arXiv:2304.06873v2 [cs.RO] 6 Jul 2023

通过消息效用估计降低去中心化多机器人团队的网络负载

Isabel M. Rayas Fern´andez[1], Christopher E. Denniston[1, †], Gaurav S. Sukhatme[1, *]

摘要——我们的研究受到湖泊中藻类浓度分位数估计问题的启发，并探讨去中心化多机器人团队如何有效地解决这一问题。我们发现，在此任务中，多机器人团队的性能优于单机器人团队，而支持通信的团队又优于不支持通信的团队；然而，真实机器人受到资源限制，通信网络也无法承载任意规模的消息负载，因此，采用朴素的恒定信息共享方式，以及复杂的建模与决策，都不可行。为此，我们提出了可在线本地计算的指标，用于确定向其他团队成员传输特定消息的效用；同时提出一种决策理论方法，仅选择传输最有用的消息，并以去中心化、相互独立的框架维护对其他队友的信念。我们在真实世界水生数据集上通过仿真验证了该方法，并表明，基于消息对队友未来行为的预期影响进行效用估计，从而限制通信，可在使分位数估计误差降低 1.84% 的同时，将网络负载减少 42%。

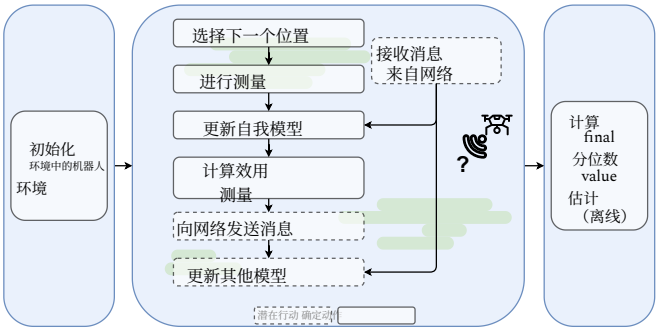

图 1：问题流程概览。该方法由三个阶段组成：初始化、探索和聚合。大部分工作在去中心化探索阶段进行，机器人会反复选择下一个要前往的位置、测量环境、选择性地通过消息广播测量结果，并更新其模型，直到预算 B_T 耗尽。

I. 引言

多机器人团队在有效探索自然环境并准确估计分位数值方面展现出潜力，有助于加深对自然环境的理解 [1]。在研究空间变化场时，分位数估计有助于描述感兴趣数据的不同分布范围，其做法是利用环境测量值估计目标分位数的数值（例如，第 25 百分位数等同于 0.25 分位数） [2]。我们希望利用一支由支持通信的移动机器人组成的团队协同探索自然环境，从而准确估计该环境中的分位数值。然而，实际的多机器人部署必须考虑通信能力的限制；本文关注降低网络负载，即减少已尝试传输的消息数量，同时尽可能少地牺牲分位数估计精度。我们探讨如何判断一条消息是否值得发送。具体而言，我们：

- 提出降低带宽和资源受限的去中心化多机器人团队网络负载的方法，这些方法利用基于效用的消息评估，且不保证队友提供的信息；
- 介绍一种决策理论方法，根据效用评估和传输成本，决定是否向网络发送消息；
- 展示结果表明，即使存在概率性传输失败和基于距离的传输概率，在真实水生环境的数据集上执行估计分位数值这一问题任务时，也能够显著降低网络负载，而性能仅略有下降；
- 结合现实世界应用讨论结果。

我们假设存在一个去中心化团队，其中每个机器人都独立规划，并且不保证能够接收到其他机器人发送的信息。然而，我们假设采用一种通信协议，使机器人能够在每一步决定是否向网络广播其新获取的测量值。网络的带宽有限，因此机器人会受到激励，只发送那些对其他机器人最有益的消息。我们通过阈值来编码使用消息加载网络的成本：效用高于该阈值的消息，其收益高于成本。在这些假设的基础上，如何评估潜在消息的效用，便成为这一问题的关键考量；在本文中，我们提出了几种解决方法。总体概览见图1。

本研究部分经费由南加州沿海水域研究项目管理局，获得加利福尼亚州水资源控制委员会通过编号为 19-003-150 的协议提供的主要经费，并部分获得 USDA/NIFA 奖项 2017-67007-26154 的资助。

本材料基于美国国家科学基金会研究生研究奖学金计划资助的工作，资助编号为 DGE-1842487。本材料中表达的任何观点、研究结果、结论或建议均属于作者个人意见，不一定反映美国国家科学基金会的观点。

