LLM 引导的自组织机器人团队分布式探索

Hiroaki K awashima[1], Shun Ikejima[2], Takeshi Takai[3],Mikita Miyaguchi[3],和 Yasuharu Kuni i[4]

[1]兵库县立大学信息科学研究生院，日本兵库县（电子邮箱：kawashima@gsis.u-hyogo.ac.jp）[2]兵库县立大学社会信息科学学院，日本兵库县（电子邮箱：fa22a005.k@gmail.com）[3]竹中工务店竹中研发研究所，日本千叶县（电子邮箱：{takai.takeshi, miyaguchi.mikita}@takenaka.co.jp）[4]中央大学理工学部，日本东京（电子邮箱：kunii@elect.chuo-u.ac.jp）

arXiv:2603.04762v1 [cs.RO] 5 Mar 2026

摘要：当单个机器人感知能力有限或容错能力不足时，探索过程中就需要多个机器人组成团队，从而扩大集体观测范围并提高可靠性。传统上，机器人群的群体编队通常由中央控制器管理；然而，从鲁棒性和灵活性的角度来看，即使没有集中式控制，机器人群也能自主运行更为理想。此外，在此类多团队探索场景中，为每个团队确定探索目标对实现高效探索至关重要。因此，本研究提出一种探索方法，将（1）一种自组织算法（能够自主、动态地组建多个团队）与（2）一种使每个团队能够自主确定下一个探索目标（目的地）的算法相结合。具体而言，对于（2），尽管基于前沿的方法和深度强化学习方法已得到广泛研究，本研究探索了一种基于大语言模型（LLM）的新策略。通过涉及数十至数百个机器人的仿真，验证了所提方法的有效性。

关键词：移动机器人、团队组建、自组织、分布式探索、大语言模型（LLM）

1. 引言

探索月球熔岩管等未知环境时，机器人可能因通信中断、传感器故障和物理损伤而面临各种失效风险。在此类场景中，部署大量小型移动机器人，而非依赖少数高性能机器人，可以增强系统的整体容错能力。

为了提高探索效率，通常希望多个机器人覆盖互不重叠的区域。然而，当单个机器人感知能力有限或容错能力不足时，多个机器人就有必要在探索过程中组建团队，以扩大其集体观测范围并提高可靠性。传统上，机器人群的编队通常由中央控制器管理；但从鲁棒性和灵活性的角度来看，即使没有集中式协调，机器人群也应能够自主运行。

本研究提出了一种探索框架，该框架融合了：（1）一种可自主、动态地组建多个团队的自组织算法；（2）一种使每个团队能够决定其探索目标的目的地选择算法。团队组建完成后，每个团队都必须确定下一个探索目的地。尽管传统的基于前沿的方法和深度强化学习方法已得到广泛研究，我们探索了一种基于大语言模型（LLM）的新策略。

具体而言，每个团队领队都会利用从概率占据栅格地图中提取的信息，以及团队的当前位置和其他团队的位置与目的地，从前沿单元格集合中选择一个目标单元格。使用这些信息，大语言模型会进行常识推理，以确定团队的下一个目的地。我们的贡献在于通过仿真展示大语言模型在分布式多机器人探索中的有效性，尤其是在涉及动态团队组建的场景中。

2. 相关工作

未知环境中的多机器人探索是一个长期存在的研究课题，研究人员已提出多种技术，包括经典的基于前沿的方法 [5],随机图搜索 [2, 4], 效用优化 [1]。近年来，深度强化学习方法也被应用于多机器人探索 [3], ，并扩展到异步和分布式场景 [6]。虽然这些方法侧重于单个机器人的行为，但我们的方法强调以分布式方式进行团队组建和目的地选择。此外，我们还探索了利用大语言模型进行目的地选择，这在该领域是一种新颖的方法。

3. 分布式探索

3.1. 机器人模型和环境表示

本研究探讨如何利用一群配备短程稀疏采样传感器的 N 移动机器人探索未知环境。所有实验均在仿真中进行。每个机器人的直径约为0.3米。为模拟有限的感知能力，每个机器人配备了九条射线，用于探测其周围的障碍物和自由空间，如

