arXiv:2602.01804v1 [eess.SY] 2 Feb 2026

促进数字交通市场中的数据协作：隐私保护机制的作用

Qiqing Wang[a]、Haokun Yu[b]和Kaidi Yang[a]*

[a]新加坡国立大学土木与环境工程系，新加坡 *117576*，*Engineering Drive 2*号[b]新加坡国立大学运筹学与分析研究所，新加坡 *117602*，*Innovation 4.0 Research Link 3*号

关键词：数字交通市场 数据协作博弈 数据隐私 隐私保护机制 交通管理与控制

在大数据时代，市政当局（MA）与移动性提供商（MPs）之间的数据协作为交通系统带来了巨大益处。参与协作可以改善这些数据所有者的服务运营（例如减少延误），但也可能引发隐私问题，降低其共享数据的意愿。具体而言，数据所有者可能担心共享数据会泄露有关其客户移动模式或商业机密的敏感信息，从而导致协作失败。本文探讨了隐私保护机制如何在此类情境中促进数据协作。我们提出一个博弈论框架，用于研究交通利益相关者之间的数据共享，尤其考虑基于扰动的隐私保护机制。数值研究表明，降低对数据质量的期望能够激励自愿数据共享，提升MA和MPs双方的交通相关福利。我们的研究结果为政策制定者和系统设计者提供了可操作的见解，说明隐私保护技术如何帮助打破数据孤岛，并推动协作型、注重隐私的交通系统的发展。

1 引言

大数据一直是推动交通运输领域持续发展的强大催化剂。具体而言，传感技术和车辆技术（如路侧传感器、网联车辆等）的快速进步，以及出行服务的广泛普及，推动了数据可用性的激增。数据的激增使包括市政当局（MAs）和出行服务提供商

* 通讯作者。电子邮箱：kaidi.yang@nus.edu.sg。

（MPs）（如网约车公司）在内的利益相关者能够制定有效的数据驱动分析和决策策略，包括交通状态估计/预测（Xu等人，2020年）、交通控制（Guo等人，2019年）、路径规划（Huang等人，2020年）和车队管理（Guo等人，2023年）。这些新兴的智能解决方案显著提升了交通系统的安全性、出行便利性和可持续性，从而推动了数字交通市场的发展。预计2025年至2030年，交通运输领域大数据分析的全球市场规模将以23.8%的复合年增长率增长，到2030年最终达到预计430亿美元的规模（Grand View Research，2025年）。

释放大数据在交通市场中所带来效益的关键在于数据协作。这是因为交通数据通常分散在多个利益相关者之间，形成彼此孤立的数据孤岛。具体而言，由于服务类型（例如，市场管理者提供的交通控制服务与移动性提供商提供的拼车服务）、服务覆盖范围的时间和空间尺度（例如，节假日与工作日、公共交通网络与道路网络）以及所采用的传感技术（例如，环形检测器、移动传感器等）存在显著差异，每个利益相关者通常只掌握能够反映交通系统有限部分的数据。因此，各利益相关者掌握的数据往往具有很强的互补性；由于市场中的利益相关者之间存在频繁互动，这些数据能够创造显著的互惠价值。这为数据协作奠定了坚实基础，使其成为一种双赢策略，并有望显著提升社会福利（对市场管理者而言）和服务质量（对移动性提供商而言）（Smichowski，2018年）。现有文献已广泛证明数据协作的益处。例如，研究表明，市场管理者整合多个移动性提供商的运营数据，能够显著改进交通状态估计算法，从而实现更高效的交通管理，提升社会福利（Zheng等人，2018年；Wang和Yang，2024年；Wang等人，2024a）。此外，不同移动性提供商之间的协作对于开发广受欢迎的出行即服务（MaaS）解决方案发挥了重要作用。借助此类方案，移动性提供商可以协调彼此的服务覆盖范围，从而吸引更多出行者使用其平台并拓展市场覆盖面（Pantelidis等人，2024年；Smichowski，2018年）。

然而，数据协作并非易事，因为数据共享可能会给共享数据的利益相关者带来隐私问题。这类隐私问题有两个方面。首先，共享数据（例如轨迹数据、调查数据等）可能包含有关个人出行模式的敏感信息；如果这些信息被第三方或对手获取，便可能被用于推断个人画像或社会经济地位。个人信息通常受到日益严格的法规保护，例如欧盟的《通用数据保护条例》（GDPR）（欧盟法规，2018年）；违反该条例可能导致高额罚款（《纽约时报》，2024年；Telecompaper，2024年；Bloomberg Law，2022年）。其次，除了个人层面的信息外，共享数据还可能被综合利用，以推断利益相关者的运营机密，例如敏感的算法参数（He和Chow，2020年；Strava，2018年）。这些参数一旦泄露，可能会削弱利益相关者的竞争优势，甚至危及运营安全（Yu等人，2025年）。这两类隐私问题对在实践中促进协作构成了重大挑战。例如，香港的私人巴士运营商拒绝向政府共享运营数据（Today，2018年）；欧洲的汽车制造商则拒绝与业内同行共享电动汽车测试数据（Martens和Zhao，2021年）。

为应对隐私问题，近年来人们提出了多种先进的隐私保护机制，以促进数据协作，同时最大限度地减少敏感信息的披露（Jin 等，2022；Gati 等，2021；Ying 等，2022；Gao 等，2022；Zhang 等，2022；Tsao 等，2022、2023；Tan 和 Yang，2024；Zhou 和 Yang，2024；Wang 和 Yang，2024）。这些机制大致可分为基于加密的技术和基于扰动的技术。基于加密的技术（如安全多方计算）使利益相关者能够协同计算函数，而无需明确共享其数据（Tsao 等，2022；Tan 和 Yang，2024；Ying 等，2022）。尽管这些技

术能够实现较高的数据准确性，但其实施通常计算密集，因此交通利益相关者可能难以高效共享大量数据。相比之下，基于扰动的技术（如差分隐私（差分隐私）和信息论方法）允许数据所有者生成原始数据的扰动版本，在降低隐私风险的同时，保留其对数据请求方的效用（Chow，2018年；Tsao 等，2023；Zhang 等，2022；Zhou 和 Yang，2024）。换言之，基于扰动的技术有望在共享原始数据与不共享之间找到“最佳平衡点”，从而优化数据协作收益与相关隐私问题之间的权衡。得益于其可解释性和计算效率，基于扰动的技术已在实践中得到广泛应用。例如，美国人口普查局使用差分隐私发布2020年人口普查数据（美国人口普查局，2020年），在保护个人隐私的同时支持统计分析。

尽管隐私保护机制具有潜力，但鲜有研究探讨其对交通利益相关者（即市场管理者和移动性提供商）在数据协作中的数据共享策略的影响。现有研究大多聚焦于个人，采用理论分析（Fainmesser等人，2023年）和调查（Berke等人，2023年；Cottrill等人，2015年），表明隐私保护机制能够提升个人共享数据的意愿。然而，从个人层面研究得出的见解无法直接推广至交通利益相关者。这是因为利益相关者之间存在紧密互动，而这种互动源于其运营活动的重叠性和相互依赖性。例如，一家移动性提供商与市场管理者共享数据以改善交通管理，可能带来系统层面的改善（如减少拥堵），从而使在同一交通网络中运营的所有移动性提供商受益。这种紧密互动意味着，一方共享的数据可能影响其他利益相关者的决策，进而影响其自身运营。近期鲜有研究考虑此类动态（Butler等人，2021年；Liu和Chow，2022年），例如竞合，即为合作目的共享的数据也可能同时被用于竞争情境，尤其是在利益相关者运营的服务或市场相互重叠时（Smichowski，2018年）。尽管有多项研究采用博弈论分析探讨交通利益相关者之间的数据协作（Liu和Chow，2022年；Chen和Keppo，2025年），但这些研究假设利益相关者只决定是否共享数据以及共享多少数据，并未考虑隐私保护机制引入的潜在扰动，如图1所示。引入扰动必然会降低共享数据的质量，由此带来两项基本挑战。首先，数据质量下降可能导致数据请求方的需求与共享数据的质量不匹配，从而使底层博弈结构更加复杂。其次，引入随机扰动会增加准确估计市场管理者和移动性提供商效用的难度。据我们所知，尚无研究在数据协作模型中纳入隐私保护机制，因此，人们对这些机制在促进数据协作中所发挥作用的理解仍存在重要空白。

贡献声明。为弥补这一关键研究空白，我们着手研究隐私保护机制在促进企业对政府（B2G）数字交通市场中的数据协作方面所发挥的作用。我们的贡献主要有四个方面。第一，我们提出了一个斯塔克尔伯格博弈论框架，用于研究交通利益相关者（如 MA 和 MPs）之间的企业对政府数据协作，并纳入基于扰动的隐私保护机制。第二，在温和且符合实际的假设条件下，我们从理论上分析了所提出的斯塔克尔伯格博弈的性质（例如均衡的存在性、成功开展数据协作的条件）。第三，我们通过设计一种将差分隐私与 LWR 理论相结合的差分隐私交通需求估计机制，在一个实际的交通数据协作场景（即协同交通信号优化）中具体应用了该斯塔克尔伯格博弈论框架。第四，我们开展了基于真实世界数据的数值研究，以验证典型数据共享场景中的分析结果，并提供管理启示

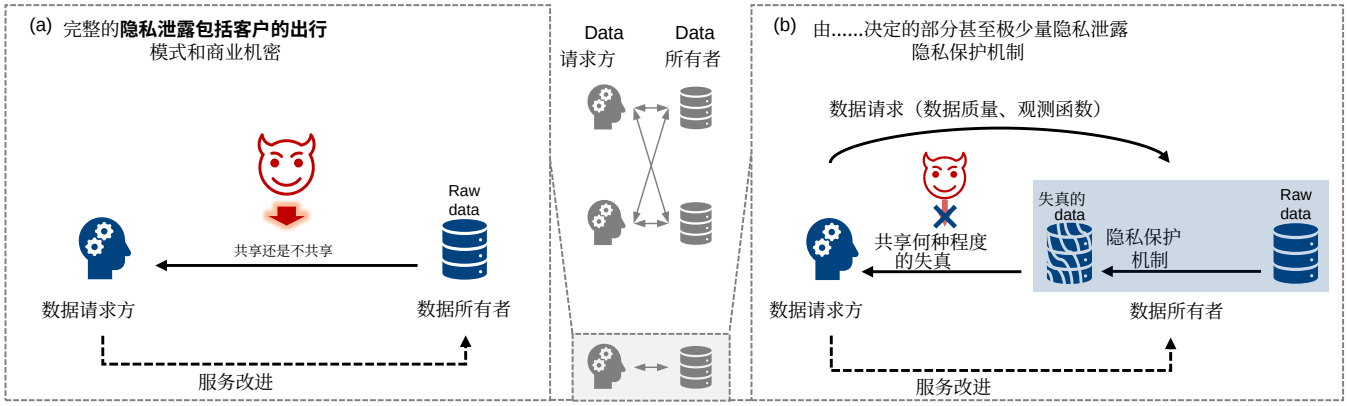


图1：交通系统中无（a）和有（b）隐私保护机制时数据协作的示意图。没有隐私保护机制（a）时，数据共享仅限于二元选择，并可能导致隐私完全泄露。有隐私保护机制（b）时，数据所有者可决定数据失真程度，在隐私泄露与数据效用之间进行权衡。

以及关于隐私保护机制何时以及如何能够促进数字交通市场中数据协作的政策指导。总体而言，我们的研究率先提供了一套方法论框架和管理启示，以推动对数字交通市场中隐私保护数据协作的认识与实践。

本文结构如下。第2节对数据共享拓扑进行形式化描述，并介绍B2G数字市场中的数据协作问题。第3节提出数据协作博弈的一般模型，并通过理论分析研究其均衡。第4节针对交通领域的特定场景——协同交通信号优化，具体说明模型。第5节进行数值实验，并提出促进数据协作的可行见解。第6节总结全文。

2 问题陈述

考虑一组 $\mathcal{K} = \{1, 2, \ldots, K\}$ 个利益相关者（在本文语境下，每个利益相关者都是市场管理者或移动性提供商），他们打算加入数字交通市场中的数据协作博弈。我们用有向图 $\mathcal{G} = (\mathcal{K}, \mathcal{E})$ 表示这些利益相关者的数据共享拓扑，其中边 $(k, l) \in \mathcal{E}$ 表示利益相关者 k 向利益相关者 l 请求数据。我们用集合 $\mathcal{H}^+k = \{l | (l,k) \in \mathcal{E}\}$ 表示利益相关者 k 考虑向其共享数据的利益相关者集合，并用集合 $\mathcal{H}^-k = \{l | (k, l) \in \mathcal{E}\}$ 表示利益相关者 k 考虑向其请求数据的利益相关者集合。根据数据共享拓扑，每个利益相关者在数据协作中的角色可以是需要其他利益相关者提供数据的数据请求方，和/或考虑与其他利益相关者共享数据的数据所有者。我们假设每个利益相关者至少承担其中一种角色，否则他们就没有参与数据协作的意义。从数学上说，数据请求方集合可表示为 $\mathcal{K}_R = \{k | \mathcal{H}^- k \neq \emptyset\}$，数据所有者集合可表示为 $\mathcal{K}_O = \{k | \mathcal{H}^+ k \neq \emptyset\}$。我们进一步假设数据共享不可转让，即请求方收到的数据不能再与他人共享。通常通过合同协议来执行这一限制，合同协议会明确禁止将数据转让给第三方。本文使用的重要变量列于附录 D 的表4中。

注 1（数据共享拓扑示例）。在K 个利益相关者参与的现实世界交通数据协作博弈中，数据共享拓扑有两种典型示例：

(1) 星型拓扑（见图 2 (a)）。在此情形下，市场管理者向一组MPs请求数据。从数学角度看，仅存在一个数据请求方k（即市场管理者），其中 $\mathscr{H}^-{}_k = \mathscr{K}\backslash\{k\}$，而其他参与者（即MPs）均为数据所有者。对于每个数据所有者$l \in \mathscr{K}\backslash\{k\}$, $\mathscr{H}^+{}_l = \{k\}$。

(2) 全连接拓扑（见图 2 (b)）。在此情形下，一组MPs（例如，公共交通运营商、网约车公司等）开展数据协作，彼此共享数据。从数学角度看，每个利益相关者既是数据请求方，也是数据所有者。对于每个利益相关者k,$\mathscr{H}^+{}_k = \mathscr{K}\backslash\{k\}$ and $\mathscr{H}^-{}_k = \mathscr{K}\backslash\{k\}$。

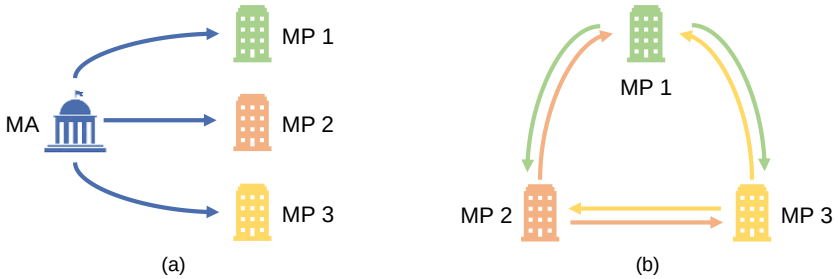


图 2：示例 数据共享拓扑。（a）星型拓扑，（b）全连接拓扑 .

