改进 协作式智能交通系统中的不可链接性：最优混淆方法 n

Yevhen Zolotavkin[1] [a]、Yurii Baryshev[2] [b]、Vitalii Lukichov[2] [c]、Jannik M¨ahn[1] [d] 和 Stefan K¨opsell[1] [e]Barkhausen Institut gGmbH，德国德累斯顿 Würzburger Straße 46[2]文尼察国立技术大学信息保护系，乌克兰文尼察赫梅利尼茨基公路 95 号{yevhen.zolotavkin，jannik.maehn，stefan.koepsell}@barkhauseninstitut.org， {yuriy.baryshev，lukichov.vitalyi}@vntu.edu.ua

arXiv:2301.04130v1 [cs.CR] 10 Jan 2023

关键词： 隐私、车联网、不可链接性、隐马尔可夫模型、网络安全、熵、混淆

摘要： 本文提出一种新方法，以高置信度保障协作式智能交通系统（C-ITS）中的隐私。我们聚焦于由协作感知基本服务支持的车联网（V2X）通信。通过分析不可链接性保障方法，我们发现了其中的不足，并据此确立了研究动机。为解决这一问题，我们提出一种隐马尔可夫模型（HMM），用于描述两辆汽车通过协作感知消息（CAM）与路侧单元（RSU）通信时的不可链接性。该HMM包含带标签的状态，用以表示被动攻击者可观测到的协作感知消息（CAM）的不同来源。随后我们证明，在假设攻击者知道车辆实际位置（例如HMM中的隐藏状态）的情况下，可以高置信度地评估攻击者对带标签状态的不确定性程度（例如熵）。我们进一步展示了如何提升C-ITS中的不可链接性：我们提出一种联合概率分布，要求两位驾驶员共同使用该分布来混淆各自的实际数据。混淆后的数据随后封装在他们的协作感知消息（CAM）中。最后，我们将研究结果融入一种混淆算法，其复杂度与HMM中的离散时间步数量呈线性关系。

1 引言

近几十年来，交通系统迅速发展，各种协作智能模式已融入其功能之中。智能交通系统（ITS）是应用先进的信息、通信、传感器和控制技术（包括互联网）的交通系统，旨在提高安全性、可持续性、效率和舒适度。协作式智能交通系统（C-ITS）是一类 ITS 技术，其服务的提供依赖于或受益于使用来自其他具备类似功能的实体的、与当前情境相关的实时动态数据/信息，和/或交通网络不同组成部分（包括车辆和基础设施）之间的数据/信息（ISO/TR 17427-7:2015(E), 2015）。

使车辆能够与基础设施、其他车辆及其他利益相关方交换附加信息的技术，在协作式智能交通系统（C-ITS）领域称为车联网（V2X）。现代协作式智能交通系统应用的诸多进展，例如协同前向碰撞预警和紧急电子制动灯，若没有车联网便无法实现。然而，这些进展也有代价：协作式智能交通系统应用依赖车辆广播信号来表明其位置，而这些信号旨在由一系列其他设备接收和处理。例如，车辆可以通过短程协作感知消息（CAMs），以 1—10 Hz 的频率向附近的车辆协作广播地理空间信息。因此，车联网引发了重要的隐私问题：i）能在多大程度上根据此类信息定位和跟踪特定车辆？ii）有哪些技术能够改善车联网的隐私保护？为回答这些问题，我们运用不可链接性这一概念来探讨隐私问题。

1.1 研究动机

尽管多份相关文件（兼具规范性和说明性）已承认协作式智能交通系统中的隐私问题（

[a] https://orcid.org/0000-0002-1875-122X
[b] https://orcid.org/0000-0001-8324-8869
[c] https://orcid.org/0000-0002-3423-5436
[d] https://orcid.org/0000-0003-0870-7193
[e] https://orcid.org/0000-0002-0466-562X

但上述隐私问题至今仍未得到令人满意的解答。例如，文件（ISO 24102-1, 2018）承认，将私人数据与无线消息中的可追踪地址元素和标识符解除关联十分重要，而这些无线消息由 ITS 站台单元（ITS-SU）发送。然而，该文件中的相关考量仅止于建议“可通过假名实现此类解除关联”，至于这些建议以及许多类似建议是否充分，仍未得到探讨。相比之下，学术界作者已认识到协同感知消息中的假名变更存在局限（Escher 等人，2021；Karim Emara，2013；Wiedersheim 等人，2010）。具体而言，由于协作式感知消息的内容经过签名但未加密，车辆追踪因此成为可能：协作式智能交通系统的相关标准要求如此（DS/ETSI EN 302 637-2 V1.4.1, 2019）。这是因为对协同感知消息进行完全加密，可能会妨碍协作式智能交通系统中对安全至关重要的功能。尽管如此，人们虽然已认识到协作式智能交通系统中影响隐私的问题，却仍未找到能够正确衡量隐私程度并消除隐私局限的解决方案。本文提出了一种应对这些挑战的方法：我们为估算协同感知消息不可关联性下界的流程提供保障，并提出一种算法，在位置精度下降这一总体约束下最大化该指标。

1.2 研究原则

为了解决车联网隐私问题，必须获得置信度较高的结果。本文提出的方法建立在鲁棒优化（RO）的原则之上（Gorissen 等人，2015；Sniedovich，2016）。以往尝试回答上述“车联网隐私问题”之所以未能成功，主要是因为针对协作式智能交通系统（C-ITS），尤其是车联网场景，所作的隐私假设不恰当。例如，假设攻击者只能观察包含经混淆（失真）地理位置测量值的协同感知消息，且没有精确的车辆运动物理模型，这种假设组合可能需要运用目前已知的最佳技术来估计此类模型（Blackman，1986）。由于这类估计存在计算复杂性，研究人员往往依赖启发式且缺乏充分依据的步骤：当需要高水平的隐私保障时，这是不可接受的（Blackman，2004；Wiedersheim 等人，2010）。为避免这种情况，我们对攻击者已知信息作出更强的假设，从而能够估计协作式智能交通系统中不可链接性的下界：这一估计具有较高的置信度。我们的这一思路体现在以下塑造本文方法的原则中：

•隐私风险评估方法不应低估风险：它应清晰、定量地评估在涉及多个用户的多个车联网通信会话中，提取属于同一实体的数据有多容易。因此，该方法应为此类数量的评估设定合理的界限；

• 该方法应提出对原始用户数据的最优修改。“最优性”的概念取决于修改类型。对于可逆修改（如加密），应保证非法方在计算上无法进行此类逆向操作（例如，此类保证的程度达到最大）。对于不可逆修改（如加噪），应量化数据质量的损失：对于每个此类数量，都应保证无法达到更高程度的不可关联性（例如，所提出的加噪方法是最优的）。

1.3 我们的贡献

本文的独特贡献源于研究目标（以不可链接性为准则）、稳健的假设以及本文提出的最优混淆技术三者的结合。

• 首先，本研究旨在保护协作式智能交通系统免受链接威胁：这与众多旨在阻碍对 ITS-SU 实际位置进行推断的混淆方法形成对比（Andr´es 等，2013年；Bordenabe 等，2014年）；

• 其次，为了对测得的不可链接性具有较高置信度，我们假设攻击者完全了解系统设计、混淆算法和质量退化（失真）约束，能够访问协同感知消息、车辆地理位置的存储状态，以及刻画车辆在任意时刻地理位置的真实状态；

• 第三，我们提出一种最优混淆算法：在给定失真约束的情况下，该算法能为试图将混淆后的协同感知消息与其来源进行链接的攻击者提供最高程度的不确定性。

本文其余部分结构如下。第 2 节奠定本研究的基础并给出基本定义。随后第 3 节将

从通用信息系统的描述入手。然后通过增加集合和关系来体现协作式智能交通系统的具体特征，由此得到用于研究不可链接性的隐马尔可夫模型（HMM）。第 4 节将对假设进行形式化，通过熵定义不可链接性，并优化联合混淆，以生成 HMM 中的可观测状态。接下来，第 5 节将介绍一种紧凑高效的算法，用于计算协作式智能交通系统中的不可链接性指标，并运用前述研究成果提升不可链接性。最后，我们将在第 6 节讨论研究结果及其重要性、创新性、优势和局限性。

2 预备知识

为更好地论证后续建模步骤，我们先介绍相关背景信息，以支撑本文的研究目标、设定和隐私假设。

2.1 研究目标

为明确研究目标，我们分析相关的网络安全要求。在此，我们借鉴复杂系统隐私领域的一些经典定义，以更精确地阐明研究目标。C-ITS 的隐私要求通常源自 ISO/IEC 15408-2。例如，(ETSI TS 102 941 V2.1.1, 2021) 指出，假名性与不可链接性的结合，可为基本的 ITS 安全消息（如 CAM）提供适当的发送者隐私保护。简单来说，假名性要求用户的身份绝不被披露或推断。然而，处理假名性时的一项重大难题是：攻击者可能根据多个会话、事件或轨迹，获知用户身份的构成。不可链接性则是指确保能够抵御对这种构成的获知（ISO/IEC 15408-2:2022(E), 2022）：

定义1（操作不可链接性）
要求用户和/或主体无法确定系统中的某些特定操作是否由同一用户导致，或这些操作是否以其他方式相关。

在拥有众多用户的协作式智能交通系统（C-ITS）车联网通信场景中，由这些用户控制的智能交通系统用户（ITS-SUs）所广播的加密签名消息应具备定义1所述的属性（Hicks和Garcia，2020；ETSI TS 102 941 V2.1.1，2021）。不过，这种解释存在一些缺点，其中最主要的是缺乏灵活性。实际上，“……无法确定……”这一表述可能为真，也可能为假，这意味着整个协作式智能交通系统（C-ITS）（涉及许多车辆，且可观测时间长达许多小时）的不可链接性是用二元值表示的。业界从业者和研究人员都已注意到这一问题，这一点反映在面向ITS工程师和管理人员的评论及最佳实践建议中。例如，（ISO/SAE 21434，2021）包含一张名为“隐私影响评级标准示例”的表格，其中列出影响评级标准，用以解释隐私影响指标各严重程度等级（如可忽略、中等、重大、严重）的含义。值得注意的是，该表中的解释围绕两个方面展开：a）关于道路使用者的信息的敏感程度；b）这些信息与个人身份信息（PII）主体建立链接的难易程度。对链接难易程度的强调促使我们将定义1修改如下：

定义2（操作不可链接性*）指用户和/或主体无法确定系统中的某些特定操作是否由同一用户导致，或这些操作是否以其他方式相关的程度。

为了便于比较不同条件下协作式智能交通系统中的不可链接性，我们采用香农熵：它是衡量系统不确定性的综合指标，能够充分体现“……无法确定的程度……”（Wagner and Eckhoff，2018年）。

