Don't be Rude! Learning Group-aware Policies for Robot Navigation

Yuxiang Gao*, Kapil Katyal†, Jared Markowitz†, I-Jeng Wang† and Chien-Ming Huang*
*Department of Computer Science, Johns Hopkins University
{yuxiang.gao, chienming.huang}@jhu.edu
†Johns Hopkins University Applied Physics Laboratory
{kapil.katyal, jared.markowitz, i-jeng.wang}@jhuapl.edu

***Abstract*—This paper explores learning group-aware navigation policies based on dynamic group formation using deep reinforcement learning. Through simulation experiments, we show that group-aware policies achieve greater robot navigation performance (e.g., fewer collisions), minimize violation of social norms and discomfort, and reduce the robot's movement impact on pedestrians when compared to the baseline.**

I. Introduction

Mobile robots that are capable of navigating crowded human environments in a safe, efficient, and socially appropriate manner hold promise in bringing practical robotic assistance to a range of applications, including security patrol, emergency response, and parcel delivery. As human movements are fast, dynamic, and following delicate social norms, enabling human-aware robot navigation has been proven to be a challenging task and sparked many interests [4, 17, 13, 18]. While prior works [10, 21, 19, 2, 14, 11] have mainly treated people as individual, independent entities in robot navigation, the majority of people walk in groups [9, 1]; an empirical study showed that up to 70% of pedestrians in a commercial environment walked in groups [15]. It is therefore important that a mobile robot respects human grouping (e.g., not to cut through a social group) during its navigation in a human environment.

In this work, we consider the problem of a robot interacting with dynamic human groups—people walking together in groups—rather than standing groups that are commonly seen in social events (e.g., [16]). While substantial efforts have been made to model and understand dynamic groups (e.g., [15, 23, 3]), how mobile robots should navigate effectively and appropriately around dynamic human groups is under-explored. For example, attention-based DRL has been demonstrated to capture human-human and human-robot interactions in crowded environments [5, 7]. Different from these prior works, we explicitly include group modeling, rather than a simple consideration of pairwise interactions between individuals in a crowd [6]. In addition, our approach uses a more compact representation of group space by computing a polygon based on the convex hull of the pedestrians instead of the F-formation as in prior works [22, 12].

Toward successful robot navigation in crowds of human groups, we propose a learning method that allows the robot to safely reach its desired goal while minimizing impact to individual and groups of pedestrians (Fig. 1). Our contributions include:

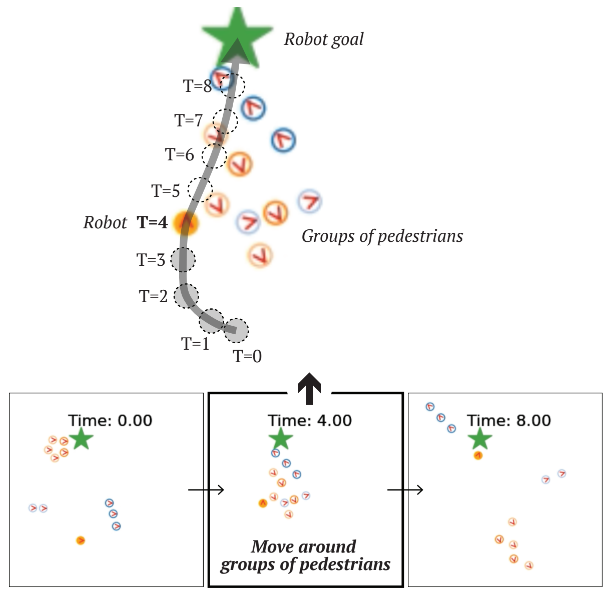

Fig. 1: The objective of this work is to learn a navigation policy that allows the robot to safely reach its goal while minimizing impact to individual and groups of pedestrians.

- A reinforcement learning (RL) algorithm that combines robot navigation performance and group-aware social norms for learning a robust policy;
- A novel reward function that uses the convex hull of a group as the group space to minimize impact to pedestrian groups and improve navigation performance;
- Software extensions to the CrowdNav simulation environment [5] to support social navigation research; and
- Experimental results that demonstrate the efficacy of our learned policy with respect to robot navigation performance, human navigation performance, and maintenance of social norms.

II. Preliminaries

A. Problem Formulation

Our main objective is to learn a controller that allows a robot to navigate to a desired goal while maintaining social norms and avoiding collisions with groups of pedestrians. We formulate our approach using reinforcement learning (RL) to learn a policy to meet the stated objectives. In this form of a Markov decision making process, the robot uses observations to generate a state vector, **S**, and chooses an action, **A**, that maximizes expectation of the future reward, **R**.

