# Strategic Searching Approaches in a Multi-Robot System

Yan Meng, Ke Cao

Department of Electrical and Computer Engineering
Stevens Institute of Technology, Hoboken, New Jersey, USA

**Abstract**. In a partially known dynamic environment, two multi-robot strategic searching approaches are proposed in this paper: utility greedy approach and game theoretic approach. It is assumed that a-priori probabilities of the targets' distributions are provided. A one-step dynamic-programming is used to formalize the utility functions for both approaches, which not only depends on the targets' distribution probabilities, but also on travel cost. Extensive simulation results shows that the proposed approaches are more efficient and robust compared to the other heuristic searching strategies, and game theoretic approach guaranteed better worst-case performance and be more robust to handle the environmental uncertainty.

## 1 Introduction

To be more efficient for the searching task in an Urban Search and Rescue (USAR), it is reasonable to assume that some partial information are available either through distributed sensors installed in the area, or based on some heuristics from human beings in the emergency situation. One natural way to capture the available information is to represent it as the likelihood of the target presence in the search space. [1] [2] [3] proposed different searching strategies based on the priori probabilities of the target distribution. However, these strategies are applied only to a single individual robot. As we know, multi-robot systems are more desirable in some scenarios, such as exploration, USAR, and hazardous environments, due to the robustness, stability, adaptability, and scalability.

Some researches have been conducted on the multi-robot searching. The interaction between the robots is relative simple in [4][5][6] due to the special configurations. An alternative approach that proved to be more efficient consists of discretizing time and partitioning the continuous space into a finite collection of cells. The search problem is then reduced to deciding which cell to visit at each time interval.

For a multi-robot system, the mutual interactions between individual robots sharing a common workspace could be much more complex in general cases. The game theory seems to be a convenient tool for modeling and solving multi-robot interaction problems. In principle game theory can be applied to solve the coordination problem and some researches have conducted using the game theory [7], [8], [9]. A pursuit-evasion problem as a Markov game is described in [9], which is the generalization of a Markov decision process to the case when the system evolution is governed by a transition probability function depending on two or more player's actions. This probabilistic setting makes it possible to model the uncertainty affecting the player's motion.

In this paper, we propose game-theory based searching strategy for a multi-robot system in a partially known dynamic searching area. The searching area is partitioned into different regions, where the initial probability of the target distribution in each

Meng Y. and Cao K. (2006).
Strategic Searching Approaches in a Multi-Robot System.
In *Proceedings of the 2nd International Workshop on Multi-Agent Robotic Systems*, pages 106-111

region is given. When the searching task starts, the probability for each region will be updated dynamically based on the new searching status. A one-step dynamic programming is applied to formalize the utility function for each robot to reduce the computation time. Based on this utility function, the decision making approach and a non-zero-sum game theory are proposed to coordinate a team of robots for the searching task. In addition, event-based discretisized approach rather that the fixed-time iteration is applied to further decrease the computation complexity. Compared to other multi-robot searching algorithms, the proposed approaches are more efficient and robust than other heuristic algorithms, especially under USAR environment where the wireless network is tend to be unreliable.

## 2 Searching Strategies

Assume there are *N* robots searching for a single target in an indoor area with *J* different regions. The initially priori probability of the target distribution is provided. The searching area is discretized and partitioned into a finite collection of cells. The robots can only be allowed to move to the adjacent ones. Initially, *N* robots start from the entrance of the searching area. The team of robots can be homogeneous or heterogeneous in terms of their searching capabilities.

To make the decision making procedure to be numerical tractable, the discretization of the searching procedure does not depend on a pre-defined fixed time interval, instead, it only depends on the event. A robot is busy when it is searching inside a region. Otherwise, it is free. Initially all robots are set as free. A new event happens when a robot enters a region or finishes searching its current region, which can trigger the update of the robot state. With this event-triggered discretization, the searching time is updated at each event. Since the robots only communicate with each other upon new event it can reduce the communication overhead significantly compared with the fixed-time-interval discretization method.

### 2.1 Utility Function

The utility can be defined as the searching payoff value by selecting which region to search on the next discrete time. For a multi-robot system, to improve the collective searching efficiency, the utility value of each robot does not only depend on its own payoff value, but also on other robots' decisions.

