javier Cañada Toledo、Boyang Wang、Won-Jae Yi 和 Jafar Saniie。。使用Android 智能手机实现面向物联网应用的近超声通信。载于2020年IEEE电气与信息技术国际会议（EIT）。 407 410 10 1109 48999 2020 9208265 – 。https://doi.org/ . /EIT . . Veronika Hromadová、 Peter Kasák、Roman Jarina 和 Peter Brída。。智能手机在可听范围上限处的频率响应。载于2022 ELEKTRO（ELEKTRO）。1–5。https://doi.org/10.1109/ELEKTRO53996.2022. 9803475[22] Kwang Myung Jeon、Hong Kook Kim 和 Myung J. Lee。2016年。具有最佳符号长度的非相干低频超声波通信系统。《国际分布式传感器网络期刊》12，5（2016年5月），9713180。 https://doi.org/ 10 1155 2016 9713180 。/ / Xiaoyu Ji、Chaohao Li、Xinyan Zhou、Juchuan Zhang、Yanmiao Zhang 和 Wenyuan Xu。2020年。通过家庭受限信道验证智能家居设备。《ACM物联网汇刊》1，4，文章24（2020年8月），24页。 https://doi.org/ 10 1145 3399432 。/ Meng Jin、Yuan He、Yunhao Liu 和 Xinbing Wang。2022年。利用声学噪声隐蔽连接物联网设备。载于2022年第21届ACM/IEEE传感器网络信息处理国际会议（IPSN）。195–207。https: //doi.org/ 10 1109 54338 2022 00023 。/IPSN . . Tero Jokela、Ming Ki Chong、Andrés Lucero 和 Hans Gellersen。2015年。连接设备以实现协作交互。《Interactions》 22, 4 2015 39 43 10 1145 2776887 （6月），– 。https://doi.org/ . / Tero Jokela 和 Andrés Lucero。。FlexiGroups：通过设备触碰，将移动设备绑定起来，以支持中型群组的协作交互。载于第16届人机交互与移动设备及服务国际会议论文集（MobileHCI ’14）。计算机协会，纽约，美国纽约州，369–378。 https://doi.org/ 10 1145 2628363 2628376 。/ . Daniel Jones。。近超声音频如何为 Sonos 系统带来空间感知。 https://tech-blog.sonos. com/posts/how-near-ultrasonic-audio-adds-spatial-awareness-to-the-sonos-system/ [访问日期：2025-03-03]。[28] Soonwon Ka、Tae H. Kim、Jae Y. Ha、Sun H. Lim、Su C. Shin、Jun W. Choi、Chulyoung Kwak 和 Sunghyun Choi。2016年。面向电视第二屏服务的近超声通信。载于第22届移动计算与网络年度国际会议论文集。ACM, 42 54 42 54 10 1145 2973750 2973774 – – – 。https://doi.org/ . / . Vijay A. Kanade。。音频互联网 A-I：一种通过声音传输比特的流媒体传输新技术：SoundData。载于 2018 年计算、通信与信息学进展国际会议（ICACCI）。1690–1694。