因此，本文的主要目标是制定一种方法，确保协作式智能交通系统中协同感知消息来源的熵具有较高的保证水平。

2.2 研究设置

在图 1 中，我们介绍了本研究的一般设置。在图 1(a) 中，两辆车（ITS-SU）分别由 Alice 和 Bob 驾驶。两辆车都以相同的频率发送 CAMs，路侧单元（RSU）无损接收这些消息。攻击者由 RSU 扮演，其试图将 Alice 的 CAMs 与 Bob 的 CAMs 区分开来，从而将属于同一实体的 CAMs 关联起来。虽然 CAMs 经过签名，但在本研究中，我们假设使用了能够提供不可链接性的签名方案。因此，区分依据是 CAMs 的内容（因为它们以明文形式传输）以及每个时间间隔内 CAMs 的到达顺序——见图 1(b)。一般而言，我们认为 CAMs 的到达顺序并非均匀分布：例如，在一种极端情况下，Alice 的 CAM 总是先到达，而 Bob 的 CAM 在任意时间间隔 i 中都后到达。如果攻击者知道这一独特属性，就可以不考虑有效载荷而关联 CAMs。然而，这类情况并不常见，这意味着攻击者还应能够根据 CAM 的内容推断其

来源（例如“来自 Alice”或“来自 Bob”）。CAMs 内容方面的要求见（DS/ETSI EN 302 637-2 V1.4.1, 2019）。具体而言，我们认为地理位置、速度和加速度是必不可少的。一方面，这些参数是 CAM 中高频容器的必备参数。另一方面，已有大量技术利用这些参数处理多目标跟踪（MTT）问题：相应的估计器对于探讨协作式智能交通系统（C-ITS）中的隐私问题大有裨益（Blackman, 1986; Karim Emara, 2013年）。

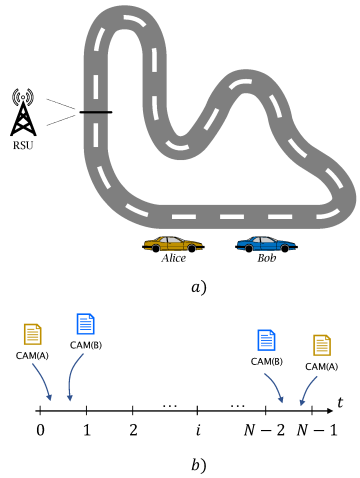


图 1：我们研究 CAM 不可链接性的场景 .

本研究关注的核心是协同感知消息（CAMs）内容的不可链接性。我们不再进一步考虑以下 CAM 载荷：1）通过密码学手段生成的真实性证明（例如签名）；2）分类数据（例如车辆角色）。之所以排除这些内容，是因为密码学界成员已对问题“1）”给予了大量关注。例如，已有研究提出了伪名不可链接方案（Camenisch等人，2020；Hicks和Garcia，2020）。然而，我们的研究仍有必要对这些工作加以补充：由于安全原因，CAMs 中没有加密，因此，实现伪名不可链接性虽有必要，却不足以确保 CAMs 的不可链接性。这是因为 CAMs 中的其他数据，例如地理位置，也可能被用于链接。问题“2）”可以忽略，因为分类数据属于低频基本车辆容器，且在 CAMs 中为可选项（DS/ETSI EN 302 637-2 V1.4.1, 2019）。我们还排除车辆长度和车辆宽度（高频容器中为必填项），否则这些信息很可能有助于区分不同车辆（Escher 等人，2021）。我们认为，这种排除有以下依据：

（ETSI TS 102 894-2V1.3.1, 2018）规定，如果相应信息不可用，长度和宽度分别可以使用代码 1023 和 62。

鉴于上述细节，我们将 CAM 建模为 $\mathbb{R}^z$ 中的一个向量，其中 $z \geq 1$。这样做很有益：我们可以采用常用的失真度量，例如平方误差（SE）。这是描述基本位置服务质量退化的一种清晰而直接的方式（Shokri等人，2016）。不过，本文不再进一步讨论所选的失真度量。

2.3 隐私假设和威胁

这里，我们将从整体上介绍系统和可链接性威胁，具体细节将在后续章节中说明。Alice 和 Bob 协同合作，将允许的总失真分配到 $N-1$ 个时间步上：因此，他们知道每个时间步 i 的失真限值。在每个时间间隔 i 开始时，Alice 和 Bob 都知道对方的真实测量值（包括位置、速度、加速度等）。为了混淆各自 CAM 中的数据，他们会随机商定每个 i 到达路侧单元的先后顺序。对于每个 i，他们会定义一个联合分布，并根据该分布更改（混淆）各自的实际测量值：从期望上看，测量值仍处于失真限值之内。

完全控制路侧单元的攻击者，会对其在时间 i 观测到的每一对协同感知消息的来源进行统计推断：此统计推断用于计算熵，并符合定义2。为此，攻击者会参考 Alice 和 Bob 在混淆过程中使用的联合分布。他还掌握其他信息，例如参与者在每个 i 时刻的原始地理位置，以及协同感知消息的到达顺序概率。系统中的不可链接性取决于：i）参与者发送的协同感知消息的到达顺序统计数据；ii）允许的失真程度；iii）Alice 和 Bob 在每个 i 时刻的实际测量值相距多远。

3 数学模型

我们将在以下章节中解释我们的数学模型。为便于阅读，表 1 概述了我们的符号。

表 1：符号

符号	描述
ITS	智能交通系统
协作式智能交通系统	协作式智能交通系统
ITS-SU	ITS站单元（包括安装在车辆中的单元）
V2X	车联网
CAM	协作感知消息
RSU	路侧单元
HMM	隐马尔可夫模型
$\mathbf{D}_u$	用户相关数据集
$\mathbf{D}_s$	信息系统相关数据集
$\mathbf{U}$	信息系统用户集
$\mathbf{P}$	信息系统中的数据处理过程集
$\mathbb{P}$	包含 Alice 和 Bob 的参与者集合
$\mathbf{x}_k^A, 1 \le k \le \mu$	Alice 的一个隐藏状态
$\mathbb{X}^A = \{\mathbf{x}_k^A\}$	Alice 的隐藏状态集合
$\mathbf{x}_j^B, 1 \le j \le \omega$	Bob 的一个隐藏状态
$\mathbb{X}^B = \{\mathbf{x}_j^B\}$	Bob 的隐藏状态集合
$\mathbb{X}^{(A,B)}$	⟨Alice 和 Bob 的联合隐藏状态集合⟩
$\mathbb{X}^{(B,A)}$	⟨Bob 和 Alice 的联合隐藏状态集合⟩
$\mathcal{R}$	玫瑰节点的索引（标签）
$\mathbb{X}_{\mathcal{R}} = \mathbb{X}^{(A,B)}$	所有玫瑰节点的集合
$\mathcal{B}$	蓝色节点的索引（标签）
$\mathbb{X}_{\mathcal{B}} = \mathbb{X}^{(B,A)}$	所有蓝色节点的集合
$\mathbb{L} = \{R, B\}$	编码 $\lvert\mathbb{P}\rvert!$ 组合的标签集合
$\mathbb{Y}$	Alice 和 Bob 的联合可观测状态集合
离散隐马尔可夫模型中的第 $i \in \{1,2,\ldots,N-1\}$ 护时间步	
$\mathbf{X}_i^A$	第 i 步的 $\mathbb{X}^A$ 变量
$\mathbf{X}_i^B$	第 i 步的 $\mathbb{X}^B$ 变量
$\mathbf{X}_i$	第 i 步的联合隐藏状态变量
ℓ_i	第 i 步的 $\mathbb{L}$ 变量
$\mathbf{Y}_i$	第 i 步的 $\mathbb{Y}$ 变量
$\Pr(\mathbf{X}_{i+1} \mid \mathbf{X}_i)$	隐藏状态之间的转移概率
$\Pr(\mathbf{Y}_i \mid \mathbf{X}_i)$	观测状态的条件概率
φ	顺序混合（标签置换）概率
ρ_i	第 i 步隐藏状态上的分布
$\rho_{i+1} \mid \rho_i$	第 i 步隐藏状态上的条件分布 + 1

3.1 模型设置

我们考虑存在被动攻击者（Eve）的信息系统的一般情况。信息系统会为某些用户处理数据集 $\mathbf{D}_u$，这些用户构成集合 U。为此，信息系统会执行一组数据处理算法 P：数据 $\mathbf{D}_s$，用于描述这些数据处理算法的配置和参数。任一数据处理算法 $\forall \mathrm{p} \in \mathrm{P}$ 都由用户或信息系统的状态触发，该状态由其配置和参数描述。执行数据处理算法 p 时，会将特定用户数据作为输入，从而创建新的用户数据或更改现有用户数据，同时也会改变信息系统的配置和参数。在某些情况下，系统中的数据处理算法甚至可以更改用户集合 U。因此，p 是一个映射 $\mathrm{p} : \mathrm{U} \times \mathrm{D} \to \mathrm{U} \times \mathrm{D}$，其中 $\mathrm{D} = \mathrm{D}_u \cup \mathrm{D}_s$。因此，信息系统可描述为一个元组 {U, D, P, Eve}。

对于 $\mathrm{u}_\iota \in$U 用户的身份识别信息，系统会执行特殊的数据处理算法 $\mathrm{identif}\, \mathrm{y}_s(\cdot) \in \mathrm{P}$，使 $\mathrm{identi\, f}\, \mathrm{y}_s(\mathrm{d}_{u\iota}) = \mathrm{u}_\iota \in \mathrm{U}$。为保护数据，信息系统可以应用混淆算法 $\mathrm{ob\, f\, uscate}(\cdot) \in \mathrm{P}$，以实现数据变换 $\mathrm{d}^*_{u_\iota} = \mathrm{ob\, f\, uscate}(\mathrm{d}_{u\iota})$，从而使 $\mathrm{identi\, fy}_s(\mathrm{d}^*_{u_\iota}) \neq \mathrm{u}_\iota$。

假设 Eve 以某种能力级别监视数据处理流程，使其能够访问信息系统中的数据——Eve 能看到 $\mathrm{d}^*_u, \mathrm{d}^*_s$，其中 $\mathrm{d}^*_u \subseteq \mathrm{D}_u$ 且 $\mathrm{d}^*_s \subseteq \mathrm{D}_s$。为了保证用户数据在所考虑系统中的不可链接性，必须满足 $\forall \mathrm{d}^*_{u_\iota} \subseteq \mathrm{D}_u$，使得 Eve 能看到 $\mathrm{d}^*_{u_\iota}$ 时，Eve 无法找到满足 $\mathrm{identify}_e(\mathrm{d}^*_{u_\iota}) = \mathrm{u}_\iota$ 的变换 $\mathrm{identif\, y}_e(\mathrm{d}_{u\iota})$。