表 1：数字交通市场中星型数据协作的示例。

类型	研究与实践中的示例
市场管理者利用交通观测数据，高效地进行交通管理并管控	交通信号控制（Wang 等，2024b,a），交通状态估计（Wang 和 Yang，2024年），出行即服务（MaaS）平台（Cottrill，2020年）
市场管理者提供便捷的多式联运服务，并结合移动性提供商的车队运营	电动汽车开发（Martens和Zhao，2021年）,自动驾驶车辆开发—展（Chen和Keppo，2025年）
市场管理者推动深科技出行创新与MPs开展研发及产品验证	

在本研究中，我们主要关注星形数据共享拓扑，即单个市场管理者向多个移动性提供商请求出行数据。这种情形符合企业对政府数字市场的结构，私营实体向公共部门提供数据，以支持治理、政策实施和基础设施优化（Chen和Keppo，2025年；Wang等人，2024b；Martens和Zhao，2021年）。这种拓扑在交通系统中尤为重要，在研究和实践中都十分常见，表 1 所总结的示例便说明了这一点。具体而言，星形情形的重要性体现在两个方面。首先，许多交通控制和管理系统由市场管理者建立并运营，因此自然需要获取移动性提供商的数据，以提高运营效率（例如，市场管理者利用移动性提供商的车辆轨迹数据实施交通信号控制（Wang等人，2024b））。其次，在许多经济体中，市场管理者通常在制定和执行监管政策方面享有优先权，因此在推动多个移动性提供商之间的数据协作时具有天然优势（例如，市场管理者汇总移动性提供商的实时导航数据，以支持科学研究并促进研发活动（Martens和Zhao，

2021年））。需要指出的是，移动性提供商是否参与企业对政府数字市场取决于其自主意愿，并由其商业利益驱动。事实上，数据共享能够为移动性提供商带来切实收益，例如通过减少延误改进服务（Wang等人，2024b），以及获得政策支持（Martens和Zhao，2021年）。因此，在隐私关键场景中，移动性提供商仍可能受到激励，权衡协作带来的潜在效用收益与隐私问题，从而就数据协作作出策略性决策。

为此，我们考虑一种星型数据协作博弈，其中一个作为数据请求方的市场管理者（记为 $\mathscr{K}_R = \{1\}$），以及$K-1$个作为数据所有者的MPs（记为 $\mathscr{K}_O = \{2,3,\cdots,K\}$）组成该博弈。

3 一般模型与分析

本节针对交通系统中企业对政府市场中单个市场管理者与多个MPs之间的数据协作博弈，介绍一般模型及其均衡分析。

3.1 数据协作博弈建模

我们介绍数据请求方和数据所有者的策略形式化（见第3.1.1节和第3.1.2节）及效用函数（见第3.1.3节）。具体而言，我们将问题建模为斯塔克尔伯格博弈：数据请求方（领导者）选择策略，以分配其数据请求/期望；数据所有者（跟随者）则对此作出响应，决定是否采用隐私保护机制参与数据协作。

3.1.1 数据请求方的策略

市场管理者 MA 确定数据请求*datarequest*的具体内容，并将其告知每个相关数据所有者$k \in \mathscr{K}_O$。这里，我们允许数据请求方针对每个数据所有者定制请求，因为不同数据所有者持有的数据可能存在异质性（例如，车辆渗透率不同的轨迹数据）。数据请求的格式见定义1。

定义1（数据请求）。市场管理者 MA 发送的数据请求由一组 $\mathrm{R}_1 = \{(\mathrm{F}\sigma_k, \Phi_k, \mathrm{d}_k)\}_{k\in\mathscr{K}_O}$构成，其中每个元组 $(\mathrm{F}\sigma_k, \Phi_k, \mathrm{d}_k)$指定发送给数据所有者 $\mathrm{k} \in \mathscr{K}_O$的请求。这里，函数 $\mathrm{F}\sigma_k(\cdot)$表示一个由 σ_k 参数化的观测函数，用于描述利益相关者 k 的系统状态 xxx_k 到所需观测 yyy_k 的映射（即 $\mathrm{yyy}_k = \mathrm{F}\sigma_k(\mathrm{xxx}_k)$）。函数 $\Phi_k(\cdot)$和标量 d_k 用于构建对来自数据所有者 k 的共享数据的数据质量要求，即 $\Phi_k(\mathrm{yyy}_k, \hat{}\mathrm{yyy}_k) \leq \mathrm{d}_k$，其中 $\Phi_k(\mathrm{yyy}_k, \hat{}\mathrm{yyy}_k)$表示共享数据 $\hat{}\mathrm{yyy}_k$ 与真实数据 yyy_k之间的距离（例如，l_2-范数），而 d_k 表示阈值。在数据请求中，函数 Φ_k 和函数 $\mathrm{F}\sigma_k$ 的形式以及参数 σ_k 均为预先确定的，且市场管理者 MA 和移动性提供商 MP k 均已知；数据请求方只需优化参数 $\mathrm{ddd} = \{\mathrm{d}_k\}_{k\in\mathscr{K}_O}$。

我们对数据请求作如下说明。首先，观测函数 $F\sigma_k(\cdot)$使数据所有者能够按照预先定义的变换共享数据，这在实践中具有广泛的适用性。例如，观测函数可以涉及时间和空间上的数据采样或数据聚合，这两种方式都常用于降低共享数据的分辨率。在此情境下，聚合级别和采样频率可以作为参数 σ_k 的一部分。其次，数据质量的决策

源于数据所有者可能采用基于扰动的隐私保护机制，从而给数据引入失真。数据质量阈值d_k代表数据请求方对数据所有者k所提供数据准确性的容忍度，因为过于不准确的数据可能无法满足请求方的用途。总体而言，观测函数与数据质量相结合，体现了数据请求方对数据的预期。

3.1.2 数据所有者的策略

基于扰动的隐私保护机制通过在数据效用与数据隐私之间提供可控的权衡，从根本上改变了数据共享方式。这类机制的定义见定义2。

定义2（基于扰动的隐私保护机制）。基于扰动的隐私保护机制定义为一个随机函数$G\varepsilon_k(\cdot)$，由隐私预算 ε_k参数化，用于将真实数据映射为其失真后的对应数据。此类函数可通过噪声注入（例如差分隐私）、数据压缩（例如联邦学习）或其他扰动操作来实现。隐私预算 ε_k 控制施加于真实数据的扰动强度（即 ε_k 值越小，隐私保护越强，因而扰动也越强），从而决定数据效用与数据隐私之间的权衡。

收到数据请求后，每个数据所有者$k \in \mathcal{K}_O$ 都会针对来自数据请求方市场管理者的请求R_1，选择采用隐私保护机制的数据共享策略数据共享策略；该策略也是一种纯策略。数据共享策略的具体说明见定义3。

定义3（数据共享策略）。数据所有者$k \in \mathcal{K}_O$ 针对数据请求方$k \in \mathcal{K}_R$所选择的数据共享策略被构造为一个元组$S_k = (a_k, G\varepsilon_k)$。其中，$a_k \in \{0,1\}$是一个二元变量，表示数据是否与请求方（市场管理者）共享。若$a_k = 1$，即数据被共享，则函数$G\varepsilon_k(\cdot)$表示一种随机化的基于扰动的隐私保护机制。$G\varepsilon_k$ 的形式是预先确定的，并被假定为公开知识，参数（a_k，ε_k）将由数据所有者进行优化。

我们对数据共享策略作如下说明。首先，我们假设隐私保护机制由数据所有者设计并投入，原因有三：i）只有数据所有者能够访问这些原始数据；ii）只有数据所有者知道他们希望保护何种敏感信息；iii）他们希望通过数据共享提高市场利润。其次，我们假设数据所有者将严格遵循数据请求中规定的要求，数据请求方可使用基于加密的方法（如零知识证明）对此进行验证（Tsao等人，2022年）。

3.1.3 数据参与者效用

数据协作中所有数据参与者的决策过程可以表述为斯塔克尔伯格博弈。具体而言，数据所有者根据数据请求方发布的数据请求规格，决定其数据共享策略，并以非合作方式展开竞争。各参与者的效用函数如下，其中我们分别使用k 和 $-k$ 表示自身数据所有者和其他数据所有者。

跟随者效用最大化。数据请求方发送数据请求后，每个数据所有者$k \in \mathcal{K}_O$旨在确定其行动a_k 和 ε_k，以最大化一个效用函数，该函数体现了以下两者之间的权衡：

交通相关福利与隐私成本，其可表示为

$$\max_{a_k,\varepsilon_k} \quad U_k(a_k,\varepsilon_k;a_{-k},\varepsilon_{-k},\boldsymbol{d}) = W_k\left(\left\{G_{\varepsilon_{k'}}(\boldsymbol{y}_{k'})\right\}_{k':a_{k'}=1}\right) - \beta_k\varepsilon_k a_k, \tag{1}$$

$$\text{s.t.} \quad \Phi_k(\varepsilon_k) = \bar{\Phi}_k(\boldsymbol{y}_k, G_{\varepsilon_k}(\boldsymbol{y}_k)) \le d_k + Z(1-a_k), \tag{2}$$

$$a_k \in \{0,1\}, \quad \varepsilon_k \in (0,1), \tag{3}$$

其中，决策变量包括一个二元变量a_k，用于表示移动性提供商k是否参与数据协作并共享数据，以及隐私预算ε_k。目标函数（1）包含：（i）由市场管理者确定的期望交通相关福利$W_k(\cdot)$，该福利取决于利益相关者之间交换的所有数据$\left\{G^{\varepsilon_{k'}}(yyyk')\right\}_{k':a_{k'}=1}$，即参与情况；以及（ii）由参数$\varepsilon_k$衡量的隐私保护机制的隐私泄露惩罚。第二项以$\beta_k$加权，以反映利益相关者对隐私与效用之间权衡的偏好。不失一般性，我们假设$W_k(/0) = 0$，因为没有数据可用于提升福利。约束（2）表示，由$\bar{\Phi}_k(\cdot)$计算得出的失真数据$G^{\varepsilon_k}(yyyk)$与真实数据$yyyk$之间的期望距离不得超过数据请求方k要求的数据质量阈值d_k。其中，Z是一个足够大的数，以确保该约束仅对$a_k = 1$的请求方k生效（即共享数据）。

为刻画跟随者的行为，我们对效用函数作出以下正则性假设。

假设 1（效用函数的正则性条件）。我们考虑满足以下条件的正则效用函数：

(1) 如果任意数据所有者j不参与数据协作，即$a_j = 0$，那么其隐私预算不应影响任何效用函数。换言之，$\forall k$，U_k独立于ε_j。这可以写为

$$U_k(\cdots,\varepsilon_j',\varepsilon_{j+1},\cdots,a_{j-1},0,\cdots) = U_k(\cdots,\varepsilon_j,\varepsilon_{j+1},\cdots,a_{j-1},0,\cdots), \forall k, j \in \mathcal{K}_O, \forall \varepsilon_j, \varepsilon_j' \tag{4}$$

(2) 将隐私预算趋近于 0 的数据报告（即扰动强度趋于无穷大），等同于不报告数据。具体而言，如果$a_j = 1$且$\varepsilon_j \to 0$，则任意数据所有者的效用都会收敛至$a_j = 0$时的效用，写作

$$\lim_{\varepsilon_j \to 0^+} U_k(\cdots,\varepsilon_j,\varepsilon_{j+1},\cdots,a_{j-1},1,\cdots) = U_k(\cdots,\varepsilon_j',\varepsilon_{j+1},\cdots,a_{j-1},0,\cdots), \forall k, j \in \mathcal{K}_O, \forall \varepsilon_j' \tag{5}$$

关于假设 1，我们作如下说明。第一个条件确保为非活跃数据所有者分配隐私预算不会影响任何效用函数，因为他们并非数据协作博弈的参与者。这可以避免效用函数中出现潜在的混淆。第二个条件指出，在市场管理者具备充分能力、能够不论数据质量如何都识别和分析数据的假设下，以最大扰动程度共享数据（即数据中除随机噪声外不包含任何有用信息）既不会使任何数据所有者受益，也不会损害其效用。

领导者效用最大化。给定所有跟随者 $= \{\text{的响应}_k\}_{k\in\mathcal{K}_O}$ 和 $\boldsymbol{\varepsilon} = \{\varepsilon_k\}_{k\in\mathcal{K}_O}$，领导者，即数据请求方 MA/利益相关者 1，通过求解以下优化问题来确定数据请求规格 $= \{_k\}_{k\in\mathcal{K}_O}$：

$$\max_{\boldsymbol{d}\in\mathcal{B}} \quad U_1(\boldsymbol{d};\boldsymbol{a},\boldsymbol{\varepsilon}) = W_1\left(\left\{G_{\varepsilon_k(\boldsymbol{d})}(\boldsymbol{y}_k)\right\}_{k\in\mathcal{K}_O:a_k=1}\right), \tag{6}$$

其中，目标函数（6）包含预期的交通相关福利收益函数$W_1(\cdot)$，该函数取决于利益相关者之间交换的所有数据$\left\{G^{\varepsilon_{k'}}(yyyk')\right\}_{k':a_{k'}=1}$。在此，不失一般性地假设$\mathcal{B}$是有限的，因为距离阈值可以通过数值方法离散化。

3.2 均衡分析

接下来，我们分析数据协作博弈。为此，我们首先在定义 4 中确立均衡概念。

定义 4 (斯塔克尔伯格-纳什均衡（子博弈精炼纳什均衡）)。 策略组合（ddd^*，aaa^*，$\boldsymbol{\varepsilon}^*$）是数据协作博弈的子博弈精炼纳什均衡，当且仅当：
(1)（a_k^*，ε_k^*）是给定领导者策略d^*时跟随者k的纳什均衡

$$(a_k^*, \varepsilon_k^*) \in \arg\max_{a_k, \varepsilon_k} U_k\left(a_k, \varepsilon_k; a_{-k}^*, \varepsilon_{-k}^*, \boldsymbol{d}^*\right), \quad \forall k \in \mathscr{K}_O, \tag{7}$$

(2) ddd^* 是领导者在预期跟随者的最佳响应时的最佳响应，即，

$$\boldsymbol{d}^* \in \arg\max_{\boldsymbol{d} \in \mathscr{B}} \left\{ U_1\left(\boldsymbol{d}; \boldsymbol{a}^*, \boldsymbol{\varepsilon}^*\right) \mid (\boldsymbol{a}^*, \boldsymbol{\varepsilon}^*) \in \mathrm{BR}(\boldsymbol{d}) \right\}, \tag{8}$$

其中，BR(*ddd*)表示针对领导者决策*ddd*的跟随者的最佳响应，即 ，

$$\mathrm{BR}(\boldsymbol{d}) = \left\{ (\boldsymbol{a}^*, \boldsymbol{\varepsilon}^*) : (a_k^*, \varepsilon_k^*) \in \arg\max_{a_k, \varepsilon_k} U_k\left(a_k, \varepsilon_k; a_{-k}^*, \varepsilon_{-k}^*, \boldsymbol{d}\right), \quad \forall k \in \mathscr{K}_O \right\}. \tag{9}$$

根据均衡概念，我们接下来将在第3.2.1节进行均衡分析，并在第3.2.2节建立成功的数据协作条件。

3.2.1 均衡的存在性

我们首先通过提出若干符合实际的假设来开展均衡分析。假设2规定了期望距离与隐私预算之间呈单调递减关系。这一假设符合实际情况，因为较高的隐私预算意味着添加到查询统计数据中的噪声较少，从而生成的FoQ轨迹点与真实数据更加吻合。

