Communication Strategy for Efficient Guidance Providing

Domain-structure Awareness, Performance Trade-offs, and Value of Future Observations

Shih-Yun Lo and Andrea L. Thomaz
the University of Texas at Austin
yunl@utexas.edu, athomaz@ece.utexas.edu

Abstract— **Service robots are gaining capabilities to be deployed in public environments for human assistance. While robot actively providing guidance has shown great success in field study, the communication strategy (the strategy to decide whom to initiate the service for and when), and hence the performance evaluation, has been based on behavioral-based qualitative analysis. We attribute this to the challenge of accessing large-scale field data with condition control, and approach the problem with simulation from the agent-based modeling literature, to simulate pedestrian behavior in unfamiliar environments and estimate travel cost. We contribute a planning approach that uses the pedestrian behavior prediction from the model, to decide whom to initiate guidance and when for performance maximization. The results suggest that our approach is more efficient based on the measure of saved pedestrian travel time, compared to the behavioral-based strategy and a baseline that maximizes service counts.**

I. Introduction

For robots in the field to successfully initiate interaction with humans, social context awareness has been widely studied to improve the success rate, e.g., not to disturb goal-oriented pedestrians [1] and not to talk about other shops in front of a potentially competing shop [2]. For guidance service, from the identification of potential subjects [3][4], proper initiative motions to approach [5][6] and to the communication process [7][4], past research showed increasing success for robots to initiate guidance in public areas, with positive subjective evaluation of service quality [3].

However, quantitative measurement of robot service quality is relatively less addressed, especially in an objective manner. Such measurement is important for robot capacity management, and is also important for the service provider (robot) to calculate "service value" - the underlying value of each service to the service customer (human). Since there is a cost for each service providing process (time to communicate for goal clarification and activity invite, and task interruption) that is at both service provider and service customer's expense, if the service provider is able to make the decision to initiate service only when the calculated service value is worth the cost, the overall service quality can be improved. Furthermore, for guidance service, since neither the customer's destination nor her knowledge of the environment is publicly known, the calculated service value may be highly inaccurate. A decision to initiate guidance may soon be found less valuable once the customer turns and finds her destination at the nearest site, as shown in Fig. 1.

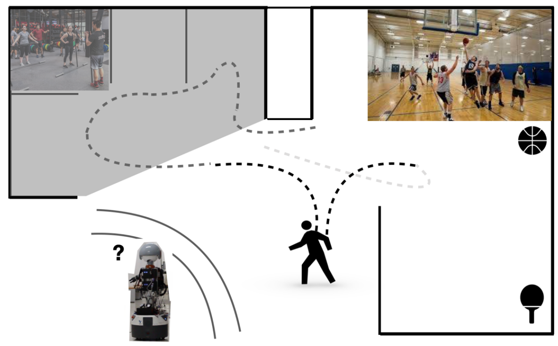

Fig. 1: The robot evaluates pedestrian potential routes (marked in dashed lines) and destinations to decide if it should initiate guidance service. As it knows that the group session (upper-left) is under restricted access and the configuration is complex to search around, an (indecisive) pedestrian reaching to the left has high probability (in dark gray) spending long time not finding her destination. If she reaches to the right, the expected search time is short; with low probability (in light gray) she will need to search another area.

In this work, we propose 1) to utilize the agent-based simulation literature to sample pedestrian travel behaviors for performance evaluation, and contribute 2) a method that uses the agent-based model for cost-performance estimation, to generate domain-structure aware communication strategies, to decide if to initiate service or to collect more observations. To afford the computation for long-horizon prediction, the contributed algorithm decouples the prediction and action selection process, to make the search complexity polynomial in agent number and horizon length, and then iteratively improves plan quality.

The validation simulates continual pedestrian flows in real-world collected maps. We also consider a cost for task switching, e.g., to interrupt public-area sanitizing and resume with calibration to continue, for performance validation in coordination with other tasks. We show superior performance in improved pedestrian travel efficiency while initiating less number of guidance service and therefore more undeployed time for other task capacity, compared to baseline strategies in initiative guidance and intent communication.

*This work was supported by the US NSF (IIS-1564080, IIS-1724157) and US ONR (N000141612835, N000141612785).

II. RELATED WORK

Communication plays an essential role in interaction. Utterance, in combination with other communication channels, e.g., gesture, has been widely studied to produce socially competent robot behaviors to interact with humans [8][9][10]. Technical advances in dialog systems, e.g., online clarification strategies to deal with noisy inputs [11][12] and knowledge groundings [13][14], facilitate smooth verbal communication.

Guidance, as a service that robots supply, requires the ability to provide human-understandable instruction, e.g., through landmarks and local directions [9][15]. To initiate interaction, robot behaviors that reveal the intent have been proposed, through approaching motions [16][17], adjusting relative orientations [18], and eye contacts [3]. Initiative guidance to pedestrians with "indecisive" behavior features, which are learnt from field data, showed high success rate [4]; such service was also suggested in field study as useful and preferable in shopping malls [7][3].