在协作式智能交通系统的不可链接性语境下，可能有以下解释：$\mathrm{d}_u = \{\mathrm{registrationNumber}, \mathrm{manu\, f\, acturer}\ldots, \text{路径}\}$，其中 $\forall \mathrm{path}_\iota \in \mathrm{PAT\, HS}, \mathrm{path}_\iota = \{\langle \mathrm{x}_j, \mathrm{y}_j \rangle\}$。用户链接的一些方式包括：

- $\mathrm{search(registrationNumber)} = \mathrm{u}_\iota$：通常由执法机构执行；
- $\mathrm{recon(manu\, f\, acturer, color)} = \mathrm{u}_\iota$：例如，可由私人侦探执行；
- 驾驶建模（路径$_\iota$）$= \mathrm{u}_\iota$：可通过收集路侧单元的数据来实现。

本研究聚焦于用户链接的后一种方式。因此，我们考虑用户车辆在不同时间点的状态（通过协同感知消息传送），并评估协作式智能交通系统中的不可链接性，具体而言，是考察 Eve 利用例如驾驶建模（·）作为识别$_e$（·）的手段来推断信息的能力。

3.2 用于不可链接性分析的马尔可夫模型

为了研究车联网中的不可链接性，我们使用隐马尔可夫模型，其图示如图 2 所示。描述该模型需要以下集合。参与者集合为 $\mathbb{P} = \{\text{Alice、Bob } \ldots\}$。对于每位参与者，其车辆都有一个隐藏状态集合。例如，Alice 的集合为 $\mathbb{X}^A = \{\mathrm{x}_1^A、\mathrm{x}_2^A、\ldots、\mathrm{x}_k^A、\ldots、\mathrm{x}_\mu^A\}$；Bob 的集合为 $\mathbb{X}^B = \left\{\mathrm{x}_1^B、\mathrm{x}_2^B、\ldots、\mathrm{x}_j^B、\ldots、\mathrm{x}_\omega^B\right\}$。每个状态（例如 x_1^A）都可以是一个向量，其中包含特定时间点 Alice 车辆的位置、速度、加速度及其他适用特征。本文始终假设 $\mathbb{X}^A \cap \mathbb{X}^B$ 通常非空。

$|\mathbb{P}|$ 参与者系统由隐藏和可观测的联合状态表征。当时间步 i 推进至 $\mathrm{i} + 1$ 时，隐藏状态 $\mathbf{X}_i$ 与 $\mathbf{X}_{i+1}$ 之间会发生转移，其中联合状态 $\mathbf{X}_i = (\mathbf{X}_i^A, \mathbf{X}_i^B)$ 是变量 $\mathbf{X}_i^A \in \mathbb{X}^A$ 和 $\mathbf{X}_i^B \in \mathbb{X}^B$ 的组合（串接）。因此，$\forall \mathrm{k, j}(\mathrm{x}_k^A, \mathrm{x}_j^B) \in \mathbb{X}^{(A,B)}$，其中 $|\mathbb{X}^{(A,B)}| = |\mathbb{X}^A| \times |\mathbb{X}^B|$（为简化表示，我们进一步假设 $|\mathbb{P}| = 2, |\mathbb{X}^A| = \mu = 2, |\mathbb{X}^B| = \omega = 2$）。

从 $\mathbf{X}_i$ 到 $\mathbf{X}_{i+1}$ 的可能转移用索引 $1 - 16$ 表示（见图 2）：这些转移由相应的概率决定。例如，从 $\mathbf{X}_i = (\mathbf{X}_i^A = \mathrm{x}_1^A, \mathbf{X}_i^B = \mathrm{x}_1^B)$ 到 $\mathbf{X}_{i+1} = (\mathbf{X}_{i+1}^A = \mathrm{x}_2^A, \mathbf{X}_{i+1}^B = \mathrm{x}_2^B)$ 的转移用索引 4 表示。这种转移的概率为 $\Pr\left(\mathbf{X}_{i+1}^A = \mathrm{x}_2^A, \mathbf{X}_{i+1}^B = \mathrm{x}_2^B \mid \mathbf{X}_i^A = \mathrm{x}_1^A, \mathbf{X}_i^B = \mathrm{x}_1^B\right)$。在实践中，这些概率可以根据

关于车辆的成熟物理模型获得（Blackman，1986）。

对于每个隐藏联合状态 X_i ，其来自用户的串联分量有$|\mathbb{P}|!$种可能的排列。攻击者试图标记 Alice 和 Bob 的组合 CAM 时，这些排列是不确定性的主要来源。实际中，这是由每次扫描（或会话）i 内 CAM 的排列不可预测所导致的。因此，应从可能的转移中随机选择一个排列。例如，当系统处于联合状态 $(X_i^A = x_1^A, X_i^B = x_1^B)$ 时，如果发生索引为 17 的转移，则应考虑排列 (x_1^A, x_1^B)（玫瑰色节点）；如果发生转移 18，则应考虑 (x_1^B, x_1^A)（蓝色节点）（见图 2）。我们将分别用符号 X_i,R 和 X_i,B 表示玫瑰色和蓝色节点，其中 X_i,R $\in \mathbb{X}$R，X_i,B $\in \mathbb{X}$B，且 $\mathbb{X}R = \mathbb{X}^{(A,B)}$，$\mathbb{X}B = \mathbb{X}^{(B,A)}$。下文中，我们将把彩色节点所表示的状态称为“带标签状态”。为简单起见，且不失一般性，对于隐藏状态 X_i 的所有实现值，我们考虑 $\Pr(X_i,R \mid X_i) = \varphi \le 0.5$ 和 $\Pr(X_i,B \mid X_i) = 1 - \varphi$。

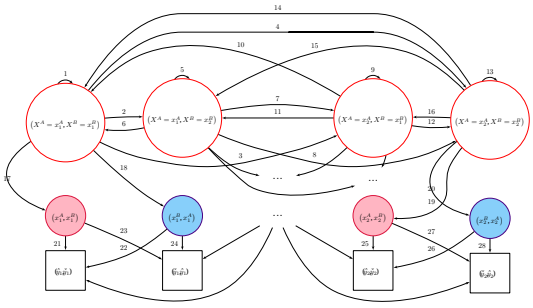


图2：两名参与者发送协同感知消息的隐马尔可夫模型。

为了表示所有隐藏的置换联合状态，我们使用集合 $\mathbb{X}_{\{R,B\}} = \mathbb{X}R \cup \mathbb{X}B$，其中 $|\mathbb{X}R| \le |\mathbb{X}_{\{R,B\}}| \le 2|\mathbb{X}R|$。对于每个 X_i,R 和 X_i,B，都存在指向可观测联合状态 $Y_i \in \mathbb{Y}$, $\mathbb{Y} = \{(\hat{y}_1, \check{y}_1)、(\check{y}_1, \hat{y}_1)$ ...、$(\hat{y}_q, \check{y}_q)$、$(\check{y}_q, \hat{y}_q)$ ... 、$(\hat{y}_\xi, \check{y}_\xi)$、$(\check{y}_\xi, \hat{y}_\xi)\}$ 的转移。图2中用索引 $21-28$ 标出了其中一些指向可观测状态的转移。在另行证明之前，认为 $\mathbb{Y}$ 的基数与 $|\mathbb{X}_{\{R,B\}}|$ 无关。

测量关于标签 $\ell \in \mathbb{L}$，$\mathbb{L} = \{\mathcal{R}、\mathcal{B}\}$ 的不确定性是我们主要关注的问题：这是基于可观测状态完成的。

4 模型属性

为了遵循定义2形式化地表达不可链接性，我们将使用条件熵 $H(\ell_1, \ell_2, \ldots | Y_1, Y2, \ldots)$ 来表示标签序列 $\ell_1, \ell_2, \ldots, \ell_{N-1}$，前提是攻击者观测到 Y_1, Y2,..., Y_{N-1}（Wagner and Eckhoff，2018年）。

4.1 不可链接性的通用表达式

对于所述隐马尔可夫模型，任意时间步的任意隐藏状态的概率均可用多变量离散分布 $\boldsymbol{\rho}$: $\mathbb{X}^{(A,B)} \times \{0, 1, \ldots, N-1\} \to [0,1]^{N|\mathbb{X}^{(A,B)}|}$ 表示。我们还将使用 ρ_i 个 $\boldsymbol{\rho}$ 切片，例如$_{N-1}$，其中 $\boldsymbol{\rho} = \bigcup \rho_i$，每个切片表示第 i 步隐藏状态上的分布。切片 ρ_0 定义了系统开始之前隐藏状态上的分布。由于隐马尔可夫模型此前（见图 2）是用在所有时间步均保持不变的转移概率定义的，因此可以按条件顺序完全确定每个切片： $\rho_{i+1}|\rho_i$ 表示，若给定 ρ_i，则可直接推导出 ρ_{i+1}。

由于攻击者观察到 $Y_1, Y2, \ldots, Y_{N-1}$ 并且知道 $\boldsymbol{\rho}$，对 $H(\ell_1, \ell_2, \ldots | Y_1, Y2, \ldots, \boldsymbol{\rho})$ 的分析是我们推理不可链接性的核心。我们陈述如下。

引理1 车联网系统中的不可链接性（如图 2 所示）表示为：

$$H(\ell_1, \ell_2, \ldots | \mathbf{Y}_1, \mathbf{Y}_2, \ldots, \boldsymbol{\rho}) = \sum_{i=0}^{N-2} H(\ell_{i+1} | \mathbf{Y}_{i+1}, \{\rho_{i+1} | \rho_i\}) . \quad (1)$$

证明：为简单起见，我们仅考虑 $N = 3$。首先，应该注意到

$$H(\ell_1, \ell_2 | \mathbf{Y}_1, \mathbf{Y}_2, \boldsymbol{\rho}) = H(\ell_1, \ell_2, \mathbf{Y}_1, \mathbf{Y}_2 | \boldsymbol{\rho}) - H(\mathbf{Y}_1, \mathbf{Y}_2 | \boldsymbol{\rho}) . \quad (2)$$

接下来，我们考察等式 (2) 的右侧。其中每一项都可以表示为：

$$H(\ell_1, \ell_2, \mathbf{Y}_1, \mathbf{Y}_2 | \boldsymbol{\rho}) = H(\ell_2, \mathbf{Y}_2 | \ell_1, \mathbf{Y}_1, \boldsymbol{\rho}) + H(\ell_1, \mathbf{Y}_1 | \boldsymbol{\rho}) , \quad (3)$$

and

$$H(\mathbf{Y}_1, \mathbf{Y}_2 | \boldsymbol{\rho}) = H(\mathbf{Y}_2 | \mathbf{Y}_1, \boldsymbol{\rho}) + H(\mathbf{Y}_1 | \boldsymbol{\rho}) , \quad (4)$$