假设 2（期望距离单调递减）。 失真数据 $\tilde{}yyy_k = G\varepsilon_k(yyy_k)$与真实数据$yyy_k$之间的期望距离，记为 $\Phi_k(\varepsilon_k) = \bar{\Phi}_k(yyy_k, G\varepsilon_k(yyy_k))$，随着 ε_k的增加而单调递减。

根据假设2，我们可以将约束（2）改写为

$$\phi_k \le \varepsilon_k \le 1,\ k \in \mathscr{K}_O, \tag{10}$$

其中 $\phi_k = \Phi^{-1}(d_k)$，由于单调性，该表达式定义良好。

接下来，我们介绍关于效用函数的假设。先从其表达式的简化形式说起。根据假设 1，如果数据所有者k 选择不参与（$a_k = 0$），效用$U_k(\boldsymbol{\varepsilon}, aaa)$与 ε_k无关，即$U_k(\varepsilon_k, \varepsilon_{-k}, aaa) = U_k(0_k, \varepsilon_{-k}, aaa)$。为简化记号，我们定义$z_k = a_k\varepsilon_k$。于是，变量$z_k$ 衡量数据所有者k所贡献的信息水平，同时表征参与情况和噪声水平。因此，我们将简化效用函数定义为

$$\bar{U}_k(\boldsymbol{z}) = U_k(\boldsymbol{\varepsilon}, \boldsymbol{a}) \tag{11}$$

其中$z_k = 0$表示$a_k = 0$。反过来，给定$z_k > 0$，我们可以还原$a_k = 1$和$\varepsilon_k = z_k$。

不失一般性，假设 $\bar{U}(\cdot)$ 二阶可微。在许多现实交通系统中，数据边际价值会随着信息供给的增加而降低。因此，随着隐私预算增加，增加参与者数量和提高数据准确性的边际收益也会降低。以下假设概括了这一点。

假设 3（数据共享边际收益递减）。 数据共享的边际收益随变量 zzz 递减，具体而言，$\partial^2 \bar{U}(zzz)/\partial z_k \partial z_j < 0$。

此外，无论是否参与系统，数据所有者的效用往往都会随着信息供给增加而提高，因为更大的数据池通常能改善系统服务。

假设 4（对他人行动的单调性）。 数据所有者 k 的效用U_k关于其他任何数据所有者的行动均为非递减，即 $\partial U_k / \partial z_j \geq 0, \quad \forall j \neq k$。

基于这些假设，我们可以将数据所有者k的模型简化如下：

$$\max_{a_k, z_k} \quad \bar{U}_k(z_k, z_{-k}), \tag{12}$$

$$\text{s.t.} \quad \phi_k a_k \leq z_k \leq a_k, \tag{13}$$

$$a_k \in \{0,1\}, z_k \in [0,1] \tag{14}$$

其中$\bar{U}_k$关于z_k是凹函数。对于任意给定的a_k，该优化问题是一个凸优化问题。

由于二元行动aaa与连续行动zzz相结合，直接分析跟随者博弈颇具挑战。因此，我们采用两阶段方法：上层确定aaa，下层则在给定aaa的情况下计算纳什均衡。

定义 5（跟随者模型的两阶段均衡寻求过程）。我们采用两阶段过程：上层确定 aaa，下层则在给定 aaa 的情况下计算纳什均衡。在此情形下，纳什均衡定义为一对 $(\text{aaa}^*, \text{zzz}^*)$，满足：(i) 对于固定的 aaa^*，没有参与者能够通过偏离 z_k^* 单方面提高其效用，即 $U_k(z_k^*, z_{-k}^*, \text{aaa}^*) \geq U_k(z_k, z_{-k}^*, \text{aaa}^*)$；(ii) 在上层，没有参与者能够通过偏离 aaa^* 单方面提高其效用，即 $V_k(a_k^*, a_{-k}^*) \geq V_k(a_k, a_{-k}^*)\ \forall a_k \in \{0,1\}, \forall k$，其中 $V_k(\text{aaa})$ 表示参与者 k 在与 aaa 对应的下层均衡处的效用。

命题1（两阶段过程的有效性）。 通过两阶段过程得到的任何均衡都是原始跟随者博弈的纳什均衡，反之亦然。

这一结论可由纳什均衡的定义直接得出：在两阶段过程的均衡中，任何参与者都无法通过单方面改变其参与决策或隐私预算来提高自身效用。反向结论同理可证。

这一两阶段过程得出了一个用于求解原始跟随者博弈均衡的算法。由于移动性提供商通常很少（例如，2或3个），我们可以先枚举所有可能的上层决策aaa，据此计算给定aaa时的纳什均衡。

定理2（下层均衡的存在性）。 固定 $aaa \in \{0,1\}^K$。则存在唯一的下层跟随者纳什均衡。

证明。 我们注意到，(i) 对于具有$z_k \in [\phi_k, 1]$和$a_k = 1$的活跃参与者，以及具有$z_k = 0$和$a_k = 0$的非活跃参与者，其集合非空、紧且凸。效用函数关于zzz连续，且关于自身行动严格凹。因此，根据 Glicksberg（1952），存在唯一的纳什均衡。 □

如前所述，效用函数$V_k(aaa)$表示参与者k 在对应于aaa的下层均衡中的效用。接下来，我们分析上层博弈，该博弈由有限参与者构成，每位参与者均采取二元行动。我们首先在定理3中证明效用函数是子模的（证明见附录B）。

定理3 （上层二元博弈的次模性）。 在下层均衡 zzz^*（aaa）下，当 aaa 固定时，上层效用函数 V_k（aaa）：$=\bar{U}_k(zzz^*$（aaa）$)$ 在（$a_i,a\,j$）上具有递减差异，对任意 $i \neq j$ 均成立；即，

$$V_k(1,1_\ell,a_{-k\ell}) - V_k(1,0_\ell,a_{-k\ell}) \leq V_k(0,1_\ell,a_{-k\ell}) - V_k(0,0_\ell,a_{-k\ell}), \quad \forall a_{-k\ell} \in \{0,1\}^{K-2}.s \tag{15}$$

请注意，在双参与者子模博弈中，纯纳什均衡总是存在的。然而，一般而言，当参与者超过三名时，此类博弈可能不存在纯纳什均衡。我们在附录C中提供了一个反例。因此，对于参与者超过三名的情况，我们计算上层博弈的混合均衡。

获得跟随者博弈的均衡后，领导者MA通过简单枚举确定最优距离阈值d_i 。为此，我们可以证明所提出的数据协作博弈存在SNE。

定理4 （数据协作的子博弈精炼纳什均衡的存在性）。 数据协作博弈存在一个子博弈精炼纳什均衡。

证明。 根据定理2和定理3，对于任意领导者决策 ddd，跟随者博弈至少存在一个均衡。随后，领导者可以通过在有限集合 $\mathscr{B}$ 中进行简单枚举，确定距离阈值 $bbb \in \mathscr{B}$ 。 □

最后，我们对求解斯塔克尔伯格博弈均衡的过程作如下说明。

注 2（子博弈精炼纳什均衡的求解）. 我们作两点说明。首先，由于在典型的交通数据协作博弈中MPs的数量较少，我们通过枚举确定领导者的决策以及每个跟随者的上层决策。由于每个跟随者的下层问题是强凸的，相应的下层均衡是唯一的。因此，对于任意固定的领导者决策和跟随者的上层决策，下层均衡可通过求解相应的原始-对偶KKT系统得到。其次，在数据提供方更多的更一般情形中，计算子博弈精炼纳什均衡等价于求解带均衡约束的数学规划（MPEC），在我们的设定中它可化为带互补约束的数学规划（MPCC）。由于MPCC通常违反标准约束资格条件，一种成熟的补救方法是对互补约束进行松弛或正则化，并求解一系列收敛到MPCC解的光滑非线性规划（NLPs）（Scholtes，2001年；Hoheisel等人，2013年）。每个松弛后的非线性规划随后可由标准的大规模非线性规划算法处理，例如序列二次规划（SQP）或内点法。在交通应用中，一种常见做法是通过分解和交替更新来处理MPEC，并利用已有的变分不等式（VI）求解器（例如与SDMSA/MSA相结合的对角化）求解下层交通均衡，而不是将所得MPCC作为单个大规模非线性规划来求解（Liu和Chow，2023年）。

3.2.2 成功的数据协作的条件

接下来，我们分析成功的数据协作的条件（即至少有一个移动性提供商/跟随者决定共享数据），这意味着 $aaa = \mathbf{0}$ 不是该数据协作博弈的子博弈精炼纳什均衡的策略组合。我们假设 $\bar{U}_k(000) = 0$ 对应于无人决定共享数据的情形（即$aaa = \mathbf{0}$）。

我们首先在定义 6中定义每个移动性提供商k 的协作收益。

定义 6 （每个移动性提供商k的协作收益）。 对于任意固定的z_{-k}，移动性提供商k 的协作收益定义为

$$\Delta_k(\phi_k, z_{-k}) = \max_{z_k \in [\phi_k, 1]} \bar{U}_k(z_k, z_{-k}) - \bar{U}_k(0, z_{-k}) \tag{16}$$

该协作 每个 MP k 的协作收益具有命题 5 和命题 6 所确立的两个性质 .

命题 5 （市场管理者的数据质量要求的影响d_k）。 对于任意固定的 z_{-k}，协作收益Δ_k（ϕ_k,z_{-k}）关于 ϕ_k 非递增。等价地，放宽市场管理者的最低质量要求（即降低 ϕ_k，这可通过增大 d_k 实现）会增强 *MP k* 共享数据的激励。

证明。 对于 $\phi'_k \geq \phi_k$，可行集满足 $[\phi'_k\ 1] \subseteq [\phi_k, 1]$。因此，$\max_{z_k \in [\phi'_k, 1]} \bar{U}_k(z_k, z_{-k}) \leq \max_{z_k \in [\phi_k, 1]} \bar{U}_k(z_k, z_{-k})$，因此，$\Delta_k(\phi'_k, z_{-k}) \leq \Delta_k(\phi_k, z_{-k})$。 □

命题 6 （隐私成本权重的影响 β_k）。 对于任意固定的 （ϕ_k,z_{-k}），协作收益 Δ_k（ϕ_k,z_{-k}）随 β_k增大而递减。

证明。由于 $\bar{U}_k(z_k, z_{-k}; \beta_k) = \bar{U}_k(z_k, z_{-k}) - \beta_k z_k$，有 $\partial \bar{U}_k(z_k, z_{-k}; \beta_k) / \partial \beta_k = -z_k < 0$.因此，$\bar{U}_k(z_k, z_{-k}; \beta_k)$ 关于 β_k 单调递减。因此，协作收益 $\Delta_k(\phi_k, z_{-k})$ 关于 β_k 单调递减。 □

于是，我们可以得到如下充分条件。

命题7 （充分 条件：成功的数据协作）。 如果存在一个 *MP k*，使得 *t*

$$\Delta_k(\phi_k, \mathbf{0}) > 0, \tag{17}$$

则全零策略组合 $aaa = \mathbf{0}$ 不可能是跟随者的最佳响应。

证明。 考虑由任意领导者策略 *ddd*诱导的跟随者上层博弈。当 $aaa = \mathbf{0}$时，MP k获得收益 $\bar{U}_k(000) = 0$。如果 $\Delta_k(\phi_k, 000) > 0$，则存在一个可行的 $z_k \in [\phi_k, 1]$，使得 $\bar{U}_k(z_k, 000) > \bar{U}_k(000) = 0$。因此，MP k 存在从 $a_k = 0$ 到 $a_k = 1$的正向偏离。因此，$aaa = \mathbf{0}$ 不可能是跟随者的最佳响应。 □

我们针对命题7作如下说明，以满足条件 $\Delta_k(\phi_k, 000) > 0$。首先，这在实际中是可行的。由于市场管理者无法获取交通观测，来自一家MP的少量信息就足以显著改善U_k（仅需10%的轨迹数据即可进行交通状态估计）。其次，市场管理者在请求数据时可以提供更高的d_k，从而使 MP k 能够最优地选择其信息水平$z_k \in [\phi_k, 1]$，以满足命题5中所述条件。第三，MPs在准备数据时可能会产生额外的固定成本（例如员工时间和运营投入），这可以建模为加到参与条件左侧的负常数。为补偿此类成本，市场管理者可以向参与的MPs提供专属福利（例如货币激励），这可以表示为加到该条件左侧的正常数。这种机制设计将作为未来的研究工作。

4 交通信号优化的模型设定

本节针对协同交通信号优化问题给出一般模型。在该问题中，单个市场管理者旨在利用多个MP提供的轨迹数据，更新其固定时制交通信号配时方案。

4.1 协同信号优化问题

我们考虑一条由交叉口集合 $\mathscr{I}$ 构成的双向干道，其基数为$I = |\mathscr{I}|$。每个交叉口$i \in \mathscr{I}$按照包含相位集合 $\mathscr{P}_i$的周期信号方案运行，其中每个相位$p \in \mathscr{P}_i$控制该交叉口处由m索引的一组交通流向 $\mathscr{M}_p$。由于协调信号控制的要求，信号配时方案包括所有交叉口共用的公共周期长度C、每个相位的绿信比分配$ggg = \{g_{ip}\}_{p\in\mathscr{P}_i, i\in\mathscr{I}}$（即交叉口$i$中的相位$p$），以及每个交叉口$i$的相位差$ooo = \{o_i\}_{i\in\mathscr{I}}$。

市场管理者旨在将过时的信号配时方案$(C^{\text{old}}, ggg^{\text{old}}, ooo^{\text{old}})$更新为新的优化方案（$C^*, ggg^*, ooo^*$），具体是通过应用信号控制策略 π，该策略基于所有交通流向的估计交通需求$\hat{}qqq = \{\hat{}q_m\}_{m\in\mathscr{M}_p, p\in\mathscr{P}_i, i\in\mathscr{I}}$，即（$C^*, ggg^*, ooo^*$）$= \pi$（$\hat{}qqq$）。为简便起见，我们在后续各节中直接使用$m \in \mathscr{M}$表示整个路网中的所有交通流向。本文采用一种简单且常用的方法，将（1）用于计算信号周期和绿信比分配的韦伯斯特公式与（2）用于计算相位差的基于绿波带的相位差优化方法相结合。该组合方法已被许多文献采用（Little等人，1981年；Gartner等人，1991年；Zhang等人，2015年），并继续用于近期的研究与实践。需要说明的是，其他干道级信号控制器也可集成到我们的方法框架中。

二者C 和ggg 均可使用韦伯斯特公式计算得出

$$C^* = \max_{i\in\mathscr{I}} \left(\frac{1.5L_i + 5}{1 - \sum_{p\in\mathscr{P}_i} \max_{m\in\mathscr{M}_p} \left(\frac{\hat{q}_m}{q_m^c} \right)} \right), \tag{18}$$

$$g_{ip}^* = \frac{\max_{m\in\mathscr{M}_p} \left(\frac{\hat{q}_m}{q_m^c} \right)}{\sum_{p\in\mathscr{P}_i} \max_{m\in\mathscr{M}_p} \left(\frac{\hat{q}_m}{q_m^c} \right)} C, \quad \forall i \in \mathscr{I}, p \in \mathscr{P}_i, \tag{19}$$