Obviously, the utility associated with each robot depends on the probability of the target at each region. The higher the probability, the higher the utility value should be. When the searching task starts, the probabilities of all regions are updated dynamically based on the current searching results. For example, if one robot finishes searching in region 1 without detecting the target, then the initial probability of the target in region 1 is evenly distributed by all of unsearched regions on next discrete time, which can be expressed in the following equation.

$$p_i^{n+1} = 0;\ p_j^{n+1} = \frac{p_j^n}{1 - p_i^n},\ j \neq i,\ j=1, 2, \ldots, J. \tag{1}$$

where $p_i$ represents the priori probability of the target in each region, *J* is the maximum region number and *n* is the current discrete time. However, this priority-only based approach may tend to achieve the highest priority irrespective of the difficulty of that goal. Therefore, we add the travel cost to the utility calculations.

The set of decisions made by the robot from 1 to $N$ is denoted by $\mathbf{D} = [d_1, ..., d_n, ..., d_N]$. The set of probabilities of the target from region 1 to region $J$ is denoted by $\mathbf{P} = [p_1, p_2, ..., p_J]$, where $p_i$ represents the priori probability of the target in each region, and $i$ represents the region number. To obtain the optimal solution, a Dynamic Programming Equation (DPE) is applied to define the utility function for robot $n$ as follows.

$$U_n(d_1, d_2, ..., d_N) = f_n(\mathbf{D}, \mathbf{P}, \mathbf{T}_n, \mathbf{T}_c) =$$

$$\begin{cases} 0, & \text{if } p_{d_n} = 0 \\ g(\mathbf{D}, p_{d_n}, T_{nd_n}, T_{d_n}) & \text{if } p_{d_n} = 1 \\ h(\mathbf{D})[g(\mathbf{D}, p_{d_n}, T_{nd_n}, T_{d_n}) + (1 - p_{d_n}) \max\limits_{\hat{d}_n} \{ f_n(\hat{\mathbf{D}}, \hat{\mathbf{P}}, \hat{\mathbf{T}}_n, \mathbf{T}_c) \}] & \textit{otherwise} \end{cases} \tag{2}$$

where $U_n$ is the utility function of robot $n$, $d_1, d_2, ..., d_N$ represent the decisions made by robot 1 to robot $N$. $p_{d_n}$ represents the probability of target detection in region $d_n$ by robot $n$. $g(\mathbf{D}, p_{d_n}, \mathbf{T}_{nd_n}, \mathbf{T}_{d_n})$ represents the payoff gain of robot $n$ searching the region $d_n$, which is defined as follows:

$$g(\mathbf{D}, p_{d_n}, \mathbf{T}_{nd_n}, \mathbf{T}_{d_n}) = \frac{p_{d_n}}{k_1 T_{nd_n} + k_2 T_{d_n}}, \tag{3}$$

where $T_{nd_n}$ and $T_{d_n}$ represent the time required for robot $n$ to navigate from its current position to region $d_n$, and the time required for a robot to cover region $d_n$, respectively. $k_1$ and $k_2$ are scale factors which can be adjusted based on different environmental structures. $(1 - p_{d_n}) \max\limits_{\hat{d}_n} \{ f_n(\hat{\mathbf{D}}, \hat{\mathbf{P}}, \hat{\mathbf{T}}_n, \mathbf{T}_c) \}$ represents the maximum expected utility of robot $n$ by selecting different $\hat{d}_n$ for the rest of the unsearched regions after finishing the region $d_n$ with the assumption that other robots keep their current decision during this recursive procedure.

In general, the utility is zero if the probability of target detection in region $d_n$ is zero. If the probability of the target detection in region $d_n$ is 1, which means this is the last room need to be searched. In this case, the utility function is only related to the payoff value by searching region $d_n$. Otherwise, the utility is a recursive function defined in (2). The dynamic programming is intractable for large-scale region numbers. To reduce the computational time, one-step dynamic programming solution is applied in Equation (2). The average expected teammate contribution is computed as the contribution that the teammate would make from its current pose. This approximation is reasonable when each step is relatively small.

Considering the situation that several robots may choose the same region simultaneously based on their own utility functions, which may decrease the overall searching performance. We define a factor $h(\mathbf{D})$ as follows:

$$h(\mathbf{D}) = \begin{cases} \dfrac{T_{ni}}{T_{ni} + \sum T_{mi}}, & \textit{if} \quad d_m = d_n, m \neq n, m = 1,2,...,N. \\ 1 & \textit{otherwise} \end{cases}, \tag{4}$$

where $T_{ni}$ is the travel time for robot *n* from its current position to the selected region *i*, $\Sigma T_{mi}$ is the total travel time for robot *m* (*m* can be multiple robots from 1 to *N*, except *n*) from their current positions to region *i*. The definition of $h(\mathbf{D})$ actually embeds the coordination between the multiple robots by cutting down the utility value, eventually it helps to prevent multiple robots picking up the same region simultaneously and improves the overall searching efficiency.