The state space, **S**, consists of observable state information for each pedestrian i, represented as $\mathbf{Ped}_i$ as well as internal state of the robot represented as **Rob** as described by Eq. 1. Here, p_x and p_y are the x and y coordinates of the position, v_x and v_y are the x and y coordinates of the velocity, rad is the radius of the pedestrian or the robot, g_x and g_y represent the x and y goal positions, v_pref is the preferred velocity and θ is the turn angle.

$$\begin{aligned} \mathbf{Ped}_i &= [p_x, p_y, v_x, v_y, rad], \\ \mathbf{Rob} &= [p_x, p_y, v_x, v_y, rad, g_x, g_y, v\_pref, \theta], \\ \mathbf{S}_i &= [\mathbf{Ped}_i, \mathbf{Rob}] \end{aligned} \tag{1}$$

B. Group Aware Social Force Model

We developed a custom Python implementation of the extended Social Force Model [1] for the CrowdNav environment using an extended Social Force Model (SFM) proposed by Moussaid et al. [15] to simulate dynamic social groups. In the extended SFM, each individual's motion, as defined in Eq. 2, is driven by a combination of an attractive force $\overrightarrow{f_i}^{des}$ that drives them to a desired goal, the obstacle repulsive forces $\overrightarrow{f_i}^{obs}$, the sum of social repulsive forces from other agents $\sum_j \overrightarrow{f}_{ij}^{social}$, and a new group term $\overrightarrow{f}_i^{group}$ defined by Eq. 3.

$$\frac{d\overrightarrow{v_i}}{dt} = \overrightarrow{f}_i^{des} + \overrightarrow{f}_i^{obs} + \sum_j \overrightarrow{f}_{ij}^{social} + \overrightarrow{f}_i^{group} \tag{2}$$

The group term is defined as the summation of the attractive forces between group members $\overrightarrow{f}_i^{att}$, the repulsive force between group members $\overrightarrow{f}_i^{rep}$, and a gaze force $\overrightarrow{f}_i^{gaze}$ that steers the agent to keep the center of mass of the social group within their vision field to simulate with-in group social interactions:

$$\overrightarrow{f}_i^{group} = \overrightarrow{f}_i^{att} + \overrightarrow{f}_i^{rep} + \overrightarrow{f}_i^{gaze} \tag{3}$$

III. Approach

To evaluate our group aware policy, we extend the existing CrowdNav simulation environment [5] to represent pedestrian motion in groups. We accomplish this by stochastically sampling the number of groups per episode using a Poisson distribution ($\lambda = 1.2$) [8] and then randomly assigning pedestrians to the groups. The average number of groups and group size for five pedestrians are 2.5 and 1.96, respectively. For ten pedestrians, the number of groups and group size increase to 4.9 and 2.0, respectively.

[1]https://github.com/yuxiang-gao/PySocialForce

A. Policy based on Convex Hull of Group

To train the policy, we use a multi-term reward function that encourages the robot to reach its goal while maintaining social norms and avoiding collisions with groups of pedestrians. In particular, we focus on social norms that minimize discomfort to individuals and discourage intersections with a group of pedestrians. Our reward function is given by Eq. 4, where d_{goal} is the distance from the robot to the goal, $d_{\text{coll.}} = 0.6$ is the distance between the centers of entities beneath which a collision is considered to have occurred, d_i is the distance between the robot and pedestrian i, $d_{\text{disc.}} = d_{\text{coll.}} + 0.2$ is the minimum "comfortable" distance between a robot and a pedestrian (as in [5]), and d_j is the distance from the robot to the edge of the convex hull surrounding group j:

$$\begin{aligned} R(t) = & C_{\text{prog.}}(d_{\text{goal}}(t-1) - d_{\text{goal}}(t)) \\ & + C_{\text{goal}}\delta(d_{\text{goal}}(t) < d_{\text{coll.}}) \\ & - C_{\text{disc.}}\sum_i (d_{\text{disc.}} - d_i(t))\delta(d_{\text{coll.}} \le d_i(t) \le d_{\text{disc.}}) \\ & - C_{\text{coll.}}\sum_i \delta(d_i(t) < d_{\text{coll.}}) \\ & - C_{\text{group}}\sum_j \delta(d_j(t) < d_{\text{coll.}}). \end{aligned} \tag{4}$$