As a communication task, to provide information that complements the travelers' knowledge of the environment, guidance is expected to improve travelers' efficiency, by means of communicative actions. While guidance service received positive subjective evaluation, due to the difficulty to conduct objective measures and condition control in field study, performance analysis is relatively less addressed. Such analysis is however crucial for robot task management, as robots increasingly gain capabilities, to evaluate the task-switching trade-off based on the expected outcome. While prior literature focused the guidance robots on guidance task, either prepared the robot as always available [18], or to initiate service to whoever appears indecisive and closest [16][3], such strategy is validated in this work as inefficient to robot capacity usage and ineffective in its assistance to travellers, in terms of their travel efficiency.

From the perspective of task performance, what to communicate, when, and how, have been formulated as a sequential optimization process in the human-robot teamwork literature [19][20], which is suitable for guidance and service scheduling applications. We build upon such multi-agent planning formulation, and propose our solutions to deal with long-horizon pedestrian trajectory prediction and partial observability in the initiative service setting.

III. PROBLEM FORMULATION

For interactions to take place in the wild, there are a few prerequisites. The initiation process requires attention and can be time- and effort-consuming. A joint activity, θ_{jnt}, parametrized by task specifications including the objective and constraints, is defined when agents under their individual objectives can act and affect the (foreseeable) future values of others' actions. The *awareness of potential joint activity* with others is needed and to be grounded, such that the engaged agents coordinate and expect others to coordinate. *Awareness grounding* can be completed by simply ensuring mutual awareness, e.g., through eye contact to avoid collision, or may require more steps, to confirm about each agent's own objective, the joint activity (among multiple options $\theta_{jnt} \in \Theta_{jnt}$), and the subtasks accordingly.

Depending on the individual objective and potential subtasks, preparative behaviors to initiate interaction were proposed in the human-robot interaction literature, e.g., to actively acquire attention [16], signal availability [2], and prepare for future coordinating actions [6], and they are at the expense of the robot's own ongoing task progress; we denote such cost as $C^R_{init,P}$, and the cost of joint activity grounding among both human and robot as $C^{HR}_{init,J}$.

A. Expected Deployment Value

To manage the robot's capacity to effectively provide service, we consider the overall initiation cost, denoted as C_{init}, as a soft constraint, $C_{init}(x^R_t, x^H_t, a^R_t, a^H_t) < C_{MAX}$, for the robot to coordinate its tasks in the queue (in a preference-based manner), while maximizing the expected performance to the service customer,

$$a^R_t = \underset{a^R_t}{\operatorname{argmax}} \mathbb{E}_{\theta_{jnt}}[V^{R|\pi^H_{jnt}}(x^R_t, x^H_t, a^R_t, a^H_t|\theta_{jnt}) - V^{\pi^H}(x^H_t)]. \tag{1}$$

Before joint activity grounding, the individual acts according to their single-agent policy π^H, with expected value $V^{\pi^H}(x^H_t)$. After joint activity grounding θ_{jnt}, the participant acts according to $a^H_t \sim \pi^H_{jnt}(x^H_t, x^R_t)$. The robot maximizes its action value $V^{R|\pi^H_{jnt}}$ given human policy, conditioned on the joint activity θ_{jnt}. Taking into account the initiative cost C_{int}, joint performance may or may not outperform the single-agent policy; to initiate or not then depends on the value function, which is domain-structure dependent and service-activity dependent.

For guidance service initiation, θ_{jnt} specifies the human subject's objective θ^H: to guide towards her destination (unknown to the robot) following an efficient route. The robot's subtask is to provide such route information, and the value to provide such information depends on the subject's prior domain knowledge, e.g., their cognitive map and its annotation. We denote such partial knowledge as their belief b^H_t, which guides their route to either explore or exploit: $a^H_t \sim \pi^H(x^H_t, b^H_t|\theta^H)$. b^H_t evolves over time as information is collected, e.g. through map exploration and explicit instructions, $b^H_{t+1} \sim f(b^H_t, a^H_t)$. As b^H_t transits and contains sufficient route information for θ^H, π^H is modeled to efficiently navigate to the destination. The resultant value V^{π^H} of self exploration determines the benefit to initiate guidance, which we denote as $V^{R|\pi^H}$:

$$V^{R|\pi^H} = \max_{a^R_t} \mathbb{E}_{\theta_H} \mathbb{E}_{b^H_t|\theta_H}[V^{R|\pi^H}(x_t, b^H_t, a_t|\theta_H) - V^{\pi^H}(b^H_t|\theta_H)]. \tag{2}$$