分别。我们指出，式（2）和式（3）中分别为 $H(\ell_2, Y_2 | \ell_1, Y1, \boldsymbol{\rho}) = H(\ell_2, Y_2 | \boldsymbol{\rho})$ 和 $H(Y_2 | Y1, \boldsymbol{\rho}) = H(Y_2 | \boldsymbol{\rho})$。这是因为 ℓ_i, Y_i 的实现值不会影响 ℓ_{i+1}, Y_{i+1}。最后我们强调，对于确定 $\boldsymbol{\rho}$ 而言，

确定 ℓ_{i+1},Y_{i+1} 是多余的，因为只有 $\rho_{i+1}|\rho_i$ 相关：$\mathrm{H}(\mathrm{Y}_1|\boldsymbol{\rho}) = \mathrm{H}(\mathrm{Y}_1|\{\rho_1|\rho_0\}), \mathrm{H}(\ell_1,\mathrm{Y}_1|\boldsymbol{\rho}) = \mathrm{H}(\ell_1,\mathrm{Y}_1|\{\rho_1|\rho_0\}), \mathrm{H}(\mathrm{Y}_2|\boldsymbol{\rho}) = \mathrm{H}(\mathrm{Y}_2|\{\rho_2|\rho_1\}), \mathrm{H}(\ell_2,\mathrm{Y}_2|\boldsymbol{\rho}) = \mathrm{H}(\ell_2,\mathrm{Y}_2|\{\rho_2|\rho_1\})$。因此，式（2）可以改写为

$$\begin{aligned} &H(\ell_1,\ell_2|\mathbf{Y}_1,\mathbf{Y}_2,\boldsymbol{\rho}) = \\ &H(\ell_1,\mathbf{Y}_1|\{\rho_1|\rho_0\}) + H(\ell_2,\mathbf{Y}_2|\{\rho_2|\rho_1\}) - \\ &\Big(H(\mathbf{Y}_1|\{\rho_1|\rho_0\}) + H(\mathbf{Y}_2|\{\rho_2|\rho_1\})\Big). \end{aligned} \tag{5}$$

后一个式（5）可以重新分组

$$\begin{aligned} &H(\ell_1,\ell_2|\mathbf{Y}_1,\mathbf{Y}_2,\boldsymbol{\rho}) = \\ &\Big(H(\ell_1,\mathbf{Y}_1|\{\rho_1|\rho_0\}) - H(\mathbf{Y}_1|\{\rho_1|\rho_0\})\Big) + \\ &\Big(H(\ell_2,\mathbf{Y}_2|\{\rho_2|\rho_1\}) - H(\mathbf{Y}_2|\{\rho_2|\rho_1\})\Big) = \\ &H(\ell_1|\mathbf{Y}_1,\{\rho_1|\rho_0\}) + H(\ell_2|\mathbf{Y}_2,\{\rho_2|\rho_1\}). \end{aligned} \tag{6}$$

□

4.2 最坏情况不可链接性

我们的目标是获得一种计算上可行的不可链接性估计方法。直接利用引理1的结果需要计算 $\{\rho_{i+1}|\rho_i\}$，但这有几个缺点：a）必须指定隐藏状态的转移概率（这通常需要研究用户移动的物理模型）；b）因此，定义隐藏状态分布的总计算复杂度为 $O(N\mu^2\omega^2)$。为避免这些复杂问题，我们基于理性下界 $\mathcal{H}_{\mathrm{r}}$ 为 $\mathrm{H}(\ell_1,\ell_2,...|\mathrm{Y}_1,\mathrm{Y}2,...,\boldsymbol{\rho})$ 构建不可链接性保证。以下假设阐释了理性下界的概念（Sniedovich，2016）。

假设 1（最坏情况不可链接性）要求攻击者知道隐藏状态、带标签状态和可观测状态的集合。他知道所有转移以及顺序混合概率 φ。对于时刻 i 的每个可观测状态，他随后定义与其所知信息不矛盾的最坏可能隐藏状态。

尽管假设 1 可能被视为过于苛刻，我们仍要强调，在时刻 i，攻击者并不知道标记状态 ℓ_i（也无法强制其选择）。

假设2（理性下界 $\mathcal{H}_{\mathrm{r}}$）
要求用户是理性的，并使最坏情况不可链接性最大化：可观测状态是通过对攻击者所考虑的最坏标记状态进行理性混淆而获得的。

影响此类$\mathcal{H}$计算任务的因素有几个方面：1）隐藏状态之间的转移概率（例如，定义{v1}）；2）隐藏状态转移到带标签状态的概率（例如，$\varphi, 1-\varphi$），以及带标签状态转移到可观测状态的概率。此外，我们考虑这样一种情况：在总失真受约束 D、历经 $N-1$ 步的条件下，1）取最坏情况 $\boldsymbol{\rho}$（使熵最小化），而 2）则取最优概率˜（使熵最大化）。

我们利用引理1的结果提出以下要求：

$$\begin{aligned} &\mathcal{H}_{\mathrm{r}} = \min_{\boldsymbol{\rho}}\Big[H(\ell_1,\ell_2,...|\mathbf{Y}_1,\mathbf{Y}_2,...,\boldsymbol{\rho})\Big] = \\ &\sum_{i=0}^{N-2}\min_{\{\rho_{i+1}|\rho_i\}}\Big[H(\ell_{i+1}|\mathbf{Y}_{i+1},\{\rho_{i+1}|\rho_i\})\Big]. \end{aligned} \tag{7}$$

为了以使 $\mathcal{H}_{\mathrm{r}}$ 最大化的方式混淆隐藏状态，我们需要确定

$$\rho_{\min,i+1} = \arg\min_{\{\rho_{i+1}|\rho_i\}}\Big[H(\ell_{i+1}|\mathbf{Y}_{i+1},\{\rho_{i+1}|\rho_i\})\Big]. \tag{8}$$

概率 $\Pr\left(\ell_{i+1}=\mathcal{R},\mathrm{Y}_{i+1}\middle|\{\rho_{i+1}|\rho_i\}\right)$ 和 $\Pr\left(\ell_{i+1}=\mathcal{B},\mathrm{Y}_{i+1}\middle|\{\rho_{i+1}|\rho_i\}\right)$ 将用于后续推导。为简化符号，我们分别记为 $\Pr(\ell_{i+1}=\mathcal{R},\mathrm{Y}_{i+1})$ 和 $\Pr(\ell_{i+1}=\mathcal{B},\mathrm{Y}_{i+1})$。这些概率定义如下：

$$\begin{aligned} &\Pr(\ell_{i+1}=\mathcal{R},\mathbf{Y}_{i+1}) = \\ &\sum_{\mathbf{x}_{i+1,\mathcal{R}}\in\mathbb{X}_{\mathcal{R}}}\Pr\left(\mathbf{Y}_{i+1}\mid\mathbf{X}_{i+1,\mathcal{R}}\right)\Pr\left(\mathbf{X}_{i+1,\mathcal{R}}\right) = \\ &\sum_{\mathbf{x}_{i+1,\mathcal{R}}\in\mathbb{X}_{\mathcal{R}}}\Pr\left(\mathbf{Y}_{i+1}\mid\mathbf{X}_{i+1,\mathcal{R}}\right)\varphi\Pr(\mathbf{X}_{i+1}), \end{aligned} \tag{9}$$

$$\begin{aligned} &\Pr(\ell_{i+1}=\mathcal{B},\mathbf{Y}_{i+1}) = \\ &\sum_{\mathbf{X}_{i+1,\mathcal{B}}\in\mathbb{X}_{\mathcal{B}}}\Pr\left(\mathbf{Y}_{i+1}\mid\mathbf{X}_{i+1,\mathcal{B}}\right)\Pr\left(\mathbf{X}_{i+1,\mathcal{B}}\right) = \\ &\sum_{\mathbf{X}_{i+1,\mathcal{B}}\in\mathbb{X}_{\mathcal{B}}}\Pr\left(\mathbf{Y}_{i+1}\mid\mathbf{X}_{i+1,\mathcal{B}}\right)(1-\varphi)\Pr(\mathbf{X}_{i+1}). \end{aligned} \tag{10}$$

接下来我们指出，

$$\Pr(\ell_{i+1}=\mathcal{R}\mid\mathbf{Y}_{i+1}) = \frac{\Pr(\ell_{i+1}=\mathcal{R},\mathbf{Y}_{i+1})}{\Pr(\mathbf{Y}_{i+1})}, \tag{11}$$

$$\Pr(\ell_{i+1}=\mathcal{B}\mid\mathbf{Y}_{i+1}) = \frac{\Pr(\ell_{i+1}=\mathcal{B},\mathbf{Y}_{i+1})}{\Pr(\mathbf{Y}_{i+1})}, \tag{12}$$

其中

$$\Pr(\mathbf{Y}_{i+1}) = \Pr(\ell_{i+1}=\mathcal{R},\mathbf{Y}_{i+1}) + \Pr(\ell_{i+1}=\mathcal{B},\mathbf{Y}_{i+1}). \tag{13}$$

以下结果确立了 $\rho_{\min,i+1}$ 的一项重要性质。

引理2 对所有 $i \in [1, N-1]$ 分布 $\rho_{\min,\,i}$ 均为退化的。

证明：我们假设 $\rho_{\min,\ i}$ 是非退化的。为简单起见，且不失一般性，我们考虑两点分布 $\rho^{*}_{\min,\ i}$ ： $\{x_1, x_2\} \to [0,1]^2$。例如，可以使用实现值 $x_1 = (x_1^A, x_1^B)$、$x_2 = (x_2^A, x_2^B)$。这里 $\Pr(X_i = x_1) = \xi$，且 $\Pr(X_i = x_2) = 1 - \xi$。