https://doi.org/10.1109/ICACCI.2018.8554953
[30] Soohyun Kim、Hyunsu Mun 和 Youngseok Lee。2019 2019 20 1 4 10 23919。《一种声音数据应用：考勤簿》。载于第届亚太网络运营与管理研讨会（APNOMS）。–。https://doi.org/。/APNOMS。2019.8892996[31] Greg Kumparak。2013。《SlickLogin 旨在通过向智能手机唱一首无声歌曲来终结密码》。https://techcrunch.com/2013/09/09/slicklogin-wants-to-kill-the-password-by-singing-a-silent-song-to-your-smart- phone/ [访问日期：2025-03-03]。[32] Chulyoung Kwak、Seongwong Kim、Soonwon Ka、Jihwan Lee 和 Sunghyun Choi。2019年6月。《禁止进入：基于啁啾的声学通信抗噪能量检测器》。载于2019年第16届IEEE传感、通信与网络国际会议（SECON）。1 9 10 1109 2019 8824986 –。https://doi.org/。/SAHCN。。[33] Hyewon Lee、Tae H. Kim、Jun W. Choi 和 Sunghyun Choi。2015/四月。《面向智能设备的基于啁啾信号的空中声学通信》。载于INFOCOM。2407 2415 10 1109 2015 7218629 –。https://doi.org/。/INFOCOM。[34]。Jihwan Lee、Chulyoung Kwak、Seongwon Kim 和 Saewoong Bahk。。《可靠且低复杂度的基于啁啾扩频的空中声学通信》。IEEE Access 8（2020），151589–151601。https://doi.org/10.1109/ ACCESS.2020.3017097[35] Min-Chun Lin、Fu-Yu Huang 和 Tzi-Dar Chiueh。2015。《A-NFC：通过不可听声学实现双向近场通信（NFC）》。载于2015年第6届信息、智能、系统与应用国际会议（IISA）。1 6 10 1109 2015 7388085 –。https://doi.org/。/IISA。。Qiongzheng Lin、Lei Yang 和 Yunhao Liu。2017。[36]《TagScreen：通过隐藏式声音标记同步社交电视》。载于 IEEE INFOCOM 2017——IEEE计算机通信会议。1–9。https://doi.org/10.1109/ INFOCOM.2017.8057170[37] LISNR, Inc. 2025。《Radius 3.0超声波软件开发工具包》。https://lisnr.com/radius-ultrasonic-sdk-3/ [访问日期：2025-03-03]。[38] Cristina V. Lopes 和 Pedro M. Q. Aguiar。2001。《空中声学通信》。载于2001年IEEE信号处理在音频与声学中的应用研讨会论文集（Cat. No.01TH8575）。219–222。https://doi.org/10.1109/ASPAA.2001.969582[39] Cristina V. Lopes 和 Pedro M. Q. Aguiar。2003年7月。《面向普适计算的声学调制解调器》。IEEE普适计算 2, 3 2003 62 71 10 1109 2003 1228528 （/七月），–。https://doi.org/。/MPRV。。[40]Anil Madhavapeddy、David Scott 和 Richard Sharp。2003。《利用声音实现上下文感知计算》。载于 UbiComp 2003：普适计算（计算机科学讲义），Anind K. Dey、Albrecht Schmidt 和 Joseph F. McCarthy（编）。施普林格，柏林，海德堡，315 332 10 1007 978 3 540 39653 6_25 –。https://doi.org/。/ - - - - Anil Madhavapeddy、Richard Sharp、David Scott 和 Alastair Tse。/七月。2005。《音频联网：被遗忘的无线技术》。IEEE普适计算 4, 3 2005 55 60 10 1109 2005 50 （/七月），–。[41]https://doi.org/。/MPRV。。Adib Mehrabi、Antonella Mazzoni、Daniel Jones 和 Anthony Steed。2019。《声学数据传输的用户体验评估》。《Personal and Ubiquitous Computing》（2019年12月），1–14–1–14。https://doi.org/10.1007/s00779-019-01345-7[43] Jonathon Miegel、Philip Branch 和 Peter Blamey。2018。《利用高频音频和超声波在个人电子设备与助听器之间进行无线通信》。《美国声学学会杂志》 144, 4 2018 2598 2604 10 1121 1 5063813 （10月），–。[42]https://doi.org/。/。Satoshi Moriyama、Junji Takahashi 和 Yoshito Tobe。2017。《利用智能手机声音创建通信群组》。载于《智能环境》 *2017* 8 14 10 3233 978 1 61499 796 2 8 。IOS Press，–。[44]https://doi.org/。/ - - - - - Rajalakshmi Nandakumar、Krishna K. Chintalapudi、Venkat Padmanabhan 和 Ramarathnam Venkatesan。2013。《Dhwani：安全点对点声学近场通信》。ACM SIGCOMM计算机通信评论 43，4（2013年8月），63 74 63 74 10 1145 2534169 2486037 – – –。https://doi.org/。/。B. A. Nosek、G. Alter、G. C. Banks、D. Borsboom、S. D. Bowman、S. J. Breckler、S. Buck、C. D. Chambers、G. Chin、G. Christensen、M. Contestabile、A. Dafoe、E. Eich、J. Freese、R. Glennerster、D. Goroff、D. P. Green、B. Hesse、M. Humphreys、J. Ishiyama、D. Karlan、A. Kraut、A. Lupia、P. Mabry、T. Madon、N. Malhotra、E. Mayo-Wilson、M. McNutt、E. Miguel、E. Levy Paluck、U. Simonsohn、C. Soderberg、B. A. Spellman、J. Turitto、G. VandenBos、S. Vazire、E. J. Wagenmakers、R. Wilson 和 T. Yarkoni。2015。《促进开放研究文化》。《科学》348，6242（2015年6月），1422–1425。https://doi.org/10.1126 2374 。/science.aab Ed Novak、Zhuofan Tang 和 Qun Li。2019。《面向物联网应用的智能移动设备超声波近距离组网》。IEEE Internet of Things Journal 6, 1 2019 399 409 10 1109 2018 2848099 （2月），–。[45][46][47]https://doi.org/。/JIOT。。Matthew J. Page、Joanne E. McKenzie、Patrick M. Bossuyt、Isabelle Boutron、Tammy C. Hoffmann、Cynthia D. Mulrow、Larissa Shamseer、Jennifer M. Tetzlaff、Elie A. Akl、Sue E. Brennan、Roger Chou、Julie Glanville、Jeremy M. Grimshaw、Asbjørn Hróbjartsson、Manoj M. Lalu、Tianjing Li、Elizabeth W. Loder、Evan Mayo-Wilson、Steve McDonald、Luke A. McGuinness、Lesley A. Stewart、James Thomas、Andrea C. Tricco、Vivian A. Welch、Penny Whiting 和 David Moher。2021。《PRISMA 2020声明：系统综述报告指南更新》。《英国医学杂志》372（2021年3月），n71。https://doi.org/10.1136/bmj.n71[48]