The multiple objectives are weighted via the following constants: $C_{\text{prog.}} = 0.1$, $C_{\text{goal}} = 1.0$, $C_{\text{disc.}} = 0.5$, $C_{\text{coll.}} = 0.25$, and $C_{\text{group}} = 0.25$. The first term encourages the robot to progress toward the goal, allowing us to remove the initial imitation learning phase in [5]. The second, third, and fourth terms encourage the robot to reach the goal, avoid close encounters with pedestrians, and avoid collisions, respectively. The last term encourages the robot to adhere to group social norms by penalizing any incursion into a group's "space." To determine the d_j terms, we first compute a polygon representing the convex hull of the positions of all members of the pedestrian group. We then calculate the minimum distance between the robot and the polygon and penalize the robot for intruding into this space.

B. Neural Network Design

Our overall network architecture is depicted in Fig. 2. Our architecture matched that of [5] without the interaction module and (1) with a softmax layer being added to produce a categorical policy output and (2) a single fully-connected layer with 100 neurons connecting to a scalar value head. This configuration enabled actor-critic learning. Our agents were trained using proximal policy optimization (PPO; [20]), a leading model-free, actor-critic approach. Hyperparameters were chosen to mimic those used for Atari in [20], with the exceptions of shorter windows (16 steps) and more windows per batch (64). This change was made to accommodate the shorter episodes of CrowdNav while maintaining the number of experiences per batch.

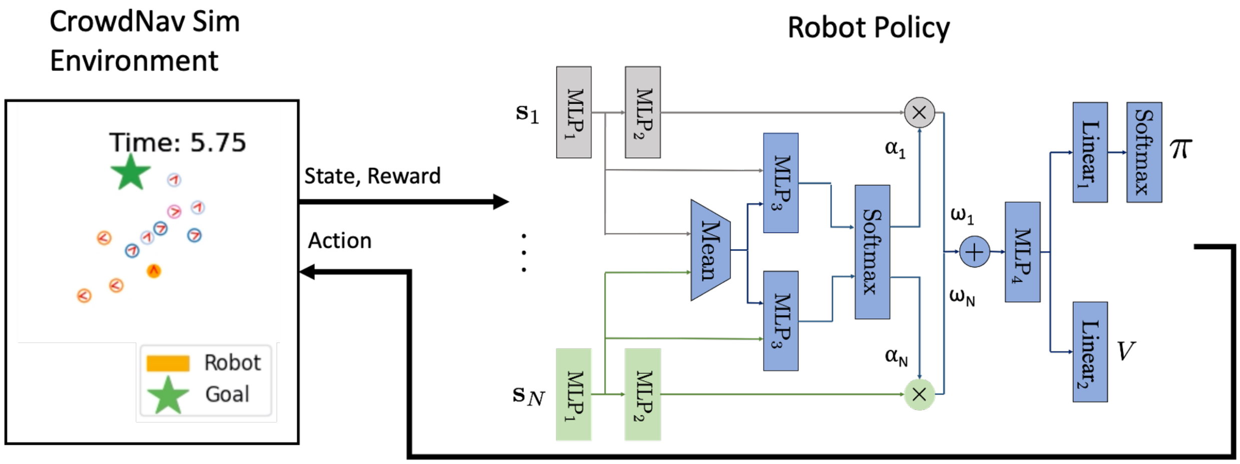

Fig. 2: This is the neural network architecture used for our attention-based, actor-critic policy. The CrowdNav Simulation Environment [5] on the left provides agent states information and the reward to the policy. The pedestrian and robot state vectors are concatenated to represent a pairwise combined state vector; the network outputs the policy π over potential actions and value V of the current state. The gray and green blocks indicate features from individual pedestrians. The blue blocks indicate aggregate features across pedestrians. The argmax of the policy is chosen as the action, which subsequently is sent to the CrowdNav Simulation Environment to control the robot.