[1]计算机科学系，南加州大学。

*G.S. Sukhatme 同时担任南加州大学教授和亚马逊学者。本文介绍的工作是在南加州大学开展的，与亚马逊无关。

†C.E. Denniston 现任职于 OffWorld公司。本文介绍的工作由南加州大学完成，与 OffWorld公司无关。

II. 相关工作

关于基于效用的通信，已有大量研究，而我们的方法与先前的工作存在相似之处。考虑如何决定通信内容或通信时机的研究，大多出于同样的动机：遵守现实网络必须面对的带宽限制，并降低通信成本。然而，具体方法在解决方案的处理方式以及问题所处的背景方面存在差异。基于效用的决策在机器人间通信中的研究背景各不相同，其中包括持续数月、需要路由至基站的数据收集 [3], 灾害响应 [4], 非合作式多智能体路径规划 [5], 以及采用请求-应答协议的场景 [5], [6]。

一种常见方法是利用消息类别、类型或数据内在特征，判断哪些消息最值得发送 [3], [4], [7]。这些方法通常会利用特定应用领域的领域知识。以多机器人定位问题为目标的先前研究，通过在传输和不传输每条消息两种情况下对机器人团队进行前向仿真，构建有向无环图 [8]。其目标是在尽可能降低成本（即 DAG 中所需的通信次数）的同时，达到最佳群体定位精度。

消息对机器人动作的预期影响也被用于确定效用 [5], [6]。一种限制通信次数的方法适用于机器人需要选择要访问的目标航点的任务：采用启发式方法限制新路径的长度，并且只有当新路径长度在当前路径长度的特定倍数范围内时，才向其他机器人请求信息 [9]。蒙特卡洛树搜索（MCTS）已在通用中间件中用于决定是否向网络发送消息 [7]。该方法考虑了消息效用可能随时间变化这一时间因素；在每个决策点构建的搜索树都包含所有过去和未来的消息。不过，该方法没有明确编码网络约束，并假设机器人始终处于彼此的通信范围内。

另一类解决方案提出了一些指标，用于根据纳入消息 [10], [11], [12] 后期望奖励的变化来确定效用。在 [13], 中，作者提出了一种效用函数，它由两个相互竞争的因素之和计算得出：最大化整个团队获得的效用，同时尽量减少消息发送次数。在该方案中，网络采用点对点连接，因此每一步都要决定是将收到的消息转发给随机接收方，还是停止转发。 [7] 和 [8]也纳入了基于期望奖励的效用度量，并且团队在规划和奖励计算方面均采用去中心化方式。这些工作与我们的主要区别在于，我们的最终目标是以能获得最佳分位数估计的方式探索环境，而它们则为机器人设定了一个静态目标位置。此外，出于计算效率和可处理性的考虑，我们专门采用贪心算法进行规划，这一做法的动机来自资源受限系统，而且不会对其他机器人的模型进行前向模拟。

值得注意的是，除了通信决策本身，我们还采用了更现实的随机通信假设：机器人决定发送消息，并不意味着所有接收方或任何接收方都一定能收到消息。我们不假设测量感知是完美的。尽管我们的规划和建模方法相对简单，通信假设也更容易导致失败，但我们仍能显著减少网络负载，同时保持良好的任务性能。

III. 方法

我们的建模方案扩展了 [1] 中的设置，详情请参阅该文献；此处简要介绍相关定义。团队由 N 个空中机器人组成，每个机器人都在离散化环境 $\boldsymbol{G}^{\#} \subset \mathbb{R}^2$中以固定高度运行，并具有固定的规划预算 B_T 。每一步中，机器人可执行四种可能动作之一： $\pm x$ 或 $\pm y$。由于我们关注的是机器人间通信，因此假设存在低层控制器、避碰程序和完美定位。人工操作员提供一组感兴趣的任意分位数 Q，最终目标是˜准确估计 V Q ，其中真实值定义为 $V =$ 分位数$(\boldsymbol{Y}_{\#}Q)$；机器人可用传感器测量的所有可能位置集合为 $\boldsymbol{X}^{\#}$，这些位置对应的真实值为 $\boldsymbol{Y}^{\#}$。每个机器人都配备一台相机，可在环境中的这些像素位置获取代表藻类浓度的含噪测量值 $\boldsymbol{Y}$ $\boldsymbol{X}$。在整个部署过程中，每个机器人都会收集测量值 $\boldsymbol{X}$，$\boldsymbol{Y}$ 既包括通过自身相机直接获取的测量值，也包括通过网络从其他机器人接收的消息。机器人利用收集到的所有测量值构建并更新各自的环境高斯过程（GP）模型。机器人 i对分位数值的估计为

$$\tilde{V}_i = quantiles(\mu_{GP^i}(\boldsymbol{X}^{\#}), Q) \tag{1}$$

其中， $\mu_{GP^i}(\boldsymbol{X}^{\#})$是基于 i迄今收集的测量值 $\boldsymbol{X}^i$ 和 $\boldsymbol{Y}^i$，使用高斯过程估计的所有可能位置。