[49] Prasad Patil、Roger D. Peng 和 Jeffrey T. Leek。2016。《可重复性与可复现性的统计定义》。，066803 10 1101 066803 页。https://doi.org/。/ Florentin Putz、Flor Álvarez 和 Jiska Classen。[50] 2020。《声学完整性代码：利用短距离声学通信进行安全设备配对》。载于第13届ACM无线与移动网络安全和隐私会议论文集。31 41 10 1145 3395351 3399420 2005 08572 –。https://doi.org/。/。arXiv:。[51] Florentin Putz、Philipp Fortmann、Jan Frank、Christoph Haugwitz、Mario Kupnik 和 Matthias Hollick。2025。《复现包：评估用于附近智能设备间自组织通信的声学数据传输方案》。https://doi.org/10.5281/zenodo.17661991 [数据集]。[52] Florentin Putz、Steffen Haesler 和 Matthias Hollick。2024。《听起来不错？群组中的快速安全联系人交换》。载于《ACM人机交互论文集》8（纽约，纽约州，美国）（CSCW2）。计算机协会。https://doi.org/10 1145 3686964 。/ Florentin Putz、Steffen Haesler、Thomas Völkl、[53] Maximilian Gehring、Nils Rollshausen 和 Matthias Hollick。2024。《PairSonic：帮助群组安全地交换联系信息》。载于2024年计算机支持的协同工作与社会计算会议伴随出版物（CSCW '24 Companion）。计算机协会，纽约，纽约州，美国。https://doi.org/10 1145 3678884 3681818 。/。Nafis Sadeq 和 Md. Shohrab Hossain。。《通过不可听声音侵害 Android 智能手机隐私[54]》。载于2022年IEEE无线通信与网络会议（WCNC）。2441–2446。https://doi.org/10.1109/WCNC51071. 2022.9771721[55] Kavous Salehzadeh Niksirat、Lahari Goswami、Pooja S. B. Rao、James Tyler、Alessandro Silacci、Sadiq Aliyu、Annika Aebli、Chat Wacharamanotham 和 Mauro Cherubini。2023。《CHI 2017至CHI 2022年间实证研究的研究伦理、开放性与透明度变化》。载于2023年CHI计算系统中的人为因素会议论文集（CHI '23）。计算机协会，纽约，纽约州，美国，1–23。https://doi.org/10. 1145/3544548.3580848 [56] Shigeru Sasano、Fumihito Sasamori、Osamu Takyu 和 Shiro Handa。2014。《采用重复编码的声学 OFDM 性能评估》。载于2014年国际信息与通信技术融合会议（ICTC）。798 799 10 1109 2014 6983292 –。https://doi.org/。/ICTC。。Utkalika Satapathy、Soumyajit Chatterjee 和 Sandip Chakraborty。2020。《SilentInformer：利用不可闻声学信号传播信息的框架[57]》。载于2020年国际通信系统与网络会议（COMSNETS）。543 546 10 1109 48256 2020 9027493 –。https://doi.org/。/COMSNETS 。。[58] Markus Scheck、Florentin Putz、Frank Hessel、Hermann Leinweber、Jonatan Crystall 和 Matthias Hollick。2023。《演示：利用声学通信安全引导智能音箱》。（2023）。https://doi.org/10.26083/tuprints- 00024180 在第16届ACM无线与移动网络安全和隐私会议（WiSec 2023）上展示。[59] Liming Shi、Xinheng Wang、Limin Yu、Wenwu Wang、Zhi Wang、Muddesar Iqbal、Charalampos C. Tsimenidis 和 Shahid Mumtaz。2023。《一种远距离空中声学通信方案》。《物理通信》60（2023年10月），102135 10 1016 2023 102135 。https://doi.org/。/j.phycom。。Shopkick, Inc.。Shopkick。https://www.shopkick.com/[60] [访问日期：2025-03-03]。[61] Patrick E. Shrout 和 Joseph L. Rodgers。2018。《心理学、科学与知识建构：从可重复性危机拓宽视角》。《心理学年度评论》69，第69卷，2018年（2018年1月），487–510。https://doi.org/10 1146 122216 011845 。/annurev-psych- - Stimshop。。Stimshop 的 Wi-Us。通过超声波实现无线连接[62]。https://wi-us.eu/ [访问日期：2024-11-06]。[63] 美国商业奖。2014。VeriFone 的 Way2ride。https://stevieawards.com/aba/verifones-way2ride [访问日期：2025-03-03]。[64] TrillBit Inc. 2025。《TrillBit——非接触式、快速且安全的身份验证协议》。https://www.trillbit.com/technology- trill bit [访问日期：2025-03-03]。[65] Qian Wang、Kui Ren、Man Zhou、Tao Lei、Dimitrios Koutsonikolas 和 Lu Su。2016。《声音背后的消息》。载于第22届移动计算与网络年度国际会议论文集——MobiCom '16。ACM Press。https://doi.org/10 1145 2973750 2973765 。/。Tianben Wang、Daqing Zhang、Yuanqing Zheng、Tao Gu、Xingshe Zhou 和 Bernadette Dorizzi[66]。2018。《基于 C-FMCW 声学信号的非接触式呼吸检测》。《ACM交互、移动、可穿戴与泛在技术论文集》1，4（1月 2018 170 1 170 20 10 1145 3161188 ），：–：。https://doi.org/。/ Weiguo Wang、Jinming Li[67]、Yuan He、Xiuzhen Guo 和 Yunhao Liu。2022年3月。《MotorBeat：通过可变脉冲宽度调制实现家用电器声学通信》。《ACM交互、移动、可穿戴与泛在技术论文集》6，1（2022 1 24 1 24 10 1145 3517255 /三月），– – –。https://doi.org/。/ Yu Wang、Xiaojun Zhu 和 Hao Han[68]。。《ChirpMu：利用音乐进行基于啁啾的不可感知信息广播》。载于2021年IEEE/ACM第29届国际服务质量研讨会（IWQOS）。1–9。https://doi.org/10.1109/ IWQOS52092.2021.9521353