作者版本。载于 AROB 第31届与 ISBC 第11届联合研讨会论文集（AROB-ISBC 2026），第923–927页，2026年。

图1a所示。感知扇区从前进方向向左右两侧各延伸70度，每条射线的最大感知范围为1米。

如图1b所示，每台机器人都根据其传感器观测构建概率占据栅格地图。栅格地图由 0.5 m × 0.5 m 的栅格单元组成，每个栅格单元存储被障碍物占据的概率。机器人使用贝叶斯滤波更新这些占据概率：基于不确定性较高的距离测量以及被认为更可靠的自身位置，在对数几率表示中进行加性更新。

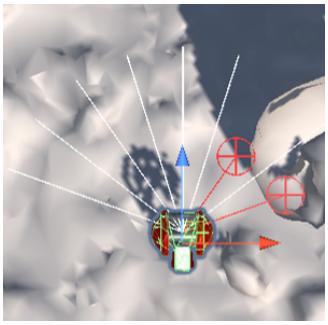
由于通信范围限制也是现实场景中的一个重要因素，每台机器人都会持续维护自己的局部地图，并在遇到附近的机器人时，使用对数几率表示中的加性更新，将接收到的地图与自己的地图融合。然而，在本研究中，我们假设所有机器人之间都能充分通信，以便专注于主要目标：团队组建和目的地选择。在这一假设下，每一步都会将所有局部地图融合为一张全局地图，并将其共享给每台机器人。为简化起见，我们还假设每台机器人都知道自己和其他机器人的位置；不过在现实场景中，则需要利用机载传感器和通信来估计相对位置。

3.2. 任务描述与团队行为

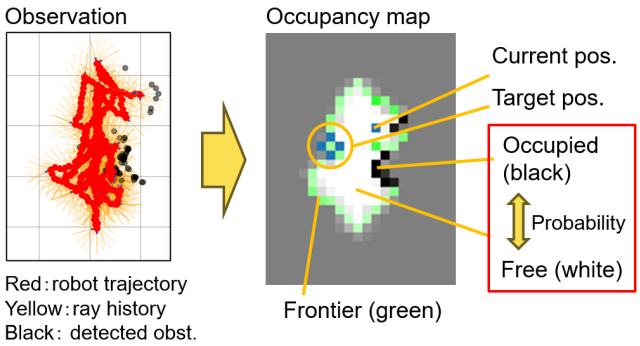
探索的目标是在固定时间内最大化覆盖面积。起初，机器人对环境一无所知。因此，前沿单元格（定义为已探索自由空间与未探索空间之间的边界）被确定为潜在的探索目标，并从中选取一个作为下一个探索目标。在此处，通过设置阈值，根据占据概率将栅格单元分类为自由、占据或未知。每个机器人根据当前占据栅格地图，使用 A* 算法计算前往分配目标的路径，其中概率值作为代价。尽管地图以栅格单元表示，机器人每一步都在连续空间中通过离散动作移动：停止、固定前进运动以及固定角度的左转/右转。每一步的前进距离小于栅格单元尺寸（0.2 米/步），旋转角度为 40 度/步。

在探索过程中，我们引入由多个机器人组成的团队，以扩大集体观测范围并提高可靠性。这些团队以去中心化方式动态组建，每个团队自主确定其探索目标。在每个团队中，会指定一台领航机器人代表团队选择一个目标栅格单元，并将该目标共享给所有团队成员。随后，每个成员独立规划前往共享目标的路径，并据此移动。任一成员到达指定目标后，领航机器人便会选择一个新目标，重复这一过程。

我们还假设，每个机器人都有有限的电池容量，并以两种模式之一运行：探索（*EXP*）或充电（*CHR*）。当电量低于指定阈值时，机器人会切换到 CHR 模式，离开当前团队，并前往



(a) 射线传感器



(b) 栅格地图生成

图 1. 传感器模型和占据栅格地图生成。(b) 中，灰度表示占据概率值：灰色栅格单元表示未探索区域，白色栅格单元表示已探索自由区域（自由空间），黑色栅格单元表示已探索的含有障碍物的区域（占据空间）。绿色栅格单元表示前沿单元格。

充电站进行充电。完成充电后，机器人会切换回 EXP 模式，并通过自主组建新团队或加入现有团队来继续探索。

以下小节将详细介绍这两个组成部分：去中心化团队组建和基于大语言模型的团队目的地确定。

3.3. 去中心化团队组建

最佳团队组建取决于多种参数，例如待执行的任务（如探索、物体运输）、运行条件以及机器人的状态（如位置、相对配置、电池电量）。因此，我们从单个机器人的行为策略角度提出一种自主组建算法。

