

DETC2018-86295

ROBOT EVIDENCE BASED SEARCH FOR A DYNAMIC USER IN AN INDOOR ENVIRONMENT

Shayne Lin and Goldie Nejat

Autonomous Systems and Biomechatronics Laboratory (ASBLab)
Department of Mechanical and Industrial Engineering
University of Toronto
Toronto, Ontario, Canada

ABSTRACT

In this paper we present the development of an evidence-based search planner for a mobile assistive robot to autonomously search for a dynamic person in a multi-room home environment in order to provide assistance. We solve the dynamic person search problem by uniquely considering evidence of household objects along with a user spatial-temporal model to increase the probability of finding the user. Our planner utilizes a Partially Observable Markov Decision Process (POMDP) to plan optimal robot search paths in the environment as the user and evidence locations are partially observable. Extensive simulated experiments in a home environment were conducted to compare our proposed evidence-based search approach with 1) a search technique without prior user information, and 2) a search technique that only uses a user model. The results show that our proposed search technique has higher success rates for finding the user and is more robust to the dynamic behaviors of the user.

INTRODUCTION

More than 1.2 million Canadians over the age of 65 are living alone, accounting for 25% of the senior population in Canada [1]. Socially assistive robots can be used to assist these individuals with activities of daily living such as providing reminders, aiding with meal preparation and eating, dressing, and hygiene activities [2]–[5], and facilitating recreational activities such as exercise and memory games [6]–[8]. In particular, these robots provide social and cognitive support to promote aging-in-place. In order to initiate such assistance, a socially assistive robot needs to first search for and find a user in his/her home [9].

In this paper, we address the search problem of a mobile social robot searching for a dynamic user in a multi-room home environment before a time limit. This task can be challenging for several reasons. The robot only uses its onboard sensors to

detect the user in the environment. Furthermore, the user can be moving from region to region during the search and also revisiting previous locations. As the robot is also sharing the environment with people, it would need to consider social etiquette rules [10].

In the context of robot searching an indoor environment for people, many probabilistic search techniques have been designed that use knowledge of past information to construct a user model and generate the search policy, e.g. [2], [11]–[18]. Examples of past information include frequency of past observations, last known locations, and activity schedule, which are combined to generate a prior user location model. These prior models are usually represented as a probability density function (PDF) over the entire search space or a set of selected navigation points, each assigned with a probability of the person being there. Based on the user models, the search policies used for a single target prioritize locations in order to maximize the probability of finding the user [2], [11], [13], [19], [20]. In multi-target scenarios, common policies used are to maximize the number of users found [15], [17] or maximize the probability of encountering at least one person [14].

In this paper, we address the dynamic person search problem by uniquely considering the evidence around the home to increase the probability of finding the dynamic user. Such evidence is based on common household objects such as lights and TVs, whose presence or state are often associated with the location or activity a user is engaging in [9], [10]. Therefore, if the robot can observe such evidence during a search for a user, this information can also be used with a prior user model to aid in the search. The proposed approach can also handle changes in the user's behaviors that deviate from the prior user model. We present a novel probabilistic search strategy that involves checking evidence around the home while searching for a single dynamic user in order to find them more quickly. The search problem is modelled as a Partially Observable Markov Decision Process (POMDP) where the objective is to maximize

the probability of finding the user within a time limit. Information on past user locations, activities and their associations with evidence is used to generate the prior user model for the POMDP.

PERSON SEARCH IN INDOOR ENVIRONMENTS

Existing search approaches use one or more robots to search for people in indoor structured environments. These approaches can be categorized as either searching for: 1) static users [2], [11], [12], [20], or 2) dynamic users [13]–[19], [21].

For example in [11], the HOMER robot was developed to deliver messages to multiple static recipients one at a time. The robot used a location likelihood function created manually for each message recipient. Upon receiving a request to deliver a message, the robot searched for the recipient by choosing the closest region with the maximum location likelihood that it had not yet searched. The search plan was terminated when the recipient was found or when all regions had been visited. In [12], a dynamic programming algorithm was used to search for multiple static targets with a single robot. The objective was to minimize the ratio of total search time to number of targets found. Past target location data was used to determine the expected number of targets in each region.

To account for target dynamic behaviors during search, in [13], a greedy search technique was used for searching for a single target at home using a robot. The planner used past frequencies of target locations to create a person occurrence map. The map was used to set navigation points where the robot would visit to find the target. Navigation points with the highest probability of finding the target were searched first.