其中，L_i 表示交叉口i每个周期的总损失时间，$\max_{m\in\mathscr{M}_p} \left(\hat{}q_m^c \right)_q$表示相位$p$的关键交通流向的流量比。$\hat{}q_m$ 和q_m^c 分别表示交通流向m的估计交通需求和通行能力流量。

更新后的相位差ooo^* 可通过求解基于绿波带的优化问题得到，该问题旨在最大化出城与入站车流的绿波带宽，从而减少车辆停车。具体而言，这可以表述为混合整数线性规划问题（Little等人，1981年；Gartner等人，1991年；Zhang等人，2015年）。我们采用一种经典且成熟的方法MAXBAND（Little等人，1981年），并利用Gartner等人（1991年）；Zhang等人（2015年）中出城与入站权重的调整系数来优化相位差。在该优化问题中，我们有$3I + 1$个决策变量，包括出城方向绿波带宽b、入站带宽 $\bar{}b$、I 个出城方向相对相位差 $\boldsymbol{\zeta} = \{\zeta_i\}_{i=1}^I$、$I$个入站方向相对相位差 $\bar{\boldsymbol{\zeta}} = \{\bar{\zeta}_i\}_{i=1}^I$，以及$I-1$ 个循环整数变量$MMM = \{M_i\}_{i=1}^{I-1}$。该

问题可以写为

$$\max_{b,\bar{b},\boldsymbol{\zeta},\bar{\boldsymbol{\zeta}},\boldsymbol{m}} \quad \left(\max_{m\in\mathscr{M}^{out}} \frac{\hat{q}_m}{q_m^c}\right)^{\alpha} b + \left(\max_{m\in\mathscr{M}^{in}} \frac{\hat{q}_m}{q_m^c}\right)^{\alpha} \bar{b}, \tag{20a}$$

$$\text{s.t.} \quad \zeta_i + b \le 1 - r_i, \quad \forall i \in \{1,\dots,I\}, \tag{20b}$$

$$\bar{\zeta}_i + \bar{b} \le 1 - \bar{r}_i, \quad \forall i \in \{1,\dots,I\}, \tag{20c}$$

$$\begin{aligned}&\left(\zeta_i + \bar{\zeta}_i\right) - \left(\zeta_{i+1} + \bar{\zeta}_{i+1}\right) + \left(t_i + \bar{t}_i\right) + \Delta_i - \Delta_{i+1} \\ &= -\frac{1}{2}\left(r_i + \bar{r}_i\right) + \frac{1}{2}\left(r_{i+1} + \bar{r}_{i+1}\right) + \left(\bar{e}_i + e_{i+1}\right) + M_i \quad \forall i \in \{1,\dots,I-1\},\end{aligned} \tag{20d}$$

$$M_i \in \mathbb{Z}, \quad \forall i \in \{1 \dots I-1\}, \tag{20e}$$

$$b, \bar{b}, \zeta_i, \bar{\zeta}_i \ge 0, \quad \forall i \in \{1,\dots,I\}, \tag{20f}$$

其中，目标函数（20a）同时包含出站带宽和入站带宽，并分别以所有出站和入站交通流向中关键交通流向的ˆ ˆ流量比$\max_{m\in\mathscr{M}^{out}} q_m/q_m^c$和$\max_{m\in\mathscr{M}^{in}} q_m/q_m^c$进行加权。– α 表示指数幂，默认值为1。约束（20b-20f）确保出站和入站绿波带在每个交叉口均可行。r_i（$\bar{r}_i$）表示交叉口i处的出站（入站）红灯时间，单位为周期。Δ_i表示从 $\bar{r}_i$的中心到最近的r_i中心的时间，单位为周期。t_i（$\bar{t}$）表示从交叉口i到交叉口$i+1$的出站（入站）行程时间，单位为周期。e_i（$\bar{e}_i$）表示排队清空时间（单位为周期），即离开交叉口i时出站（入站）带宽的提前量。更多细节，读者可参阅Little等人（1981）。

随后，交叉口o_i之间的相位差i和$i+1$可按如下方式计算 s

$$o_{i,i+1} + \frac{1}{2}r_{i+1} + \zeta_{i+1} + c_{i+1} = \frac{1}{2}r_i + \zeta_i + t_i \tag{21}$$

由于我们假设两个方向的绿信比分配保持不变，因此驶入方向的相位差与驶出方向的相位差相同。为校准相位差系统，将o_1设置为0。

总而言之，市场管理者采用了一种控制策略（C^*，ggg^*，ooo^*）$=\pi$（ˆqqq），将韦伯斯特公式与基于绿波带的相位差优化方法相结合，如公式（18）—（21）所示。所有沿主干道行驶的车辆所产生的平均延误可表示为

$$W_1(C^*, \boldsymbol{g}^*, \boldsymbol{o}^*, \boldsymbol{q}) = W_1\left(\pi(\hat{\boldsymbol{q}}), \boldsymbol{q}\right). \tag{22}$$

其中，$W_1(\cdot)$表示由于市场管理者的定时信号控制策略而导致的所有车辆的延误函数，该函数可通过解析方法（Manual，2000；Yao等人，2019）或基于仿真的方法计算。需要注意的是，由于控制策略（C^*，ggg^*，ooo^*）$=\pi$（ˆqqq），所得的平均延误也可以表示为真实需求qqq和估计需求ˆqqq的函数。

在此信号优化问题中，市场管理者需要必要的信息来估计关键的交通需求参数 ˆ$qqq = \{\hat{q}_m\}_{m\in\mathscr{M}}$，这些参数对于公式（18）至（20a）至关重要。关于市场管理者与MPs建立数据协作的动机，我们作如下说明。

备注 3（市场管理者与移动性提供商建立数据协作的动机）。 市场管理者寻求与移动性提供商开展协作的原因有两点。首先，由于缺乏传感设备（例如环形检测器），市场管理者无法直接观测这些关键交通需求参数。这种情况十分普遍，因为许多城市干道由于安装和维护成本高昂，仍缺乏环形检测器和交通摄像头等可靠覆盖的路侧传感器，因而采用定时信号控制器。这一限制

自然促使市场管理者向交通网络中的其他利益相关者寻求/请求数据，从而发起数据协作博弈。其次，移动性提供商可以在交通网络中运营车队，并拥有不同的市场份额和渗透率。因此，每家移动性提供商都能够获取其自有车队在近期主干道服务运营期间收集的轨迹数据。这些轨迹数据捕捉了车辆通过信号交叉口时细粒度的时空运动，提供了丰富的行为洞见，例如排队模式和消散率。利用这些关键指标可以推断出重要的交通需求参数；相关数据用于交通估计的有效性已在研究和实际部署中得到验证（Wang et al., 2024b）。值得注意的是，与受限于固定安装点的环形检测器相比，轨迹数据能够为交通观测提供显著更广泛且更灵活的空间覆盖范围。

4.2 数据请求方（市场管理者）的观测函数

然后，我们可以明确市场管理者在其策略中使用的观测函数$F\sigma_k(\cdot)$。具体而言，市场管理者请求与队列前端（FoQ）对应的轨迹点，以估计关键的交通需求参数。我们用xxx_k表示轨迹数据，其中轨迹的索引为$n \in \mathcal{N}_k = \{1, 2, \cdots, N_k\}$。根据定义7，这些FoQ轨迹点表示车辆因红灯信号而在交叉口加入队列的时刻。由于FoQ代表停止状态与自由流交通状态之间的边界，根据Lighthill–Whitham–Richards（LWR）理论（Lighthill and Whitham, 1955；Richards, 1956），这些轨迹点包含有关交叉口到达流的关键信息。

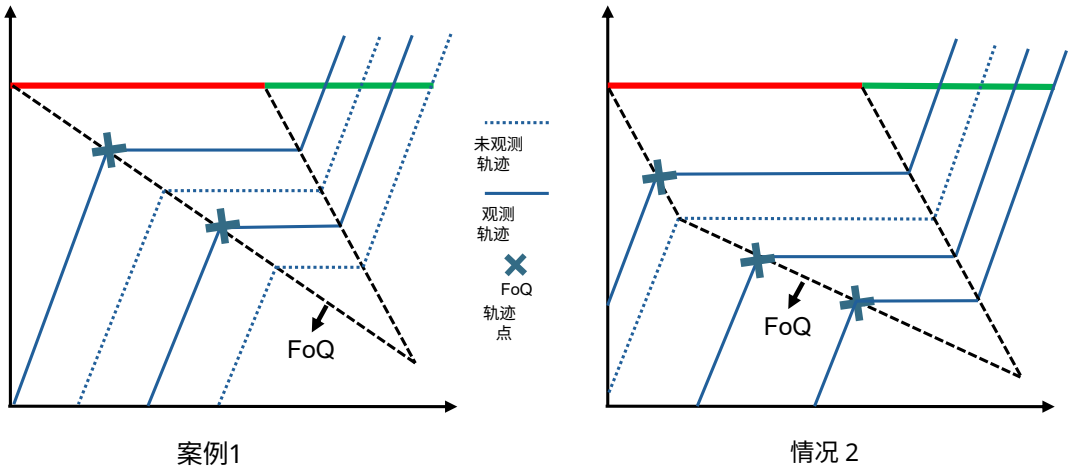


图3：时空图中的激波示意图。蓝色虚线表示未观测到的车辆轨迹，蓝色实线表示已观测到的车辆轨迹。蓝色叉形点表示 FoQ轨迹点。

定义7（FoQ轨迹点作为观测函数）。对于交叉口 i，FoQ轨迹点定义为一个元组 $(\mathrm{h}_{kn}^{m}, \mathrm{t}_{kn}^{m})$，其中 h_{kn}^{m} 和 t_{kn}^{m}分别表示由数据所有者 k 运营的车辆 n 在交通流向 $\mathrm{m} \in \mathscr{M}_p$, $\mathrm{p} \in \mathscr{P}_i$, $\mathrm{i} \in \mathscr{I}$中加入队列时的位置和时间。具体而言，时间 t_{kn}^{m} 是当前信号周期内该交通流向 m 红灯开始后经过的时间，h_{kn}^{m} 表示停车位置到该交通流向 m 停止线的距离，以负值表示。然后，我们定义观测函数 $\mathrm{F}\sigma_k$，将车辆轨迹数据 xxx_k映射为相应的 FoQ轨迹点 $\mathrm{yyy}_k = \left\{ \left(\mathrm{h}_{kn}^{m}, \mathrm{t}_{kn}^{m}\right) \right\}_{\mathrm{n} \in \mathcal{N}_k, \mathrm{m} \in \mathscr{M}}$, i.e., $\mathrm{yyy}_k = \mathrm{F}\sigma_k(\mathrm{xxx}_k)$。请注意，只有在交叉口完全停车的车辆轨迹才会产生 FoQ轨迹点。

利用收集到的 FoQ轨迹点 yyy，市场管理者将运用 LWR 理论估计到达流 $\hat{}qqq = B(yyy)$，其中 $B(\cdot)$ 表示基于 LWR 理论的流程，该流程涉及 FoQ曲线的估计。对于每个交通流向m，我们假设交通动态由三角形基本图 $Q_m(k) = \min\{v_{f,m}k + w_m(k_{j,m} - k)\}$来刻画。基于该模型，我们可以在时空图上示意激波（见图3）。这里有两种典型情形：（1）到达流在一个信号周期内保持恒定，形成线性FoQ曲线；（2）到达流因上游放行而以饱和流开始（第 a 段），随后降至恒定到达流（第 b 段），形成分段线性FoQ曲线，其第一段的斜率为 w_m。在这两种情况下，都可以根据 LWR 理论，从 FoQ曲线的斜率 ψ_m 重建到达流。

因此，我们旨在估计 FoQ曲线的斜率 ψ_m。为此，我们考虑一种简化情形：交通流向m 的需求在所有信号周期[1]中保持不变。在此假设下，可将多个周期的 FoQ轨迹点汇总为一个具有代表性的周期，从而增加可用于估计的数据。随后，我们采用线性回归拟合两种情形下的 FoQ曲线，结果汇总于表2。请注意，表2不仅给出了系数的点估计，还给出了根据标准线性回归理论推导的真实参数分布。分布估计对于以统计上协调一致的方式整合从多个 MPs 收集的异质数据至关重要。值得注意的是，由于 MPs 通常拥有很长的数据序列，样本量N 往往较大。因此，Student$t$分布渐近趋近于高斯分布。我们采用贝叶斯框架整合来自不同 MPs 的估计值，以获得真实斜率的估计值 $\psi_m \sim \mathcal{N}(\hat{}\psi_m, \hat{}\sigma_m)$，并进一步用于估计各交通流向$m$的到达流 $\hat{}q_m$。估计得到的到达流 $\hat{}qqq$将用于推导策略 $(C^*, ggg^*, ooo^*) = \pi(\hat{}qqq)$对应的信号配时方案。

4.3 数据所有者（MPs）的隐私保护机制

接下来，我们将说明MPs在其策略中采用的隐私保护机制。具体而言，我们分别在第4.3.1节和第4.3.2节介绍每个跟随者MP的隐私考量和隐私保护机制。我们将采用针对轨迹级和计数级数据敏感性定制的差分隐私（DP）机制，对隐私目标进行形式化并加以保护。

4.3.1 隐私考量

在交通信号优化问题中，我们使用差分隐私（DP）来刻画隐私性。差分隐私是对隐私的数学严谨定义，能够构建针对数据共享或分析所导致的隐私泄露提供强有力保障的机制。考虑数据集 $D = \{\tau_1, \tau_2, \cdots, \tau_N\}$，其中每个 τ_i表示一条数据记录（例如车辆轨迹或FoQ轨迹点）。假设外部方希望通过查询函数 $f(\cdot)$查询该数据集，以提取某些统计洞见。典型查询包括数据点总数、数据点均值等。差分隐私背后的核心思想是对这些统计输出进行适当扰动，使相邻数据集（即略有不同的数据集）产生统计上不可区分的输出，从而降低推断攻击的风险。差分隐私已被广泛

[1]这一假设可以放宽：将每个信号周期，c 的斜率 ψ_{mc}视为高斯分布随机变量，并使用线性混合模型（LMM）构建问题。例如，在案例1中，LMM可以表示为$h = \psi_{m,c}t + \xi_m$，$\psi_{m,c} \sim N(\psi_m, \sigma^2)$。这使模型能够捕捉到达流在不同信号周期之间变化的实际情况。不过，本文不对这一泛化进行探讨，因为这超出了本文的主要研究范围。

表2：FoQ曲线估计。如图3所示，FoQ曲线在案例1中建模为线性函数，在情况 2中建模为分段线性函数。在这些公式中，$ttt_k^m = \{t_k^{mn}\}_{n\in\mathcal{N}}$和$hhh_k^m = \{h_k^{mn}\}_{n\in\mathcal{N}}$分别表示FoQ点的时间和空间坐标。参数 ψ 和 γ 表示待估计的斜率和截距，而 Λ 汇总了直接根据FoQ点计算得到的聚合量。带帽变量，即$\hat{\sigma}$ 和$\hat{\psi}$，表示根据FoQ点获得的相应统计估计。变量 $\mathcal{T}_n$ 表示具有t 分布的学生分布，其自由度为n 。