式（8）中条件熵的最小化等价于对 $p_{\ell_i,\min}$ 进行最小化，其中

$$p_{\ell_i,\min} = \min\left\{\Pr(\ell_i = \mathcal{R} \mid \mathbf{Y}_i), \Pr(\ell_i = \mathcal{B} \mid \mathbf{Y}_i)\right\}, \tag{14}$$

and without loss of generality, we assume that $p_{\ell_i,\min} = \Pr(\ell_i = \mathcal{R} \mid \mathbf{Y}_i)$. To express $p_{\ell_i,\min}$ we then use eqs. (9) to (11) and (13) with the following substitutions (simplifying expressions): $\alpha_1 = \varphi \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_{i,\mathcal{R}} = (x_1^A, x_1^B)\right)$, $\beta_1 = \varphi \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_{i,\mathcal{R}} = (x_1^A, x_1^B)\right) + (1-\varphi)\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_{i,\mathcal{B}} = (x_1^B, x_1^A)\right)$, $\alpha_2 = \varphi \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_{i,\mathcal{R}} = (x_2^A, x_2^B)\right)$, $\beta_2 = \varphi \Pr\left(\mathbf{Y}_i \mid \mathbf{X}_{i,\mathcal{R}} = (x_2^A, x_2^B)\right) + (1-\varphi)\Pr\left(\mathbf{Y}_i \mid \mathbf{X}_{i,\mathcal{B}} = (x_2^B, x_2^A)\right)$. The minimization task is then

$$\min_{\xi} p_{\ell_i,\min} = \min_{\xi} \frac{\xi\alpha_1 + (1-\xi)\alpha_2}{\xi\beta_1 + (1-\xi)\beta_2}. \tag{15}$$

通过分析 $\frac{\partial}{\partial \xi} p_{\ell_i,\min}$，我们得出结论：$\xi \in (0, 1)$ 内不存在局部极值，因此最小值出现在某个端点，例如 $\xi \in \{0, 1\}$。□

根据引理2的结果，对于每个 Y_i，都存在且仅存在一个最坏情况隐藏状态 X_i （因为 $\Pr(\tilde{X}_i \mid \rho_{\min,i}) = 1$）。这意味着：

推论 1：HMM 的设计满足：对于 $\mathbb{Y}$ 中的每个状态（实现），都存在且仅存在一个从 $\mathbb{X}^{(A,B)}$ 出发的转移，因此明确满足假设 1。

因此，我们将继续遵循这一设计˜原则，并用 X_i 表示隐藏状态。接下来，我们将详细讨论：a) 对于每个 X_i，不同可观测状态 Y_i 的最优数量是多少？b) 应如何定义最优可观测状态？c) 从带标签状态到可观测状态的转移概率是多少？

4.3 对可观测状态的要求

这里我们从代表 Alice 和˜Bob 混淆隐藏状态（例如，系统产生可观测状态）的系统角度进行分析，因此假设 X_i 已知。转移 $X_{i,\ \mathcal{R}} \to Y_i$ 和 $X_{i,\ \mathcal{B}} \to Y_i$ 的可能性意味着会产生非零失真 $\mathbb{E}[D_i]$：

$$\mathbb{E}[D_i] = \sum_{\mathbf{y}_j^{(i)} \in \mathbb{Y}^{(i)}} D_{i,j} \Pr\left(\mathbf{Y}_i = \mathbf{y}_j^{(i)} \mid \tilde{\mathbf{X}}_i\right), \tag{16}$$

其中

$$D_{i,j} = \Pr\left(\ell_i = \mathcal{R} \mid \mathbf{Y}_i = \mathbf{y}_j^{(i)}\right) d\left(\tilde{\mathbf{X}}_{i,\mathcal{R}}, \mathbf{y}_j^{(i)}\right) + \Pr\left(\ell_i = \mathcal{B} \mid \mathbf{Y}_i = \mathbf{y}_j^{(i)}\right) d\left(\tilde{\mathbf{X}}_{i,\mathcal{B}}, \mathbf{y}_j^{(i)}\right). \tag{17}$$

这里，$\mathbb{Y}^{(i)}$ 是所有可观测状态的集合；从 $X_{i,\ \mathcal{R}}$ 和 $X_{i,\ \mathcal{B}}$ 在时间步 i 的实现值出发，存在转移到这些状态的˜ ˜；$y^{(i)}{}_j$ 是 $\mathbb{Y}^{(i)}$ 中的一个元素；d (·,·) 是某种失真度量（例如平方误差）。

优化工作分为两部分：i）如何获得可观测状态 $\mathbb{Y}^{(i)}$，使 $\mathcal{H}_{\ i}$ 在约束 $\tilde{D}_i \geq \mathbb{E}[D_i]$ 下达到最大？ii）如何˜定义每个时间步 i 的 D_i，使 $\mathcal{H}_{\ i}$ 达到最大，并满足总失真约束˜$D \geq \sum_i \mathbb{E}[D_i]$？我们先回答问题 i），这将有助于回答问题 ii）。

对于混淆，我们采用以下原则： $\mathbb{Y}^{(i)}$ 中的每个元素 $y^{(i)}$ j 都可以完全˜ ˜由 $X_{i,R}$、$X_{i,B}$ 的实现值以及参数 λj 指定。概率 $\Pr\left(\ell_i = R \mid Y_i = y^{(i)}j\right)$、$\Pr\left(\ell_i = B \mid Y_i = y^{(i)}j\right)$ 会影响 $H_{r,i,j}$。所有这些参数都会影响 $D_{i,j}$。说明所有这些参数之间关系的示意图见˜图 3。在这个例子中，带标签状态为 $X_{i,R} = (x^A, x^B)$，˜$X_{i,B} = (x^B, x^A)$；集合 $\mathbb{Y}^{(i)}$ 仅包含两个元素 $y_1^{(i)} = (\hat{y}^{(i)}, \check{y}^{(i)})$ 和 $y_2^{(i)} = (\check{y}^{(i)}, \hat{y}^{(i)})$。例如，要指定 $y_1^{(i)}$，除了带标签状态外，我们只需要 λ_1 （Δ 是它们之间的距离）。要得到 $y_2^{(i)}$，我们应采用类似的步骤，其中 λ_2 已知（在我们的例子中为$\lambda_2 = 1 - \lambda_1$）。概率 $\Pr\left(\ell_i = R \mid Y_i = y_1^{(i)}\right)$ 记作 p_1：其值会影响攻击者的不确定性 $H_{r,i,j}$ 以及失真 $D_{i,j}$。

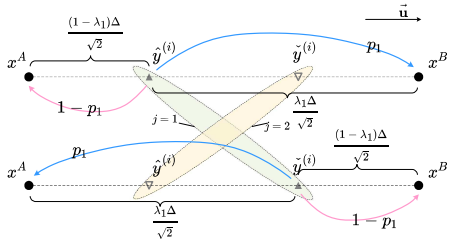


图 3：混淆原理示意图。

为了最大化 $\mathcal{H}_{\ i}$ 在 $\tilde{D}_i \geq \mathbb{E}[D_i]$ 条件下，我们考虑 Y_i 的实现值以及 λ 的最优调整：这种调整随后使我们能够增大 p_1 和 $1 - p_2$。

我们注意到，Y_i 应属于连接 $X_{i,\mathcal{R}}$ 和 $X_{i,\mathcal{B}}$ 的线段（位于$\tilde{}$ $\tilde{}$多维空间中）。这一性质显而易见（无需证明），考虑三角形 $\triangle X_{i,\mathcal{R}} Y_i X_{i,\mathcal{B}}$ 即可更好地理解。由此可得：

$$\forall \mathbf{Y}_i \left(\mathbf{Y}_i \in \mathbb{Y}^{(i)} \Longrightarrow (\exists \lambda \in [0,1]) \wedge \left(\vec{\mathbf{Y}}_i = \overrightarrow{\tilde{\mathbf{X}}_{i,\mathcal{R}}} + \lambda \overrightarrow{\tilde{\mathbf{X}}_{i,\mathcal{R}} \tilde{\mathbf{X}}_{i,\mathcal{B}}} \right) \right). \quad (18)$$

接下来我们建立如下结论：

引理 3 要使 D 最小$_{i,j}$必须满足 $\lambda_j = 1 - \Pr\left(\ell_i = \mathcal{R} \mid \mathbf{Y}_i = \mathbf{y}_j^{(i)}\right)$.

证明：由式（17）和式（18）可得

$$D_{i,j} = p_j \Delta_i^2 \lambda_j^2 + (1 - p_j) \Delta_i^2 (1 - \lambda_j)^2 ,$$

其中 $p_j = \Pr\left(\ell_i = \mathcal{R} \mid Y_i = y^{(i)}{}_j\right) \le 0.5$，以及 $\Delta_i^2 = d\left(\tilde{}X_{i,\mathcal{R}}, \tilde{}X_{i,\mathcal{B}}\right) = \left\| \overrightarrow{\tilde{}_i \quad \tilde{}_i} X_{i,\mathcal{R}} X_{i,\mathcal{B}} \right\|^2$。

接下来我们分析 $\frac{\partial}{\partial \lambda_j} D_{i,j}$ 并发现 $\lambda_j = 1 - p_j$ 是 $D_{i,j}$ 的极值（最小值）。□

推论 2 最小失真为 $D_{i,j} = \Delta_i^2 p_j (1 - p_j) \le \frac{\Delta^2}{4}$，其中 $p_j = \Pr\left(\ell_i = \mathcal{R} \mid Y_i = y^{(i)}{}_j\right)$，$\Delta_i^2 = d\left(\tilde{}X_{i,\mathcal{R}}, \tilde{}X_{i,\mathcal{B}}\right)$。

推论 3 对于每个 i，最高下界（最大最小熵）为：

$$\mathcal{H}_{r,i} = -\nu_i \log_2 \nu_i - (1 - \nu_i) \log_2 (1 - \nu_i) , \quad (19)$$

$$where\ \nu_i = \min\left\{ \varphi, \frac{\Delta_i - \sqrt{\Delta_i^2 - 4\mathbb{E}[D_i]}}{2\Delta_i} \right\} .$$

证明：需要：1）确定 $\mathbb{Y}^{(i)}$ 及其概率分布；2）确定 $\mathbb{Y}^{(i)}$中每个元素对应的 $\Pr(\ell_i \mid Y_i)$。为此，我们证明在$\mathbb{E}[D_i]$ 上的失真约束下，$\left|\mathbb{Y}^{(i)}\right| \le 2$可使熵达到最大：我们分析 $\mathbb{Y}^{(i)} = \left\{ y_1^{(i)}, y_2^{(i)} \right\}$ 的情形，其中

$$\Pr\left(\ell_i = \mathcal{R} \mid \mathbf{Y}_i = \mathbf{y}_1^{(i)}\right) = \Pr\left(\ell_i = \mathcal{B} \mid \mathbf{Y}_i = \mathbf{y}_2^{(i)}\right) . \quad (20)$$

为了证明此类设置的最优性，我们考虑若干个 $\mathbb{E}[D_i] = \tilde{}D_i$ 固定的替代情形。首先考虑一种替代情形，其中 $\left|\mathbb{Y}^{(i)}\right| = 2$，但

$$\begin{aligned} &\Pr\left(\ell_i = \mathcal{R} \mid \mathbf{Y}_i = \mathbf{y}_1^{(i)}\right) \neq \Pr\left(\ell_i = \mathcal{R} \mid \mathbf{Y}_i = \mathbf{y}_2^{(i)}\right) ; \\ &\Pr\left(\ell_i = \mathcal{R} \mid \mathbf{Y}_i = \mathbf{y}_1^{(i)}\right) \neq \Pr\left(\ell_i = \mathcal{B} \mid \mathbf{Y}_i = \mathbf{y}_2^{(i)}\right) . \end{aligned} \quad (21)$$