Markov Decision Process (MDP) planners have been commonly used to search for targets by assigning rewards based on probabilities of target locations. For example, in [14], a finite horizon MDP approach was used to plan an optimal path in an office environment that maximizes the probability of encountering at least one dynamic target within a time limit. Each state was a non-occupied cell at a given time of day and the possible actions included moving to any of the 8 neighboring cells. The reward for an action was the probability of encountering at least one target after moving to the new location at the next time step. The user location model was designed from survey responses regarding activities and activity durations throughout a day. The information was then used to fit a Poisson model where the random variable represented the rate of encountering each target. In [15], [16], an MDP planner was also used to maximize the expected number of dynamic targets found within a given time in a long-term care environment. Rewards were assigned based on the spatial-temporal likelihood function for each target, which was determined as a function of the last known locations, user schedules, observation frequency and room types. In [18], the ARTOS robot searched for an elderly user at home by using a MDP planner to plan a sequence of reference points to visit. The reference points were locations where the elderly user can be found. The reward for searching a reference point was the probability of finding the user at that location divided by the

cost of navigating to the reference point from the robot's current location.

In [19], a POMDP planner was used to directly handle uncertainty of this world state, e.g. target location. A multi-robot search plan was used for finding a moving target in both a museum and office environment. The fully observable states were the positions of all robots while the target location was partially observable. The POMDP policy mapped each state to an action for each robot at every time step. The initial target location was unknown, but the robots used a target motion model whereby targets were assumed to move at 1 m/s in a random direction. The reward for the POMDP was assigned to maximize the probability of finding the target within a time limit. In [21], a search planner was proposed where the belief over target location was represented as a probability distribution over discretized cells in a retirement home. The dimension of the belief states was reduced by performing Exponential Family Principal Component Analysis. A belief space MDP with the reduced state space was then used to minimize the time to find the target. The reward function maximized the expected total future reward for finding people in the fastest search times.

In [17], the search problem of finding multiple dynamic users in a long-term care center by multiple robots was formulated as a travelling thief problem. The planner specified the order of regions to search and the amount of time spent to search each region. The reward for the optimization problem was determined based on user location and joint activity patterns, which were acquired by having the robots follow the users collectively for several days.

In the aforementioned search approaches, prior target location PDFs are used to make an initial prediction on target location such that the planner can find a target in a shorter expected time. However, if user behaviors change over time or during a specific day (for example, due to illness), the search plans can become sub-optimal.

Only a handful of papers have considered explicitly using last known target locations to predict the current target location. For example, in [2], a Hidden Markov Model (HMM) was used to model location based on activity sequences in a home environment. Given a sequence of recent known locations (which were observed), the HMM predicted the possible sequence of activities (hidden states) that described the past locations observed. These activities were then used to predict the current activity and location of the user, who is assumed to be static during the search. Once the target PDF was computed, a robot used a greedy approach to visit different rooms.

In [20], a socially assistive robot used information from motion sensors to infer the current target location within a home. In cases where the target was not detected by a motion sensor, the robot would search regions in the order where the target was most recently detected by the sensors. For both the approaches presented in [2], [20], user locations must be directly observed. However, there are scenarios where the user has not yet been seen or the user does not follow a specific sequence of activities. Furthermore, using information from

motion sensors placed all over the environment may also not be feasible.

Our novel search planner considers direct observations of evidence by the robot in the environment during the search in order to address such limitations. Namely, evidence can be directly correlated with the user's location, thus increasing the probability of finding the target by visiting fewer regions. For example, if the robot observes that the living room light is on and the kitchen light is off after the user eats a meal, it can predict that the person is in the living room before searching that specific room. Furthermore, our proposed planner considers the trade-off of checking for evidence or directly looking for the user during the search in order to reduce the number of regions that need to be searched.

EVIDENCE-BASED ROBOT SEARCH OF A DYNAMIC USER

The objective of our proposed search approach is to determine the search plan that a robot should execute in order to find a dynamic user in a multi-room environment within a time limit. The problem consists of the following details:

Activities: An activity is a specific task that the user is doing. The set of activities performed by the user is $A = \{a_1, a_2, \dots, a_I\}$.

Environment: The home environment consists of a set of regions $R = \{r_1, r_2, \dots, r_J\}$, which can be occupied by the user. Examples of regions are rooms and hallways.

Evidence: Evidence in the environment is represented by common household objects such as lights, dishes, TV. The set of evidence that can be observed by the robot during the search is $E = \{e_1, e_2, \dots, e_K\}$. Each evidence, e_k , has a binary state, θ_k , i.e. on/off or present/absent. Some evidence states are correlated with activities while others are correlated with user locations. These are referred to as activity-based evidence and location-based evidence, respectively. For example, TV on/off is directly associated with the activity "watching TV" since the user cannot watch TV if it is not on. On the other hand, "kitchen light" on/off is associated with whether the user is in the kitchen. Let M represent the total number of possible combinations of evidence states, and $m \in [1, M]$ represent a specific combination, then the state of each evidence for the m^{th} combination can be expressed as $\{\theta_{1,m}, \theta_{2,m}, \dots, \theta_{K,m}\}$.