下文从给定机器人的角度讨论问题时，我们将其称为 i 或自我机器人，而将任何其他 N 机器人称为 j。上标为 j 的变量（例如 GP^j ）表示从的角度来看的高斯过程。

受资源受限系统的计算约束启发，我们选择使用一种可在本地计算的在线贪心规划策略。在每个规划步骤中， t i在可行的下一个测量位置上最大化目标函数 f 。我们使用经过修改的分位数标准误差目标函数来评估时间 t [2] 的候选位置；该函数通过使用机器人模型的两个版本来衡量估计分位数值标准误差的差异：当前模型 $\mu_{GP^i_{t-1}}$，以及使用候选位置处假设的新值更新后的模型 $\mu_{GP^i_t}$。我们对 f 进行修改，使其不仅以

候选测量位置$\boldsymbol{X}$为条件，还以与这些位置相关的值（通常记为 $\mathcal{Y}$）以及所使用的高斯过程为条件。其原因在于，f 之后还会用于计算消息效用，这样便能将其推广到$\mathcal{Y}$和 GP 可能取不同值的情况。当 f 在规划步骤中被 i 用于选择 i的下一个位置时， $\mathcal{Y}$ 是候选位置处的期望值 $\boldsymbol{X}$: $\mu_{GP_{t-1}^i}(\boldsymbol{X})$（与 [2]中相同）。然而，当 f 被 i 用于确定 j 接收 i 已经获取的真实测量值的效用时， $\mathcal{Y}$ 是已获取的 $\boldsymbol{Y}$，而高斯过程可以是GP^i 或 GP^j （见第III-C节）。最后一项鼓励探索高方差区域。

$$f(\boldsymbol{X}_i;\mathcal{Y},GP)=\frac{d}{|Q|}+\sum_{x_j\in\boldsymbol{X}_i}c\sigma^2(x_j) \tag{2}$$

$$d=\|se(\mu_{GP_{t-1}}(\boldsymbol{X}^{\#}),Q)-se(\mu_{GP_t}(\boldsymbol{X}^{\#}),Q)\|_1 \tag{3}$$

A. 过程概述

该方法分为三个阶段：初始化、探索和聚合。此处对这些阶段作高层次介绍；更多细节见后续章节。

1) 初始化：机器人根据初始位置分布 α 分散在环境中（[1] 提供了详细说明）。我们假设每个机器人都知道其他机器人的数量及其对应的初始位置。

2) 探索：主要工作在去中心化探索阶段进行：机器人反复选择下一个要前往的位置、测量环境、酌情通过消息广播测量结果，并更新模型，直到预算 B_T 耗尽。在 t处的下一个测量位置按照以下方式选择

$$\boldsymbol{X}_t^i=\arg\max_{\boldsymbol{X}\subset\boldsymbol{X}_{\text{next}}^i} f(\boldsymbol{X};\mu_{GP^i}(\boldsymbol{X}),GP^i) \tag{4}$$

其中， $\boldsymbol{X}_{\text{next}}$ 包含从当前位置 g_t^i出发一步之内可以感知的位置集合。在此规划周期之外， i 还可能随时收到来自另一台 j 的消息，从而进一步更新模型（第III-C节）。

3) 聚合：机器人达到分配的预算后，将收集它们的数据。操作员可以离线利用这些数据生成环境的聚合高斯过程模型，并据此获得分位数值的最终估计：

$$\tilde{V}_{\text{final}}=quantiles(\mu_{GP_{\text{aggregate}}}(\boldsymbol{X}^{\#}),Q) \tag{5}$$

B. 通信框架

通信网络由 N 个机器人组成，每个机器人都能在每个时间步广播一条消息。由 i 在时间 t 发送的消息 $\mathcal{M}_t^i=(i,\ g_t^i,\ \boldsymbol{X}_t^i,\ \boldsymbol{Y}_t^i)$ 由消息头和消息体组成：消息头包含发送机器人编号和当前位置，消息体包含当前的测量集合。发送的消息会（独立地）成功传输给网络中的每个其他机器人 j，传输概率取决于距离，并由衰减率 η 和通信半径 r [1] 参数化：

$$p_{\text{success}}(distance)=\frac{1}{1+e^{\eta(distance-r)}} \tag{6}$$

成功传输的消息会完整、无损且无延迟地接收。如果消息未能成功传输给 j，则该消息对该机器人而言视为丢失。一般而言，我们假设 i 不会收到关于其消息是否被 j 成功接收的握手或确认信息。我们假设机器人能够自主定位，并且在半径 r 范围内能够准确测量 g^j；每个机器人都会在每一步测量 j 。