Method	**# Groups**	**# Peds.**	**Succ. ↑**	**Ped. Coll. ↓**	**TO ↓**	**Mean Time (s) ↓**		**Mean Robot Vel. (m/s) ↑**		**Mean Ped. Vel. (m/s) ↑**		**Mean Ped. Angle (°) ↓**	
Baseline	1	5	**237**	11	**2**	**8.24**	$t(471) = 9.62$	0.962	$t(498) = 0.60$	1.170	$t(498) = 5.18$	3.76	$t(498) = 1.65$
Group Aware	1	5	236	**9**	5	8.92	$p < .001$	0.964	$p = .551$	**1.183**	$p < .001$	3.59	$p = .100$
Baseline	2.548	5	238	12	0	**8.23**	$t(478) = 7.73$	0.964	$t(498) = 0.51$	1.136	$t(498) = 1.32$	5.99	$t(498) = 3.23$
Group Aware	2.548	5	**242**	**8**	0	8.81	$p < .001$	0.961	$p = .610$	1.146	$p = .186$	**5.59**	$p = .001$
Baseline	1	10	222	23	5	**8.59**	$t(452) = 11.19$	0.955	$t(498) = 0.21$	1.161	$t(498) = 2.06$	4.11	$t(498) = 1.86$
Group Aware	1	10	**232**	**14**	**4**	9.87	$p < .001$	0.956	$p = .833$	**1.174**	$p = .040$	3.93	$p = .064$
Baseline	4.884	10	239	10	1	**8.72**	$t(478) = 18.83$	**0.960**	$t(498) = 6.39$	1.089	$t(498) = 3.53$	8.09	$t(498) = 8.41$
Group Aware	4.884	10	**241**	**9**	**0**	10.21	$p < .001$	0.918	$p < .001$	**1.108**	$p < .001$	**7.07**	$p < .001$

TABLE I: This table summarizes the pedestrian and robot navigation performance across 5 and 10 pedestrians. Bold text indicates statistically significant results. We show that our group aware policy is able to achieve comparable or better robot navigation performance while allowing pedestrians to achieve faster velocities with less deviation from their desired goal.

Method	**Num. Groups**	**Num. Peds.**	**Group Intersections ↓**	**Individual Discomfort ↓**		**Ped. Social Force ↓**		**Robot Social Force ↓**	
Baseline	1	5	143	3.10	$t(498) = 3.48$	0.375	$t(498) = 3.43$	0.523	$t(498) = 1.95$
Group Aware	1	5	**15**	**1.29**	$p < .001$	**0.351**	$p < .001$	0.482	$p = .051$
Baseline	2.548	5	151	2.87	$t(498) = 0.49$	0.522	$t(498) = 3.95$	0.716	$t(498) = 2.78$
Group Aware	2.548	5	**22**	2.63	$p = .625$	**0.485**	$p < .001$	**0.657**	$p = .006$
Baseline	1	10	176	4.20	$t(498) = 2.92$	0.395	$t(498) = 3.99$	0.707	$t(498) = 3.85$
Group Aware	1	10	**29**	**2.31**	$p = .004$	**0.366**	$p < .001$	**0.597**	$p < .001$
Baseline	4.884	10	258	4.94	$t(498) = 4.99$	0.681	$t(498) = 7.32$	0.964	$t(498) = 4.80$
Group Aware	4.884	10	**20**	**2.29**	$p < .001$	**0.599**	$p < .001$	**0.849**	$p < .001$

TABLE II: This table summarizes metrics associated with social compliance across 5 and 10 pedestrians. Bold text indicates statistically significant results. We show our group aware policy leads to more socially compliant navigation indicated by fewer instances of group intersection while reducing individual discomfort and overall social forces on the pedestrians and the robot.

IV. Experimental Evaluation

A. *Experimental Setup*

The goal of our experiments is to assess the efficacy of our group-aware navigation policy. Our experiments involved four settings determined by two factors: the number of pedestrians and the number of groups. We explored both 5- and 10-person settings as well as a single group and a stochastic number of groups (Sec. III). We used the Circle Crossing scenario where groups of pedestrians started and ended around the perimeter of a circle (radius = 4 m) during training and evaluation. We evaluated our trained policy on 250 trials with randomly initialized starting and ending pedestrian positions for the four settings. Lastly, our comparison baseline was based on [5], without inclusion of the group-aware reward term.