B. Expected Observations and their Impact on Decisions

To provide service that is effective, as described above, we evaluate service value considering potential joint activity specifications, and compare to the performance of the single-agent policy to decide if to initiate service. Intuitively, with more observations, the more accurately the robot can identify customer private states and evaluate its service value, to prevent inefficiency caused by uncertainty, e.g. to deploy to provide guidance but later realize the subject is heading for the most close-by destination. We therefore further incorporate observations along with the robot's sequential action evaluation, to plan for initiation taking into account the impact of potential observations $o_{0:t} \in O$ up to a certain time t. The policy is then conditioned on the potential future observations and the value is evaluated conditioned on those observations. Given time $t \geq 0$, it is then to decide if pending for more observations $o_{0:t} \in O$ yields better expected performance to the observation-conditioned policy, deciding whether to initiate, and when:

$$V_t^{R|\pi^H} = \max_{a_t^R} \mathbb{E}_{\theta_H} \mathbb{E}_{o_{0:t}|\theta_H} \mathbb{E}_{b_t^H|\theta_H, o_{0:t}} [V^{R|\pi^H}(x_t^R, b_t^H, a_t^R, a_t^H|\theta_H) - V^{\pi^H}(b_t^H|\theta_H)]. \tag{3}$$

By comparing $V_t^{R|\pi^H}$ over time $t \geq 0$, the robot then decides if to defer the initiation decision to a future time t. We later denote this policy as π-$V_t^{R|\pi^H}$.

C. Hidden Parameters, Observability, Expected Performance

In guidance service, θ_H and b_t^H influence human behaviors and therefore service value, we are then concerned with the inference and observability of θ_H and b_t^H, and to solve for Eq.3. The robot may partially observe θ_H and b_t^H through human behavioral patterns, e.g., waiting at a corner, wandering around, or going straight to a direction. Here we consider the pedestrian's navigation state as the source of observations to infer $\theta_H \in \Theta_H$ and $b_t^H \in B^H$, which we truncate as $\theta_{guide,t}^H \in \Theta_H \times B^H$. As we assume x_t^H transition to be markovian, we maintain robot belief b_t^R over $\theta_{guide,t}^H$ and update through Bayes' rule:

$$b_{t+1}^R = \eta \sum_{\theta_{guide,t+1}^{H-} \in \Theta_{guide}^H} \Omega(o_t|\theta_{guide,t+1}^{H-}) b_t^R(\theta_{guide,t+1}^{H-}), \tag{4}$$

where η is a normalizing constant. θ_H is assumed time-invariant; b_t^H is updated throughout x_t^H transition over time and incorporated as: $\theta_{guide,t+1}^{H-} = g(\theta_{guide,t}^H, x_t^H, a_t^H)$. $x_{t+1}^H = \mathcal{T}(x_t^H, a_t^H)$ follows a first-order motion model.

Per Eq. 3, the expected observability to the hidden parameters affects the robot's decision: if to initiate service to one customer now, and to decide later. Consider a robot serving in a one-way one-outlet corridor, past which all pedestrians reach their later-on destinations. The observability of the pedestrian's goal and route certainty is limited in this case. Passive observations are then not informative, and therefore Eq. 3 would perform similarly to Eq. 2. Such decision is then not only dependent on how strong the state detector Ω is, but also on the domain structure.

In summary, we evaluate the expected decision quality $V_t^{R|\pi^H}$ over $\theta_{guide,t}^H$, conditioning on potential future observations. The robot therefore may defer its decision on whether to initiate, if decision quality improvement is expected, $V_t^{R|\pi^H} > V_0^{R|\pi^H}$. This is of common practice in human interaction, e.g., a stadium receptionist initiates help after a person went past common facilities yet still appeared searching. Yet, the marginal value to provide service decreases as initiation timing is delayed. The robot is then to assess the trade-off bwtween preparation overhead $C_{init,P}^R$, interaction joint cost $C_{init,J}^{HR}$, and expected service value. The higher $C_{init,P}^R$ and $C_{init,J}^{HR}$ are, e.g., with its task queue tightly scheduled and complex domain knowledge to clarify before providing service, the robot more tend towards waiting until service appears critical to initiate.

IV. METHODOLOGY

Here we first detail π^H, our choice of model for V^{π^H} evaluation, and function g of pedestrian belief transition during their exploration. We then detail our planning technique and action space to solve for Eq. 3.

A. Modeling: Human Exploratory Navigation

In the literature of cognitive psychology, Gibson's ecological theory of perception was formulated to model an agent conjoined with its environment; by the contents *perceived*, an agent uses the affordances to guide its action, here our π^H. Natural vision was proposed to characterize pedestrian visual factor affecting behaviors, moving in a direction that provides the potential of further movement: "we look around" and "walk up to something of interest" [21].

1) Isovist: The idea of isovist was introduced in architecture to calculate viewable areas at a location, and was further applied in ecological psychology, in assessment of walkable surface that *affords* movements. The theory of natural movement [22] suggested that the majority of pedestrian movements occur along lines of sight, and the more a line of sight is connected with others, the more movement exists along. Turner et.al applied the concept of natural movement in agent-base modeling [23], and simulated exploratory pedestrian motions based on the isovist; the farther the pedestrian can see along a perceivable direction, computed based on range of view and current walking direction, the higher probability she will be guided, visually, to move towards.

