Probabilistic Recharging Model in Uncertain Environments (Extended Abstract)

Zhao Song
The University of Texas at Austin
Texas, USA
zhaos@utexas.edu

Wen Sun
The University of North Carolina at Chapel Hill
North Carolina, USA
wens@cs.unc.edu

Categories and Subject Descriptors

I.2 [**Artificial Intelligence**]: Miscellaneous

General Terms

Algorithms

Keywords

Recharging, Uncertainty, Planning

1. INTRODUCTION

In a long-term work or experiment, a robot is usually required to dock on a charging station and recharging itself. However, docking on a recharging station might not be always reliable. Even the most meticulous developed robot system may have the possibility that sometimes would execute abnormally. In our work, we focus on the uncertainty in the docking procedure. More formally, we use a Bernoulli distribution to model the events of successfully docking or not in one trial. Because of the uncertainty in a docking procedure, the robot may need to try multiple times till successfully docking the station. Hence it is very important for a robot to *a priori* consider the uncertainty in the docking procedure so that it can optimally decide when it needs to recharge and how much energy it needs to charge.

Previous research work that related to recharging includes [4], [5], [2], [3], and [1]. They used a battery threshold as a flag to make decisions on when to go back for recharging. Instead of using a static threshold, another strategy is the utilize adaptive method [7], and [6]. All previous methods do not consider uncertainty in the procedure when the robot is trying to docking the recharging station. In our work, we assume there is uncertainty in docking such that a robot may need several attempts before successfully docking the recharging station.

The main contribution of our work is that we designed a probabilistic recharging model, which considers the uncertainty in the recharging procedure. With our probabilistic recharging model, we can fast and accurately estimate the average reward that the robot may gain in real execution. To the best of author's knowledge, this is the first model that tolerates imperfect docking mechanism. When modelling the rewards, we apply a discounting parameter with respect to time. Namely, the earlier the work is done by the robot, the more reward the robot can gain. To see how accurate our probabilistic estimation can be, we conducted several simulations where we saw that our method can accurately estimate the reward that a robot would gain within constant computational time.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

2. APPROACH

We model the environment as three distinct, spatially separated stations: (1) work station, (2) pre-stand station and (3) recharging station. All three stations are static. The pre-stand station is a stepping stone of recharging station. If the robot fails to dock, it will return to the pre-stand station before trying docking again. Traveling from recharging station to pre-stand station takes time τ_3 while traveling from pre-stand station to recharging station takes time τ_2. Transition between working station and pre-stand station takes τ_1 for a single trip. In the working station, we model robot's work consisting of sequence of discrete steps, where τ_w stands for the duration for each step. For a piece of work that consists k steps, the time for working will be $k\tau_w$. At the recharging station, we denote the recharging time as τ_r.

We also define $\dot{e}_r$ as the charging rate in the recharging station, $\dot{e}_d$ as the energy consuming rate when robot travelling between any two stations, $\dot{e}_w$ as the energy consuming rate when the robot is working. The robot gain reward during its working. The reward can be computed as $\int_{t_0}^{t_0+\Delta} \beta^t \, \mathrm{d}t$, where t_0 is the time when robot starts working, Δ is the total working time and β is a discounting parameter and $\beta \in (0,1)$. We define a *Cycle* as a period between two successive successful docking in the recharging station. The time length for a i'th cycle can be computed as $t_c^i = \tau_r^i + \tau_3 + \tau_1 + k_w^i \tau_w + \tau_1 + \tau_2 + (k_d^i - 1)(\tau_2 + \tau_3)$, where $k_w^i \in \mathbb{R}^+$ is the number of steps of working and $k_d^i \in \mathbb{N}^+$ is the number of attempts to dock. We define a *Trail* as a sequence of cycles such that the trail ends up if the robot uses up all its energy before successfully docking. The total time duration of a trail can be computed as $t_t^j = \sum_{i=1}^{M} t_c^i$. Finally, we define *Round* as a sequence of trails of which the time duration can be defined as $t_r^l = \sum_{j=1}^{N} t_t^j$.