Case	模型	模型估计值
案例1	$h = \psi_{m,k}t + \xi_{m,k}, \quad \xi_{m,k} \sim \mathcal{N}(0, \sigma^2_{m,k})$	来自 FoQ 点的汇总数量： $\Lambda^{\mathrm{TH}}_{mk} = \sum_{n=1}^{N} t^m_{kn} h^m_{kn}, \quad \Lambda^{\mathrm{T}}_{mk} = \sum_{n=1}^{N} (t^m_{kn})^2, \quad \Lambda^{\mathrm{H}}_{mk} = \sum_{n=1}^{N} (h^m_{kn})^2$ 斜率的统计估计 $\psi_{m,k}$ 和标准残差偏差 $\sigma_{m,k}$： $\hat{\psi}_{m,k} = \frac{\Lambda^{\mathrm{TH}}_{mk}}{\Lambda^{\mathrm{T}}_{mk}}, \quad \hat{\sigma}^2_{m,k} = \frac{1}{N-1}\left(\Lambda^{\mathrm{H}}_{mk} - \frac{(\Lambda^{\mathrm{TH}}_{mk})^2}{\Lambda^{\mathrm{T}}_{mk}}\right)$ $\psi_{m,k} \sim \mathcal{T}_{N-1}\left(\hat{\psi}_{m,k}, \frac{\hat{\sigma}^2_{mk}}{\Lambda^{\mathrm{T}}_{mk}}\right) \approx \mathcal{N}\left(\hat{\psi}_{m,k}, \frac{\hat{\sigma}^2_{mk}}{\Lambda^{\mathrm{T}}_{mk}}\right)$,
情况 2 （第 a 段）	$h = wt + \gamma_{m,k} + \xi_{m,k}, \quad \xi_{m,k} \sim \mathcal{N}(0, \sigma^2_{m,k})$	截距 γ_m 的统计估计 $\gamma_{m,k}$ 以及标准差残差的标准差 $\sigma_{m,k}$： $\gamma_{m,k} = \frac{1}{N}\sum_{n=1}^{N}(h^m_{kn} - wt^m_{kn})$ $\hat{\sigma}^2_{m,k} = \frac{1}{N-1}\sum_{i=1}^{n}(h^m_{kn} - wt^m_{kn} - \gamma_{m,k})^2$ $\gamma_{m,k} \sim \mathcal{T}_{N-1}\left(\hat{\gamma}_{m,k}, \frac{\hat{\sigma}^2_{m,k}}{N}\right) \approx \mathcal{N}\left(\hat{\gamma}_{m,k}, \frac{\hat{\sigma}^2_{m,k}}{N}\right)$
情况 2 (第 b 段)	$h = \psi_{m,k}t + \gamma_{m,k} + \xi_{m,k}, \quad \xi_{m,k} \sim \mathcal{N}(0, \sigma^2_{m,k})$	斜率的统计估计值 $\psi_{m,k}$、截距 $\gamma_{m,k}$、以及残差标准差 $\sigma_{m,k}$： $ZZZ_{m,k} = [ttt^m_k, \mathbb{I}_{N\times 1}]$ $[\hat{\psi}_{m,k}, \hat{\gamma}_{m,k}]^\top = (ZZZ^\top_{m,k} ZZZ_{m,k})^{-1} ZZZ^\top_{m,k} hhh^m_k$ $\hat{\sigma}^2_{m,k} = \frac{1}{N-2}\sum_{i=1}^{n} \|hhh^m_k - ZZZ_{m,k}[\hat{\psi}_{m,k}, \hat{\gamma}_{m,k}]^\top\|^2$ $\psi_{m,k} \sim \mathcal{T}_{N-2}\left(\hat{\psi}_{m,k}, \hat{\sigma}^2_{m,k}(ZZZ^\top_{m,k} ZZZ_{m,k})^{-1}\right) \approx \mathcal{N}\left(\hat{\psi}_{m,k}, \hat{\sigma}^2_{m,k}(ZZZ^\top_{m,k} ZZZ_{m,k})^{-1}\right)$

应用于隐私保护数据生成，即基于经过扰动的关键统计量生成合成数据，以确保隐私，同时保留效用。

在这项工作中，每个数据所有者 $k \in \mathcal{K}_O$ 都关注两种隐私概念：（1）轨迹参与隐私，即对手是否能够推断共享数据集中是否存在某辆特定车辆的轨迹；（2）车队规模或分布隐私，即对手是否能够从共享数据中推断数据所有者车辆车队的聚合特征，例如特定时段内活跃车辆的数量或其空间分布。值得注意的是，若车队规模或车队空间分布泄露，可能会暴露移动性提供商的运营规模、市场份额或需求模式。

为应用差分隐私，我们首先通过在定义8和定义9中指定相应的邻接关系，将隐私概念形式化。

定义 8（轨迹级邻接）。 两个数据集 D 和 D' 是轨迹级相邻数据集，记作$D \sim_{\text{traj}} D'$，如果D可以通过从D'中移除一条数据记录得到（即$D \subset D',\ \ |D| = |D'| - 1$），反之亦然（即$D' \subset D\ \ |D'| = |D| - 1$）。

关于定义 8 中的轨迹级邻接，我们作两点说明。首先，这种邻接定义对应于轨迹参与隐私，在这一概念下，任何差分隐私机制都旨在防止对手推断个体轨迹是否包含在数据集中。其次，经过简单计算，该条件也可以写成 $|D \backslash D'| + |D' \backslash D| = 1$。

定义 9（计数级相邻数据集）。 两个数据集D 和 D' 是计数级相邻数据集，记作$D \sim_{\text{cnt}} D'$，如果D 可以通过从D' 中移除b 条数据记录得到（即$D \subset D',\ \ |D| = |D'| - b$），反之亦然（即$D' \subset D\ \ |D'| = |D| - b$）。

我们针对定义 9 中的计数级邻接作两点说明。首先，这种邻接概念自然对应于车队规模或车队空间分布的变化，即增加或移除固定数量的车辆轨迹。其次，该邻接关系假设车队规模变化 b 时，所有其他数据条目均保持不变。对于提供网约车或地图服务的 MP 而言，这一假设是合理的，因为当一部分用户在较短时间内（例如一个月）离开或加入平台时，其余用户往往会因较强的用户黏性而继续留存。

接下来，我们引入针对这两种隐私要求的差分隐私概念，具体定义见定义 10 和定义 11。特别地，我们采用 (ε, δ)-差分隐私（DP），其要求随机机制的输出在统计上不可区分。此处，参数 ε 通常被视为隐私预算，且该值越低，隐私保护越严格。

定义 10（轨迹级 $(\varepsilon_{\text{traj}}, \delta_{\text{traj}})$-差分隐私（DP））。 对于 $\varepsilon_{\text{traj}}, \delta_{\text{traj}} > 0$，将数据集 $D \in \mathscr{D}$映射到输出空间 $M(D) \in \mathscr{X}$的随机机制 M： $\mathscr{D} \to \mathscr{X}$，如果对于任意 $D \sim_{\text{traj}} D'$ 和任意事件 $E \subset \mathscr{X}$，均有

$$\mathbb{P}[M(D) \in E] \leq e^{\varepsilon_{\text{traj}}} \mathbb{P}\left[M\left(D'\right) \in E\right] + \delta_{\text{traj}} \tag{23}$$

定义11 (计数级 $(\varepsilon_b, \delta_b)$-差分隐私)。 对于 $\varepsilon_b, \delta_b > 0$，一个随机机制 M: $\mathscr{D} \to \mathscr{X}$将数据集 $D \in \mathscr{D}$映射到输出空间 $M(D) \in \mathscr{X}$的机制满足 $(\varepsilon_b, \delta_b)$-差分隐私，当且仅当对于任意 $D \sim_{\text{cnt}} D'$以及任意事件 $E \subset \mathscr{X}$，都有

$$\mathbb{P}[M(D) \in E] \leq e^{\varepsilon_b} \mathbb{P}\left[M\left(D'\right) \in E\right] + \delta_b \tag{24}$$

现在，我们通过定理 8 联系差分隐私的这两种定义。

定理 8（从轨迹级差分隐私转换为计数级差分隐私）。 如果随机机制 M 满足轨迹级（ε, δ）-差分隐私，那么它也满足计数级 $\left(b\varepsilon, \delta\sum_{b'=0}^{b} e^{b'\varepsilon}\right)$-差分隐私。

证明。 参见附录 8。 □

定理 8 表明，任何轨迹级差分隐私机制同时也满足计数级差分隐私。因此，我们只需以更严格的隐私预算保证轨迹级差分隐私。隐私保护机制的详细设计将在下一小节讨论。

4.3.2 隐私保护机制

基于上一节定义的隐私概念，我们设计了一种隐私保护机制，使市场管理者能够生成并与市场管理者共享 FoQ轨迹点。

定义12（隐私保护机制 $G\varepsilon_k$）。 隐私保护机制 $G\varepsilon_k$ 对于每个数据所有者 k，都是一种将原始 FoQ轨迹点 $yyy_k = \left\{ \left(h_{kn}^m、 t^m_{kn}\right) \right\}_{n\in\mathcal{N}_k、\ m\in\mathcal{M}}$转换为失真版本 $\hat{}yyy_k = \left\{ \left(\hat{}h_{kn}^m、 \hat{}t^m_{kn}\right) \right\}_{n\in\hat{}\mathcal{N}_k、\ m\in\mathcal{M}}$的映射，用于保护轨迹级隐私和计数级隐私，可描述为

$$\hat{\boldsymbol{y}}_k = G_{\varepsilon_k}(\boldsymbol{y}_k) = R \circ L_{\varepsilon_k} \circ f(\boldsymbol{y}_k) \tag{25}$$

其中，符号 ∘ 表示算子的复合。$f(\cdot)$表示查询函数，当市场管理者使用时，该函数返回关键统计量以保留数据效用，如表2所示。算子$L\varepsilon_k$ 表示噪声注入，用于扰动这些查询以保护隐私，而R 表示根据扰动后的关键统计量重构共享的FoQ轨迹点。

本文采用的隐私保护机制基于命题9所述的高斯机制。该机制根据查询对相邻关系的敏感度，向查询$f(\cdot)$添加经过精心设计的高斯噪声。

命题*9*（高斯机制）。设 $f: \mathcal{X}^n \to \mathbb{R}^d$ 为查询函数，其全局 l_2 敏感度$\Delta f = sup_{D\sim D'} \|f(D) - f(D')\|_2$。对于给定的数据集 $D \in \mathcal{X}n$ 以及任意 $\varepsilon, \delta \in (0,1)$，发布 $f(D) + \mathcal{N}(0, \sigma^2 I_{d\times d})$ 的机制，其中 $\sigma = \Delta f \sqrt{2log(1.25/\delta)}/\varepsilon$，满足 (ε, δ)-差分隐私。

命题9的证明见 Dwork（2006）。高斯机制通过添加按查询函数敏感度缩放的高斯噪声，实现 $(\varepsilon,\ \delta)$-差分隐私。接下来，我们将展示移动性提供商如何按照式（25）中介绍的流程应用高斯机制，生成隐私保护的 FoQ轨迹点。为避免重复，以下讨论将聚焦于表2中的案例1，因为其他案例的机制与之类似。

查询函数 $f(\cdot)$。第一步是仔细设计查询函数 $f(\cdot)$，以保留与数据生成任务相关的关键统计量。受表2中计算的启发，我们确定了N N个移动性提供商 MP k 希望保留的关键变量，包括 $\Lambda_{mk}^{\mathrm{T}} = \sum_{n=1} (t_{kn}^m)^2$、 $\Lambda_{mk}^{\mathrm{TH}} = \sum_{n=1} t_{kn}^m h_{kn}^m$，

$_N\Lambda^{\mathrm{H}}_{mk} = \sum (h^m_{kn})^2$, 以及 N。换言之，我们可以定义 $f(D) = [\rho^{\mathrm{T}}\Lambda^{\mathrm{T}}_{mk}, \rho^{\mathrm{TH}}\Lambda^{\mathrm{TH}}_{mk}, \rho^{\mathrm{H}}\Lambda^{\mathrm{H}}_{mk}, \rho^{\mathrm{N}}N],$，其中 $\boldsymbol{\rho} =_{n=1}[\rho^{\mathrm{T}}, \rho^{\mathrm{TH}}, \rho^{\mathrm{H}}, \rho^{\mathrm{N}}]$ 是市场管理者 MA 和移动性提供商 MP 公开已知的权重参数，用于平衡它们在敏感度中的贡献。

高斯机制 $L_k(\cdot)$。为设计高斯机制，我们首先计算纳入单个数据点对查询结果的敏感度。令 $t^m_{\max}$ 和 $h^m_{\max}$ 分别表示 t^m_{kn} 和 h^m_{kn} 的上界。然后，可以计算每个查询变量的敏感度，如下所示：

$$\Delta_{TH} = t^m_{\max} h^m_{\max}, \quad \Delta_T = (t^m_{\max})^2, \quad \Delta_H = (h^m_{\max})^2, \quad \Delta_N = 1 \tag{26}$$

然后，我们可以计算敏感度 $\Delta_f = \sqrt{\Delta^2_{TH}\rho^2_{TH} + \Delta^2_T\rho^2_T + \Delta^2_H\rho^2_H + \Delta^2_N\rho^2_N}$，并应用命题9生成满足 $\sigma_f = \Delta_f \log(1.25/\delta)/\varepsilon$ 的高斯噪声。

令 $\sigma^2_l = \frac{\rho^2_l}{\rho^2_{TH}+\rho^2_T+\rho^2_H+\rho^2_N}, l \in \{\mathrm{N}, \mathrm{T}, \mathrm{TH}, \mathrm{H}\}$ 表示各查询分量所占的方差份额。然后，我们通过添加以下噪声，对这些查询分量实施高斯机制。

$$\tilde{N} = N + \eta_{\mathrm{N}}, \quad \eta_{\mathrm{N}} \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{N}}\right), \quad \tilde{\Lambda}^{\mathrm{T}}_{mk} = \Lambda^{\mathrm{T}}_{mk} + \eta_{\mathrm{T}}, \quad \eta_{\mathrm{T}} \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{T}}\right) \tag{27}$$

$$\tilde{\Lambda}^{\mathrm{TH}}_{mk} = \Lambda^{\mathrm{TH}}_{mk} + \eta_{\mathrm{TH}}, \quad \eta_H \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{TH}}\right), \quad \tilde{\Lambda}^{\mathrm{H}}_{mk} = \Lambda^{\mathrm{H}}_{mk} + \eta_{\mathrm{H}}, \quad \eta_n \sim \mathcal{N}\left(0, \sigma^2_{\mathrm{TH}}\right) \tag{28}$$

FoQ点重构 $R(\cdot)$。利用扰动后的统计数据，我们可以生成隐私保护的合成FoQ轨迹点，其表达式为 $h = \tilde{\psi}_{m,\,k}t + \xi_{m,\,k}, \quad \xi_{m,\,k} \sim \mathcal{N}(0, \ \tilde{\sigma}^2_{m,\,k})$，其中 $\tilde{\psi}_{m,\,k}$ 和 $\tilde{\sigma}^2_{m,\,k}$ 表示为

$$\tilde{\psi}_{m,k} = \frac{\tilde{\Lambda}^{\mathrm{TH}}_{mk}}{\tilde{\Lambda}^{\mathrm{T}}_{mk}} \approx \hat{\psi}_{m,k} + \frac{1}{\hat{\Lambda}^{\mathrm{T}}_{mk}}\left(\eta_{TH} - \hat{\psi}_{m,k}\eta_T\right) \tag{29}$$

$$\tilde{\sigma}_{m,k} = \frac{1}{\tilde{N}-1}\left(\tilde{\Lambda}^{\mathrm{H}}_{mk} - \frac{(\tilde{\Lambda}^{\mathrm{TH}}_{mk})^2}{\tilde{\Lambda}^{\mathrm{T}}_{mk}}\right) \tag{30}$$