[69] Katsuyoshi Yamamoto 和 Hiroshi Kubo。2022年。采用发射预编码和信道插值的单载波块传输实现 32kbps 地面声学通信。*IEICE CommunicationsExpress* 11, 3 (2022年), 141 147 10 1587 2021 0206 – 。https://doi.org/ 。/comex. XBL[70]Bingsheng Zhang、Qin Zhan、Si Chen、Muyuan Li、Kui Ren、Cong Wang 和 Di Ma。2014年。PriWhisper：为智能手机实现无密钥安全声学通信。*IEEE Internet of Things Journal* 1, 1 (2014年2月), 33–45。https://doi.org/ 10 1109 2014 2297998 。/JIOT. 。Xiao Zhang、Jiqiang Liu、Si Chen、Yongjun Kong 和 Kui Ren。2019[71]年。PriWhisper+：一款面向智能手机的增强型短距离声学通信系统。*IEEE Internet of Things Journal* 6, 1 (2019年2月), 614–627–614–627。https://doi.org/ 10 1109 2018 2850524 。/jiot. 。Yanchao Zhao、Yiming Zhao、Si Li、Hao Han 和 Lei Xie。2023[72]年。UltraSnoop：通过基于智能手机的超声波声呐实现与位置无关的击键窃听。*ACM*物联网汇刊 4, 4, 文章 22（2023年11月），24 页。https: //doi.org/ 10 1145 3614440 。/Man Zhou、Qian Wang、Kui Ren、Dimitrios Koutsonikolas、Lu Su 和 Yanjiao Chen。[73]2019年。Dolphin：利用智能手机实时捕获隐藏声学信号。*IEEE Transactions on Mobile Computing* 18, 3（2019年3月），560–573–560–573。https://doi.org/10.1109/tmc.2018.2842771