The search plan involves a sequence of actions that the robot executes until the user is found. An action can be to search for the user in region r_j or check the state θ_k of an evidence. We formulate the search problem as a finite horizon POMDP to handle the uncertainty associated with the user location and evidence states.

To address the dynamic user search problem, we have developed the robot search architecture presented in Fig. 1. The user activity database consists of the prior information of the user in the environment. In particular, it consists of the user's activity, user's location, the time of day, and the state of the

evidence. This information is used to generate four PDFs: 1) user activity-time, 2) user activity-location, 3) evidence-location, and 4) evidence-activity.

The *activity-time PDF*, $P(a_i | t_i)$, represents the probability that at discrete time t_i , the user performs activity a_i . The *location-activity PDF*, $P(r_j | a_i)$, represents the probability of the user being in region r_j while performing activity a_i . The *evidence-activity PDF*, $P(\theta_k | a_i)$, is the probability that when the user is performing activity a_i , evidence e_k 's state θ_k is present. The *evidence-location PDF*, $P(\theta_k | r_j)$, is the probability that when the user is in r_j , evidence e_k 's state θ_k is present.

These PDFs are then used by the Evidence-Location-Activity (ELA) model, represented as a Bayesian Network, to provide the relationship between evidence, location and activity. The output of the ELA model is the initial joint user evidence belief, $b_0(r_j, \theta_{1,m}, \theta_{2,m}, \dots, \theta_{K,m}, t_i)$, which is the set of probabilities that, at discrete time t_i , the user is in region r_j ; and the evidence states are $(\theta_{1,m}, \theta_{2,m}, \dots, \theta_{K,m})$.

At the start of the search, the joint user evidence belief b is equal to b_0 . This allows the search planner to reason about the current user location and evidence states using past user information. The POMDP solver then uses b to find the optimal action that maximizes the probability of finding the user before the time limit. The action chosen by the POMDP solver is executed by the robot. The robot's perception system then obtains an observation of the environment, which is used to update b . The process is repeated until the user is found or the time limit is reached. The detailed formulation of the ELA model and POMDP model are discussed below.

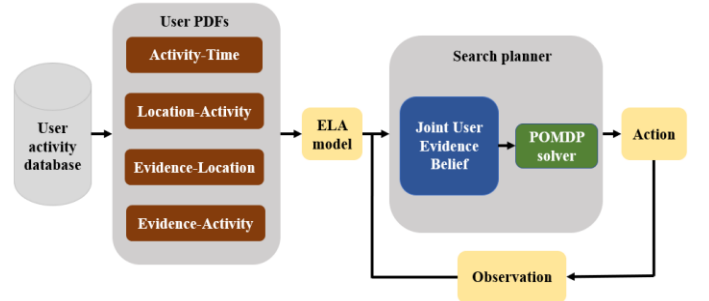


Figure 1. Search architecture.

EVIDENCE LOCATION ACTIVITY (ELA) MODEL

As previously mentioned, the overall ELA model is represented by a Bayesian network, Fig. 2. A Bayesian network is used as it can relate multiple conditional probabilities (i.e., the 4 PDFs in the user probability model) to derive the belief state that is used by the POMDP model. The evidence considered in the ELA model is shown in Table 1. All evidence used in Table 1 have two states, i.e. on/off or present/absent. Dishes are dependent on the activity "meal eating" and "TV" is dependent on the activity "watching TV". Therefore, both "dishes" and "TV" are activity-based evidence. The presence of

slippers outside the bedroom indicates that the user is in the bedroom; the presence of shoes in the hallway indicates that the user is at home. Therefore, both slippers and shoes are location-based evidence. The light in each room is dependent on whether the user is in that room, which is also location-based evidence.

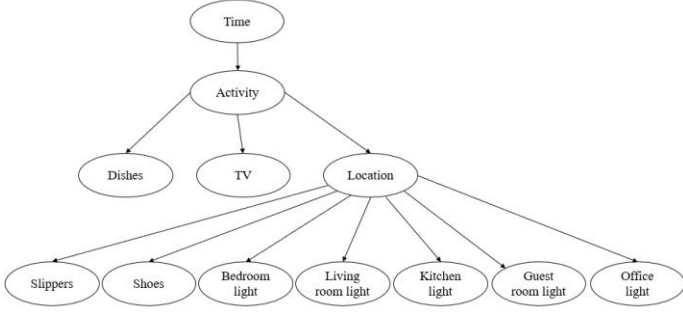


Figure 2. Bayesian network expressing the relationship between target locations, activities and evidence.