当样本量 N 较大时，近似成立。我们可以进一步分析FoQ斜率的分布，如下所示：

$$\tilde{\psi}_{m,k} \sim \mathcal{N}\left(\hat{\psi}_{m,k}, \frac{1}{(\hat{\Lambda}^{\mathrm{T}}_{mk})^2}\left(\sigma^2_{\mathrm{TH}} + \hat{\psi}^2_{m,k}\sigma^2_{\mathrm{T}}\right)\right) \tag{31}$$

由于 $\hat{\Lambda}^{\mathrm{T}}_{mk}$ 随样本量 N 呈二次增长，随着 N 增大，差分隐私噪声对 ψ 准确性的影响会减弱。

因此，该隐私保护机制为所有数据所有者提供轨迹级和计数级隐私保证。

4.4 用于确定效用函数的基于采样的方法

总之，协同交通信号优化的跟随者与领导者问题可写为如下形式。

(i) 跟随者效用最大化。每个跟随者求解与式 (1) 至式 (3) 相同的优化问题。请注意，在 (ε, δ)-差分隐私机制中，我们假设 δ为给定的小值，不将其视为决策变量。 ε_k 指定义 10 和定义 11 中定义的隐私预算。这些新的设定不会改变假设 1。

（ii）领导者效用最大化。领导者的效用最大化可以写成以下优化问题：

$$\max_{\boldsymbol{d}} \quad U_1(\boldsymbol{d};\boldsymbol{a},\boldsymbol{\varepsilon}) = \mathbb{E}_{\boldsymbol{\eta}}\left[W_1(C^*,\boldsymbol{g}^*,\boldsymbol{o}^*,\boldsymbol{q},\boldsymbol{\eta})\right], \tag{32}$$

$$\text{s.t.} \quad \tilde{Y}YY(\boldsymbol{\eta}) = \{G\varepsilon_k(yyy_k,\boldsymbol{\eta}_k)\}_{k\in\mathcal{K}_O:a_k=1}, \tag{33}$$

$$\tilde{q}qq(\boldsymbol{\eta}) = B(\tilde{Y}YY(\boldsymbol{\eta}_k)), \tag{34}$$

$$C^*(\boldsymbol{\eta}),\ ggg^*(\boldsymbol{\eta}),\ ooo^*(\boldsymbol{\eta}) = \pi(\tilde{q}qq(\boldsymbol{\eta})), \tag{35}$$

其中，目标函数 (32) 包含交通相关福利收益函数$W_1(\cdot)$，该函数取决于最优解 $(C^*,\ ggg^*,\ ooo^*)$，其中 $\boldsymbol{\eta}_k$表示数据所有者k添加的差分隐私噪声。为了获得最优解，市场管理者首先根据所有利益相关者的ˆ FoQ轨迹点 (YYY)，通过约束 (34) 估计所有交通流向的到达流；这些轨迹点可通过约束 (33) 计算得出。随后，可通过求解使用约束 (35) 的子优化问题获得最优解；该问题是在优化问题 (20) 中定义的混合整数线性规划定时交通信号控制问题。

效用函数的确定。在现实中，要获得上述斯塔克尔伯格博弈中的具体函数形式可能颇具挑战。具体而言，确定市场管理者和 MPs 的效用函数面临两项挑战。首先，市场管理者无法观测真实交通需求，只能接收来自 MPs 的经过扰动（含噪声）数据，因此难以计算交通相关福利。其次，MPs 的数据共享策略受概率机制（例如差分隐私）支配，这给期望效用的评估带来了不确定性。

为应对这两项挑战，我们采用以下技术：

(i)真实需求采样： 我们对真实交通需求与 MP 共享的扰动观测值之间的关系进行建模。具体而言，我们构建从共享的扰动观测值到潜在真实需求的概率映射（见式 (31)）。这样，MA 就能根据收到的（含噪）数据，对真实交通需求可能的实现值进行采样。基于这些采样结果，MA 随后可以估算预期交通相关福利，并量化其随不同隐私级别而产生的变化。

(ii) 领导者信息简报： 基于上一节定义的效用函数，MA 可以发布一个参考效用函数，用于量化不同数据协作条件下的预期交通相关福利（即该条件可表示为一对真实交通需求和扰动交通需求）。该函数还可以用表格形式表示，称为数据-效用映射，用以展示在真实交通需求与扰动交通需求的不同组合下，MA 可获得的潜在平均延误，并帮助 MA 将相关信息传达给各 MP。

$$W_0 = f(\hat{q}_m, q_m), \quad i \in \mathcal{I}, p \in \mathcal{P}_i, m \in \mathcal{M}_p \tag{36}$$

这一构造与前文提出的定义一致，因为在所研究的数据协作场景中，市场管理者和 MPs 的基础福利函数保持不变。需要注意的是，本表中使用的真实交通需求值和估计交通需求值是通过网格搜索而非实际博弈过程获得的。因此，所得函数或表格可作为基准效用图，供每个 MP 用于评估共享特定数据子集是否能显著提升交通效率。

5 数值实验

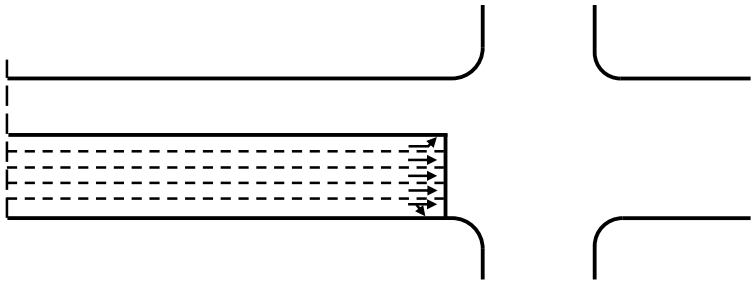
本节基于真实世界数据开展数值案例研究，考察单个数据请求方（市场管理者）与两个数据所有者（MPs）之间的协作。本节旨在提供数值分析结果，以验证所提出博弈的假设、分析结果及均衡行为。具体而言，第 5.1 节介绍基本实验设置，包括网络布局和交通需求。第 5.2 节分析结果，并就如何促进数据协作提供可行的见解。

5.1 实验设置

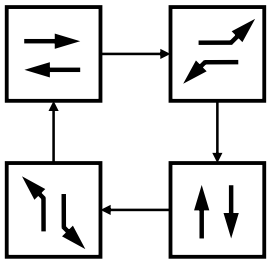
协作设置：我们考虑一个包含三个利益相关者的B2G数字交通市场，其中一名数据请求方（市场管理者）拟向两名数据所有者（即两名活跃的MPs，各自运营渗透率为20%的车辆车队）请求数据，以更新城市主干道网络内的定时信号配时方案。



(a) 包含三个交叉口的真实主干道网络



(b) 单个交叉口的示意布局



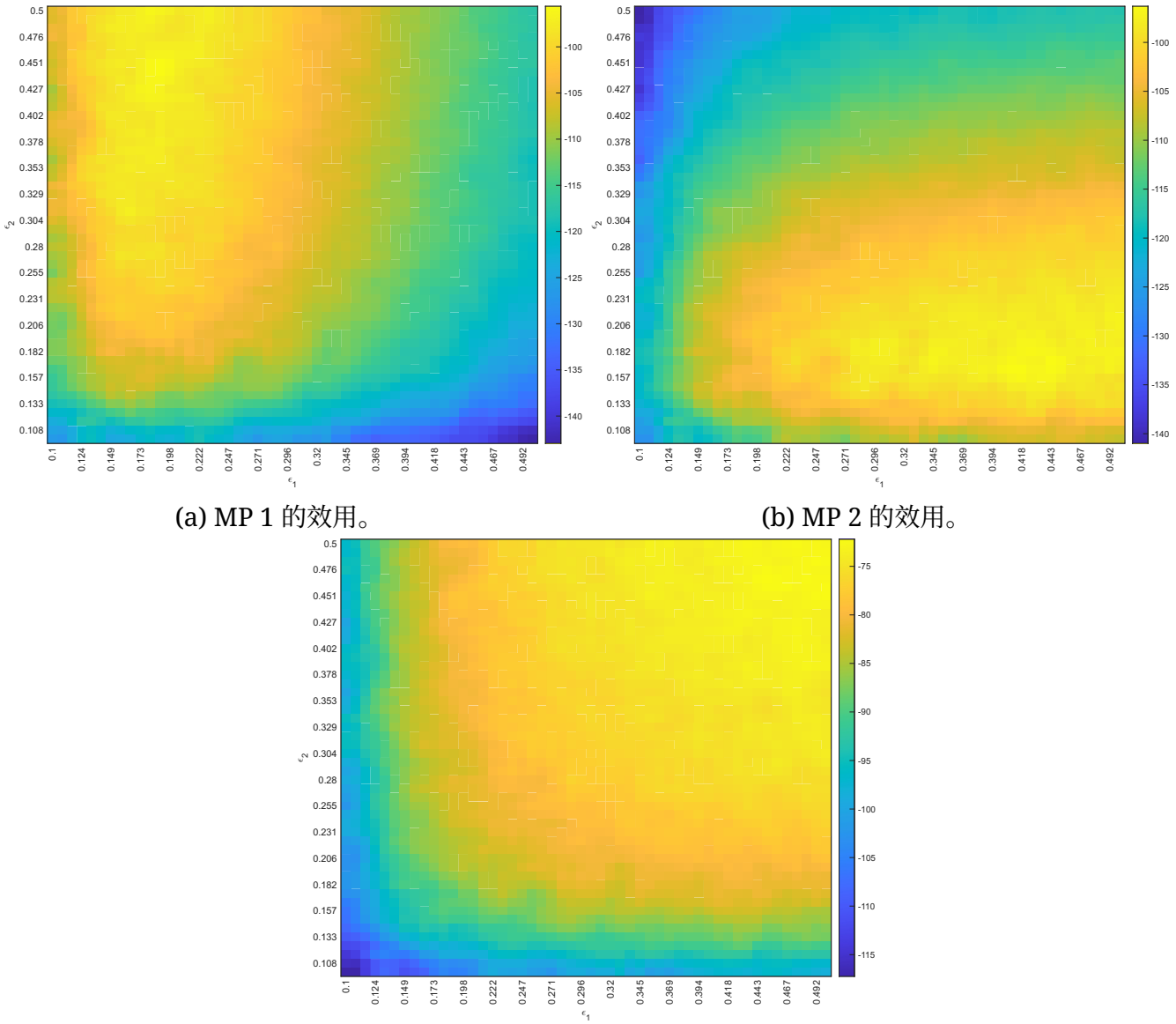
(c) 信号相位

图4：包含三个交叉口的城市主干道网络示意图。

网络与需求设置：我们在中国杭州的一个真实案例上开展实验。包含三个交叉口的城市主干道网络的布局、车道配置和相位顺序如图4所示。各路段长度分别为922、989、1552和550米。我们按照第4.2节所述的案例1对FoQs进行仿真。我们将出城方向（即由东向西）直行交通需求设为3400辆/小时，进城方向（即由西向东）交通量设为1800辆/小时。支路沿线的其他左转及直行交通均配置为200辆/小时。在该需求设置下，各交叉口的v/s（交通量/饱和流）比为0.93。交通需求数据被配置为代表真实的高峰时段需求水平，依据交通需求数据

收集并处理自交通摄像头观测数据（Wei 等人，2019年；Zhu 等人，2024年；Wei 等人，2020年）。

隐私设置：市场管理者确定的数据质量直接对应于一组由MPs选择的隐私参数 ε。因此，我们假设市场管理者可以根据与每个数据所有者相关联的 ε 的规模来指定期望的数据质量。每个 MP 的 β 均设为 90。



(a) MP 1 的效用。

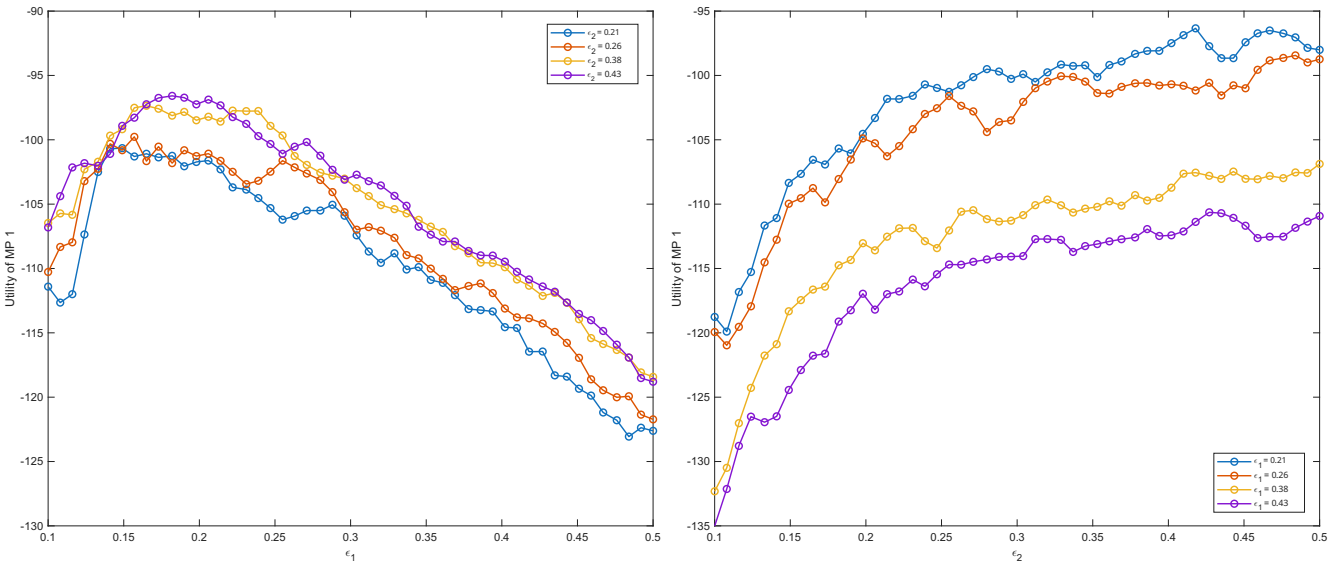
(b) MP 2 的效用。

(c) 市场管理者的效用。

图5：MP 1和MP 2相对于分别由MP 1和MP 2选择的不同隐私参数 ε_1 和 ε_2 的效用。

5.2 结果分析

在这个斯塔克尔伯格博弈中，关键是确定跟随者（MPs）对领导者（市场管理者）的策略所作出的最佳响应。我们从两个角度分析数值结果。首先，我们探究所有



(a) MP 1 相对于 ε_1 的效用。 (b) MP 1 相对于 ε_2的效用。

FigMP 1 的效用随其中一个 MP 的隐私参数固定时的变化。

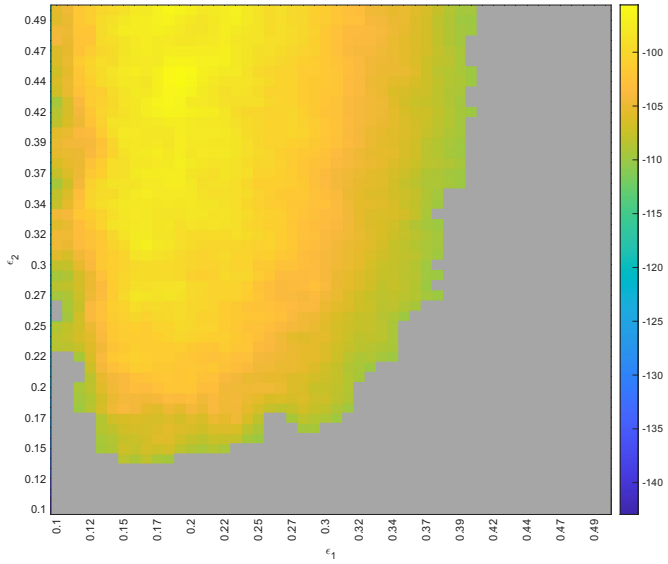
MPs 的完整策略空间，以刻画效用函数的性质。其次，我们计算 MPs 的所有可能最佳响应，从而为确定领导者（市场管理者）的最佳策略（即均衡）提供见解。