C. 对其他机器人的建模与模型更新

一个关键问题是，如何在信息不完整的情况下对其他机器人进行建模，因为只有建立这样的模型，才能充分评估任何给定消息对其他机器人的效用。自我机器人 i 会为每个 j维护一个独立的模型（$g^j,\ GP^j$）。 g^j 是其所掌握的最新位置i ，并会在两种情况下更新。首先，如果 j 位于半径 r范围内，则根据 i在每一步的感知能力进行更新；其次，每当 i 通过消息头收到来自 j的消息时，也会更新。GP^j 是一个基于 i 认为 j 所掌握的测量值的高斯过程。我们将其记为 $M^j:=(\boldsymbol{X}^j,\ \boldsymbol{Y}^j)$。 M^j以及由其定义的 GP^j也可以在两种情况下更新。第一种情况是i 从 j收到消息 $\mathcal{M}$ 时，因为显然 j 必定拥有其中包含的数据。第二种情况则不太确定。如果i 向网络广播 $\mathcal{M}$ ，它无法知道消息是否已成功传输给任何 j；不过请记住， i可以感知 r范围内 j 的位置。我们利用这一点来确定传输成功的合理置信度代理：如果 j 位于 r范围内， i 就会认为 $\mathcal{M}$ 已被接收，从而更新 M^j。

D. 计算消息效用

我们介绍几种计算 $\mathcal{M}^i=$的效用u^j的方法（为便于阅读，省略上标和下标）$(i,\ g^i,\ \boldsymbol{X},\ \boldsymbol{Y})$。第一种方法奖励将 u_{rw} 定义为 i认为 j 会获得的奖励，使用目标函数 f:

$$u_{rw}(\boldsymbol{X},\boldsymbol{Y},j)=\begin{cases}\infty, & \text{if}\quad |\boldsymbol{X}^j|=0\\ f(\boldsymbol{X};\boldsymbol{Y},GP^j), & \text{otherwise}\end{cases} \tag{7}$$

我们将第二种方法称为动作。直观而言，如果 i 认为接收这些测量值会导致 j 采取不同于不接收这些测量值时的动作，那么该方法就认为消息有用。于是，如果根据方程（4）计算出的动作$\boldsymbol{X}_{\text{next, w/o}}$ 与$\boldsymbol{X}_{\text{next, w/}}$ 不同，则 u_{ac} 定义为 ∞；如果相同，则 u_{ac} 取值为 u_{rw}:

$$u_{ac}(\boldsymbol{X},\boldsymbol{Y},j)=\begin{cases}\infty, & \text{if } |\boldsymbol{X}^j|=0\\ \infty, & \text{if } \boldsymbol{X}_{\text{next,w/o}}\neq\boldsymbol{X}_{\text{next,w/}}\\ u_{rw}(\boldsymbol{X},\boldsymbol{Y},j), & \text{otherwise}\end{cases} \tag{8}$$

u_{rw} 和 u_{ac} 都包含这样一个限定条件：如果 i 认为 j 尚未收集任何测量值，就应发送该消息。

最后介绍的三种方法较为简单，因为它们不使用 j 的模型。在第三种方法 自我奖励 中， i 使用自己通过测量实际获得的奖励，并直接将其设为所有其他 j的效用：

$$u_{ego}(\boldsymbol{X},\boldsymbol{Y},j)=f(\boldsymbol{X},\boldsymbol{Y};GP^i) \tag{9}$$

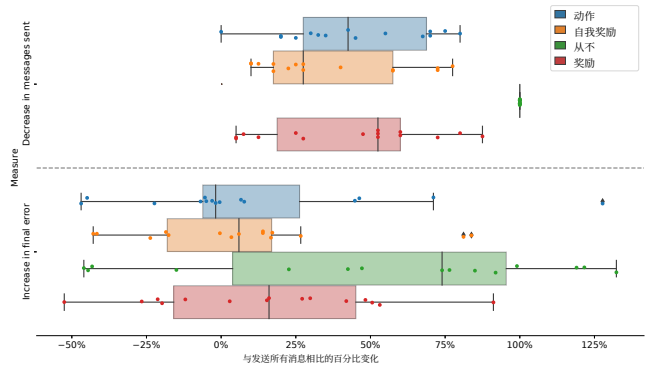

图 2：与始终发送消息相比，不同效用方法在网络负载与任务性能之间的权衡。上：发送的消息数量减少百分比（网络负载），越高越好。下：最终分位数估计误差的增加百分比，越低越好。