Table 1. List of evidence and their locations

Evidence	Location	Influence Type
Dishes	Kitchen sink	Activity
TV	Living room	Activity
Slippers	Bedroom entrance	Location
Shoes	Back hallway	Location
Bedroom light	Bedroom	Location
Living room light	Living room	Location
Kitchen light	Kitchen	Location
Guest room light	Guest room	Location
Office light	Office	Location

The ELA model which represents the initial joint user evidence belief b_0 is represented as a combination of the four PDFs:

$$b_0 = P(r_j, \theta_{1,m}, \theta_{2,m}, \dots, \theta_{K,m} | t_l) = \sum_{i=1}^I P(a_i | t_l) \left(\prod_{k=1}^{K'} P(\theta_{k,m} | a_i) \right) P(r_j | a_i) \left(\prod_{k=K'+1}^K P(\theta_{k,m} | r_j) \right), \quad (1)$$

where $\{e_1, e_2, \dots, e_{K'}\}$ represents the set of evidence associated with activities, and $\{e_{K'+1}, e_{K'+2}, \dots, e_K\}$ represents the set of evidence associated with locations. The first term is given by the activity-time PDF; the second term by the evidence-activity PDF; the third term by the location-activity PDF, and the last term by the evidence-location PDF, respectively.

POMDP FORMULATION FOR THE SEARCH PLANNER

A POMDP represents a tuple of $\langle S, \alpha, \Omega, F, O, W \rangle$, where S is a finite set of states, $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_V\}$ is a finite set of actions, Ω is a finite set of observations, F is the transition function, O is the observation function, and W is the reward function [22]. For our POMDP model, every state in S consists of four components: 1) time elapsed, $t_{elapsed}$, since the beginning of the search, 2) robot location, 3) user location, and 4) the state of each evidence. The user location and evidence states are partially observable. There is also an absorbed goal state, which corresponds to the scenario where the user is found. As previously mentioned, the possible actions in α are searching for the user in a region and checking for evidence.

Observation function:

The observation function, O , calculates the probability of receiving an observation o , where $o \in \Omega$, after taking action α and entering a new state s' [22]:

$$O(s', \alpha, o) = p(o | s', \alpha) \quad (2)$$

The possible observations for searching the user are “user present” and “user absent”, and the observations for checking evidence are “evidence present/on” and “evidence absent/off”. Within our model, we have also included perception errors ϵ_U and ϵ_E , respectively. These errors are with respect to detecting users and evidence in the environment.

Transition function:

The transition function F maps the current state s and action α_i to a new state s' [22]:

$$F(s, \alpha, s') = p(s' | s, \alpha) \quad (3)$$

If the action is to search the region that the user is occupying, the probability of finding the user and thus transitioning into the absorbed goal state is $(1 - \epsilon_U)$. When the action is to check for evidence, the evidence state will transition into the observed evidence state with a probability of $(1 - \epsilon_E)$. After taking an action, $t_{elapsed}$ is incremented by the time required by the action.

Rewards:

The reward function W assigns a reward for each state-action pair and the observation received as a result of performing the action. Given the time limit, t_{limit} , the reward function is represented as:

$$W(s, \alpha) = \begin{cases} t_{limit} - t_{elapsed} & \text{if } o = \text{"user present"} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Namely, a reward equal to the remaining search time is given only when the user is found. No reward is given for any intermediate actions. The reward is assigned in this manner so

that the robot is encouraged to find the user before the time limit.

Belief Update:

The initial belief, b_0 , is the joint user evidence belief from the ELA model. After the robot takes action α and receives an observation o , the information obtained from the observation is incorporated into the search planner by updating the belief, $b_\alpha^o(s')$. This belief represents the probability that the robot transitions to s' after action α and receives observation o , and is updated using Bayes Rule:

$$b_\alpha^o(s') = \frac{p(o|s', \alpha) \sum_s p(s'|s, \alpha) b(s)}{\sum_{s'} [p(o|s', \alpha) \sum_s p(s'|s, \alpha) b(s)]}, \quad (5)$$

where $\sum_s p(s'|s, \alpha) b(s)$ is the sum of the probabilities that belief state $b(s)$ will transition into the new state s' .

Given a policy π , the expected total discounted reward at belief $b(s)$ can be computed using the value function [13]:

$$V_\pi(b) = \sum_s W(s, a) b(s) + \gamma \sum_{o, s'} p(o|s', \alpha) p(s'|s, \alpha) b(s) V_\pi(b_\alpha^o(s')), \quad (6)$$

where γ is the discount factor (we use $\gamma = 1$). The first term is the expected immediate reward and the second term is the expected discounted future reward. Our aim is to maximize the probability of reaching the absorbed goal state to obtain an overall positive reward.