为实现去中心化团队组建，我们引入期望团队规模 $\tilde{n}_i$ 这一概念，将其作为每个机器人 i的内部状态参数。这样，每个机器人都能根据当前情况自主确定组建团队所需的机器人数量。例如，进入尚未探索的区域时，较大的期望团队规模更为合适，因为这需要一定数量的机器人。在这种情况下，当期望规模超过实际团队规模n_i，即 $\tilde{n}_i > n_i$时，机器人便进入招募状态。它会尝试与同样处于

招募状态且符合特定条件（例如距离较近）的其他个体或团队加入。相反，当较小的团队更有优势时，机器人会降低期望规模。例如，离开团队前往充电站时，最好单独行动，因此机器人会将期望团队规模设为1。如果期望规模小于实际团队规模，即 $\tilde{n}_i < n_i$，机器人便会尝试离开团队。离队条件取决于具体场景，我们将在第4.2节介绍一个简单的实现示例。

3.4. 基于大语言模型的团队目的地确定

在目的地确定方面，虽然已有经典的基于前沿的方法和深度强化学习方法，但本研究探索了一种使用大语言模型（LLM）的新方法。

大语言模型是一个未经微调的预训练模型，输入内容为单元格坐标列表，其中包含以下信息的占用状态：

1. 已探索的非障碍单元格 2. 已探索的障碍物单元格 3. 前沿单元格

以及团队领队的当前位置和其他团队的位置及目的地。领航机器人的位置用作各团队的当前位置，但也可以采用其他团队位置定义（例如团队中心）。

列表中的每个栅格单元都包含其与团队领队当前位置之间的距离以及邻域特征：每个栅格单元的信息由（坐标、标签、邻域特征、距离）组成。此处，标签表示该栅格单元是已探索的非障碍单元格（1）、已探索的障碍物单元格（2）还是前沿单元格（3）。对于邻域特征，我们采用八连通邻近单元格内前沿单元格和占用单元格的数量。

利用这些信息和给定的提示模板，大语言模型会进行常识推理，从前沿单元格中确定团队的下一个目的地。具体而言，在每个决策步骤中，团队领队会为大语言模型构建提示，其中包括当前地图信息（即上文说明的栅格单元列表）、团队的当前位置，以及其他团队的位置和目的地，并要求大语言模型选出下一个目标前沿单元格，以实现“高效探索”。

4. 实验

4.1. 设置

探索环境根据未来的月球熔岩管建模，使用了从地球熔岩管结构中获取的三维网格数据。大语言模型采用 Azure OpenAI gpt-4o。如果返回的栅格单元不是前沿单元格，则会再次调用 API，直到获得有效的前沿单元格。API 调用的最大重试次数设为 5 次，之后将使用下文介绍的基线方法作为备用方案。

基线目标选择： 作为目的地选择的基线方法，我们实现了一种概率采样方法，为每个团队选择附近的一个前沿单元格，并结合

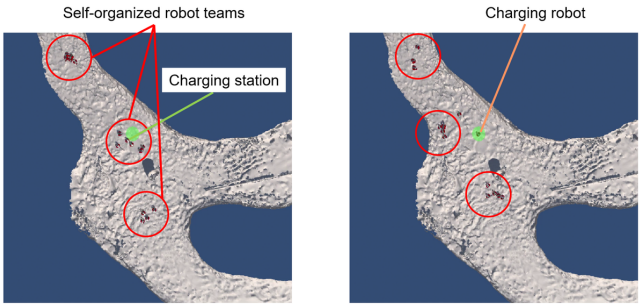


图 2. 自组织团队与充电机器人的示例

分配给每个机器人的个体偏差；因此，目的地选择会受到当前团队领导者具体特征的影响。具体而言，首先计算领队位置到每个前沿单元格的距离。随后，从具有机器人特定有偏均值和标准差为 0.02 的正态分布中采样一个分位数值，并将其限制在范围 $[0, 1]$ 内。对于每个机器人，在每次仿真开始时，均值都会从范围 $[0.10, 0.25]$ 中均匀采样。然后选择与相应距离分位数对应的前沿单元格作为目标。这种方法允许偶尔选择距离更远的前沿单元格，而不是始终选择最近的单元格，从而促进探索多样性。