2) Search-based exploratory behavior: Pedestrian wayfinding in unfamiliar environments was also studied for disaster simulation. Search heuristics and information from the environment, e.g. signs and information boards, were incorporated for strategic wayfinding [24], [25], [26]. As for the times when no clear direction information is available, people explore the environment. Here we consider such situation and model the search process as if s/he has to follow the biological instinct of walking during the exploration. During the process they search for their destination, at the same time for potential visual attraction, if not for urgent purposes. We refer to such performance as search-based exploratory behavior, and target this behavior for robot guidance service initiation.

We follow the design in Turner et.al. to simulate search-based exploratory behavior: a pedestrian's visible range is applied to cover 170 degree wide along their walking direction; each step takes the duration of 0.5s, with a probability to

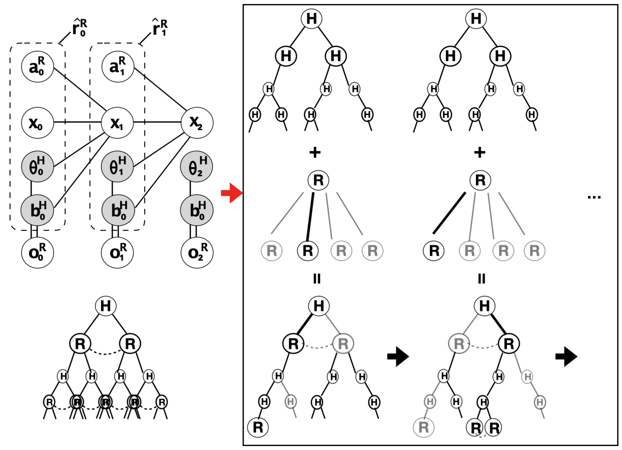

Fig. 2: [Top-left] The graphical model of the multi-agent decision making formulation, to plan with partial observability to human hidden parameters. [Bottom-left] The full search tree, v.s. [Right] the sparse and decoupled tree, separating robot actions from human actions, for robot action search and iterated evaluation.

change walking direction following the Poisson distribution $Poisson(\lambda)$, $\lambda = \frac{1}{3}$, validated in a gallery tour dataset. If to change walking direction, a new direction is sampled, with the probability proportional to the viewable surface area along that direction. Example trajectories are shown in Fig. 3, marked in solid lines from designated entrances to goals. The path colors fade over time.

3) Pedestrian belief transition: The cognitive map of a pedestrian in construction of routes in the perceived environment is of active research [27], [28]. Here we apply the belief representation b_t^H upon a map-labeling mechanism, to model cognitive map construction in unfamiliar environments. Through the computation of isovist and visible angle, a distance of $5m$ is applied as the range of survey and labeling. Upon the assumption that labeled/surveyed sub-areas as well as their spatial connections were constructed along the process, we assume an agent to know their route once labels containing their destination and subgoals (from current location) are available. We therefore implement wayfinding as a map labeling process; route is found once the destination is labeled either through self exploration or robot guidance.

B. Planning to Initiate Guidance with Partial Observability to Pedestrian Hidden Parameters

The problem formulation in Eq. 3 involves the robot's trade-off between exploration and exploitation, based on the action's long-term value. Therefore, belief planning serves well as our solution basis, interleaving decision value with expected information gain. In belief planning, solution complexity is exponential to the search horizon H, based by the action space A and observation space O. For our application, long-horizon rollouts are needed for wayfinding performance evaluation. Search complexity then becomes a challenge for real-time computation. Past research using samples for belief representation were effective in reducing the computational growth based by observation space [29], [30]; yet, the search grows exponentially by action space. We therefore contribute a solution to deal with long-horizon search by decoupling actual information-gathering (navigational) actions from the planning of communication actions, and iteratively updating optimal plan value till convergence.

1) Decoupled information-gathering from search: Since $o_{0:t}$ are concerned with *passive* observations to pedestrian motions for belief update, the robot's information-gathering actions do not affect the pedestrian's state transition [1]. We therefore disassociate the observation collection process from the communication-action planning, and divide the planning into two phases, in a backward-forward fashion: 1. solving for Eq. 3, using only communication actions to plan the guidance initiation timing (and potential targets), while considering potential future observations $o_{0:t}$ assuming "perfect visibility" to collect observations; and then, given planned initiation timings to potential targets (if any, otherwise stay idle for other task assignment), 2. to plan navigation actions, to physically maintain visibility to target pedestrians to collect more observations (that potentially improve decision quality), while meeting the spatial preconditions to initiate interaction at planned timings. With updated navigation actions and the true collected observations along the navigation process, the estimated guidance value in Eq. 3 should be updated, as shown in Fig. 2. This process should iterate, to make sure the observations which impact the belief update (and therefore the optimal expected guidance value) are indeed measurable given navigation actions, and the updated navigation costs are worth it. This search is initialized with an optimistic value estimate with perfect observability; with an Euclidean-distance heuristic initialized for navigation costs, the overall hybrid search is admissible.