The search planner constructs an optimal policy V_π^* that satisfies Bellman's Equation, which corresponds to choosing the best actions to maximize the value function in Eqn. 6. Several online and offline techniques have been developed to find an optimal policy. For our search task, an online solver is required since it updates the belief after each observation to aid in planning the next action. In order to compute the policy and perform belief updates in real-time, we use the online Determinized Sparse Partially Observable Tree (DESPOT) technique [23]. A common technique to solve the POMDP problem is to express all possible sequence of actions and observations as a belief tree. In our case, a sequence corresponds to a possible search scenario where the robot either searches a region or looks for evidence until the user is found. DESPOT samples several sequences from the belief tree and finds the optimal policy from the sampled tree, thus giving a near-optimal policy with reduces computational complexity. Once the policy is generated, the robot executes the first action from the policy and receives an observation. The belief is updated based on the observation and the process is repeated until the robot finds the user or the time limit is reached.

EXPERIMENTS

To validate the performance of our proposed person search technique, we conducted simulated experiments in a personal

home environment consisting of 5 separate rooms and two hallways, Fig. 3. The red lines represent boundaries between regions. The simulations consisted of a simulated mobile robot finding a single person living in the home. The robot is allowed to enter all regions except for the bathroom. The simulations were run on a computer with Intel® Core™ i5-6600 CPU and 16GB of RAM.

The Aruba dataset for an elderly female [24] was used to obtain the *activity-time PDF* and *location-activity PDF*. The dataset includes time-stamped location and activity events obtained from a smart home for the individual over the course of 220 days.

To generate the *evidence-location PDF* and *evidence-activity-PDF*, we need to obtain the probability of evidence being present given either the user location or user activity. Since the Aruba dataset does not contain evidence information, evidence was added to the dataset based on existing user location information and annotated activities, e.g. “meal preparation”, “washing dishes”, and “relaxing”. The conditions in which the evidence states are assigned are described in Table 2. Common household objects used during activities of daily living were chosen to demonstrate that our model can be generalizable to any home environment with a single elderly resident. For example, dishes used to prepare food are present in the kitchen sink during meal eating. The user was assumed to be watching TV if she was “relaxing” for more than 20 minutes in the living room. The user wears slippers around the house except when going into the bedroom, for which she leaves the slippers outside the bedroom door. It was also observed from the Aruba dataset that the user leaves home using the door in the back hallway, therefore, it is assumed that when she is home her shoes are beside the door in this hallway.

The conditions for lights being on or off in each of the rooms are the same. As the user may not always have the light on in a room she is in, i.e., when there is natural light from outside, or she may not always be in a room for which the light is on, we use probabilities to represent the light evidence state as shown in Table 2.

Once the *evidence-location PDF* and *evidence-activity-PDF* are generated, they are used with the *activity-time PDF* and *location-activity PDF* to compute the ELA model using Eqn. 1.

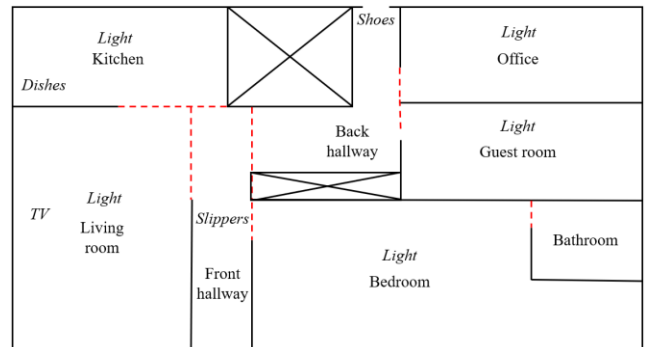


Figure 3. Home environment used in Aruba dataset.

Table 2. Evidence States and Conditions

Evidence	States and Conditions
Dishes	<i>Present</i> in the kitchen sink during “meal time”; <i>absent</i> otherwise
TV	<i>On</i> if the person is relaxing for more than 20 minutes in the living room; <i>off</i> otherwise
Slippers	<i>Present</i> if the user is in the bedroom or bathroom; <i>absent</i> otherwise.
Outdoor shoes	<i>Present</i> in the back hallway when user is at home.
Lights in the Bedroom, Living room, Kitchen, Guest room, and Office	<u>During the day</u> (assumed to be before 6pm): 80% <i>on</i> if user is in the same room, 20% <i>on</i> if user is in a different room. <u>At night</u> : always <i>on</i> if user is in the same room, 20% <i>on</i> if user is in a different room.

A total of 1633 trials were conducted during a 10 am to 11 pm time interval, with at least one trial at the start of each hour. The search limit was set to 4 minutes based on the small size of the environment. The robot moved in the environment with a speed of 0.5m/s. The recognition time needed for the evidence was 5 seconds and for the user was 30 seconds, respectively. The same trials were simulated under two scenarios with: $\epsilon_U = \epsilon_E = 0$, and $\epsilon_U = \epsilon_E = 0.1$. If there is less than 30 seconds remaining for the robot to search a region, the probability of finding the user is reduced to the time remaining divided by the time required to search a region.