为简单起见，我们使用以下符号：$\Pr\left(Y_i = y_1^{(i)} \mid \tilde{}X_i\right) = \alpha$，以及 $\Pr\left(Y_i = y_2^{(i)} \mid \tilde{}X_i\right) = 1 - \alpha$；$\Pr\left(\ell_i = \mathcal{R} \mid Y_i = y_1^{(i)}\right) = p_1 \le 0.5$，以及 $\Pr\left(\ell_i = \mathcal{R} \mid Y_i = y_2^{(i)}\right) = p_2 \ge 0.5$。结合条件熵的表达式，我们要求：

$$\begin{cases} \max\left[\mathcal{H}_{r,i}\right] = \max\left[\alpha H_1 + (1 - \alpha) H_2\right] ; \\ \tilde{D}_i = \alpha D_{i,1} + (1 - \alpha) D_{i,2} , \end{cases} \quad (22)$$

其中 $H_1 = H\left(\ell_i \mid Y_i = y_1^{(i)}\right)$, $H_2 = H\left(\ell_i \mid Y_i = y_2^{(i)}\right)$。根据公式 (21)，$D_{i,1} \neq D_{i,2}$。现在我们说明 H_1 和 H_2 分别是 $D_{i,1}$ 和 $D_{i,2}$ 的函数。为此，我们只需指出，p_1（p_2也可得到类似结果）是 $D_{i,1}$ 的单调递增函数：由推论 2 可知，$p_1 = \frac{\Delta_i - \sqrt{\Delta_i^2 - 4 D_i 2\Delta}}{2\Delta_i}$。要证明无法同时达到公式 (21) 和公式 (22)，~~只需说明以~~下内容（凹性）：

$$\alpha F(x) + (1 - \alpha) F\left(\frac{\tilde{D}_i - \alpha x}{1 - \alpha}\right) \le F(\tilde{D}_i) , \quad (23)$$

其中 $x = D_{i,1}$，且 $F(x) = -p_1(x) \log(p_1(x)) - (1 - p_1(x)) \log(1 - p_1(x))$。式 (23) 的有效性源自

$$\begin{aligned} &\frac{\partial}{\partial x} F(x) = \frac{1}{\Delta_i \theta} \log\left(\frac{\Delta_i + \theta}{\Delta_i - \theta}\right) \ge 0 ; \\ &\frac{\partial^2}{\partial x^2} F(x) = -\frac{2}{\Delta_i \theta^2} \left(\frac{1}{\Delta_i + \theta} + \frac{1}{\Delta_i - \theta} - \frac{1}{\theta} \log\left(\frac{\Delta_i + \theta}{\Delta_i - \theta}\right) \right) \le 0 , \end{aligned}$$

where $\theta = \sqrt{\Delta_i^2 - 4x}$, and $x \in \left[0, \frac{\Delta_i^2}{4}\right]$.

接下来，我们指出另一种 $\left|\mathbb{Y}^{(i)}\right| > 2$ 的情况，并证明它是非最优的。为此，我们考虑 $\left|\mathbb{Y}^{(i)}\right| = 3$，而关于 $\left|\mathbb{Y}^{(i)}\right| > 3$ 的结论随后可以归纳地推导出来。与式 (22) 类似，我们要求

$$\begin{cases} \max\left[\mathcal{H}_{r,i}\right] = \max\left[\alpha H_1 + \beta H_2 + (1 - \alpha - \beta) H_3\right] ; \\ \tilde{D}_i = \alpha D_{i,1} + \beta D_{i,2} + (1 - \alpha - \beta) D_{i,3} . \end{cases}$$

接下来的任务是证明存在 $y_4^{(i)}$，使得 $D_{i,4} = \frac{\alpha}{D_{i1} + \beta D_i} \alpha + \beta^2$，且 $\max H_4 \ge \max\left[\frac{\alpha}{\alpha + \beta} H_1 + \frac{\beta}{\alpha + \beta} H_2\right]$。因此，我们接下来始终假定 $\left|\mathbb{Y}^{(i)}\right| \le 2$ 代表最优设置。

要使式（22）中的 $\max\left[\alpha H_1 + (1 - \alpha) H_2\right]$ 成立，只需满足 $H_1 = H_2$ 以及 $D_{i,1} = D_{i,2} = D_i$。后者要求 $\lambda_1 = \lambda_2$ 或 $\lambda_1 = (1 - \lambda_2)$：第一个条件意味着 $p_1 = p_2 = 0.5$，并导致

一种平凡情况，即 $y_1^{(i)} = y_2^{(i)} = 0.5\,(\tilde{X}_{i,\mathcal{R}} + \tilde{X}_{i,\mathcal{B}})$，这意味着 $\left|\mathbb{Y}^{(i)}\right| = 1$。第二个条件意味着 $p_1 = 1 - p_2$，并导致 $y_1^{(i)} \neq y_2^{(i)}$，如果 $\tilde{D}_i < 0.25\Delta_i^2$。

要求 $\alpha \in [0,1]$ 必须与顺序混合概率 φ 一致：

$$\alpha p_1 + (1-\alpha) p_{2_1} = \varphi\,, \tag{24}$$

由此我们推导出 $\alpha = {}^{\varphi+}{}_{p2_____p_1-1}$ 要求 $\varphi \geq p_1$。或者，也可以根据以下事实来理解这一要求：$H(\ell_i) \geq H(\ell_i \mid Y_i)$：设置 $p_1 > \varphi$ 会导致更大的失真，但不会增加熵。 □

推论 3 的证明˜有几个重要结论。首先，对于每个隐藏状态 X_i，根据公式（18）可以得到两个可观测状态，其中 $\lambda_1^{(i)} = 1 - \nu_i$ 用于定义实现 $y_1^{(i)}$，而 $\lambda_2^{(i)} = 1 - \lambda_1^{(i)}$ 用于 $y_2^{(i)}$。其次，在步骤 i 应使用允许的最大失真，这意味着 $\mathbb{E}[D_i] = \tilde{D}_i$。第三，从带标签状态到可观测状态的转移概率为

$$\begin{aligned}
&\Pr\left(\mathbf{Y}_i = \mathbf{y}_1^{(i)} \mid \ell_i = \mathcal{R}\right) = \tfrac{\nu_i}{\varphi}\tfrac{\varphi+\nu_i-1}{2\nu_i-1}\,;\\
&\Pr\left(\mathbf{Y}_i = \mathbf{y}_2^{(i)} \mid \ell_i = \mathcal{R}\right) = 1 - \Pr\left(\mathbf{Y}_i = \mathbf{y}_1^{(i)} \mid \ell_i = \mathcal{R}\right)\,;\\
&\Pr\left(\mathbf{Y}_i = \mathbf{y}_1^{(i)} \mid \ell_i = \mathcal{B}\right) = \tfrac{1-\nu_i}{1-\varphi}\tfrac{\varphi+\nu_i-1}{2\nu_i-1}\,;\\
&\Pr\left(\mathbf{Y}_i = \mathbf{y}_2^{(i)} \mid \ell_i = \mathcal{B}\right) = 1 - \Pr\left(\mathbf{Y}_i = \mathbf{y}_1^{(i)} \mid \ell_i = \mathcal{B}\right)\,.
\end{aligned} \tag{25}$$

4.4 N − 1 个时间步的最优混淆

现在，对于每个 i，我们定义 $\tilde{D}_i$，使得 $\mathcal{H}_t = \sum_i \mathcal{H}_{t,i}$ 在总失真约束 $\tilde{D} \geq \sum_i D_i$ 下最大化。因此，我们得到了所有时间步的最优可观测状态及其对应的转移概率（从带标签状态出发）。根据推论 3 的证明，我们使用 $\frac{\partial}{\partial \tilde{D}_i}\mathcal{H}_{t,i} \geq 0$ 和 $\frac{\partial^2}{\partial \tilde{D}_i^2}\mathcal{H}_{t,i} \leq 0$。因此，要最大化 $\mathcal{H}_t$，我们要求

$$\begin{cases}
\forall i\ \frac{\partial}{\partial D_i}\mathcal{H}_{t,i} = \frac{1}{\Delta_i^2\sqrt{1-\kappa_i}}\log\left(\frac{1+\sqrt{1-\kappa_i}}{1-\sqrt{1-\kappa_i}}\right) = C\,;\\
\tilde{D} = \sum_{i=1}^{N-1}\tilde{D}_i = \frac{1}{4}\sum_{i=1}^{N-1}\kappa_i\Delta_i^2\,,
\end{cases} \tag{26}$$

其中 C 是某个常数，$\kappa_i = {}^{4\tilde{}}{}_{D_i\,\overline{\Delta^2_i}}$。然后，我们对所有 κ_i 求解系统方程 (26)，$i \in [1, N-1]$，，并根据推论 3 得到 $\nu_i = \min\left\{\varphi,\ 0.5 - \sqrt{0.25 - 0.25\kappa_i}\right\}$。

5 混淆算法

这里，我们将上述发现表示为混淆算法的形式（见算法1）。它切实可行，可在实际环境中实现：其复杂度（不包括 solve 过程的复杂度）仅为 O(N)。算法的输入为数组（大小为 N）X^A、X^B，以及标量 $\tilde{D}$、φ。这些数组的元素是 X_i^A 和 X_i^B 的标量/向量实现值，分别表示 Alice 和 Bob 在时间 i 的地理位置。实际应用中，这些数组可能包含基于历史数据和重复模式进行外推的值。例如，Alice 和 Bob 可能每天都大致在同一时间沿相同路线通勤。过程 solve 为式（26）提供解：数组 $\boldsymbol{\kappa}$ 包含定义混淆状态 Y_i 的实现值所需的元素 κ_i。它还用于计算依赖于混淆过程的不可链接性判据（熵）$\mathcal{H}_{t,\,i}$。过程 send 将在时间 i 混淆的数据封装为一种车联网通信格式，并发送至最近的路侧单元。因此，该算法的输出是数组 Y，其中包含所有混淆记录，以及系统在 $N-1$ 个步骤内总体不可链接性的指标 $\mathcal{H}_t$。

算法1：混淆算法

输入：$X^A, X^B, \tilde{D}, \varphi$;
输出：$Y, \mathcal{H}_t$;