A 方法论透明度与可复现性附录

本节详细介绍我们的实验设置，旨在提高研究工作的透明度和可重复性。我们记录了发送与接收所用的准确设置。 表 5 列出了我们使用的所有硬件和软件，包括版本号。

A.1 概述

对于每项实验，如果受到外部因素影响（例如人声或噪声干扰），或在传输的播放和录音过程中出现问题，我们都会重复测量。

A.2 传输

按照 图 3所示的设置，我们使用了 表 5中列出的智能手机。首先，我们为每种方案生成传输 WAV 文件，并将其纳入数据集。消息长度设为论文默认值；如果实现支持自定义消息大小，我们则对其进行调整，使传输时间约为五秒，我们认为这对于用户等待时间而言是切实可行的上限。随后，将所有音频信号除以整个信号中的最大绝对采样值进行归一化，并缩放至 −3dBFS。

我们使用开源 Android 应用 *Vinyl*在每部智能手机上传输这些文件，该应用支持通过播放列表自动播放多个 WAV 文件。我们编制了一个播放列表，其中包含所有方案的传输文件，并在文件之间插入时长两秒的静音 WAV 文件，以防止混响影响后续传输。为方便起见，我们还提供了一个 WAV 文件，其中包含所有方案的重复传输以及相应的静音间隔，便于播放。我们使用开源 Android 应用 *VolumeControl*控制输出音量；该应用可精确控制 Android 音量索引，如表 3所示。

A.3 接收

在接收端智能手机上，我们使用开源 Android 应用录音机录制每次传输，该应用支持对麦克风源进行详细控制。除 Google Pixel 4a 外，我们为所有智能手机选择了 MIC 源；Google Pixel 4a 还支持未处理源，可绕过 Android 的后处理。录音以 24 位脉冲编码调制浮点单声道 WAV 文件格式保存，采样率至少为 44.1 千赫，软件增益设为 100%，并禁用所有额外滤波。

我们连续录制了多次传输，之后在*REAPER* 数字音频工作站中手动分段。为简化这一流程，我们使用了一个脚本[9] 进行批量剪切

[9]用于分割的 REAPER 脚本：https://stash.reaper.fm/v/25677/SplitX.lua

以及一个用于生成唯一文件名的自定义迭代器。这些片段以至少 44.1 千赫的采样率渲染为 32 位脉冲编码调制单声道 WAV 文件。最后，我们使用各方案的接收器实现对这些文件进行解码，以计算传输错误率。如果解码后的消息长于预期，我们就将其截断；如果短于预期，则将缺失的比特计为错误。