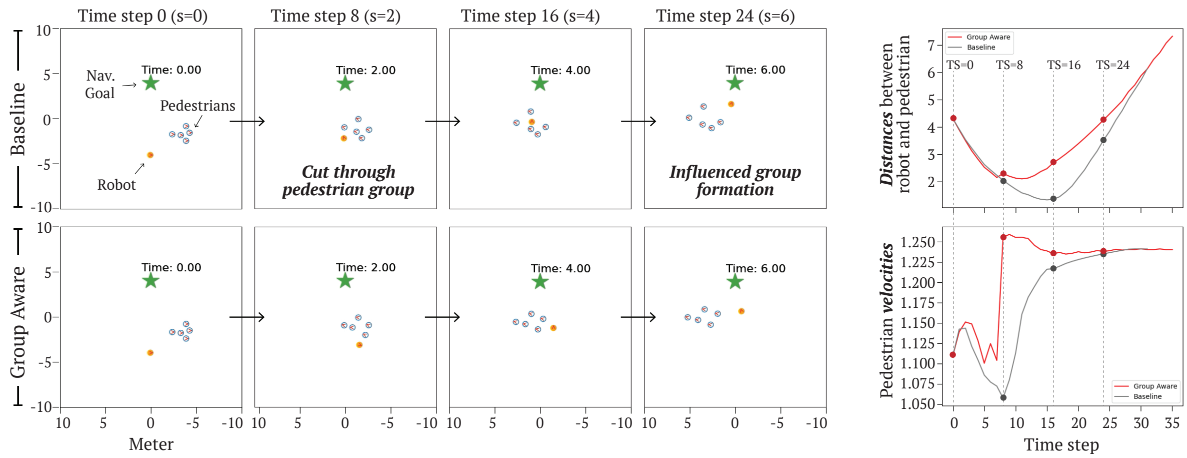

Fig. 3: The left figures give some representative examples of the robot navigating through the crowd of pedestrians over time using both the baseline and the group aware policy. The baseline policy chooses actions that cut through the group of pedestrians and influences the group formation, while the group aware policy chooses actions that move around the group with minimal disturbance. The figure on the right shows the average distance between the pedestrian and the robot (top) and the average pedestrian velocity (bottom) over time. Here, we show the group aware policy results in increased distance to the pedestrians while allowing the pedestrians to maintain faster speeds.

B. Metrics

Our evaluation was focused on 1) robot navigation performance, 2) pedestrian navigation performance, and 3) social compliance. For robot navigation performance, our metrics represent the quality of the robot's ability to navigate to the goal quickly without collision. Specifically, we measured the number of success, collisions, timeouts, the average time to goal, and the average velocity of the robots.

To assess pedestrian performance, we measured the impact of the robot's navigation behavior on the desired pedestrian motion. Specifically, we measured the average velocity the pedestrians and the average angular deviation between the pedestrian's observed motion and the direct vector to the pedestrian's goal, which reflects the pedestrians' disturbance from the optimal trajectory to the goal caused by the robot.

Finally, to assess social norms, we quantified how the robot maintained social distance among individual pedestrians and limited intersections with groups of pedestrians. For this, we considered the number of groups intersected by the robot, the mean social force applied to each pedestrian, the average social force applied to the robot, and the*individual discomfort* caused by the presence of the robot, which is defined as the mean distance between the robot and the pedestrians aggregated over all pedestrians when the robot violates the discomfort threshold (0.2 m).

C. Results

We conducted independent two-tailed t-tests to compare our group-aware and the baseline policies. For all the statistical tests, we used an α level of .05 ($p < .05$) for significance. Table I summarizes the robot and pedestrian navigation performance as well as their corresponding statistical test results. Overall, the group aware policy generally led to higher number of successful trials, while allowing the pedestrians to travel at faster speeds with less disturbance towards the goal.

In Table II, we summarize the social compliance results with their corresponding statistical test results. As indicated by this table, the group aware policy yielded an 88% improvement in reducing the number of instances where the robot navigated through a group. Moreover, the group aware policy resulted in a 43% reduction in individual discomfort. In addition, we observe that the group aware policy improved the overall social forces applied to the pedestrians and robot.

We note that the robot with the group aware policy took longer to reach the goal. Fig. 3 illustrates an example of such behavior. The resulted group-aware behavior ultimately enabled greater group cohesion and less disruption while improving group and individual discomfort.

V. Discussion

This paper explores group-aware behaviors that respect pedestrian group formations and trajectories while minimally sacrificing robot navigation performance. Our results show that the learned policy is able to achieve higher number of successful trials, fewer collisions, and less impact to the pedestrian's motion towards their goal. Additionally, our learned policy not only reduced the number of group violation but also decreased the individual discomfort and social forces applied to the pedestrians and robot. Our approach, however, resulted in an increase of the robot's total time to goal when compared to the baseline. This increase was expected as the robot sought to move around groups as opposed to navigate through them (Fig. 3). However, our results show that even though the total time to goal increased, the average velocity of the robot was mostly unaffected by the group aware policy.

