Robust Following with Hidden Information in Travel Partners

Shih-Yun Lo
University of Texas at Austin
yunl@utexas.edu

Elaine Schaertl Short
Tufts University
elaine.short@tufts.edu

Andrea L. Thomaz
University of Texas at Austin
athomaz@ece.utexas.edu

KEYWORDS

Robot Navigation with Humans; Robust Multi-agent Path Planning

ACM Reference Format:
Shih-Yun Lo, Elaine Schaertl Short, and Andrea L. Thomaz. 2020. Robust Following with Hidden Information in Travel Partners. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

Pedestrians often travel with one another [9]. Travel partners share destinations or temporary subgoals. Such information is often communicated on-the-fly, when observations are received and local direction decisions are needed. At decision points, e.g. intersections, we observe that there is often a leader (or more) in a group who actively decides where to go, and the follower(s), without explicit communication, can adapt their motions quickly to catch up [10]. In robotics, human-following is achieved by performing online subgoal inference, for the robot to adapt the motion while maintaining desired group shapes [5], such as side-by-side walking. This capability has been realized by applying maximum-likelihood subgoal estimates upon the robot path planning problem [8, 10]. While assertively following towards the most likely subgoal has shown success in relatively simple environments, such strategy can lead to bad states under false inference, e.g. poor visibility to other route options or blocking group member path options, yielding poor performance under inference delay.

To enable robust robot following towards real-world application, which involves durable performance in complex environment configurations, we present a formulation for robot navigation with humans as a partial-information multi-agent planning problem, and incorporate subgoal estimation *into* the planning process for action evaluation. Under this multi-agent formulation, individuals plan to accommodate paths planned for the human travel partners, while maximizing the expected outcome towards possible subgoals to reach. We show that the proposed planner enables more efficient following when subject to inference delay, including more efficient path planning and robust tracking of the travel partner, and retains the behavior features for natural co-navigation as suggested in the state-of-the-art robot following approaches [5, 8] and agent-based pedestrian models [4, 9].

We also noticed an emergent human-like behavior feature: hesitation. When uncertain, the planner delays its action selection at states where actions have distinctive values under different subgoal specification. Our framework then serves as an improved model for realistic pedestrian group simulation, which has gained attention for interactive agent design, focusing on long-duration small-group (or often one-to-one) interaction [7].

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

2 PROBLEM FORMULATION

We formulate the dynamics of group motions as a sequential optimization problem. We use a game formulation to lay out the mutual-adaptability behavioral features in the literature of crowd simulation [2, 4, 9] and robot planning for human-following [5, 8, 10]. We then point out the issue in their behavioral assumption and propose our group follower model using stochastic Bayesian game.

Collaborative Bayesian Stochastic Games. Pedestrian groups often maintain certain *shapes* to facilitate communication between participants, such as to easily see the faces of each other, stay aware of the focus or attention of others. Group shapes are affected by environmental conditions; agents adapt group shapes under environment changes, or in coordination with other pedestrians sharing the space [9]. These features describe the *macro* pedestrian group behaviors: shape formulation and mutual adaptation.

To ensure all agents have feasible paths, group members need to collaboratively plan for others. We therefore first model the macro pedestrian group behavior as a stochastic game. In stochastic games, N agents act at a time, here illustrated at time t: the joint-action $a_t = (a_t^1, a_t^2, ..., a_t^N) \in A$ is defined by the action spaces of all agents $A = A^1 \times A^2 ... \times A^N$; the joint-state $x_t = (x_t^1, x_t^2, ..., x_t^N) \in X$ is defined by the state spaces of all agents $X = X^1 \times X^2 ... \times X^k$. Time is discretized, and game periods are defined: at the start of each period t, each agent selects an action $a_t^i, i = 1 : N$, then the transition function $\mathcal{T} : X \times A \rightarrow X$ takes in the current state x_t and determines (probablistically) the state at the beginning of the next period x_{t+1}. The reward $\mathbf{r}_t^i$ of an agent i at time t is defined as follows: $\mathbf{r}_t^i = r^i(x_t, a_t^i, a_t^{-i}|\theta) \in \mathbb{R}$, where r^i is the agent's reward function, and $\theta \in \Theta$ is the possible destinations to capture the goal-driven characteristics of group navigation; and $-i$ denotes all agents except i. Here we consider all agents to have the same collaborative reward function r, to maximize the group social welfare.

The optimal policy is then conditioning on θ, given partner policy π^{-i}, which is a function that maps joint state x_t to action a_t^{-i},

$$a_t^{i*} = \underset{a_t^i}{\operatorname{argmax}}\, \mathbb{E}_{a_t^{-i}|\pi^{-i},\theta}[Q^{i|\pi^{-i}}(x_t, a_t^i, a_t^{-i}|\theta)]. \tag{1}$$

This equation follows the stochastic Bayesian game formulation, in which agent utilities (rewards) are parametrized by their types. Agents only have partial observability to the types of other agents.