2) Communication action abstraction: In the belief planning process, we consider abstracted communication action representation for guidance: the action has the expected effect to update the subject's cognitive map, whose policy transits to efficient route-following behavior towards the destination. Researchers have demonstrated such action effects through modalities of verbal commands and gestures in field studies [9]; the behaviors include those to initiate interaction [16], [18], which we refer to for detailed behavior design.

Here we apply an estimated lump-summed value for $C^{HR}_{init,J}$, 20s, and same for $C^{R}_{init,P}$, for performance demonstration, and apply a 2-horizon search for communication planning. One action then takes 20s. The robot is then to evaluate who to *first* initiate guidance to maximize service value, considering service costs, hidden parameters and their observability along the planned horizons.

3) Fast convergence: Here we omit the computation of exact navigation costs for plan evaluation. This simplification saves further iteration over initiation decisions given navigation cost updates, since a lump sum preparative cost is applied; and the optimal plan converges once the optimal value is updated given measurable observations and no other plan outperforms with its optimistic value estimate (assuming perfect observability).

[1]The active observation action which directly clarifies hidden pedestrian states (here θ_H and b_t^H) is only considered along the communication process to provide accurate guidance, but not solely.

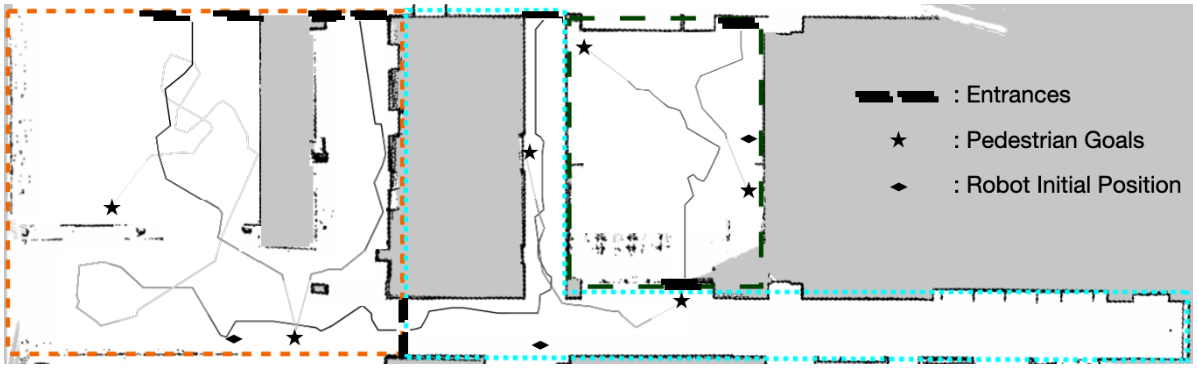

Fig. 3: Experiment domains collected in university public areas, featuring: rooms (framed in dark green), corridors (light blue), and open areas (orange) with divided spaces. Pedestrian inflows from 2-3 entrances over 1000s of test time are sampled over Poisson distribution.

C. Online Performance

1) Search complexity: Due to the decoupled planning mechanism, for navigation action planning, the (worst-case) exponential growth in action space is up to a time t, the guidance initiation timing, but not the full time horizon, H. The belief planning applies rollouts to time H, when the pedestrian exploratory navigation ends; during the process, belief b^R in $B_R = \triangle\Theta_H \times B_H$ is updated. Since the decision on whether to initiate is of one-time actions, the computation is linear to H, throughout the search process of initiation timing $t < H$. Here we consider a fixed number of sampled observations to maintain, K, reducing the exponential growth to linear growth. The overall computation of the long-horizon planning problem is reduced to $|A|^t + KH|B_R|$, $t < H$. We choose $K = 10$ to estimate pedestrian route choices. The navigation action space $|A|$ includes a static action (to not move), and 3 actions of [-15, 0, 15] degrees of angular rotation, sampled from the previous moving direction.

2) Pedestrian state detector and belief update: we here detect inconsistent walking direction and deviation from potential goals, to identify pedestrian 1) goal, and 2) certainty about their goal location. The detector of certainty about goal location is conditioned on the goal location, $o_t \sim \Omega^c(x_t^H|goal, c)$. c represents the state of certainty. This detector is well studied for robot navigation in human environment [5], [31]. Here we use a velocity-based detector: $p(o_t|goal, c) \propto exp(-\beta(\frac{v_t^H}{|v_t^H|} - \frac{v_t^{goal}}{|v_t^{goal}|}))$, where v_t^{goal} is the goal-driven velocity at current location and v_t^H is the pedestrian's velocity. As velocity estimate is noisy in real-time, we filter position tracking data, at 10hz per laser scan rate, and estimate velocity at 2hz. This also matches human gait frequency, taking two steps in 1 sec on average.