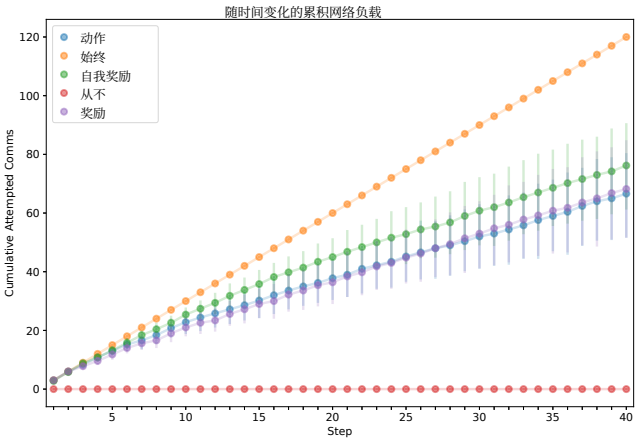

图 3：随时间变化的累积网络负载，以发送的消息数衡量。x 轴上的每一步表示机器人执行一个规划步骤。

最后，始终 和从不 是恒定的基线，它们始终将 ∞ 或 0 作为分配值：

$$u_{al}(\boldsymbol{X}, \boldsymbol{Y}, j) = \infty \tag{10}$$

$$u_{nv}(\boldsymbol{X}, \boldsymbol{Y}, j) = 0 \tag{11}$$

E. 决定通信

针对每条消息 $\{u^j\}$，在为每个 j计算出效用后，i 现在需要决定该效用是否足以抵消向网络加载消息所产生的成本。由于通信采用基于广播的方式，而非点对点方式，i 必须先为团队计算单一效用值 U，然后判断它是否超过代表成本的效用阈值 T。为此，我们介绍一种聚合一组效用的方法，该方法通过按估计的传输成功概率对效用加权来计算期望效用：

$$U_{EU}(\{u^j\}) = \frac{1}{N-1} \sum_j (p_{est}(j) \cdot u^j) \tag{12}$$

$$p_{est}(j) = \begin{cases} p_{\text{success}}(d^j), & \text{if} \quad d^j \leq r \\ p_{\text{success}}(r), & \text{otherwise} \end{cases} \tag{13}$$

其中，$d^j = ||g^i - g^j||_2$ 和 r 的定义见第 III-B 节。式（13）计算出的估计成功概率，在已知 j 时（即，如果 j 处于感知半径 r以内），会直接采用基于距离的概率；但对于超出感知半径的任何 j，则默认采用乐观概率估计，并将 r 作为假定距离。这种乐观启发式方法可避免在队友位置未知时，效用值受到过度限制。请注意，如果任何 u^j 为 ∞，则认为 U 已超过阈值 T。如果 $U \geq T$，i就决定传输 $\mathcal{M}$。

IV. 实验与讨论

现在，我们展示第 III 节所述方法的实验结果。我们在 3 组分位数 Q 上测试性能：四分位数 (0.25, 0.5, 0.75)、中位数-极值 (0.5, 0.9, 0.99) 和极值 (0.9, 0.95, 0.99)。对于每种参数组合，我们运行 5 个随机种子，共得到 60 个实例。我们固定以下参数：$N = 4$；预算 $B_T = 10$（总共 40 步）；初始分散度 $\alpha = 0.2$；$\eta = 0.4$；通信半径 $r = 15.0$；阈值 $T_{rw} = 2.8 \times 10^{-4}$ 和 $T_{ego} = 8.3 \times 10^{-5}$。$T_{rw}$ 用于奖励和动作，而 T_{ego} 用于自我奖励。我们通过观察测试期间收到的 u 值，并将 T 设为观测奖励的近似第 25 百分位数，凭经验选取了这些 T。这可以理解为：只有效用位于前 75% 的消息才值得花费成本发送到网络，而位于后四分之一的消息则不值得。

机器人使用一个真实世界数据集进行仿真，该数据集是通过安装在无人机上的高光谱相机采集自加利福尼亚州的一座湖泊。环境约为 80×60 米，离散为 $\boldsymbol{G}^{\#} = 25 \times 25$ 个位置。每个测量集合都是一幅图像，包含高光谱图像 400nm 通道的 5×5 个像素强度值，并归一化至 $[0, 1]$。我们为每个测量值添加均值为零、标准差为 0.05 的高斯噪声。

A. 效用方法比较

我们研究了不同方法在网络负载和估计分位数˜值 V_{final} 与 V之间的最终误差（均方根误差）这两方面的性能表现。单位为像素强度，归一化至范围 $[0, 1]$。表I展示了数值结果，包括将各方法与始终方法进行比较，列出所报告指标的中位数百分比变化。这些结果也以图表形式呈现在图2中，展示网络负载与任务性能之间的权衡。图3展示了随时间变化的累积网络负载，该负载以已尝试的传输次数衡量；图4则展示了随时间变化的已尝试（蓝色）和成功（橙色）消息传输次数。图5展示了不同效用方法的最终均方根误差箱线图。图6展示了叠加在湖泊高光谱图像上的示例路径。