A.4 实验设置

我们在多个室内地点和一个消声室中进行了测试。在室内环境中进行实验时，我们将所有智能手机安装在带橡胶脚垫的小型三脚架上，使其与智能手机扬声器产生的振动隔离。我们调整智能手机的位置，使其扬声器和麦克风彼此正对（同轴），并测量两者之间的距离。室内空间的尺寸（长 × 宽 × 高）如下：小型办公室 （5.2 m × 4 m × 3 m）、大型办公室 （13.5 m × 5.3 m × 2.8 m）、会议室 （7.9 m × 5.3 m × 3 m）、报告厅（13.5 m × 6.6 m × 3.6 m），以及用于长距离测试的走廊 （40m × 2 m × 3 m）。

消声室 （5.2 m × 7.7 m × 5.8 m）提供了无反射环境，其墙壁覆盖着涂有合成树脂的矿物棉锥体（锥体长度为 100 cm，锥底面积为 24 cm × 24 cm，对于高于 100Hz 的频率，反射系数低于 0.01）。

A.4.1 初步最佳情况设置。 对于初步测试（第6.2节），我们使用录音室设备而非智能手机建立了最佳情况基线。在传输方面，一台 Neumann KH 80 DSP 录音室扬声器播放信号（本地控制关闭，声学控制为“独立放置”，输出电平 94 dBSPL，输入增益为 0 dB）。我们将扬声器连接到 RME ADI-2 Pro FS R BE 音频接口（−40 dB 增益，参考电平为 24dBu）。

在录音方面，一支经过校准的 Earthworks M23R 参考麦克风采集音频。信号经 Camden EC1 前置放大器（35.5 dB 增益，无额外滤波）送入音频接口（无额外增益）。前置放大器的输出被分配到接口的两个输入端，并使用中侧处理合成为单声道，从而使信噪比提高 3dB。我们使用 REAPER 进行播放和录音。扬声器与麦克风同轴对齐，距离为 50cm。

A.4.2 环境噪声设置。在环境噪声实验（第 6.5 节）中，我们使用标准智能手机，并通过 Neumann KH 80 DSP 录音室扬声器引入噪声。扬声器放置在距接收器 1 米处，与传输轴垂直。扬声器的配置与初步测试（第 A.4.1 节）相同，唯一不同的是接口的输出增益设为 −35dB，以便在接收器所在位置主观匹配实际场景中的噪声水平。在每次实验迭代中，重复播放一段现场录音，同时依次播放传输声音。对三段录音（咖啡馆、火车站、集市）均重复此过程；这些录音可在网上免费获取[10]，采用知识共享署名许可。

A.4.3 动态环境设置。为评估时变信道中的性能（第 6.6 节），我们在报告厅进行了测试，期间有四个人在设备周围缓慢、安静地走动（不交谈）。智能手机安装在小型三脚架上，三脚架放置于房间中央的桌子上，两部手机沿轴向相距 50 厘米。作为对比，我们在家具布局相同但没有人的情况下进行了静态对照测试。

[10]咖啡馆现场录音（知识共享署名 3.0）：https://freesound.org/people/AshFox/sounds/172968/，火车站现场录音（知识共享署名 4.0）：https://freesound.org/people/Kyster/sounds/121576/，市场现场录音（知识共享署名 3.0）：https://freesound.org/people/le_abbaye_Noirlac/sounds/129677/

表 5. **我们实验中使用的硬件和软件**。

Type	制造商	产品	版本
麦克风	Earthworks	M23R	
前置放大器	Camden	EC1	
音频接口	RME	ADI-2 Pro FS R BE	
校准设备	Galaxy Audio	CM-C200	
扬声器	Neumann	KH80 DSP	
智能手机	谷歌	Pixel 4a	128GB
智能手机	谷歌	Pixel 6 Pro	5G Sub-6 128GB
智能手机	华为	Nexus 6P	32GB
智能手机	Oppo	Reno 6	双卡 5G 128GB
智能手机	三星	Galaxy S20 Ultra	5G 128GB
桌面麦克风支架	K&M	23150-3	
球形关节	Roadworx	通用球形关节	
智能手机夹	Roadworx	智能手机夹	
智能手机应用	Adrien Poupa	黑胶音乐播放器[1]	v1.5.0
智能手机应用	Stas Shakirov	音量控制[2]	v2.6.0
智能手机应用	axet	录音机[3]	v3.5.15
DAW	Cockos	REAPER[4]	v6.82