Instead of solving this stochastic Bayesian game equation based on modeling and inference of π^{-i} and θ, in previous approaches mentioned in agent-based modeling and robot navigation [2, 4, 5, 11, 12], θ is assumed shared by all agents. With this assumption, solving for the optimal collaborative policy for agent i in Eq. 1 converges to the solution of having a *centralized* system optimizing for the whole group,

$$a_t^{i*} = \underset{a_t^i}{\operatorname{argmax}}\, \underset{a_t^{-i}}{\max} \mathbb{E}_{x_{t+1} \sim \mathcal{T}(x_t, a_t)}[r^i(x_t, a_t^i, a_t^{-i}|\theta) + V^i(x_{t+1}|\theta)]. \tag{2}$$

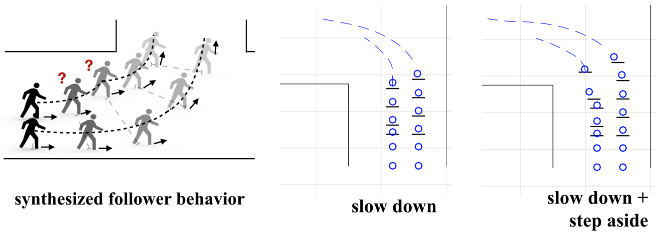

Figure 1: Synthesized human group motion at an intersection (Left), and the simulation using $\pi^{i,1}$ (Middle and Right).

This optimal policy with known θ, is referred as the *zero-inference* policy of agent i: $\pi^{i,0}$, imposing the collective agency assumption in teammate modeling in Bayesian games [1]. The converged homogeneous solution $\pi^{i,0}$ describes this collaborative mutually-adaptive macro pedestrian group behavior; they reciprocally leave room for others to avoid obstacles, and expect partners to do the same.

Proposed Follower Behavioral Model. The shared knowledge assumption for path planning is however invalid for real-world applications, and can lead to inefficient planning performance due to false parametrization. Here we describe follower behavior using the general stochastic Bayesian game formulation in Eq. 1, and model group leader(s) by $\pi^{-i,0}$; the follower makes observation $o_t \in O$ at time t, and use the observation history $o_{0:t-1}$ to compute the expected value over θ, $\mathbb{E}_{\theta|o_{0:t-1}}[Q^i(x_t, a_t^i, \pi^{i,0}(x_t|\theta)|\theta)]$. Here O is the observation space. Given online observations, type identification influences the optimal policy on-the-fly. Along with sensing capability for observation modeling, we use Ω^i to sample agent i's observation given joint state x_t and action a_t^i; $\Omega^i : X \times A^i \to O$. We then formulate followers as the following,

$$a_{0:T}^{i*} = \underset{a_{0:T}^i}{\operatorname{argmax}} \, \mathbb{E}_{x_{0:T},\theta|\Omega^i,\mathcal{T}}\Big[\sum_{t=0}^{T-1} r^i(x_t, a_t^i, \pi^{i,0}(x_t|\theta)|\theta) + Q_T^i(x_T, a_T^i, \pi^{i,0}(x_t|\theta)|\theta)\Big]. \tag{3}$$

Here we consider finite-horizon lookahead T, for the local-observation-driven human behavior rather than long-horizon planning [6, 13]. The multi-agent sequential optimization problem in Eq. 1 now becomes a tractable belief planning problem in finite type space Θ. We also use Eq. 3 to model human followers, based on a Bayesian-optimal assumption. The choice of Ω affects agent behavior by their sensing capability, which we will detail in Section. 3.

Compared to prior work in crowd modeling and robot following, here formulated by Eq. 2, with Eq. 3, the following agent would not assertively turn to a dead-end, or go in front of the leader or stay behind a corner obstacle, which leaves the leader out of its sensing range. We propose that *behaviors driven by uncertainty* as such are missing in the literature to describe the micro inter-group interactive behavior.

The reasoning in Eq. 3 involves the ability to model the individual differences in other agents' knowledge. We here refer this policy as to have *first-order* inference for teamwork planning: $\pi^{i,1}$; it is the basic ability of *theory of mind* in human behaviors [3].

3 VALIDATION

To validate the approach as a robust planner for robot following with humans, we first evaluate our approach's performance subject to inference delay, which is a common issue in robotics. We show superior path efficiency in simulation, with controlled noise and state initialization for benchmark study. We then simulate human follower behavior in comparison to synthesized group motion.

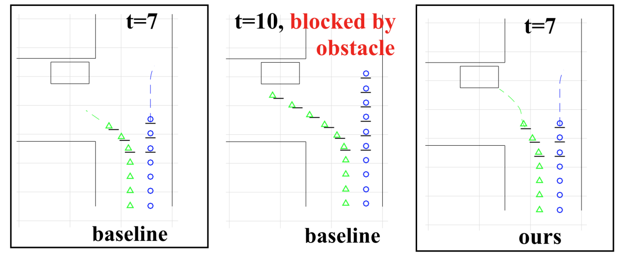

Figure 2: Robot follower behaviors under inference delay : assertive decisions (Left) lead to bad values, e.g. lost tracking of partner (Middle), which our planner prevents (Right).