For the state of uncertainty, $o_t \sim \Omega^{-c}(x_t^H|-c)$, we detect based on velocity direction changes over a short horizon h: if velocity remains unchanged, $|std(v_{t:t+h}^H) < \delta|$, the probability of being uncertain is exponential to h, based by $Poisson(\lambda), k = 1$: $p(o_t|-c) = exp(\frac{-1}{\lambda})^h$. We apply $h = 3$, collecting 3 velocity estimates for 1 pedestrian state detection. With the state detector run at 2/h hz, the belief update can be run at 2/3hz with:

$$p(goal, c|o_t) \propto \Omega^c(o_t|goal, c)p(goal, c),$$

and

$$p(goal, -c|o_t) \propto \Omega^{-c}(o_t|-c)p(goal, -c).$$

3) Replanning: As the belief update can be run at 2/3hz, prediction and replanning can also be updated at 2/3hz: to execute navigational actions $t = 0s$ up to $t = 1.5s$, update belief at $t = 1.5s$, replan, execute the new plan's actions up to $1.5s$, and repeat, following a receding-horizon fashion.

V. Experiment

We validate our proposed approach π-$V_t^{R|\pi^H}$ in comparison to the following baselines:

1) π-Now: to maximize service value by initiating right away. This baseline outperforms all when the customer is in need of guidance (with positive benefit compared to self exploration). The design follows the fashion in legible motion for intent communication [32], which encodes a "the earlier the better" implementation in the sequential optimization formulation.

2) π-UncertaintyDetect: to initiate guidance when pedestrian uncertainty over a probability threshold is detected. We choose the value to be 0.9, such that an "obvious" back-and-forth motion is detected by 4.5s, and an inefficient wandering motion in wide areas (rooms and open spaces) is detected by 7.5s. This baseline implements the behavior control approach in the literature of robot initiative interaction, to initiate when featured behaviors of the target activity are detected [7].

3) π-$V_0^{R|\pi^H}$: to initiate service when the expected value is higher than not to, based on Eq. 2. It adopts the design in Lo et.al. [19] to evaluate information value for communicative behavior generation, and incorporates it along with the belief operation to compute service value in expectation.

We evaluate the performance in real-world-collected indoor environments, shown in Fig. 3, considering three common structures: rooms (domain I), corridors (domain II), and open areas with divided spaces (domain III). We consider 3 pedestrian inflow rates: on average 1 entry every 40s, 20s, and 10s, sampled from Poisson distribution. Each domain is associated with 2-3 entrances, with pedestrian entries randomly assigned. All entries are initialized as unfamiliar with the domain, sampled from the exploratory policy π^H (independent from those sampled for robot planning). For each entry, the robot initializes even prior on the joint distribution over her goal certainty and goal. This way, relatively

	Pedestrian inflow rates in domain I			domain II			domain III		
	$0.025/s$	$0.05/s$	$0.1/s$	$0.025/s$	$0.05/s$	$0.1/s$	$0.025/s$	$0.05/s$	$0.1/s$
π-Now	27s [17]	79s [26]	132s [36]	-46s [14]	135s [26]	168s [32]	169s [20]	546s [28]	353s [31]
π-UncertaintyDetect	85s [13]	175s [22]	306s [24]	121s [5]	204s [19]	254s [11]	198s [14]	551s [22]	397s [17]
π-$V_0^{R\|\pi^H}$	78s [14]	233s [24]	270s [27]	90s [8]	325s [17]	515s [19]	240s [15]	597s [25]	878s [19]
π-$V_t^{R\|\pi^H}$	**149**s [10]	**261**s [19]	**510**s [20]	**247**s [6]	**394**s [16]	**780**s [12]	**289**s [10]	**641**s [19]	**916**s [19]

TABLE I: Measures on saved pedestrian travel time(s) and [number of guided pedestrians] in domain I, II and III.

conservative behavior and performance are reported. Given the above controlled conditions, results in accumulated saved pedestrian travel time (s) and number of reached pedestrians to provide guidance are reported over 1000s of experiment time, shown in Table. I. Number of attempted robot task switches and undeployed time (s) in domain III are reported in Table. II for discussion.

A. *Quantitative Analysis*

I) Rooms: This domain features good visibility and connectivity for pedestrian wayfinding, with an averaged travel time 38-45s, std=34-40, among the 3 pedestrian inflow rates. The service saved time is relatively few compared to other domains, for most entries can find their destinations efficiently. π-Now significantly reached more pedestrians yet has poor performance. With the ability to identify pedestrians who just missed their destinations, e.g., turned around early and missed the destination at the farther corner, π-$V_0^{R|\pi^H}$ and π-$V_t^{R|\pi^H}$ outperform the other two.

II) Corridors: This domain features poor visibility and connectivity, with an averaged travel time 53-91s, std=56-150, among the 3 pedestrian inflow rates. The travel time highly vary in this domain, depending on the traveller's entrance, her goal, and the direction she turns at the intersection, especially significant in the low-inflow rate scenario. While π-$V_0^{R|\pi^H}$ is capable of distinguishing guidance performance per entrance, this policy is incapable of inferring future motion variations, e.g., the pedestrian from the top entrance may soon reach her destination along the same corridor. π-$V_t^{R|\pi^H}$ outperforms significantly in this domain.