To model the uncertainty in docking, we define $\mathbb{P}_d$ as the probability of successful docking in each attempt. Hence the probability that the robot successfully docks in the k'th trial (fails in all previous trials) is $\mathbb{P}(k) = (1-\mathbb{P}_d)^{k-1}\mathbb{P}_d$. In each cycle, we denote $\tau_{w\to r}$ as the time duration for robot from leaving the working station to successfully docking on the recharging station. The expectation of $\tau_{w\to r}$ can be computed as $\mathbb{E}[\tau_{w\to r}] = \sum_{k=1}^{\infty}(1-\mathbb{P}_d)^{k-1}\mathbb{P}_d(\tau_1 + \tau_2 + (k-1)(\tau_3 + \tau_2))$, where k is the number of docking attempts. In each cycle, the robot gains reward by working in the working station. We denote the reward in the i'th cycle as $R_i = \int_{T^i+\tau_r^i+\tau_3+\tau_1}^{T^i+\tau_r^i+\tau_3+\tau_1+k_w^i\tau_2} \beta^t \, \mathrm{d}t = \frac{1}{\ln(\beta)}\beta^{\tau_r^i+\tau_3+\tau_1+T^i}(\beta^{k_w^i\tau_w} - 1)$, where T_i stands for the starting time of the i'th cycle and k_w^i stands for the total number of steps for working.

We use $\epsilon_{w\to r}$ to represent the remaining energy at the moment when the robot needs to leave the working station for the regarding station. We define E_i as the energy of the robot at the moment when the robot successfully docks the charging station in the i'th cycle. Assuming robot takes k_d^i trials for a successful docking, the relationship between $\epsilon_{w\to r}$ and E^i is: $E^i = \epsilon_{w\to r} - \dot{e_d}(\tau_1+\tau_2+(k_d^i-1)(\tau_2+\tau_3))$. Note that in order to make sure $E^i \geq 0$, there exists the maximum number that the robot can try to dock before using up all energy: $I_{max} = (\epsilon_{w\to r}/\dot{e_d} - \tau_1 - \tau_2)/(\tau_2+\tau_3)+1$. When $k_d^i \geq I_{max}$, we simply set $E^i = 0$. Thus, the expectation of E^i can be computed as $\mathbb{E}[E^i] = \sum_{j=1}^{I_{max}}(1-\mathbb{P}_d)^{j-1}\mathbb{P}_d(\epsilon_{w\to r} - \dot{e}_d(\tau_1 + \tau_2 + (j-1)(\tau_2+\tau_3)))$. We also define the energy that the robot has after recharging as $\epsilon_{r\to w}$. The time for recharging in i'th cycle is thus $\tau_r^i = (\epsilon_{r\to w} - E^i)/(\dot{e}_r)$. The averaging recharging time is then $\mathbb{E}[\tau_r^i] = (\epsilon_{r\to w} - \mathbb{E}[E^i])/\dot{e}_r$. With $\epsilon_{w\to r}$ and $\epsilon_{r\to w}$, we can compute the total time the robot used for working: $k_w^i\tau_w = (\epsilon_{r\to w} - \epsilon_{w\to r} - \dot{e_d}(\tau_3+\tau_1))/\dot{e}_w$.

Given E^i, E^{i+1} and T^i, we can compute T^{i+1}, the starting time of the $(i+1)^{th}$ cycle as follows: $T^{i+1} = T^i + \frac{\epsilon_{r\to w} - E^i}{\dot{e}_r} + \tau_3 + \tau_1 + \frac{\epsilon_{r\to w} - \epsilon_{w\to r} - \dot{e_d}(\tau_3+\tau_1)}{\dot{e}_w} + \frac{\epsilon_{w\to r} - E^{i+1}}{\dot{e}_d}$, which is the sum of the starting time of i^{th} cycle, the time duration for recharging, the time duration for traveling from recharging station to working station, the time duration for working, the time duration for traveling from the working station to the recharging station and the total time duration used for getting a successful docking. Thus, the expectation of $T^{i+1} - T^i$ can be computed as $\mathbb{E}[T^{i+1} - T^i] = \frac{\epsilon_{r\to w} - \mathbb{E}[E^i]}{\dot{e}_r} + \tau_3 + \tau_1 + \frac{\epsilon_{r\to w} - \epsilon_{w\to r} - \dot{e_d}(\tau_3+\tau_1)}{\dot{e}_w} + \frac{\epsilon_{w\to r} - \mathbb{E}[E^{i+1}]}{\dot{e}_d}$.

To model the average reward the robot can get in the i'th cycle, we compute $\mathbb{E}[R^i] \approx \frac{1}{\ln(\beta)}\beta^{\mathbb{E}[\tau_r^i]+\tau_3+\tau_1+\mathbb{E}[T^i]}(\beta^{k_w^i\tau_w} - 1)$, where $k_w^i\tau_w = (\epsilon_{r\to w} - \epsilon_{w\to r} - \dot{e_d}(\tau_3+\tau_1))/\dot{e}_w$.