RESULTS AND DISCUSSION

The performance of our proposed person search approach (ELA + evidence) was compared against: 1) a baseline POMDP technique without prior user information and a uniform initial user belief, and 2) a POMDP search technique using only the ELA model to generate the initial belief but not considering evidence during the search.

Fig. 4 presents the results for all three methods. Fig. 4(a) shows the results when there is no perception error (i.e., ϵ_U and ϵ_E are both zero), and Fig. 4(b) shows results when perception error is 10% (i.e., ϵ_U and ϵ_E are 0.1), respectively. Success rates are shown with respect to the robot finding a user by the 1,2,3 and 4 minute marks to investigate the performance between the three methods during the search. As can be seen in the figure, the baseline technique consistently has the lowest success rate. For both conditions, the success rate for our ELA + evidence method is higher than the ELA method with a difference of 11.3% and 9.4% at the 4 minute search time limit.

To investigate how the methods handle the dynamic behavior of the user, we also examined the number of regions the user occupied during the search. Namely, the user must be in at least two regions during the search to be considered dynamic. Fig. 5 shows the success rate of finding the user given the number of regions with ϵ_U and ϵ_E equal to 0 and 0.1, respectively. As expected, the success rate is the lowest for the

baseline method regardless of the number of regions the user occupies. Furthermore, for the no perception error condition, our ELA + evidence method performs better than the ELA method by 7.5% (8-10 regions) to 12.0% (2-4 regions). When perception error is added, our proposed ELA + evidence method still outperforms the ELA method by 7.9% (2-4 regions) to 12.9% (>10 regions). This highlights the robustness of the ELA + evidence method to dynamic user behaviors.

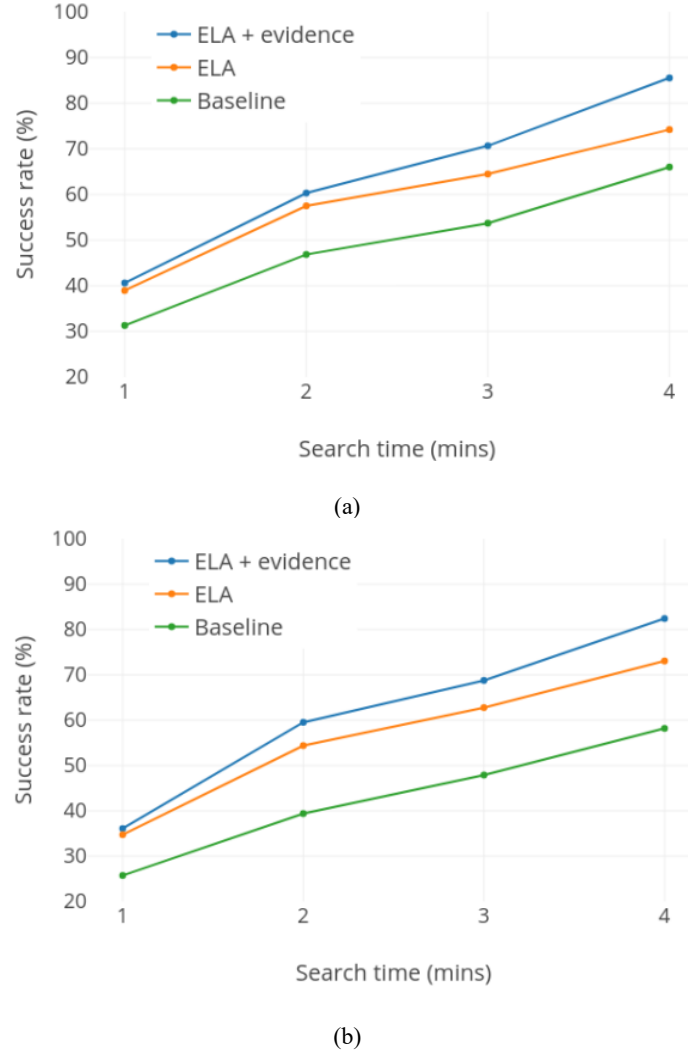
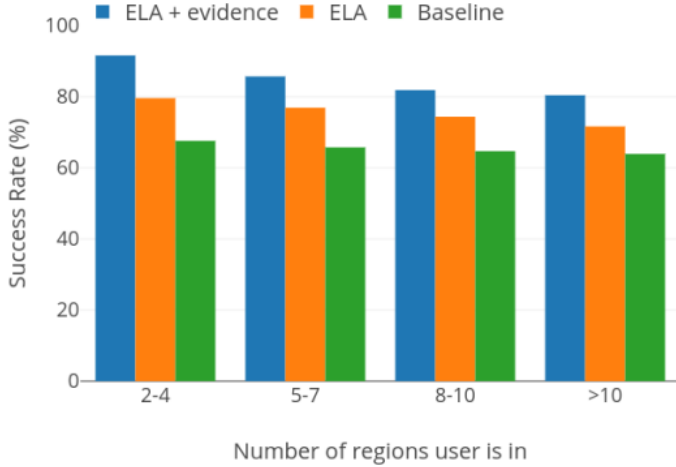
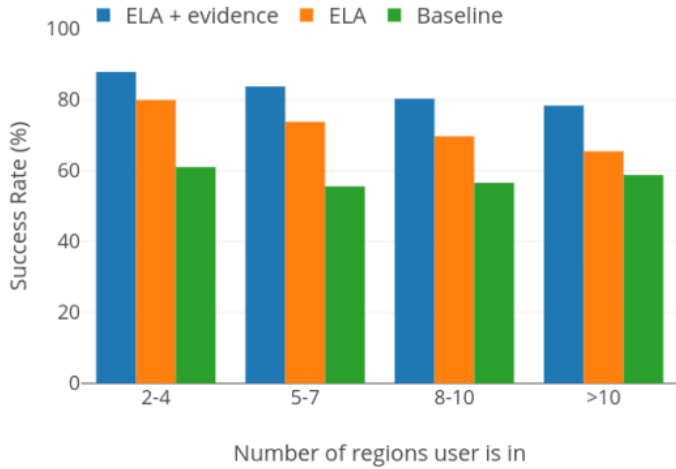