开始
- $\mathcal{H}_t \leftarrow 0, Y \leftarrow \varnothing, \boldsymbol{\kappa} \leftarrow$ `solve`$(\tilde{D}, X^A, X^B)$;
- 对于 i $\leftarrow$ 1 到 N − 1 执行
  - $\nu_i \leftarrow$ 最小值 $\{\varphi,\ 0.5 - \sqrt{0.25 - 0.25\kappa_i}\}$,
  - $\alpha \leftarrow (\varphi + \nu_i - 1)/(2\nu_i - 1)$,
  - $\mathcal{H}_{t,i} \leftarrow -\nu_i$ 对数$(\nu_i) - (1-\nu_i)$对数$(1-\nu_i)$,
  - $H_t \leftarrow H_t + H_{t,i}$, $P_{1,\mathcal{R}} \leftarrow \nu_i\alpha/\varphi$,
  - $P_{1,\mathcal{B}} \leftarrow (1-\nu_i)\alpha/(1-\varphi)$,
  - $r_1 \leftarrow$ `UniRand`$([0,1])$, $r_2 \leftarrow$ `UniRand`$([0,1])$,
  - $\Lambda_{\mathcal{R}} \leftarrow 0.5 + (0.5 - \nu_i)\,\mathtt{sign}_{\pm}(P_{1,\mathcal{R}} - r_2)$,
  - $\Lambda_{\mathcal{B}} \leftarrow 0.5 + (0.5 - \nu_i)\,\mathtt{sign}_{\pm}(P_{1,\mathcal{B}} - r_2)$;
  - 如果 $r_1 \leq \varphi$ 则 $\hat{y}^{(i)} \leftarrow X_i^A + \Lambda_{\mathcal{R}}\,(X_i^B - X_i^A)$,
    - $\check{y}^{(i)} \leftarrow X_i^B + \Lambda_{\mathcal{R}}\,(X_i^A - X_i^B)$,
    - $Y_i = \mathtt{concat}(\hat{y}^{(i)}, \check{y}^{(i)})$;
  - 否则 $\hat{y}^{(i)} \leftarrow X_i^A + \Lambda_{\mathcal{B}}\,(X_i^B - X_i^A)$,
    - $\check{y}^{(i)} \leftarrow X_i^B + \Lambda_{\mathcal{B}}\,(X_i^A - X_i^B)$,
    - $Y_i = \mathtt{concat}(\check{y}^{(i)}, \hat{y}^{(i)})$;
  - `send_RSU`(Y_i), $Y = \mathtt{concat}(Y, Y_i)$;

6 讨论

本文讨论了我们在多大程度上实现了主要目标（见第 2.1 节）——“开发一种方法，以确保 C-ITS 中 CAM 来源的熵达到较高水平”。为此，我们介绍主要成果的特点、优势、局限性，以及后续计划。

进一步的工作。

6.1 结果及其特征

本文结合了：（i）不可链接性的经典定义；（ii）对强大攻击者的假设；以及（iii）通过开发最优联合混淆技术来衡量并提升 C-ITS 中的不可链接性。学术创新之处在于这三点（i–iii）的结合。接下来，我们将更详细地讨论每一点的重要性。

首先，我们倾向于采用协作式智能交通系统应用领域相关标准所要求的不可链接性经典定义（ETSI TS 102 941 V2.1.1, 2021；ISO 24102-1, 2018年）。因此，本研究的概念更接近一些早期的位置隐私研究，例如（Shokri, 2012），该研究依托隐马尔可夫模型中的贝叶斯推断；而这与许多后续依赖 k-匿名性、差分隐私和地理不可区分性的研究形成对比（Andr´es 等人, 2013；Bordenabe等人, 2014；Corser等人, 2016）。

在本研究中，不可链接性的定义通过贝叶斯推断加以表达，并进一步由熵体现。我们还强调了这种方法的优势。根据定义2，攻击者对操作的推理起着核心作用。这种推理可能超出对象内部可观测的属性（例如，经密码签名的协同感知消息）；还可能依赖收集到的元信息，例如系统层面上的信息（如协同感知消息抵达路侧单元的顺序）。这一特点与科学和哲学中的许多概念相符。莱布尼茨曾指出，不可辨别的对象具有各自的身份（Hacking, 1975）：对这些身份的偏好可以用统计方式表达（因此可用于贝叶斯推断）。例如，在统计学中，可以借助额外的索引来辅助此类推理。相比之下，地理不可区分性不支持对不可辨别的地理数据片段作进一步推理，因此，与不可链接性相比，这一隐私概念的要求较低（因而也较弱）。要看出这些概念之间的差异，可以观察我们的隐马尔可夫模型中的顺序混合（及相应概率 φ）（见图2）：即使 $\mathrm{X}_i^A = \mathrm{X}_i^B$，模型仍将 $(\mathrm{X}_i^A, \mathrm{X}_i^B)$ 与 $(\mathrm{X}_i^B, \mathrm{X}_i^A)$ 区分开来。因此，我们赞同（Montazeri等人, 2017；Takbiri等人, 2017, 2020）的作者所述：要实现强有力的隐私保障，混淆和排列（顺序混合）缺一不可。

其次，我们关注的是不可链接性（而非位置保护），因此即使假设对手实力强大，结论依然成立：攻击者知道 Alice 和 Bob 在每个时刻 i 的实际位置。他还知道参与者使用的概率混淆和顺序混合算法，但不知道该概率算法的输出。因此，他的目标是推断这些数字签名的混淆消息源自谁。应用假设 1 后，与 Shokri（2012）相比，关于攻击者进行统计推断的推理大大简化：这类推理不需要隐马尔可夫模型隐藏状态的信息（例如实际位置、速度等）及转移概率。后一细节有助于隐私保障，因为确定隐马尔可夫模型中转移的概率是一项繁琐且往往不精确的工作，通常依赖类似卡尔曼滤波器的估计器（Blackman，1986；Lehmann 和 Pieczynski，2020）。

第三，在确定协作式智能交通系统中不可链接性的下界时，我们采用了强假设（这也简化了分析）。对于每个时间步 i，不可链接性通过熵 $\mathcal{H}_{\mathfrak{k},i}$ 表示：由攻击者进行最坏情况推断，这意味着 $\mathcal{H}_{\mathfrak{k},i}$ 是下界（见假设 1）。随后将各分量 $\mathcal{H}_{\mathfrak{k},i}$ 在 $\mathrm{N}-1$ 个时间步上求和，得到 $\mathcal{H}_{\mathfrak{k}}$（见引理 1）。这种求和是简单直观的一步；不过，这样做是合理的，因为对于任意 $\mathrm{i} \in \{1,2,\ldots,\mathrm{N}-1\}$，对已接收协作式感知消息来源（起点）的推断，与在 $\mathrm{i}-1$ 时进行的此类推断相互独立。熵（不可链接性）值 $\mathcal{H}_{\mathfrak{k},i}$ 和 $\mathcal{H}_{\mathfrak{k}}$，分别可以类比于文献（Shokri等人, 2010）中的微观隐私和宏观隐私概念。要更好地保护宏观隐私（例如轨迹），就需要对微观层面（例如地理坐标点）所报告位置的标签 ℓ_i 具有更高的不确定性。只有以参与者协作感知消息的更高期望失真$\mathbb{E}[\mathrm{D}_i]$ 为代价，才能提高 ℓ_i 的不确定性。为在失真约束下最大化标签的不确定性 $\mathcal{H}_{\mathfrak{k},i}$，我们提出了一种简单的新方法，用于定义在 Alice 和 Bob 协作进行混淆（以生成协同感知消息）时必须遵循的联合分布。许多作者此前已研究过各种失真约束下（多变量）噪声参数的最优性（Andr´es等人，2013年；Geng和Viswanath，2016年；Takbiri等人，2017年）。然而，我们的方法有所不同：为了改善 i 时刻的微观隐私，我们要求对不同参与者的样本进行联合概率混淆（见推论 3 的证明）。由于这一特性，我们的混淆方法显然依赖于数据（Croft等人，2019年；Guan等人, 2022年）。随后，我们分析了在整个协作式智能交通系统观测期间指定失真上限时，如何在 $\mathrm{N}-1$ 个时间步上最优分配失真（见式（26））。本文的所有研究结果均纳入算法1中。Procedure solve 是影响该算法时间复杂度的主要因素之一。不过，若稍微牺牲混淆最优性，这一问题便可解决。例如，只需针对若干种情况预先计算 solve：每种情况都会为随机变量 iD 生成一种不同的分布。随后，应使用最匹配的分布对实际输入数据进行近似，并使用相应的预计算 solve 输出进行混淆。这种变通方法还可以将我们的算法转变为“实时算法”：如果 Alice 和 Bob 认为其未来数据与某个预计算分布相符度较高（例如，由于日常通勤习惯），便可以“即时”对其进行混淆。因此，solve 的预计算情形可以视为参与者双方同意使用的配置文件。

测（见式（26））。本文的所有研究结果均纳入算法1中。过程 solve 是影响该算法时间复杂度的主要因素之一。不过，若稍微牺牲混淆最优性，这一问题便可解决。例如，只需针对若干种情况预先计算 solve：每种情况都会为随机变量 $\Delta^2_{\tilde{i}}$ D 生成一种不同的分布。随后，应使用最匹配的分布对实际输入数据进行近似，并使用相应的预计算 solve 输出进行混淆。这种变通方法还可以将我们的算法转变为“实时算法”：如果 Alice 和 Bob 认为其未来数据与某个预计算分布相符度较高（例如，由于日常通勤习惯），便可以“即时”对其进行混淆。因此，solve 的预计算情形可以视为参与者双方同意使用的配置文件。

6.2 局限性与未来工作

本文存在一些局限性，我们计划在后续研究中加以解决。

我们的工作仅考虑了熵（用于衡量不可链接性）与平方误差（用于衡量失真）的一种特定组合。除了熵之外，其他不确定性度量也可能有助于表达协作式智能交通系统中的不可链接性（Wagner 和 Eckhoff，2018年）。此外，过去有些研究多目标跟踪及其应用的作者分析并推荐了平方误差以外的失真度量（Gorji 等，2011年）。

本文仅研究了对最低可实现不可链接性的高度保证（例如，理性下界，见假设 1 和 2）。然而，为了进一步提升我们的方法的实用性，针对最坏情况（例如，攻击者极其强大）得出的指标，可以辅以依赖于不那么悲观情形（例如，攻击者较弱）的其他指标。

我们的模型仅考虑 2 名参与者。这大幅限制了 CAM 的排列数量：因此，每个隐藏状态只有 2 个带标签状态（见图 2）。因此，定义可观测状态的方法也很简单（见图 3）。我们计划在未来增加参与者数量。然而，参与者数量增多后，定义最优可观测状态及相应的转移概率并非易事。这是因为需要分析 $\mathbb{R}^{z}, \mathbf{z} \geq 1$ 中更复杂的结构和变换，以优化 CAM 的即时混淆和联合混淆。