III) Open areas: This domain contains areas of levels of visibility and connectivity, from the wide room space, wide hallway, to the relatively narrow corridor, with an averaged travel time 53-70s, std=53-84, among the 3 pedestrian inflow rates. Due to the wide area range of this domain, guidance is on average of good value, e.g., guiding out the large room to the hallway, or from the corridor to the room. π-$V_t^{R|\pi^H}$ also outperforms significantly in this domain, by reaching the least pedestrians yet improving the most travel efficiency.

IV) Task switching and undeployed time: Reported in Table. II, in domain III, π-$V_t^{R|\pi^H}$ has more undeployed time, given that this policy less frequently initiates guidance service, as reported in Table. I. However, the numbers of attempted task switches are relatively high. We attribute this to the noisy estimation of guidance value using π^H. Since occasionally π^H can sample unusual choice of route, e.g., repetitively entering and leaving between the hallway and the room. Such outliers significantly increase the value estimate, encouraging the robot to initiate guidance. This is significant for π-$V_0^{R|\pi^H}$, too, whereas π-UncertaintyDetect has relatively robust estimation for service initiation. We refer to potential improvements, e.g., through outlier-filtering techniques, as future work of interest.

	Pedestrian inflow rates in domain III		
	$0.025/s$	$0.05/s$	$0.1/s$
π-Now	346s [34]	308s [46]	44s [50]
π-UncertaintyDetect	716s [25]	696s [27]	**505**s [31]
π-$V_0^{R\|\pi^H}$	546s [45]	503s [44]	324s [43]
π-$V_t^{R\|\pi^H}$	**726**s [29]	**706**s [35]	**505**s [35]

TABLE II: Robot undeployed time (s) and [number of attempted task switching] in domain III.

B. *Discussion*

1) "Unexpected" behaviors: π-$V_0^{R|\pi^H}$ and π-$V_t^{R|\pi^H}$ occasionally attempt to initiate guidance way before uncertainty is detected. This occurs especially more often in Domain II, where π^H can get stuck or repeat a lengthy route, resulting in high variance of V^{π^H} estimation; and significant observations are not expected farther along the narrow corridor until the intersection, to distinguishably estimate travel uncertainty and destination. We refer to this initiative behavior as to early prevent "worst-case" loss.

2) Cost-performance of pending: π-$V_t^{R|\pi^H}$ can delay the initiation 10-20s later than π-$V_0^{R|\pi^H}$, especially in the low entry-rate scenarios, appearing more conservative.

3) Communication action and recipient response: as the resultant guidance behavior (including the presumption and timing) differs from that in previous literature, e.g., to initiate only when uncertainty is detected, the form of interaction and the recipient's response require further study. For example, the robot can provide critical information directly to the wide audience which prevents worst-case performance, as costly communication may not be expected as worth it by customers who have not realized they need guidance yet.

VI. CONCLUSION

For robots to be deployed in the field and efficiently provide service, this work contributed a solution that uses the agent-based model for guidance cost-performance estimation, and generated domain-structure aware strategies to initiate guidance. Validated in real-world maps of common indoor structures, our approach was shown to be more efficient by saving more pedestrian travel time while providing the least number of services, compared to the behavioral-based guidance strategy and the intent-communication strategy in prior arts. Improvements on service value filtering are expected for more robust and consistent decisions on service initiation. The form of communication and the subject's response require further study, as the communication strategy/behavior differs from those in the literature.

References

[1] T. Kanda, D. F. Glas, M. Shiomi, H. Ishiguro, and N. Hagita, "Who will be the customer? a social robot that anticipates people's behavior from their trajectories," in *Proceedings of the 10th international conference on Ubiquitous computing*, 2008, pp. 380–389.

[2] S. Satake, H. Iba, T. Kanda, M. Imai, and Y. M. Saiki, "May i talk about other shops here?: Modeling territory and invasion in front of shops," in *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM, 2014, pp. 487–494.

[3] S. Satake, K. Hayashi, K. Nakatani, and T. Kanda, "Field trial of an information-providing robot in a shopping mall," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2015, pp. 1832–1839.

[4] D. Brščić, T. Ikeda, and T. Kanda, "Do you need help? a robot providing information to people who behave atypically," *IEEE Transactions on Robotics*, vol. 33, no. 2, pp. 500–506, 2017.

[5] T. Kanda, D. F. Glas, M. Shiomi, and N. Hagita, "Abstracting people's trajectories for social robots to proactively approach customers," *IEEE Transactions on Robotics*, vol. 25, no. 6, pp. 1382–1396, 2009.

[6] C. Shi, M. Shimada, T. Kanda, H. Ishiguro, and N. Hagita, "Spatial formation model for initiating conversation," *Proceedings of robotics: Science and systems VII*, pp. 305–313, 2011.