4.2. 团队的自组织

为了验证去中心化团队组建的基本概念，我们根据机器人的两种运行模式 EXP 和 CHR，采用了一组简单的期望团队规模设定。在 EXP 模式下，期望团队规模设为 5；在 CHR 模式下，则设为 1。在实验中，我们具体采用了以下简单规则：合并在团队层面进行，而分裂则在单个机器人层面进行。当两个团队都处于招募状态且彼此距离足够近时，就会发生合并。具体而言，如果机器人 i 在阈值距离 d_{join}内发现机器人$j \notin T(i)$，并且满足条件 $n_i + n_j \leq \min(\tilde{n}_i, \langle \tilde{n}_j \rangle_{j \in T(j)})$，则团队$T(i)$和 $T(j)$合并。在此过程中，单个机器人被视为规模为 1 的团队。相反，离开由单个机器人执行；当电池电量较低时（即处于 CHR 模式），机器人可以无条件离开。在更一般的情况下，可以设计一条离开规则：如果团队内的平均期望团队规模小于实际团队规模，则允许机器人离队。若满足 $\langle \tilde{n}_i \rangle_{i \in T(i)} < n_i$，机器人 i 即可离开。

在评估基于大语言模型的目的地选择方法之前，我们首先定性验证采用集成自主团队组建算法的基线方法进行探索的有效性。图 2 展示了一个示例，其中 $N = 15$ 机器人动态管理团队组织和电池充电这两项任务。如图所示，处于 EXP 模式的机器人会自主组建团队（以红色圆圈标示），而处于 CHR 模式的机器人会离开团队，前往充电站（以绿色标示）充电，如图 2（右）所示。