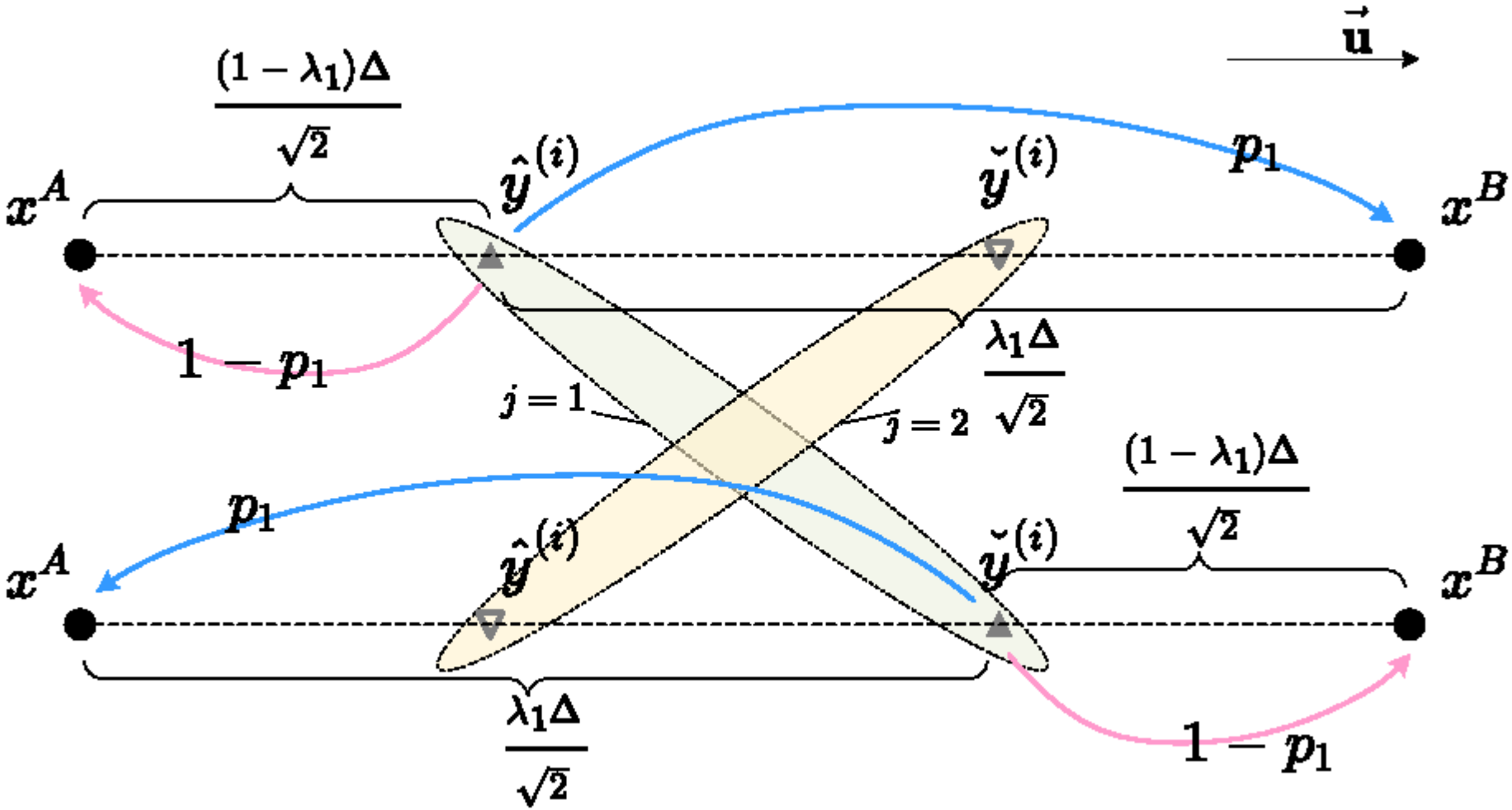
7 致谢

本研究由德国萨克森州的公共资金共同资助。

参考资料

Andr´es, M. E., Bordenabe, N. E., Chatzikoko- lakis, K., & Palamidessi, C.（2013年）。地理不可分辨性：基于位置的系统的差分隐私。《2013年ACM SIGSAC计算机与通信安全会议论文集》，901–914。Blackman, S. S.（1986年1月1日）。雷达应用中的多目标跟踪。Blackman, S.（2004）。用于多目标跟踪的多假设跟踪。《IEEE航空航天与电子系统杂志》，19(1)，5–18。Bordenabe, N. E., Chatzikokolakis, K., & Palamidessi, C.（2014年）。位置隐私的最优地理不可分辨机制。《2014年ACM SIGSAC计算机与通信安全会议论文集》，251–262。Camenisch, J., Drijvers, M., Lehmann, A., Neven, G., & Towa, P.（2020年）。用于车对车通信的区域加密与匿名认证。2020年IEEE欧洲安全与隐私研讨会（EuroS&P），405–424。合作智能交通系统——第7部分：隐私方面（技术报告（TR ISO/TR 17427-7:2015(E)））。（2015年）。国际标准化组织。瑞士日内瓦。Corser, G. P., Fu, H., & Banihani, A.（2016年）。评估车辆通信与应用中的位置隐私。《IEEE智能交通系统汇刊》，17(9)，2658–2667。Croft, W. L., Sack, J.-R., & Shi, W.（2019年）。面向人脸图像的差分隐私混淆。载于 A. Holzinger、P. Kieseberg、A. M. Tjoa 和 E. Weippl（编），《机器学习与知识提取》（第229–249页）。Springer国际出版。网络安全工程（国际标准 ISO/SAE 21434:2021）。（2021年）。国际标准化组织。瑞士日内瓦。Escher, S., Sontowski, M., Berling, K., K¨opsell, S., & Strufe, T.（2021年）。您的汽车有多容易被跟踪：欧洲合作智能交通系统的分析

假名方案。2021年IEEE第93届车辆技术会议（VTC2021-春季），1–6。IT安全评估准则——第2部分：安全功能组件（国际标准 ISO/IEC 15408-2:2022(E)）。（2022年）。国际标准化组织。瑞士日内瓦。Geng, Q., & Viswanath, P.（2016年）。近似差分隐私的最优噪声添加机制。《IEEE信息论汇刊》，62(2)，952–969。Gorissen, B. L., Yanıko ̆glu, ̇I., & den Hertog, D.（2015年）。鲁棒优化实用指南。《Omega》，53，124–137。Gorji, A. A., Tharmarasa, R., & Kirubarajan, T.（2011年）。多目标跟踪问题的性能度量。第14届国际信息融合会议，1–8。Guan, B., Takbiri, N., Goeckel, D., Houmansadr, A., & Pishro-Nik, H.（2022年）。基于超字符串的序列混淆：挫败模式匹配攻击。《IEEE物联网期刊》，1–1。Hacking, I.（1975年）。不可分辨者的同一性。《哲学杂志》，72(9)，249–256。Hicks, C., & Garcia, F. D.（2020年）。用于车联网中不可链接ECDSA假名的车辆DAA方案。2020年IEEE欧洲安全与隐私研讨会（EuroS&P），460–473。智能交通系统（ITS）；安全；信任与隐私管理；第2版（技术规范 TS 102 941）。（2021年）。欧洲电信标准化协会。法国索菲亚·安蒂波利斯。智能交通系统（ITS）；用户与应用需求；第2部分：应用与设施层通用数据字典（技术规范 ETSI TS 102 894-2）。（2018年）。欧洲电信标准化协会。法国索菲亚·安蒂波利斯。智能交通系统（ITS）；车辆通信；基本应用集；第2部分：协作感知基本服务规范（欧洲标准 DS/ETSI EN 302 637-2）。（2019年）。欧洲电信标准化协会。法国索菲亚·安蒂波利斯。ITS站管理——第1部分：本地管理（国际标准 ISO 24102-1:2018(E)）。（2018年）。国际标准化组织。瑞士日内瓦。Karim Emara, W. W.（2013年）。车载自组织网络中的基于信标的车辆跟踪。Lehmann, F., & Pieczynski, W.（2020）。利用模型阶数约简，在三元马尔可夫模型中进行次优卡尔曼滤波。《IEEE 信号处理快报》，27，1100–1104。Montazeri, Z., Houmansadr, A., & Pishro-Nik, H.（2017）。利用匿名化在无线设备中实现完美的位置隐私。《IEEE 信息取证与安全汇刊》，12(11)，2683–2698。Shokri, R.（2012）。量化并保护位置隐私。洛桑联邦理工学院。洛桑。Shokri, R., Freudiger, J., & Hubaux, J.-P.（2010）。位置隐私统一框架。第三届隐私增强技术热点研讨会（HotPETS 2010），74–94。Shokri, R., Theodorakopoulos, G., & Troncoso, C.（2016）。位置轨迹中的隐私博弈：用于优化位置隐私的博弈论框架。《ACM 隐私与安全汇刊》，19(4)，11:1–11:31。Sniedovich, M.（2016）。Wald 强大的极大极小准则：教程。《国际运筹研究汇刊》，23(4)，625–653。Takbiri, N., Houmansadr, A., Goeckel, D. L., & Pishro-Nik, H.（2017）。匿名化与混淆条件下的位置隐私极限。2017年IEEE 国际信息论研讨会（ISIT），764–768。Takbiri, N., Houmansadr, A., Goeckel, D. L., & Pishro-Nik, H.（2020）。统计匹配中相互依赖用户的隐私。《IEEE 信息论汇刊》，66(9)，5842–5865-5842–5865。Wagner, I., & Eckhoff, D.（2018）。技术隐私度量：系统综述。《ACM 计算机调查》（CSUR），51(3)，1–38-1–38。Wiedersheim, B., Ma, Z., Kargl, F., & Papadimitratos, P.（2010）。车际网络中的隐私：为什么简单的假名更换还不够。2010年第七届无线按需网络系统与服务国际会议（WONS），176–183。


$\vec{\mathbf{u}}$
$\frac{(1-\lambda_1)\Delta}{\sqrt{2}}$
x^A
$\hat{y}^{(i)}$
$\check{y}^{(i)}$
p_1
x^B
$1-p_1$
$\frac{\lambda_1\Delta}{\sqrt{2}}$
$j=1$
$j=2$
$\frac{(1-\lambda_1)\Delta}{\sqrt{2}}$
p_1
x^A
$\hat{y}^{(i)}$
$\check{y}^{(i)}$
x^B
$\frac{\lambda_1\Delta}{\sqrt{2}}$
$1-p_1$

此图“orcid.png”可从以下位置获取 PNG 格式：

http://arxiv.org/ps/2301.04130v1