[7] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita, "A communication robot in a shopping mall," *IEEE Transactions on Robotics*, vol. 26, no. 5, pp. 897–913, 2010.

[8] P. Milhorat, D. Lala, K. Inoue, T. Zhao, M. Ishida, K. Takanashi, S. Nakamura, and T. Kawahara, "A conversational dialogue manager for the humanoid robot erica," in *Advanced Social Interaction with Agents*. Springer, 2019, pp. 119–131.

[9] Y. Okuno, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita, "Providing route directions: design of robot's utterance, gesture, and timing," in *2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2009, pp. 53–60.

[10] N. Mavridis, "A review of verbal and non-verbal human–robot interactive communication," *Robotics and Autonomous Systems*, vol. 63, pp. 22–35, 2015.

[11] S. Varges, F. Weng, and H. Pon-Barry, "Interactive question answering and constraint relaxation in spoken dialogue systems," in *Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue*, 2006, pp. 28–35.

[12] S. Young, "Using pomdps for dialog management," in *2006 IEEE Spoken Language Technology Workshop*. IEEE, 2006, pp. 8–13.

[13] R. Nisimura, T. Uchida, A. Lee, H. Saruwatari, K. Shikano, and Y. Matsumoto, "Aska: receptionist robot with speech dialogue system," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2. IEEE, 2002, pp. 1314–1319.

[14] M. Tucker, D. Aksaray, R. Paul, G. J. Stein, and N. Roy, "Learning unknown groundings for natural language interaction with mobile robots," in *Robotics Research*. Springer, 2020, pp. 317–333.

[15] Y. Hato, S. Satake, T. Kanda, M. Imai, and N. Hagita, "Pointing to space: modeling of deictic interaction referring to regions," in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2010, pp. 301–308.

[16] S. Satake, T. Kanda, D. F. Glas, M. Imai, H. Ishiguro, and N. Hagita, "How to approach humans?: strategies for social robots to initiate interaction," in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM, 2009, pp. 109–116.

[17] O. A. I. Ramírez, H. Khambhaita, R. Chatila, M. Chetouani, and R. Alami, "Robots learning how and where to approach people," in *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*. IEEE, 2016, pp. 347–353.

[18] Y. Kato, T. Kanda, and H. Ishiguro, "May i help you?-design of human-like polite approaching behavior," in *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2015, pp. 35–42.

[19] S.-Y. Lo, E. S. Short, and A. L. Thomaz, "Planning with partner uncertainty modeling for efficient information revealing in teamwork," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 319–327.

[20] V. V. Unhelkar, S. Li, and J. A. Shah, "Decision-making for bidirectional communication in sequential human-robot collaborative tasks," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 329–341.

[21] J. J. Gibson, "The theory of affordances," *Hilldale, USA*, vol. 1, no. 2, 1977.

[22] B. Hillier, A. Penn, J. Hanson, T. Grajewski, and J. Xu, "Natural movement: or, configuration and attraction in urban pedestrian movement," *Environment and Planning B: planning and design*, vol. 20, no. 1, pp. 29–66, 1993.

[23] A. Turner and A. Penn, "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*, vol. 29, no. 4, pp. 473–490, 2002.

[24] E. Andresen, D. Haensel, M. Chraibi, and A. Seyfried, "Wayfinding and cognitive maps for pedestrian models," in *Traffic and Granular Flow'15*. Springer, 2016, pp. 249–256.

[25] W. Wang, S. Lo, and S. Liu, "A cognitive pedestrian behavior model for exploratory navigation: Visibility graph based heuristics approach," *Simulation Modelling Practice and Theory*, vol. 77, pp. 350–366, 2017.

[26] P. M. Kielar, D. H. Biedermann, A. Kneidl, and A. Borrmann, "A unified pedestrian routing model for graph-based wayfinding built on cognitive principles," *Transportmetrica A: transport science*, vol. 14, no. 5-6, pp. 406–432, 2018.

[27] P. Foo, W. H. Warren, A. Duchon, and M. J. Tarr, "Do humans integrate routes into a cognitive map? map-versus landmark-based navigation of novel shortcuts." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31, no. 2, p. 195, 2005.

[28] R. A. Epstein, "Parahippocampal and retrosplenial contributions to human spatial navigation," *Trends in cognitive sciences*, vol. 12, no. 10, pp. 388–396, 2008.

[29] D. Silver and J. Veness, "Monte-carlo planning in large pomdps," in *Advances in neural information processing systems*, 2010, pp. 2164–2172.

[30] A. Somani, N. Ye, D. Hsu, and W. S. Lee, "Despot: Online pomdp planning with regularization," in *Advances in neural information processing systems*, 2013, pp. 1772–1780.

[31] V. V. Unhelkar, C. Pérez-D'Arpino, L. Stirling, and J. A. Shah, "Human-robot co-navigation using anticipatory indicators of human walking motion," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 6183–6190.

[32] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 2013, pp. 301–308.