表I：每种效用方法尝试传输和成功传输的消息数量。使用动作限制传输，可使网络负载（已尝试）减少超过 40%，同时将最终平均分位数估计误差改善近 2%，而成功传输的减少幅度仅不到 25%。结果来自总计 60 次实验，每次实验长 40 步。所有实验均使用 $N = 4$，因此最多可能发送的消息总数为 $40 \times (N-1) = 120$。

方法	已尝试 (μ, σ)	中位数减少	成功 (μ, σ)	中位数减少	中位数均方根误差增加
动作	(66.6, 27.95)	42.5%	(38.2, 23.85)	24.32%	-1.84%
始终	(120.0, 0.0)	–	(44.93, 16.78)	–	–
自我奖励	(76.2, 28.33)	27.5%	(34.4, 20.27)	32.2%	5.96%
从不	(0.0, 0.0)	100.0%	(0.0, 0.0)	100.0	73.98%
奖励	(68.2, 32.1)	52.5%	(36.0, 22.77)	35.96%	16.03%

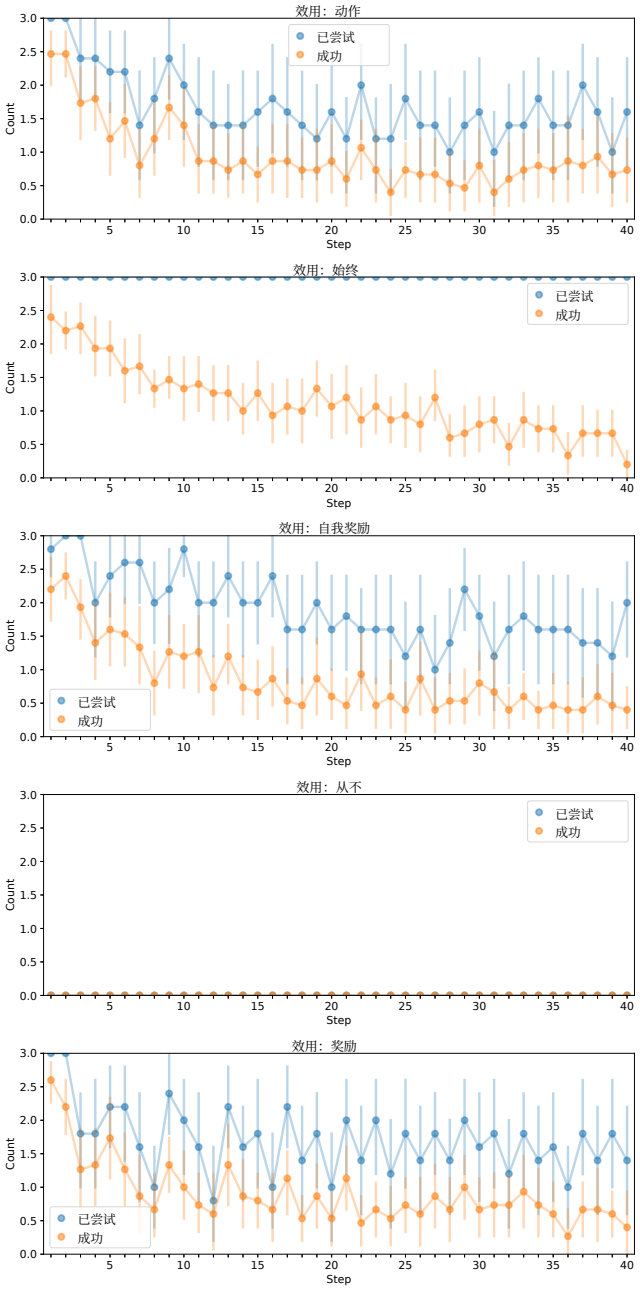

图 4：不同效用方法随时间变化的已尝试和成功传输次数。x 轴上的每一步代表一个机器人执行一个规划步。

在图4中，我们观察到，所有方法的成功传输次数总体上都会随着时间推移而减少，这是因为机器人逐渐分散开来。将图5中的从不方法与其他方法进行比较，可以确认我们此前的发现：机器人之间的通信能够提升性能。不过，我们主要关注减少网络负载所带来的影响。

分析表I和图2中的结果可知，通信量大幅减少并不一定会导致性能大幅下降。我们看到，动作方法的性能最佳：与始终发送相比，按已尝试的消息传输次数衡量，网络负载减少了 40% 以上，但实际（成功）消息传输次数仅减少不到 25%。尤其值得注意的是，就最终误差而言，均方根误差的均值降低了 1.84%，这表明即使在通信受限的情况下，平均性能也略有提升。

