Diffusion-Guided Multi-Arm Decentralized Motion Planning

Viraj Parimi, Annabel Gomez, Howard Chen, Andreas Hoffman, Brian Williams

September 25, 2024

Abstract

Multi-arm task and motion planning approaches, as described by [CLH+22], typically follow a multi-stage process. Initially, offline multi-modal decoupled roadmaps are generated for each robot in the scene, followed by the annotation of collision information. The next step leverages advancements in Multi-Agent Path Finding (MAPF), where search-based MAPF techniques use these collision-annotated roadmaps to find collision-free paths for each robot. However, the collision annotation step, which involves pairwise comparisons across all nodes in the roadmap, is highly computationally intensive and becomes increasingly impractical as the number of robots grows. While roadmap-based methods offer good asymptotic optimality guarantees, the time-consuming nature of collision annotation coupled with NP-hardness of MAPF problems can result in simple tasks, like pick-and-place, taking hours to find a feasible plan. This challenge is further compounded in dynamic settings where the environment may change frequently.

To overcome these limitations, we propose learning a combined multi-arm diffusion policy capable of handling dynamic obstacles, such as moving robots, which would eliminate the need for constructing traditional roadmaps and performing laborious collision annotations. Diffusion models, with their generative capabilities, can directly produce feasible actions for agents without relying on precomputed roadmaps. However, a key trade-off is that diffusion models typically require expert demonstrations to learn effective policies, which can be difficult to obtain in practical multi-agent contexts. More importantly, learning exclusively from successful expert demonstrations may not provide sufficient exposure to negative examples, which are crucial for understanding collision boundaries and ensuring robust behavior in near-collision scenarios.

To address this, we propose enhancing the learning process by incorporating expert demonstrations alongside an oracle mechanism that evaluates the quality of planned actions. This oracle can be realized using an actor-critic framework within reinforcement learning (RL), where value functions guide agents towards actions that optimize desired objectives. Previous work by [HXS20] explored the use of RL to develop decentralized policies for multiple agents, augmented by on-demand expert demonstrations that assisted agents in completing tasks they were unable to perform on their own. However, their approach used traditional policy models, which struggled to capture the multi-modal distributions prevalent in many manipulation tasks.

In contrast, diffusion models are adept at handling such multi-modality, particularly in single-arm manipulation contexts [CXF+24]. By integrating RL with diffusion models like [WHZ23], we aim to facilitate the transfer of learned behaviors across decentralized agents, enabling them to develop effective policies even with limited communication. Building on concepts from recent work [ZLM+24], our approach extends the diffusion model to predict not only an agent's next actions but also the anticipated actions of surrounding agents, allowing for coordinated planning in multi-agent environments. This enables agents to perform receding-horizon and closed-loop planning, which can adapt to sudden, unforeseen environmental changes, leading to quicker generation of feasible motion plans.

We hypothesize, however, that while this proposed approach may enhance adaptability and speed, it may not achieve the same level of plan near-optimality as search-based methods, particularly in terms of asymptotic performance. Therefore, a critical area for investigation is the trade-off between the optimality of traditional search-based methods and the speed and flexibility of RL-guided, diffusion-based models.

References

[CLH+22] Jingkai Chen, Jiaoyang Li, Yijiang Huang, Caelan Garrett, Dawei Sun, Chuchu Fan, Andreas Hofmann, Caitlin Mueller, Sven Koenig, and Brian C. Williams. Cooperative task and motion planning for multi-arm assembly systems, 2022.

[CXF+24] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion, 2024.

[HXS20] Huy Ha, Jingxi Xu, and Shuran Song. Learning a decentralized multi-arm motion planner, 2020.

[WHZ23] Zhendong Wang, Jonathan J Hunt, and Mingyuan Zhou. Diffusion policies as an expressive policy class for offline reinforcement learning, 2023.

[ZLM+24] Zhengbang Zhu, Minghuan Liu, Liyuan Mao, Bingyi Kang, Minkai Xu, Yong Yu, Stefano Ermon, and Weinan Zhang. Madiff: Offline multi-agent learning with diffusion models, 2024.