[1] Vinyl：https://f-droid.org/en/packages/com.poupa.vinylmusicplayer/ [2] 音量控制：https://f-droid.org/packages/com.punksta.apps.volumecontrol/ [3] 录音机：https://f-droid.org/en/packages/com.github.axet.audiorecorder/ [4] REAPER：https://www.reaper.fm/

A.5 复现实验

上述流程应有助于复现实验。不过，声学传输可靠性会受到特定设备型号、麦克风/扬声器质量以及环境声学等因素的影响。尽管我们预计总体趋势相似，但 TER 值可能会有所不同。

B 重新实现

本节详细介绍了我们在重新实现 Lee等人的方案、Nearby 和 PriWhisper 时遇到的挑战。

B.1 Lee等人

我们能够按照论文中的描述实现该方案的所有方面 [33]。他们将大多数系统参数清晰地列在表格中的做法尤其有帮助。我们将 Lee等人提到的 16比特/秒数据速率解读为毛数据速率，因为他们似乎没有在计算中考虑前导码。根据提供的默认系统参数，我们计算得出的净数据速率为 14.55 比特/秒。

B.2 Nearby

为 Nearby 系统实现发射器 [13] 较为直接，只需遵循论文所提供的信息。然而，重新实现接收器颇具挑战，因为论文中的某些方面对我们而言并不清晰。

B.2.1 同步。 为了在录音中定位传输的起始位置，系统通过遍历采样点来处理信号，从每个采样偏移处对信号进行分段，并将这些分段与一个周期的码信号进行相关。该相关结果表示一个符号在每个采样偏移处开始的可能性，因为每个 MFSK 符号都由码信号进行了扩频。在对所有时间偏移完成相关后，会对结果进行归一化，以补偿任何潜在的多普勒频移。这些结果在其论文中被称为“归一化捕获分数”（第 VI.C 节，见 [13]）。接下来，我们利用归一化捕获分数确定间隔符号的位置。理想情况下，绘制这些分数时，间隔符号的位置会出现一个明显的峰值。然而，由于相同的码信号对所有令牌符号进行扩频，每个符号都会出现峰值，而不仅仅是间隔符号。我们实现了一种方法来识别 n 个最高峰值的位置——其中n 表示令牌符号的数量——并选择最早的峰值作为起始位置。论文并未讨论这一问题，也未说明如何解决。

B.2.2 分块处理。 作者提到，他们的实现会以 100 毫秒为单位分块处理信号，这可能是为了提高时间效率，但论文并未详细说明这一功能是如何实现的。我们选择不实现这一部分，因为它不太可能影响评估结果。

B.2.3 算法 *2* 中的歧义。 论文中的算法 2 [13] 在第 2 行使用一个计数器变量 n ，该变量初始值为零，并在每次迭代时递增一，用于计算每个采样偏移 n对应的“原始捕获分数”，表示数据符号从该偏移处开始的可能性。然而，作者没有指定 n的上限。我们已在接收器函数的注释中记录了确定该上限的方法，但尚不清楚该方法是否符合作者的原意。

B.2.4 公式 *47* 中的歧义。 论文公式 47 中，表示经 sinc 插值的码序列的变量 c在位置 $(kM_p + m)/F_b$处进行索引。此处， F_b 表示 12 千赫的采样率，这意味着相除所得结果是码信号应进行索引的时间（单位为秒）。由于 c 是以 48 千赫采样的离散信号，因此无法在任意时刻进行索引。要将时间 $(kM_p + m)/F_b$ 转换为 c中对应的采样偏移，我们将其乘以 48 000（即采样率）。同样，这是否是作者的意图尚不明确。