Our exploration indicates several directions of future research. First, we would like to determine how well our learned policy reflects actual human motion. Second, we would like to investigate whether we can bootstrap our learned policy with imitation learning using observations of humans navigating groups of pedestrians. Third, we would like to investigate different representations of group space beyond the convex hull approach described in this paper. We speculate that considering additional parameters, such as social interaction during movement, the specific formation of the group, and environmental cues (e.g., social space), may contribute to learning more socially compliant navigation policies.

ACKNOWLEDGMENTS

This work was supported by the Johns Hopkins University Institute for Assured Autonomy.

REFERENCES

[1] Adrian F Aveni. The not-so-lonely crowd: Friendship groups in collective behavior. *Sociometry*, pages 96–99, 1977.

[2] Aniket Bera, Tanmay Randhavane, Rohan Prinja, and Dinesh Manocha. Sociosense: Robot navigation amongst pedestrians with social and psychological constraints. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7018–7025. IEEE, 2017.

[3] Niccoló Bisagno, Bo Zhang, and Nicola Conci. Group lstm: Group trajectory prediction in crowded scenarios. In *Proceedings of the European conference on computer vision (ECCV)*, pages 0–0, 2018.

[4] Konstantinos Charalampous, Ioannis Kostavelis, and Antonios Gasteratos. Recent trends in social aware robot navigation: A survey. *Robotics and Autonomous Systems*, 93:85–104, 2017.

[5] C. Chen, Yuejiang Liu, S. Kreiss, and Alexandre Alahi. Crowd-robot interaction: Crowd-aware robot navigation with attention-based deep reinforcement learning. *2019 International Conference on Robotics and Automation (ICRA)*, pages 6015–6022, 2019.

[6] Changan Chen, Sha Hu, Payam Nikdel, Greg Mori, and Manolis Savva. Relational graph learning for crowd navigation. *arXiv preprint arXiv:1909.13165*, 2019.

[7] Yu Fan Chen, Miao Liu, Michael Everett, and Jonathan P How. Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning. In *2017 IEEE international conference on robotics and automation (ICRA)*, pages 285–292. IEEE, 2017.

[8] James S Coleman and John James. The equilibrium size distribution of freely-forming groups. *Sociometry*, 24(1): 36–45, 1961.

[9] Marco Costa. Interpersonal distances in group walking. *Journal of Nonverbal Behavior*, 34(1):15–26, 2010.

[10] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. The dynamic window approach to collision avoidance. *IEEE Robotics & Automation Magazine*, 4(1):23–33, 1997.

[11] Kapil D Katyal, Gregory D Hager, and Chien-Ming Huang. Intent-aware pedestrian prediction for adaptive crowd navigation. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3277–3283. IEEE, 2020.

[12] Adam Kendon. *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University Press, Cambridge, U.K., 1990.

[13] Thibault Kruse, Amit Kumar Pandey, Rachid Alami, and Alexandra Kirsch. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems*, 61(12):1726–1743, 2013.

[14] Ross Mead and Maja J Matarić. Autonomous human–robot proxemics: socially aware navigation based on interaction potential. *Autonomous Robots*, 41(5):1189–1201, 2017.

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

[16] Billy Okal and Kai O Arras. Learning socially normative robot navigation behaviors with bayesian inverse reinforcement learning. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2889–2895. IEEE, 2016.

[17] Anish Pandey, Shalini Pandey, and DR Parhi. Mobile robot navigation and obstacle avoidance techniques: A review. *Int Rob Auto J*, 2(3):00022, 2017.

[18] Jorge Rios-Martinez, Anne Spalanzani, and Christian Laugier. From proxemics theory to socially-aware navigation: A survey. *International Journal of Social Robotics*, 7(2):137–153, 2015.

[19] Alexandre Robicquet, Amir Sadeghian, Alexandre Alahi, and Silvio Savarese. Learning social etiquette: Human trajectory understanding in crowded scenes. In *European conference on computer vision*, pages 549–565. Springer, 2016.

[20] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347, 2017.

[21] Pete Trautman, Jeremy Ma, Richard M Murray, and Andreas Krause. Robot navigation in dense human crowds: Statistical models and experimental studies of human–robot cooperation. *The International Journal of Robotics Research*, 34(3):335–356, 2015.

[22] F. Yang and C. Peters. Social-aware navigation in crowds with static and dynamic groups. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pages 1–4, 2019.

[23] Zeynep Yücel, Francesco Zanlungo, Tetsushi Ikeda, Takahiro Miyashita, and Norihiro Hagita. Deciphering the crowd: Modeling and identification of pedestrian group motion. *Sensors*, 13(1):875–897, 2013.