We model the expected reward in one trail as $\mathbb{E}[R_\infty] = \mathbb{E}[R_0](\sum_{j=1}^{\infty}\beta^{j(\mathbb{E}[T^{i+1}-T^i])}) = \mathbb{E}[R_0]\frac{1}{1-\beta^{(\mathbb{E}[T^{i+1}-T^i])}}$. Note that in each cycle, there is non-zero probability that the robot will use up all its remaining energy before successfully docking. For this situation, we simply assume the reward at this cycle is zero. Thus, the expectation of the total reward in each trail can be approximated as $\mathbb{E}[R_\infty] = \mathbb{E}[R_0]\frac{1}{1-\beta^{\mathbb{E}[T^{i+1}-T^i]\mathbb{P}_t}}$, where $\mathbb{P}_t = \mathbb{P}(k_d^i \leq I_{max})$. All the approximations come from the fact that we approximate $\mathbb{E}[f(x)] \approx f(\mathbb{E}[x])$ based Jensen inequality. We use $\mathbb{E}[R_\infty]$ to approximate the reward that a robot can gain during real execution.

3. EXPERIMENTS AND CONCLUSION

We evaluated our probabilistic model in simulated environments. To model the environment, we sample τ_1, τ_2, τ_3 from Gaussian distributions. We vary the successful docking probability $\mathbb{P}_d$ from 0.1 to 1.0. We also vary the discount parameter β from 0.9 to 0.99. Given a pair of $\epsilon_{w\to r}$ and $\epsilon_{r\to w}$ and the parameters (τ_1, τ_2, τ_3, $\mathbb{P}_d$ and β) that are used to model the simulated environment, we computed the expected reward $\mathbb{E}[R_\infty]$ based on the formulas in Sec. 2. For comparison, we also compute the real average reward of a trail by running 10,000 simulations and then computing the average from the simulations .

Our experiments showed that the relative error between the expected reward and the real average reward is within 5%. Hence, the expected reward $\mathbb{E}[R_\infty]$ can serve as a fast (constant computational time) and reasonable estimation of the real average reward, which usually needs thousands of simulations to estimate. In order to find the pair of $\epsilon_{r\to w}$ and $\epsilon_{w\to r}$ that results the high reward, we maximize the expected reward $\mathbb{E}[R_\infty]$ with respect to $\epsilon_{w\to r}$ and $\epsilon_{r\to w}$ using optimization algorithms. In our implementation, we used Gradient ascent. Hence, we can get a pair of $\epsilon_{r\to w}$ and $\epsilon_{w\to r}$ that can lead to high average reward in real execution. Note that our method cannot promise global optimality for $\epsilon_{r\to w}$ and $\epsilon_{w\to r}$ since $\mathbb{E}[R_\infty]$ may not be convex.

We have developed a probabilistic recharging model that can *a priori* consider the uncertainty in the docking procedure. For any given pair of $\epsilon_{r\to w}$ and $\epsilon_{w\to r}$, instead of estimating the real average from thousands of simulations, which takes a long time to compute, our method provides an accurate estimation of the average reward within constant computational time. In the future work, we will look for real applications on physical robots. We also plan to extend the current model to a adaptive model where ($\epsilon_{r\to w}$, $\epsilon_{w\to r}$) have to be learnt online (e.g., reinforcement learning). Finally, we will formulate the probabilistic recharging model in a Partial Observable Markov Decision Process (POMDP) framework, which could lead to globally optimal solutions.

4. REFERENCES

[1] R. Humza and O. Scholz. A case study on self-sufficiency of individual robotic modules in an arena with limited energy resources. *The Third International Conference on Adaptive and Self-Adaptive Systems and Application(ADAPTIVE)*, pages 29–35, 2011.

[2] A. Munoz, F. Sempe, and A. Drogoul. Sharing a charging station in collective robotics. *LIP6 research reports*, 2002.

[3] F. Sempe, A. Munoz, and A. Drogoul. Autonomous robots sharing a charging station with no communication: a case study. *Proceedings of the 6th International Symposium on Distributed Autonomous Robotic Systems(DARS)*, 2002.

[4] M. C. Silverman, B. Jungand, D. Nies, and G. S. Sukhatme. Staying alive longer: Autonomous robot recharging put to test. *Cres-03-015, Center for Robotics and Embedded Systems, Univeristy of Southren California*, 2003.

[5] M. C. Silverman, D. Nies, B. Jung, and G. S. Sukhatme. Staying alive: A docking station for autonomous robot recharging. *In IEEE International Conference on Robotics and Automation, Washington D.C.*, pages 1050–1055, 2002.

[6] J. Wawerla and R. T. Vaughan. Near-optimal mobile robot recharging with the rate-maximizing forager. *In European Conference On Artificial Life(ECAL), Listbon, Portugal*, pages 776–785, 2007.

[7] J. Wawerla and R. T. Vaughan. Opitmal robot recharging strategies for time discouted labour. *Proc. of the 11. Int. Conf. on Artificial Life, Winchester, England, Aug. 5-8*, 2008.