其他基于奖励的方法——奖励和自我奖励——表现也不错。图3和图4表明，所有三种提出的方法的消息传输率总体上都会随时间降低。我们推测，这可能有两个原因。首先，随着机器人对环境的探索逐渐深入，并从其他机器人那里获得更多信息，它们对队友的信念会不断增强，使得 $|\boldsymbol{X}^j|$增加，从而形成更准确的环境模型，因此接收额外一组测量数据的价值会降低。其次，随着探索时间增加，机器人彼此距离较远的可能性也更大。距离增加时，期望效用会根据传输概率而降低；即使认为某些消息有用，也可能因为成功送达的概率较低而不值得发送。不过我们注意到，公式（13）中的乐观假设可避免因距离较远而不发送高度有用的消息。

如图3和图4的定性结果以及表I的定量结果所示，与始终相比，动作、奖励和自我奖励的网络负载都更低，但它们的相对负载各不相同，最终估计误差也有所差异。具体而言，奖励带来的网络负载减少幅度最大（不包括从不），达到 52.5%，但最终误差的增幅也最大，超过 16%。因此，当网络带宽严重受限时，这种指标可能很有价值。另一方面，自我奖励带来的网络负载减少幅度最小，为 27.5%，但误差增幅相对较小，不到 6%。自我奖励的优点是不需要对队友信念进行任何机载建模；因此，当机载计算能力有限时，这种方法可能十分有效。总体而言，动作的表现

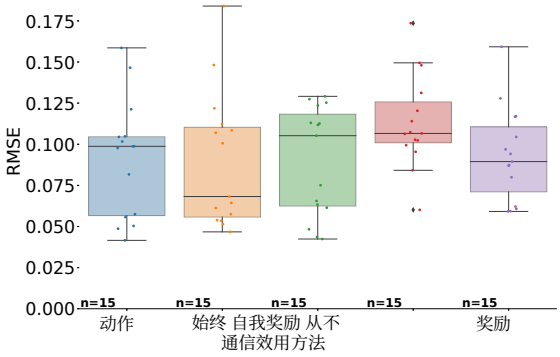

图 5：不同效用方法的最终估计误差。

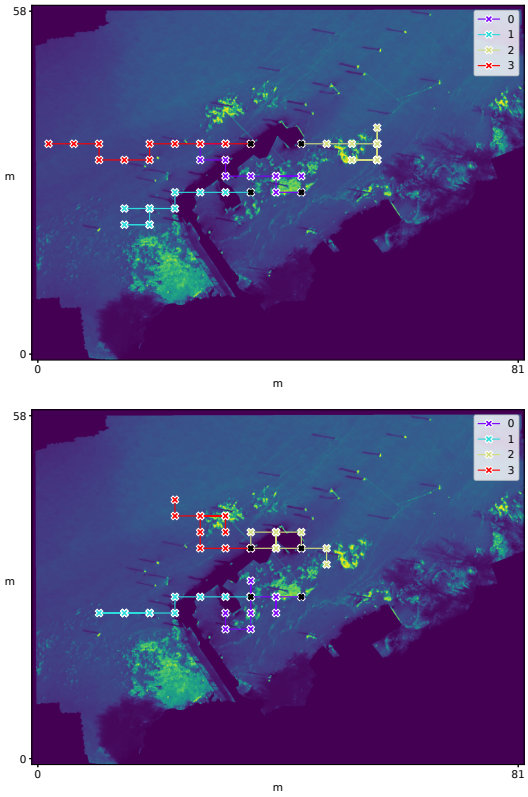

图 6：使用两种不同效用方法得到的路径示例，参数为 $Q = (0.9, 0.95, 0.99)$。上：动作。下：奖励。

最佳，网络负载减少了 42.5%，最终误差减少了 1.84%。我们推测，动作表现更好是因为它直接考虑了消息对未来行为的影响，从而在下一步能够更有针对性地获取测量，并据此做出更充分的信息判断。我们指出，在此处比较网络负载的相对值比比较绝对值更有意义，因为方程 (12) 中使用的阈值是可调参数，在实际部署中应根据真实的网络约束进行设定。