Figure 4. Search success rate for finding the user with (a) no perception error, and (b) with 10% perception error (ϵ_U and ϵ_E are 0.1) within the 4-minute time limit.



(a)



(b)

Figure 5. Success rate with respect to the number of regions the dynamic user has visited during the search with (a) no perception error, and (b) with 10% perception error (ϵ_U and ϵ_E are 0.1).

We further investigated the influence of using evidence within our proposed search approach by looking specifically at cases where evidence was explicitly used during the search. Two such cases are discussed herein in more detail, which are summarized in Table 3. In both cases, the robot initially starts in the living room.

Case 1: In case 1, the user is initially in the front hallway. The robot searches for the user in the living room first and does not find the user, so it moves on to check the kitchen light to see if there is evidence that the user would be there. In the meantime, the user walks into the living room. After the robot sees that the kitchen light is off, it goes to the living room to see if the light is on, since the ELA model suggests that the user has a 54% probability of being in the living room. When it detects

that the light is on, it searches the living room again and finds the user in a total time of 1.3 minutes.

Case 2: In case 2, the ELA model suggests that the probability of the user being in the living room is over 80%, so the robot first searches the living room directly without checking for any evidence. However, the user is not found there. As the probabilities in the ELA model for all the other rooms are under 3%, instead of searching these rooms, the robot looks for evidence including if the light is on in each of these rooms or the slippers are in front of the bedroom in a counter-clockwise search trajectory. This allows the robot to gather information on most of the regions in a very short period of time without having to search each room. In this case, when the robot detects that the light is on in the office, it starts to search the office and finds the user in a total time of 1.6 minutes, even though the probability of the user being in the office is the second lowest among all regions.

Table 3. Two scenarios where the evidence-based planner benefited from searching for evidence

Scenario	Top three region with highest belief	Sequence of actions and observations
1	Living room, kitchen, guest room	Living room (target absent), kitchen light (off), living room light (on), living room (user present)
2	Living room, bathroom, bedroom	Living room (target absent), slippers (absent), bedroom light (off), office light (on), office (user present)

CONCLUSION

In this paper, we present a unique POMDP search planner for addressing the problem of an assistive robot searching for a single dynamic user in a multi-room home environment. The novelty of our approach is in the inclusion of checking for evidence during the search. Evidence is represented as common household objects, and is directly correlated with user activities and locations. The observation of evidence during the search is used to inform the planner about the current user location. Our planner uses past user location, activity and evidence data to generate the initial joint user evidence belief for the POMDP model. Then the POMDP model is used to determine which region to search for the user or which evidence to check such that the probability of finding the user within the time limit is maximized. Comparison results of our evidence-based search planner with other POMDP-based planners show that we are able to achieve a higher success rate of finding a dynamic user. Furthermore, our approach is more robust to dynamic users. Our future work consists of testing our approach in varying environment sizes and configurations, and integrating it onto our socially assistive robot platform.

ACKNOWLEDGMENTS

This work was funded by the Natural Sciences and Engineering Research Council of Canada, the Ontario Centres of Excellence, the Canada Research Chairs Program, and CrossWing Inc.