图5展示了市场管理者和 MPs 的效用如何随 MP 1 和 MP 2 分别选择的隐私保护参数ε_1 和 ε_2 变化。该图显示，在市场管理者选择任意一组给定的数据质量参数（d_1,d_2）的条件下，所有利益相关者——市场管理者和 MPs——的效用如何随多个 MP 选择的不同隐私级别组合而变化。从数值结果中，我们观察到以下性质：

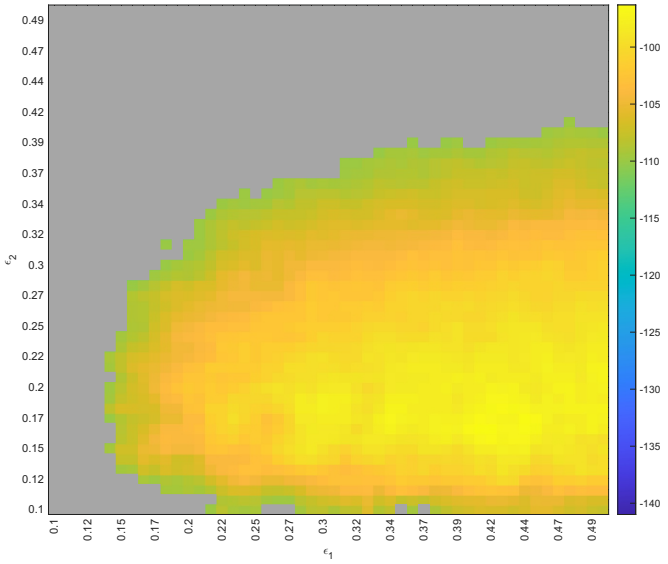
更准确的数据共享能提升市场管理者的效用。如图5c所示，市场管理者的效用随 ε_1 和 ε_2 单调递增。市场管理者的效用仅反映预期交通相关福利，该福利是基于估计交通需求得到的最优解。由于较大的ε 意味着共享数据中的噪声更少，因此估计准确性的提升会直接带来更高的福利。这与市场管理者从更高质量的数据中获益更多的直觉一致。相应地，随着 ε_1 和 ε_2 增大，市场管理者的热图呈现出更暖的色彩渐变，表明其效用更高。

MPs 的效用随其他 MPs 的隐私级别提高而呈现边际递减。如图5a和图5b所示，每个 MP 的效用都会随其他 MP 的隐私参数 ε 增大而提高，这表明一位参与者共享更准确的数据能够提升其他参与者的效用。为进一步说明这一效应，图6a和图6b展示了在对方隐私参数取不同固定值时，MP 1 的效用随 ε_1 和 ε_2 的变化情况。我们观察到，尽管 MP 2 的 ε_2 提高时 MP 1 的效用也随之上升，但改善幅度逐渐减小。这一趋势反映出，其他参与者提供的数据准确度提高所带来的边际收益递减，与假设 3 一致。换言之，来自同伴的更优质数据始终有利于 MP，但 ε_j 每增加一个单位，效用的提升幅度都会小于上一次增加所带来的提升。

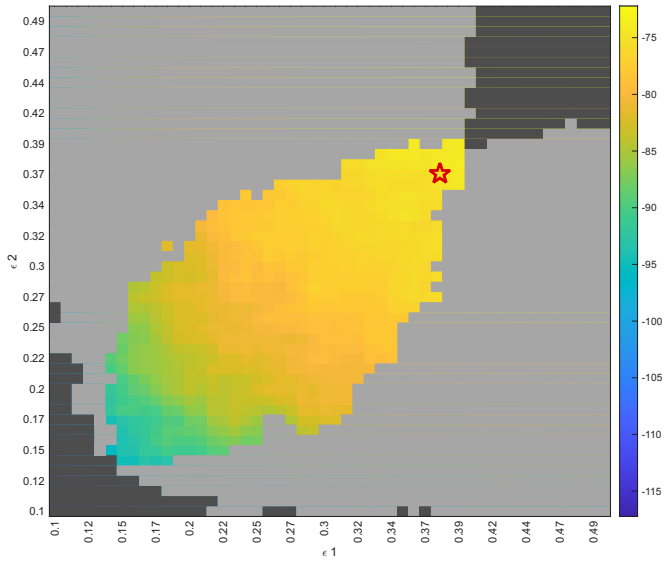
图7展示了市场管理者和 MPs 的效用，我们借此考察 MP 是否有足够的激励来共享其数据。对于图7中的所有子图，灰色区域表示至少存在一个



(a) MP 1 的效用。



(b) MP 2 的效用。



(c) 市场管理者的效用。

图7：MP 1 和 MP 2 在分别选择的不同隐私参数 ε_1 和 ε_2 下的效用。深灰色区域表示不进行数据共享。灰色区域表示至少有一个 MP 共享数据。其余区域表示两个 MP 均共享数据。红色五角星表示数据协作博弈的均衡。

不愿意共享数据的MP，其通过协作获得的交通相关收益无法抵消相关隐私成本。总结出以下两项主要发现：

过高的数据质量期望（即，较大的d）可能导致数据协作失败。如图7a和图7b所示，每个MP的 ε 越大，共享数据带来的效用就越低，从而降低MP参与协作的意愿。接着，我们在图7c中绘制了至少有一个MP不愿共享数据的区域与市场管理者效用热图之间的交叉部分。图中，浅灰色区域表示一个MP选择不共享数据的情形，而深灰色区域表示两个MP都选择不共享数据的情形。然后，我们可以根据市场管理者所选择的不同d 评估这两个MP的最佳响应。具体而言，当市场管理者将d 设为较高值（即，较高的数据质量要求）时，MP必须从 ε的一个高值子集中进行选择，这会产生显著的隐私成本，导致其效用远低于不共享数据时的效用。因此，MP选择不共享数据，而市场管理者由于缺乏数据协作，无法获得任何收益。

适中或较低的数据质量期望（较低的 d）可以促进数据协作。当 MA 设定适中或较低的d时，MP 分享数据的效用会高于不分享数据的效用，从而激励数据协作（如图7c中的黄色区域所示）。尽管 MP 分享的数据存在失真，MA 仍能获得足够的数据，从而提升效用。对于两家 MP 而言，隐私成本也相对可以接受。在这种情况下，均衡位于图7c中红色叉号附近；在 MPs 作出最佳响应的前提下，领导者（MA）在此处实现其效用最大化。总之，通过适当平衡数据质量要求与 MPs 的隐私保护策略，MA 能够确保博弈双方互利共赢。

6 结论

本研究探讨了隐私保护机制对企业对政府数字交通市场中数据协作的影响。本研究是首批探讨多个利益相关者（包括 MA 和 MPs）之间此类问题的研究之一。我们首先提出一个通用模型，将问题描述为斯塔克尔伯格博弈：数据请求方（领导者/MA）选择策略来分配其数据请求/期望，数据所有者（跟随者/MPs）则通过决定是否采用基于扰动的隐私保护机制参与数据协作来作出响应。该通用模型刻画了隐私保护、数据质量与利益相关方效用之间的复杂相互作用。随后，我们针对交通领域的一项具体应用——协同交通信号优化——对模型进行了具体化。在该应用中，MA 利用多家 MP 提供的轨迹数据更新其定时交通信号配时方案。我们设计了一种差分隐私交通需求估计机制，将差分隐私与 LWR 理论相结合，以此对隐私保护的轨迹数据共享进行建模。

接下来，我们将基于一项真实案例研究，开展均衡分析和数值实验。在理想化设定下，我们提出若干假设，并证明均衡的存在性，同时给出保证成功的数据协作的充分条件。我们的分析表明，市场管理者（MA）适度或较低的数据质量期望（即更小的 d）能够促进数据协作，而过高的数据质量期望（即更大的 d）可能导致协作失败。在更现实的设定下，我们提出一种求解方法，并利用真实案例研究对模型进行验证。数值结果验证了理想化分析中提出的假设和理论洞见。总体而言，本研究的成果为隐私保护机制如何以及何时能够促进数字交通市场中的数据协作提供了具有实践指导意义的洞见。本研究的发现预期将为

政策制定者设计有效策略、打破数据孤岛并促进可持续的数据协作提供参考。

未来有几个研究方向。首先，值得探讨能够促使市场管理者鼓励更多MPs参与的机制设计（例如货币激励）。其次，我们有兴趣将所提出的斯塔克尔伯格框架推广到更一般的情形，例如分析大量跟随者的混合策略均衡（杜等人，2014年），以及在贝叶斯博弈框架下考虑市场管理者与MPs之间的不完全信息（Liu和Chow，2022年）。第三，我们希望将当前的博弈论框架扩展到多个MPs之间的全连接数据协作拓扑。在这种拓扑中，利益相关者之间的互动更为复杂，因为协作可能会影响MPs的运营决策（例如路径选择）。第四，值得将该框架扩展到交通控制以外的其他交通应用。

致谢

本研究由新加坡教育部通过其学术研究基金第二层级（A-8003064-00-00）资助。本文仅反映作者的观点和结论，并不代表新加坡教育部或任何其他实体的观点。

定理 8 的证明

我们首先证明一个关于从计数级邻接转换为轨迹级邻接的引理：

引理 1（轨迹级邻接与计数级邻接之间的关系）**.** 对于任意计数级相邻数据集 D 与 D'，存在一个序列 $D_0 \sim_t D_1 \sim_t \cdots \sim_t D_b$，其中每两个连续的数据集都是轨迹级相邻的，$D = D_0$ 且 $D' = D_b$。

证明。 不失一般性，我们假设 $D' \subset D$ 且 $|D \backslash D'| = b$ （另一种情形是对称的）。然后，我们可以构造一个序列

$$D_0 = D, \quad D_i = D_{i-1} \backslash \{\tau_i\}, \quad i = 1, 2, \ldots, b, \tag{37}$$

其中每个 τ_i 均从$D_{i-1} \backslash D'$中任意选取。

根据定义8，我们有

$$D_{i-1} \sim_{\text{traj}} D_i, \quad i = 1, 2, \ldots, b, \tag{38}$$

由于$D' \subset D$，我们还得到$D_b = D \backslash (D \backslash D') = D'$。因此

$$D = D_0 \sim_{\text{traj}} D_1 \sim_{\text{traj}} \cdots \sim_{\text{traj}} D_b = D' \tag{39}$$

证明至此完成。 □

接下来我们证明定理 8。

定理 *8* 的证明。我们将满足定义 9 所述关系D和D'的任意计数级相邻数据集记为$D \sim_{\text{traj}} D'$。因此，根据引理 1，存在一条由 b 个轨迹级相邻数据集组成的链，如下所示。

$$wD = D_0 \sim_{\text{traj}} D_1 \sim_{\text{traj}} \cdots \sim_{\text{traj}} D_b = D' \tag{40}$$

由于对于任意轨迹级相邻数据集和任意事件 $E \in \mathscr{X}$，M均满足 (ε, δ)-差分隐私，因此我们有

$$\begin{aligned}
\mathbb{P}[M(D_0) \in E] &\leq e^{\varepsilon}\mathbb{P}[M(D_1) \in E] + \delta && (41)\\
&\leq e^{2\varepsilon}\mathbb{P}[M(D_2) \in E] + \delta e^{\varepsilon} + \delta && (42)\\
&\vdots \\
&\leq e^{b\varepsilon}\mathbb{P}[M(D_b) \in E] + \delta \sum_{b'=0}^{b-1} e^{b'\varepsilon}, && (43)
\end{aligned}$$

由于$D_0 = D$且$D_b = D'$，机制M满足计数级 $(b\varepsilon, \delta\sum^{b-1}{}_{b'=0} e^{b'\varepsilon})$-差分隐私。 □

B 定理3的证明

证明。 根据假设 3，在低阶段博弈中，每位参与者的最佳响应$BR_k(z_{-k})$ 唯一；根据隐函数定理，该最佳响应

$$\frac{\partial BR_k(z_{-k})}{\partial z_\ell} = -\frac{\partial^2 \bar{U}_k}{\partial z_\ell \partial z_k}\left(\frac{\partial^2 \bar{U}_k}{\partial z_k^2}\right)^{-1} < 0 \tag{44}$$

因此，我们有 $z_k^*(a_k,\ 1_\ell,\ a_{-k\ell}) \leq z_k^*(a_k,\ 0_\ell,\ a_{-k\ell})$。显然也可得 $z_k^*(1,\ a_{-k}) \geq z_k^*(0,\ a_{-k})$。

Then, 令

$$\begin{aligned}
V_{\Delta,1} &= V_k(1, 1_\ell, a_{-k\ell}) - V_k(1, 0_\ell, a_{-k\ell}) && (45)\\
V_{\Delta,0} &= V_k(0, 1_\ell, a_{-k\ell}) - V_k(0, 0_\ell, a_{-k\ell}) && (46)
\end{aligned}$$

这可以看作是当 a_l 从 0 变为 1 时，$\partial U \partial {}^{k} z_\ell \geq 0$ 沿 z_ℓ 的路径进行的积分，分别对应 $a_k = 1$ 和 $a_k = 0$ 的情况。根据假设 3，$\partial U \partial {}^{k} z_\ell(1, a_{-k}) \leq {}^{\partial} U \partial {}^{k} z_\ell(0, a_{-k})$。由于积分范围 $z_k^*(a_k, 1_\ell, a_{-k\ell}) \leq z_k^*(a_k, 0_\ell, a_{-k\ell})$，我们可以证明 $V_{\Delta,1} \leq V_{\Delta,0}$。 □

C 上层跟随者模型存在纯纳什均衡的反例

我们给出一个由三位参与者组成的示例，对于以下形式的上层跟随者模型，他们不存在纯纳什均衡：

$$U_1(x_1, x_2, x_3) = \log\left(x_1 + x_2 + x_3 + b\right) - c_1 x_1 - \gamma_{12} x_1 x_2 - \gamma_{13} x_1 x_3, \tag{47}$$

$$U_2(x_1,x_2,x_3) = \log\big(x_1+x_2+x_3+b\big) \;-\; c_2 x_2 \;-\; \gamma_{21} x_2 x_1 \;-\; \gamma_{23} x_2 x_3, \tag{48}$$

$$U_3(x_1,x_2,x_3) = \log\big(x_1+x_2+x_3+b\big) \;-\; c_3 x_3 \;-\; \gamma_{31} x_3 x_1 \;-\; \gamma_{32} x_3 x_2. \tag{49}$$

其中U_1, U_2, 和U_3 分别表示每个跟随者的效用函数，x_1, x_2 和x_3 分别表示每个跟随者的二元决策变量。b $= 1.0099$，$c1 = 0.161$，$c2 = 0.1545$，$c3 = 0.1102$，$\gamma_{12} = 1.0622$，γ_{13} =0。0979，$\gamma_{21} = 0.0521$，$\gamma_{23} = 0.5048$，$\gamma_{31} = 0.9145$，$\gamma_{32} = 0.2694$。