B.2.5 算法 *3* 中的歧义。 论文 [13]中的算法 3 使用由算法 2 生成的原始捕获分数矩阵 a^{raw}。在算法 3 的第 8 行，它访问 a^{raw} 的第 $n - d$ 行，其中n 是一个从零开始、每次迭代加一的计数器变量。论文未指定 n的上界，但我们假设 n 最多运行到 a^{raw} 中的行数减一。变量 d 被设为 215，导致算法在前 215 次迭代中试图访问负的行索引，这是无效的。我们通过在 $n - d$ 产生负索引的迭代中省略第 8 行来解决此问题。然而，其结果是，归一化捕获分数矩阵的行数比 a^{raw}少 215 行。目前尚不清楚这种减少是否符合作者的意图。

B.2.6 噪声抑制期间的歧义。 用于噪声抑制的带通滤波器使用变量η，该变量应为“一个较小的正参数”（ [13]中的第 VI.A 节）。由于论文未为 η指定确切值，我们将其设为 1。

B.2.7 数据速率。 我们将所述的 94.5 比特/秒数据速率理解为毛数据速率。系统可以在不使用包含额外间隔符和校验符号的令牌结构的情况下运行，此时 94.5 比特/秒也是净数据速率。然而，当使用论文中所述并在我们的评估中采用的完整令牌结构时，净数据速率降至 84 比特/秒。

B.3 PriWhisper

在重新实现 PriWhisper 时 [70], 主要困难在于设计同步与校准机制，因为论文并未清晰说明这些方面。此外，论文中对交织和校验和过程的描述含糊不清。所指定的数据速率也不明确，进一步增加了我们实现工作的难度。

B.3.1 同步。 在原始实现中，干扰信号中使用正弦波前导码来辅助同步。然而，在我们的研究中，我们没有实现系统的安全功能，而是专注于数据传输。为了将对数据速率的影响降至最低，我们将同步和校准序列合并。因此，我们在每次传输的开头引入了一个特殊符号，其中包含所有可用数据波形 $s(t)$（ [70] 中的 III.A 节）之和。为了进行同步，我们使用这个已知的同步符号，并将其与录音进行互相关。输出中最高峰的位置被视为传输的起点。

B.3.2 校准。 为了确定 MFSK 符号的频率，我们使用以下 [70] 中 III.A 节的公式，将录音中的每个符号持续时长与所有可用频率进行相关：

$$R_m = \left| \int_0^T r(t) e^{i2\pi(f_c + m\Delta f)t} dt \right|, m \in [0, M-1] \tag{1}$$

遍历所有频率后，我们选择相关值最高的频率，并将其映射到该频率所表示的二进制数。对于所有符号，我们重复这些步骤，以解码录音中的整个数据传输过程。在解调之前，使用同步符号进行校准至关重要，因为麦克风在不同频率下的灵敏度可能存在差异。若不进行校准，公式 1 可能会因麦克风的频率选择性，对符号中实际存在的频率返回较低的相关值。

为解决这一问题，我们将公式 1 应用于包含所有频率的同步符号。这样即可为每个频率建立基准相关值。然后，我们将这些基准值中的最高值分别除以其他值，从而计算出每个频率的校准因子。在解调过程中，该因子会调整相关值，确保尽管麦克风存在偏差，每个频率仍能得到适当表示。然而，这一校准步骤假设信道特性在整个传输持续时间内保持不变。

B.3.3 交织和校验和。 论文提到对数据进行交织并添加校验和，但没有明确说明如何实现交织过程以及这些步骤的顺序。在我们的实现中，我们先执行常规的随机交织，然后添加校验和。

B.3.4 数据速率。我们计算得出 $M = 8$ 的毛数据速率为 771 比特/秒，与作者报告的 1027bps 不同。数据速率的计算如下：使用 M 个不同频率，可编码 $log_2(M)$ 比特/符号。对于 $M = 8$，这相当于每个符号 3 比特。符号时长为 2 毫秒时，每秒可传输 500 个符号，相当于每秒 1500 比特。由于每个编码数据块包含 255 个 BCH 编码比特，因此每秒可传输 $\frac{1500}{255} \approx 5.88$ 个数据块。每个数据块包含 131 比特纯数据，由此得到毛数据速率为 $5.88 \times 131 \approx 771$bps，不包括用于同步或校准的信号分量。最终净数据速率为 729 比特/秒。