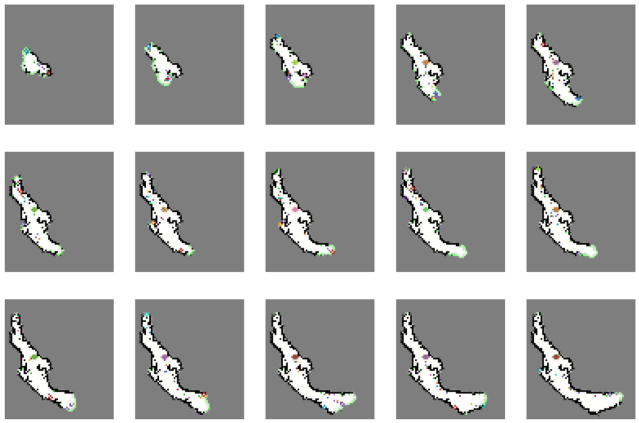


图 3. 使用基线方法的探索示例

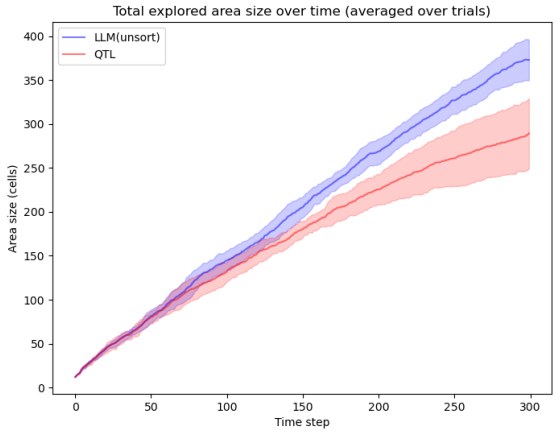


图 4. 基线方法与基于大语言模型的方法之间已探索区域的比较（N = 15）

```
==================== [LLM Response] ==============================
# REASON
To decide the NEXT_TARGET frontier cell for our group (located at (23, 42)), we aim for efficient
exploration while avoiding conflicts with other groups. Following these steps:

1. **Frontier Cell Candidate Selection**:
   Our task specifies that the target cell must be a frontier cell with a label of "3" (indicating
an unvisited neighboring area).

2. **Distance Consideration**:
   Prioritize the closest frontier cells to avoid unnecessary movement costs and ensure efficient
exploration. Distances to all frontier cells are given in the CELL_LIST.

3. **Conflict Avoidance**:
   Avoid selecting frontier cells already targeted by other groups. Checking the target cells of
Group[1] ((18, 40)) and Group[2] ((16, 43)), we remove any conflicting options. Additionally,
ensure that the next target does not disrupt ongoing exploration by being too close to other
groups.

4. **Exploration Efficiency**:
   Frontier cells with higher numbers of frontier neighbors (8-neighbor frontier count) are more
promising, as they are likely to lead to larger unexplored areas. Secondly, avoid targeting cells
with high obstacle counts for smoother navigation.

5. **Selection Based on CELL_LIST**:
   Sorting by distance and prioritizing high frontier-neighbor counts, the best-fit frontier cell
for our group is selected.

**Candidate Analysis** (near (23, 42)):
- (24, 41): Distance 1.41, 1 frontier, 3 obstacles → Close but not ideal due to high obstacle
count.
- (23, 44): Distance 2.00, 3 frontiers, 0 obstacles → Close with a high frontier count and no
obstacles.
- (23, 45): Distance 3.00, 4 frontiers, 0 obstacles → Slightly farther but the highest frontier
count.

After removing conflicts and analyzing the CELL_LIST, **(23, 45)** is the optimal next target,
balancing proximity to our group, lack of obstacles, and high frontier neighbor count.

# NEXT_TARGET
NEXT_TARGET: (23, 45)
==================================================================
```

图 5. 用于目的地选择的大语言模型推理示例

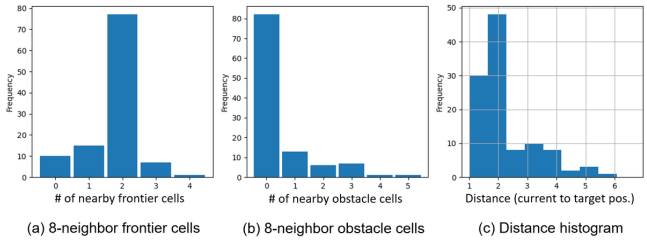


图 6. 基线方法与基于大语言模型的方法之间目标单元格统计数据的比较

图 3 展示了在与上述相同条件下，采用基线方法进行探索时，已探索区域如何随时间扩大的示例。图中清楚地显示，已探索的自由区域（白色）逐渐增加，而未探索区域（灰色）则随着探索的进行而减少。在此过程中，多个团队动态组建并重组，机器人加入或离开团队，最终各自返回充电站。

4.3. 基于大语言模型的目的地选择

为定量评估所提出的基于大语言模型的目的地选择方法的有效性，我们比较了使用 $N = 15$台机器人在固定时长300步内实现的总已探索区域。基于大语言模型的方法在20步基线探索之后启动；这一“预热期”确保有足够的地图信息可供大语言模型做出明智决策。结果取五次独立试验的平均值，如图4所示。如图所示，所提出的基于大语言模型的方法相比基线方法，已探索区域大约增加了20%。

图 5 展示了一个大语言模型推理的示例。如图中所示，大语言模型成功综合考虑了多个因素，包括与自身团队的接近程度、避免与其他团队的目标单元格重叠，以及附近前沿单元格（前沿单元格数量越多越好）和障碍物单元格的分布（障碍物单元格数量越少越好），从而生成一组候选目标单元格。在总结这些候选单元格的特征后，大语言模型选择了最合适的目标单元格，而该单元格不一定是距离最近的前沿单元格。该示例展示了大语言模型基于所提供的占据栅格地图和团队级信息进行常识推理的能力。

图6展示了附近前沿单元格的分布（图6a）、附近障碍物单元格的分布（图6b），以及从团队领队到目标单元格的距离直方图（图6c）。这些分布清楚地说明了大语言模型所选目标单元格的特征：倾向于选择被更多前沿单元格和更少障碍物单元格包围的目标单元格，同时允许稍远一些的目标单元格。