B. 预言机握手

我们对通信协议进行了修改，假设采用预言机握手例程，该例程可为每个 j indi- 提供 i 准确的确认握手

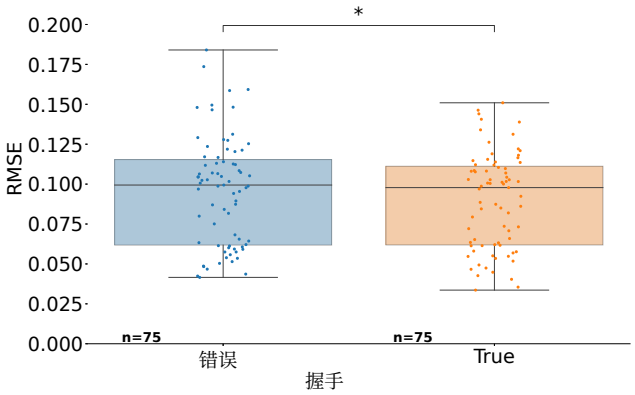

图 7：右侧：消息成功传输后，发送方会收到预言机握手。柱状条表示单侧 Wilcoxon符号秩检验的显著性；*:$p \leq 5\text{e}{-2}$。

cating，说明消息是否已成功接收；i 随后仅在成功时才更新 M^j。图 7 显示了两者最终均方根误差的对比；可以看到，采用预言机握手后性能显著提升（Wilcoxon 符号秩检验 $T{=}1143$，$p{\leq}0.05$）。虽然在实践中几乎不可能实现，但这表明，在网络中实现此类（概率性）例程或许能进一步提升性能；不过，也必须考虑由此产生的消息开销。

V. 结论

我们首次研究了环境分析中分位数估计场景下的受限通信。我们证明，根据发送方对其他机器人在下一状态下会采取何种动作的判断是否会因发送消息而改变，来限制消息传输，与尽可能传输每条消息相比，可使网络负载减少 42%，同时使分位数估计误差降低 1.84%。这表明，在机器人及其网络资源受限时，分位数估计期间的定向通信可为现实世界中的多机器人部署带来切实益处。

除了接收确认握手之外，未来工作还有几个值得探索的方向。我们认为，与将发送方限制为只能发送最新测量值相比，允许其从收集到的数据中选择最有价值的测量值，可能会提升性能。然而，随着对环境的探索不断深入，这将带来可扩展性方面的挑战；对此，可以通过对之前的测量值应用滚动窗口来缓解这一问题。此外，我们还希望研究更好地近似机器人为其每个队友维护的模型的方法；一种实现方式是为消息头添加元数据，例如该机器人的当前数据集大小或当前分位数估计。

参考文献

[1] I. M. Rayas Fern´andez、C. E. Denniston 和 G. S. Sukhatme，《自然环境中的多机器人分位数估计研究》，*arXiv*电子印本，第 arXiv–2303 页，2023年。[2] I. M. Rayas Fern´andez、C. E. Denniston、D. A. Caron 和 G. S. Sukhatme，《用于环境分析的分位数估计信息路径规划》，*IEEE*机器人与自动化快报，第7卷第4期，第10 280–10 287页，2022年。[3] P. Padhy、R. K. Dash、K. Martinez 和 N. R. Jennings，《一种用于无线传感器网络的基于效用的自适应感知与多跳通信协议》，*ACM*传感器网络汇刊（*TOSN*），第6卷第3期，第1–39页，2010年。[4] P. Lieser、N. Richerzhagen、S. Luser、B. Richerzhagen 和 R. Stein- metz，《理解消息优先级排序对灾后自组织网络的影响》，载于*2019*年网络系统国际会议（*NetSys*）。IEEE，2019年，第1–8页。[5] Z. Ma、Y. Luo 和 J. Pan，《面向多智能体路径规划的选择性通信学习》，*IEEE*机器人与自动化快报，第7卷第2期，第1455–1462页，2021年。[6] S. Q. Zhang、Q. Zhang 和 J. Lin，《通过基于方差的控制实现多智能体强化学习中的高效通信》，神经信息处理系统进展，第32卷，2019年。[7] M. Barci´s、A. Barci´s、N. Tsiogkas 和 H. Hellwagner，《多机器人系统中的信息分发：通用的效用感知优化中间件》，机器人与人工智能前沿，第8卷，第685105页，2021年。[8] J. Anderson 和 G. A. Hollinger，《面向基于地形的协同水下定位的通信规划》，传感器，第21卷第5期，第1675页，2021年。[9] G. Best 和 G. A. Hollinger，《用于多机器人信息收集的分布式自组织映射》，载于*2020*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*）。IEEE，2020年，第4790–4797页。[10] A. Simonetto、L. Kester 和 G. Leus，《分布式时变随机优化与基于效用的通信》，*arXiv*预印本 *arXiv:1408.5294*，2014年。[11] R. J. Marcotte、X. Wang、D. Mehta 和 E. Olson，《带宽约束下的多机器人通信优化》，自主机器人，第44卷第1期，第43–55页，2020年。[12] V. Unhelkar 和 J. Shah，《Contact：在未知的确定性领域中决定是否在时间关键型协作任务期间进行通信》，载于*AAAI*人工智能会议论文集，第30卷第1期，2016年。[13] P. Velagapudi、O. Prokopyev、K. Sycara 和 P. Scerri，《在大型多智能体团队中维护共享信念》，载于*2007*年第*10*届国际信息融合会议。IEEE，2007年，第1–8页。