各策略组合的效用计算见表3。本例满足定理3中的次模性。不存在纯纳什均衡。

表3：所有二元策略组合下全体参与者的效用

x_1	x_2	x_3	U_1	U_2	U_3
0	0	0	0.01	0.01	0.01
0	0	1	0.70	0.70	0.59
0	1	0	0.70	0.54	0.70
0	1	1	1.10	0.44	0.72
1	0	0	0.54	0.70	0.70
1	0	1	0.84	1.10	0.08
1	1	0	-0.12	0.90	1.10
1	1	1	0.07	0.68	0.09

D 重要变量列表

表4：重要变量列表

Sets	
$\mathscr{K}$	由$k = 1,\cdots,K$索引的利益相关者集合
$\mathscr{H}_k^+$	利益相关者k 正在考虑向其共享数据的利益相关者集合
$\mathscr{H}_k^-$	利益相关者k 正在考虑向其请求数据的利益相关者集合
$\mathscr{K}_R$	数据请求方集合
$\mathscr{K}_O$	数据所有者集合
$\mathscr{I}$	交叉口集合，以$i = 1,\cdots,I$为索引
$\mathscr{P}_i$	交叉口i的信号相位集合，以$p_i = 1,\cdots,P_i$为索引
$\mathscr{M}_p$	相位 p的受控交通流向集合
$\mathscr{M}$	主干道中所有受控交通流向的集合
$\mathscr{N}_k$	数据所有者 k 在主干道网络中于某个 最近服务期内观察到的车辆
数据请求方策略中的变量（市场管理者）	
R_1	市场管理者 1 发送的数据请求集合
F_{σ_k}	市场管理者对利益相关者的观测函数k 由 σ_k 参数化，用于描述 从系统状态到所需观测的映射

Φ_k	用于衡量共享数据与真实数据之间距离的函数
d_k	一个距离阈值
$\boldsymbol{d}$	请求方发送给所有数据所有者的距离阈值向量
$\boldsymbol{y}_k$	真实数据
$\hat{\boldsymbol{y}}_k$	带扰动的共享数据
数据所有者策略的变量（MPs）	
S_k	数据所有者k所选择的数据共享策略集合
a_k	表示数据是否被共享的二元变量
$\boldsymbol{a}$	所有数据所有者的二元变量向量
G_{ε_k}	表示基于扰动的隐私保护机制的函数
ε_k	数据所有者k的隐私保护机制参数
$\boldsymbol{\varepsilon}$	所有数据所有者的隐私保护机制参数向量
参与者效用的变量	
$U_k(\cdot)$	数据所有者k的效用函数
$W_k(\cdot)$	数据所有者k的服务质量提升
β_k	隐私成本的权重k
$U_1(\cdot)$	数据请求方 1 的效用函数
$W_1(\cdot)$	数据请求方k的交通相关福利
均衡分析的变量	
$(ddd^*, aaa^*, \boldsymbol{\varepsilon}^*)$	构成该博弈均衡的策略组合
$\bar{U}_k(z)$	$U_k(\boldsymbol{\varepsilon}, aaa)$的简化表示
ϕ_k	$\Phi^{-1}(d_k)$的简化表示
z_k	简化表示 $a_k\varepsilon_k$
协同定时交通信号优化的变量	
$(C^{\text{old}},\ ggg^{\text{old}},\ ooo^{\text{old}})$	过时的信号配时方案
$(C^*,\ ggg^*,\ ooo^*)$	新的优化信号配时方案
C	所有交叉口的公共周期长度
g_{ip}	交叉口i中相位p 的绿信比分配
o_i	交叉口i相对于i的相位差 + 1
$\zeta_i(\bar{\zeta}_i)$	交叉口i的相对相位差
$e_i(\bar{e}_i)$	排队清空时间（周期）
q_m^c	交通流向m的通行能力流量
$\hat{q}_m$	交通流向m的估计交通流量
q_m	交通流向m的真实交通流量
$\hat{\boldsymbol{q}}$	所有交通流向的估计交通流量向量
$\boldsymbol{q}$	所有交通流向的真实交通流量向量
指定的观测函数和隐私保护机制的变量	
$\boldsymbol{x}_k$	数据所有者k的系统状态
$\boldsymbol{y}_k$	利益相关者k的所需观测
$\hat{\boldsymbol{y}}_k$	利益相关者k的扰动后所需观测
τ_{kn}^m	从数据所有者k 在交通流向m上的轨迹n中收集的排队点

h_{kn}^m, t_{kn}^m	利益相关者k的第n辆车辆在时空图上交通流向m 的排队点位置和时间
$\hat{h}_{kn}^m$, $\hat{t}_{kn}^m$	利益相关者k的第n条轨迹的扰动后排队点位置和时间-时空图上的交通流向m示
D, D'	相邻数据集

参考文献

Berke, A., Ding, G., Chin, C., Gopalakrishnan, K., Larson, K., Balakrishnan, H., 和 Li, M. Z.（2023年）。无人机配送与客户隐私的价值：针对美国消费者的离散选择实验。《交通研究*C*辑：新兴技术》，157:104391。

Bloomberg Law（2022年）。谷歌与印第安纳州就位置跟踪达成2000万美元和解。https://news.bloomberglaw.com/privacy-and-data-security/google-indiana-reach-20-million-location-tracking-settlement。

Butler, L., Yigitcanlar, T., and Paz, A.（2021年）。城市中出行即服务（maas）的采用壁垒与风险：文献系统综述。《城市》，109：103036。

Chen, Z. and Keppo, J.（2025年）。新产品开发中的研发数据共享。《制造与服务运营管理》。

Chow, J.（2018年）。知情的城市交通系统：面向智慧城市的经典与新兴出行方式。Elsevier。

Cottrill, C. D.（2020年）。出行即服务监控：出行即服务中的隐私考量。《交通研究*A*辑：政策与实践》，131：50–57。

Cottrill, C. D. et al.（2015年）。位置隐私偏好：基于调查的消费者认知、权衡与决策分析。《交通研究*C*辑：新兴技术》，56：132–148。

杜, L., Han, L., and Li, X.-Y.（2014年）。采用混合策略拥堵博弈的分布式协同车载在线路径规划。《交通研究*B*辑：方法论》，67：1–17。

Dwork, C.（2006）。差分隐私。载于自动机、语言与编程国际研讨会，第1–12页。Springer。

欧洲联盟　条例（2018年）。《通用数据保护条例》。https://gdpr-info.eu　/.

Fainmesser, I. P.、Galeotti, A. 和 Momot, R.（2023年）。数字隐私。《管理科学》，69(6)：3157–3173。

Gao, H.、Li, Z. 和 Wang, Y.（2022年）。面向联网车辆的隐私保护协同估计及其在协同道路剖面估计中的应用。*IEEE*智能交通系统汇刊。

Gartner, N. H., Assman, S. F., Lasaga, F., and Hou, D. L.（1991年）。干道交通信号优化的多频带方法。《交通研究*B*辑：方法论》，25(1):55–74。

Gati, N. J., Yang, L. T., Feng, J., Nie, X., Ren, Z., and Tarus, S. K.（2021年）。面向信息物理社会系统的差分隐私数据融合与深度学习框架：最新进展与展望。信息融合。Glicksberg, I. L.（1952年）。对Kakutani不动点定理的进一步推广及其在纳什均衡点中的应用。美国数学会会刊，3(1):170–174。Grand View Research（2025年）。交通分析市场（2025-2030）。htt
ps://www.grandviewresearch.com/industry-analysis/transportation-analytics-market。Guo, Q., Li, L., and Ban, X. J.（2019年）。联网与自动驾驶车辆环境下的城市交通信号控制：综述。《交通研究*C*辑：新兴技术》，101:313–334。Guo, Z., Yu, B., Shan, W., and Yao, B.（2023年）。需求不确定条件下按需出行服务中情境化车辆再平衡的数据驱动鲁棒优化。《交通研究*C*辑：新兴技术》，154:104244。He, B. Y. and Chow, J. Y.（2020年）。交通网络数据共享的最优隐私控制。《交通研究*C*辑：新兴技术》*style>，113:370–387。Hoheisel, T., Kanzow, C., and Schwartz, A.*（*2013*年）。带互补约束的数学规划松弛方法的理论与数值比较。数学规划，137(1):257–288。Huang, F., Xu, J., and Weng, J.（2020年）。采用灵活深度学习框架的多任务出行路线规划。《*IEEE*智能交通系统汇刊》，22(7):3907–3918。Jin, F., Hua, W., Francia, M., Chao, P., Orowska, M., and Zhou, X.（2022年）。隐私保护轨迹数据发布的综述与实验研究。*IEEE*知识与数据工程汇刊。Lighthill, M. J. and Whitham, G. B.（1955年）。运动波II：关于拥挤长路上交通流的理论。伦敦皇家学会会刊*A*辑：数学与物理科学，229(1178):317–345。Little, J. D., Kelson, M. D., and Gartner, N. H.（1981年）。MAXBAND：一种用于在主干道和三角形网络上设置信号灯的多功能程序。Liu, Q. and Cho
w, J. Y.（2022年）。作为竞合博弈的公共交通寡头市场中的高效稳定数据共享。《交通研究*B*辑：方法论》，163:64–87。Liu, Q. and Chow, J. Y.（2023年）。利用停靠站计数数据估算拥堵环境下基于时刻表的动态公交客流。*Transportmetrica B*：交通动态，11(1):231–256。Manual, H. C.（2000年）。《公路通行能力手册》。华盛顿特区，2(1):1。Martens, B. and Zhao, B.（2021年）。数据访问与制度竞争：中国汽车数据共享案例研究。大数据与社会，8(2):20539517211046374。《纽约时报》（2024年）。荷兰监管机构因优步保护司机数据不力对其处以3.24亿欧元罚款。
https://www.nytimes.com/2024/08/26/business/uber-netherlands-fine-driver-data.html。

Pantelidis, T. P., Chow, J. Y., 和 Cats, O.（2024年）。移动运营商服务能力共享合同设计：通过风险共担应对网络中断。*Transportmetrica A: Transport Science*，20(3):2210229。Richards, P. I.（1956年）。公路上的冲击波。运筹学，4(1):42–51。Scholtes, S.（2001年）。带互补约束的数学规划中正则化方案的收敛性质。*SIAM*优化杂志，11(4):918–936。Smichowski, B. C.（2018年）。MaaS中通过数据共享实现竞合的决定因素。*Management&Data Science*。Strava（2018年）。Strava热力图与秘密的终结。在线获取：https://www.wired.com/story/strava-heat-map-military-bases-fitness-trackers-privacy/。
Tan, C. 和 Yang, K.（2024年）。网联车辆环境中的隐私保护型自适应交通信号控制。《交通研究*C*辑：新兴技术》，158:104453。Telecompaper（2024年）。Vodafone NL因未经授权与运输公司共享交通数据而被罚款。
https://
www.telecompaper.com/news/vodafone-nl-fined-for-unauthorised-sharing-of-traffic-data-with-transport-company–1513914。Today（2018年）。香港领导人抨击公共交通私人运营商不共享数据。
https://
www.todayonline.com/world/hong-kong-leader-blasts-private-operators-public-transport-not-sharing-data。Tsao, M., Gopalakrishnan, K., Yang, K., 和 Pavone, M.（2023年）。面向网络路由应用的差分隐私随机凸优化。载于*2023*年第*62*届*IEEE*决策与控制会议（*CDC*），第7475–7482页。IEEE。
Tsao, M., Yang, K., Zoepf, S., 和 Pavone, M.（2022年）。信任但需验证：出行管理中的密码学数据隐私。*IEEE*网络系统控制汇刊，9(1):50–61。美国人口普查局（2020年）。人口普查局采用一种称为差分隐私的统计方法，在屏蔽个人信息的同时，让我们能够共享有关社区的重要统计数据，以用于2020年人口普查。
https://www.census.gov/content/dam/Census/library/factsheets/2021/differential-privacy-and-the-2020-census.pdf。Wang, Q., Tan, C., 和 Yang, K.（2024a）。在不损害隐私的前提下协作：通过纵向联邦强化学习实现交通信号控制。发表于希腊伊拉克利翁举行的交通系统新兴技术*30*周年会议。Wang, Q. 和 Yang, K.（2024年）。面向交通状态估计的隐私保护型数据融合：一种纵向联邦学习方法。《交通研究*C*辑：新兴技术》，168:104743。选自第25届国际交通与交通理论研讨会（ISTTT25）的论文集。
Wang, X., Jerome, Z., Wang, Z., Zhang, C., Shen, S., Kumar, V. V., Bai, F., Krajewski, P., Deneau, D., Jawad, A., 等（2024b）。利用低渗透率车辆轨迹数据优化交通信号灯。《自然·通讯》，15(1):1306。
Wei, H., Li, Z., 和 Gayah, V.（2020年）。用于交通信号控制的强化学习。在线获取：https://traffic-signal-control.github.io/#open-datasets。

Wei, H., Xu, N., Zhang, H., Zheng, G., Zang, X., Chen, C., Zhang, W., Zhu, Y., Xu, K., 和 Li, Z.（2019年）。Colight：学习交通信号控制中的网络级协作。载于第*28*届*ACM*信息与知识管理国际会议论文集，第1913–1922页。Xu, D., Wei, C., Peng, P., Xuan, Q., 和 Guo, H.（2020年）。Ge-GAN：一种用于道路交通状态估计的新型深度学习框架。《交通研究*C*辑：新兴技术》，117:102635。Yao, J., Tan, C., 和 Tang, K.（2019年）。基于采样车辆轨迹的干线协调控制优化模型：stream模型。《交通研究*C*辑：新兴技术》，109:211–232。Ying, Z., Cao, S., Liu, X., Ma, Z., Ma, J., 和 Deng, R. H.（2022年）。Privacysignal：面向智能交通系统的隐私保护型交通信号控制。《*IEEE*智能交通系统汇刊》。Yu, H., Zhou, J., 和 Yang, K.（2025年）。通过粒子滤波强化学习实现耦合系统中交互感知的参数隐私保护数据共享。载于第*7*届动力学学习年会\与控制会议，第1525–1536页。PMLR。Zhang, C., Xie, Y., Gartner, N. H., Stamatiadis, C., 和 Arsava, T.（2015年）。AM-Band：用于主干道交通信号协调的非对称多带模型。《交通研究*C*辑：新兴技术》，58:515–531。Zhang, W., Jiang, B., Li, M., 和 Lin, X.（2022年）。隐私保护型聚合移动数据发布：一种信息论深度强化学习方法。*IEEE*信息取证与安全汇刊，17:849–864。Zheng, J., Sun, W., Huang, S., Shen, S., Yu, C., Zhu, J., Liu, B., 和 Liu, H. X.（2018年）。利用众包车辆轨迹数据进行交通信号优化。发表于华盛顿特区举行的第*97*届交通研究委员会年会。Zhou, J. 和 Yang, K.（2024年）。一种用于混合自主车队控制的参数隐私保护策略。《交通研究*C*辑：新兴技术》，169:104885。Zhu, H., Sun, F., Tang, K., Han, T., 和 Xiang, J.（2024年）。基于协调图的网络交通信号控制框架。《*IEEE*智能交通系统汇刊》，25(10):14298–14312。