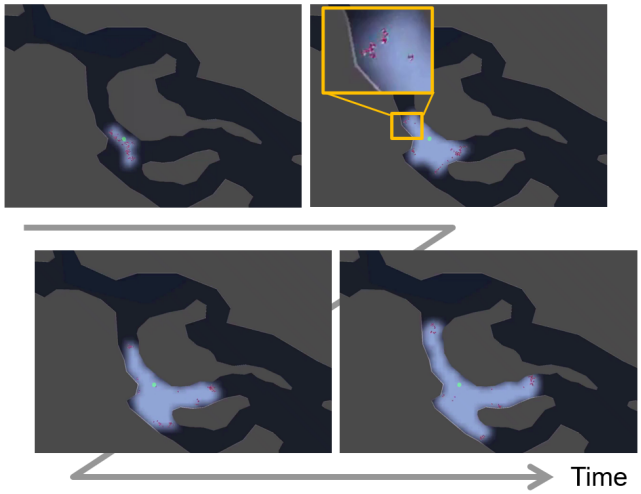


图 7. 采用自主团队组建进行探索，人数为$N = 50$。明亮区域表示已探索区域，小红点表示各个机器人。

4.4. 更多机器人数量下的可扩展性

我们结合第4.2节所述的自主团队组建机制，使用更多数量的机器人验证了所提出的方法。在 $N = 50$ 台机器人的情况下，我们确认结合自主团队组建的基线探索能够有效运行，从而实现对广阔区域的探索（图7）。黑色区域表示熔岩管，而明亮区域表示机器人已探索的区域。若干小红点对应于在三维模拟器中运行的各个机器人，它们在推进探索的同时，在机器人群体内自主组建团队。

此外，我们通过将机器人数量增加至 $N = 100$，验证了所提出的基于大语言模型的自组织团队组建探索方法的可扩展性。图 8 展示了仿真的屏幕截图。左侧窗口显示三维模拟器，其中许多小红色物体表示正在熔岩管环境中探索的各个机器人。右下方窗口显示通过探索获得的占据栅格地图，其配色方案与图 1 一致。在该地图中，被紫色点围绕的栅格单元表示相应团队的目标单元格，而其他颜色的点则表示各个机器人的位置。

5. 结论

本文提出了一种去中心化探索方法，该方法整合了（1）自主团队组建算法和（2）基于大语言模型的目的地选择算法。通过仿真，将所提方法与采用概率前沿采样的基线探索方法进行比较，验证了其有效性。此外，我们定性确认，配备团队级大语言模型推理的大规模机器人群能够高效扩展已探索区域。未来工作包括experiments在有限通信条件下进行，根据环境条件和机器人的内部状态开展策略学习以调整团队规模，以及去中心化任务切换至其他目标，例如物体运输。

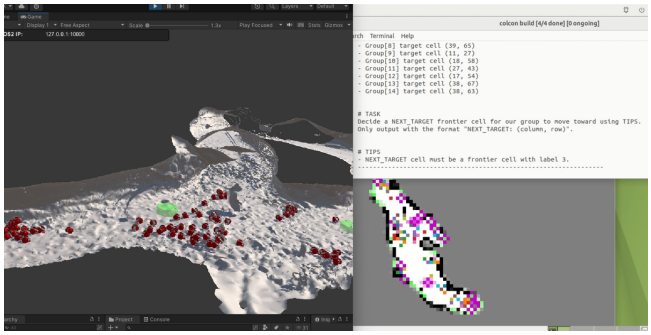


图 8. 基于大语言模型的探索仿真截图，人数为 $N = 100$

致谢

本研究得到了日本JST Moonshot研发计划的支持，项目编号为JPMJMS2238。

参考文献

[1] Wolfram Burgard、Mark Moors、Cyrill Stachniss和Frank Schneider。协同多机器人探索。*IEEE*机器人学汇刊，21(3):376–386，2005。[2] Antonio Franchi、Luigi Freda、Giuseppe Oriolo和Marilena Vendittelli。一种用于协同机器人探索的随机策略。*IEEE*国际机器人与自动化会议论文集，第768–774页，2007年。[3] Haoran Li、Qichao Zhang和Dongbin Zhao。基于深度强化学习的未知环境导航自动探索。*IEEE*神经网络与学习系统汇刊，31(6)，2020年。[4] G. Oriolo、M. Vendittelli、L. Freda和G. Troso。SRT方法：用于探索的随机策略。*IEEE*国际机器人与自动化会议论文集，第4688–4694页，2004年。[5] Brian Yamauchi。使用多台机器人进行基于前沿的探索。国际自主智能体会议论文集，第47–53页，1998年。[6] Chao Yu、Xinyi Yang、Jiaxuan Gao、Jiayu Chen、Yunfei Li、Jijia Liu、Yunfei Xiang、Ruixin Huang、Huazhong Yang、Yi Wu和Yu Wang。用于高效实时多机器人协同探索的异步多智能体强化学习。第*22*届国际自主智能体与多智能体系统会议（*AAMAS 2023*）论文集，2023年。