REFERENCES

- [1] Statistics Canada, "Living arrangements of seniors," Ottawa, 2012.
- [2] P. Bovbel and G. Nejat, "Casper: An Assistive Kitchen Robot to Promote Aging in Place1," *J. Med. Device.*, vol. 8, no. 3, p. 030945, Jul. 2014.
- [3] D. McColl and G. Nejat, "Meal-Time with a Socially Assistive Robot and Older Adults at a Long-term Care Facility," *J. Human-Robot Interact.*, vol. 2, no. 1, pp. 152–171, 2013.
- [4] L. Woiceshyn, Y. Wang, G. Nejat, and B. Benhabib, "Personalized clothing recommendation by a social robot," in *IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, 2017, pp. 179–185.
- [5] Chih-Hung King, T. L. Chen, A. Jain, and C. C. Kemp, "Towards an assistive robot that autonomously performs bed baths for patient hygiene," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2010, pp. 319–324.
- [6] D. McColl, W.-Y. G. Louie, and G. Nejat, "Brian 2.1: A socially assistive robot for the elderly and cognitively impaired," *IEEE Robot. Autom. Mag.*, vol. 20, no. 1, pp. 74–83, Mar. 2013.
- [7] D. McColl and G. Nejat, "Affect detection from body language during social HRI," in *IEEE International Symposium on Robot and Human Interactive Communication*, 2012, pp. 1013–1018.
- [8] L. Lewis, T. Metzler, and L. Cook, "Evaluating Human-Robot Interaction Using a Robot Exercise Instructor at a Senior Living Community," in *International Conference on Intelligent Robotics and Applications*, 2016, pp. 15–25.
- [9] H.-M. Gross *et al.*, "Robot companion for domestic health assistance: Implementation, test and case study under everyday conditions in private apartments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 5992–5999.
- [10] K. Qian, X. Ma, X. Dai, and F. Fang, "Robotic Etiquette: Socially Acceptable Navigation of Service Robots with Human Motion Pattern Learning and Prediction," *J. Bionic Eng.*, vol. 7, no. 2, pp. 150–160, Jun. 2010.
- [11] P. Elinas, J. Hoey, and J. J. Little, "HOMER: Human Oriented MESSenger Robot," *AAAI*, 2003.
- [12] H. Lau, S. Huang, and G. Dissanayake, "Optimal search for multiple targets in a built environment," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 3740–3745.
- [13] M. Volkhardt and H.-M. Gross, "Finding people in home environments with a mobile robot," in *European Conference on Mobile Robots*, 2013, pp. 282–287.
- [14] G. D. Tipaldi and K. O. Arras, "I want my coffee hot! Learning to find people under spatio-temporal constraints," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 1217–1222.
- [15] M. Schwenk, T. Vaquero, and G. Nejat, "Schedule-Based Robotic Search for Multiple Residents in a Retirement Home Environment," *AAAI*, pp. 2571–2577, 2014.
- [16] M. Schwenk, T. Vaquero, G. Nejat, and K. O. Arras, "Planning a Robot's Search for Multiple Residents in a Retirement Home Environment," in *International Conference on Automated Planning and Scheduling ICAPS Workshop on Planning and Robotics*, 2014, pp. 136–144.
- [17] S. C. Mohamed and G. Nejat, "Autonomous Search by a Socially Assistive Robot in a Residential Care Environment for Multiple Elderly Users Using Group Activity Preferences," in *International Conference on Automated Planning and Scheduling Workshop on Planning and Robotics (PlanRob)*, pp. 58–66, 2016.
- [18] K. Berns and S. A. Mehdi, "Use of an Autonomous Mobile Robot for Elderly Care," in *Advanced Technologies for Enhancing Quality of Life*, 2010, pp. 121–126.
- [19] G. Hollinger, S. Singh, J. Djughash, and A. Kehagias, "Efficient Multi-robot Search for a Moving Target," *Int. J. Rob. Res.*, vol. 28, no. 2, pp. 201–219, 2009.
- [20] H.-M. Gross *et al.*, "Progress in developing a socially assistive mobile home robot companion for the elderly with mild cognitive impairment," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2011, pp. 2430–2437.
- [21] N. Roy, G. Gordon, and S. Thrun, "Planning under Uncertainty for Reliable Health Care Robotics," in *Field and Service Robotics*, Berlin/Heidelberg: Springer-Verlag, pp. 417–426.
- [22] M. T. J. Spaan and N. Spaan, "A point-based POMDP algorithm for robot planning," in *IEEE International Conference on Robotics and Automation*, 2004, p. 2399–2404 Vol.3.
- [23] N. Ye, A. Somani, D. Hsu, and W. S. Lee, "DESPOT: Online POMDP Planning with Regularization," Sep. 2016.
- [24] D. J. Cook, "Learning Setting-Generalized Activity Models for Smart Spaces," *IEEE Intell. Syst.*, vol. 2010, no. 99, p. 1, Sep. 2010.