Follower under Inference Delay. Assertive decisions under false inference can lead to bad values; this is even more significant in cluttered environments, where bad decision can lead to unrecoverable states, given robot non-holonomic constraints and limited sensing capabilities. We choose narrow corridor intersections to demonstrate this scenario, with an obstacle at the corner which potentially blocks the view, leading to lost tracking of the partner shown in Figure. 2. We implemented the baseline with a state-of-the-art robot following approach [10], in comparison to our planner under sensing range of [-120,120] *deg* from forward direction, which is achievable with a common Lidar. Experiments are conducted with 20 trials of randomized initial locations. Even priors are initialized on each direction. Example path comparison can be seen in Figure. 2-Left and Right. Our approach is aware of the observability when planning for future motion and results in more robust motion.

We evaluate path quality when the belief converged to the correct value, and calculate the delay comparing to that predicted by the zero-inference policy $\pi^{i,0}$ (that is, with a perfect prior). Our planner had an averaged delay of $1.19s$, whereas the baseline had $3.53s$. Among the trials, our planner experienced *zero* lost tracking, while the baseline experienced 10. An example is shown in Figure. 2-Middle: when belief converged at $t = 10$, the baseline had gone close to the obstacle, making it lost track of its travel partner.

Human Follower Simulation. We simulate a human sensing model with a range of [-75,75] *deg* from head orientation, using $\pi^{i,1}$. The limited range of view causes the simulated follower to slow down and stay slightly behind the leader, to actively sense the leader's goal, as shown in Fig. 1-Middle; with larger safety margin specified for collision check, the follower steps slightly aside to prevent blocking the leader from potential turning, as shown in Fig. 1-Right, which is a behavior feature described in Murakami, et al. [10].

4 CONCLUSION

We presented a framework for robust planning for group agents, implemented in navigation. The framework erased the need for unrealistic behavioral assumptions as used in prior agent-based modeling and robot planning approaches. We showed superior following performance subject to decision uncertainty, and simulated pedestrians with real-world pedestrian motion features suggested in the literature.

REFERENCES

[1] Michael Bacharach. 1999. Interactive team reasoning: A contribution to the theory of co-operation. *Research in economics* 53, 2 (1999), 117–147.

[2] Dirk Helbing and Peter Molnar. 1995. Social force model for pedestrian dynamics. *Physical review E* 51, 5 (1995), 4282.

[3] Thomas Hellström and Suna Bensch. 2018. Understandable robots-what, why, and how. *Paladyn, Journal of Behavioral Robotics* 9, 1 (2018), 110–123.

[4] Lin Huang, Jianhua Gong, Wenhang Li, Tao Xu, Shen Shen, Jianming Liang, Quanlong Feng, Dong Zhang, and Jun Sun. 2018. Social Force Model-Based Group Behavior Simulation in Virtual Geographic Environments. *ISPRS International Journal of Geo-Information* 7, 2 (2018), 79.

[5] Markus Kuderer and Wolfram Burgard. 2014. An approach to socially compliant leader following for mobile robots. In *International Conference on Social Robotics*. Springer, 239–248.

[6] Shih-Yun Lo, Katsu Yamane, and Ken-ichiro Sugiyama. 2019. Perception of pedestrian avoidance strategies of a self-balancing mobile robot. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS'19*. IEEE.

[7] Owen Macindoe, Leslie Pack Kaelbling, and Tomás Lozano-Pérez. 2012. POMCoP: Belief Space Planning for Sidekicks in Cooperative Games.. In *AIIDE*.

[8] Luis Yoichi Morales Saiki, Satoru Satake, Rajibul Huq, Dylan Glas, Takayuki Kanda, and Norihiro Hagita. 2012. How do people walk side-by-side?: using a computational model of human behavior for a social robot. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM, 301–308.

[9] Mehdi Moussaïd, Niriaska Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. 2010. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one* 5, 4 (2010), e10047.

[10] Ryo Murakami, Luis Yoichi Morales Saiki, Satoru Satake, Takayuki Kanda, and Hiroshi Ishiguro. 2014. Destination unknown: walking side-by-side without knowing the goal. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 471–478.

[11] Sébastien Paris, Julien Pettré, and Stéphane Donikian. 2007. Pedestrian reactive navigation for crowd simulation: a predictive approach. In *Computer Graphics Forum*, Vol. 26. Wiley Online Library, 665–674.

[12] Adrien Treuille, Seth Cooper, and Zoran Popović. 2006. Continuum crowds. In *ACM Transactions on Graphics (TOG)*, Vol. 25. ACM, 1160–1168.

[13] Alasdair Turner and Alan Penn. 2002. Encoding natural movement as an agent-based system: an investigation into human pedestrian behaviour in the built environment. *Environment and planning B: Planning and Design* 29, 4 (2002), 473–490.