### 2.2 Searching Strategies

Utility greedy (UG) approach is to select the next searching region with the highest utility value calculated by (2). If more than one region has the same highest utility values, the robot will randomly pick one from them. For the game-theory based (GT) approach, it is modeled as a multi-player cooperative non-zero-sum game since the coordination is embedded into the utility function through (4). The players choose their strategies simultaneously at the beginning of the game. Although the overall process of the searching is dynamic, we can treat it as a sequence of static game at each discrete time.

According to the current positions of robots, obstacles, and the probability of the target in each region, the utility matrix is calculated. Then the Nash Equilibrium (NE ) is applied for this nonzero-sum game. When no pure Nash equilibrium strategy exits, a max-min method is applied to calculate the mixed-strategy equilibrium. Let $p_n(j)$ denotes the probability of robot *n* choosing region *j*, and $\mathbf{P}_n = [p_n(1) \quad p_n(2) \quad \ldots \quad p_n(J)]^T$, we have $\sum_{j=1}^{J} p_n(j) = 1, \quad n = 1,2,\ldots,N$.

Utility matrix $\mathbf{U}_n$ is a *N*-dimensional matrix for *N* robots, where there are *J* (region number) units at each dimension, and each cell of the matrix consists of *N* utility values for each robot at the corresponding position. $\mathbf{P}_n$ can be estimated by solving the above linear equation, which can be simplified as $\mathbf{U}_n^m \times \mathbf{P}_n = [0 \quad 0 \quad 0 \quad 1]^T$. Since the game is a finite strategic-form game, the existence of the mixed-strategy equilibrium can be guaranteed. The region with the highest probability is chosen for the next step.

## 3 Simulation Results

To evaluate the performance of the proposed game strategic searching approaches, the simulations with two robots using MATLAB have been conducted. The searching area is sketched as a square of 100 x 100 cells with multiple regions distributed. It is assumed that each robot takes 1 time unit to traverse one cell in the simulation. The searching time is calculated by the time units from the starting time till the target is found.

Two heuristic searching strategies are proposed for comparison. First one is called *randomly selection (RS)* approach, where each robot randomly selects the next region to search from the unsearched regions. Second one is called *probability-based (PB)* approach, where each robot only picks the region with the highest probability at any discrete time as its next objective region. If more than one region has the same highest probability, the robot randomly picks one from them.

First set of simulation is conducted, where 1000 runs for all four approaches are implemented. The simulation results of mean value and the corresponding variance of searching times are shown in Fig. 1. It is obviously that the searching time of the UG and GT approaches are much less than that of the PB approach since the travel and searching time was ignored in the latter case, and the RS has the worst performance.

The performance of GT and UG approach are competitive, which is mainly depends how the game theory is applied in the simulation. When both robots are free, the utility matrix is calculated and their searching decisions are computed based on the game strategy. However, if one robot is busy and the other one is free, the free robot will make its decision only based on the utility value instead of starting a new game. The motivation for this simplified procedure is to reduce the computational time. In addition, since the other robot is busy in searching a region, it would make more sense to let it finish its current searching instead of reselecting the searching region again due to the new status. If the second case happens very often, then the overall performance of game strategy tends to close to that of the utility greedy.

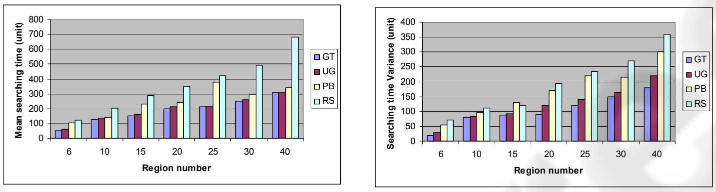


**Fig. 1.** Searching time *(mean and variance)* with different configurations.

In the real world, the prior probability of the target distribution is not accurate. To explore the system robustness to the prior probability variations, another 1000 runs of simulations are conducted for each strategy, where the target is distributed in the searching regions with the variations of the priori probability from 10% to 50%. The simulation results in the configuration of 10-room are shown in Fig. 2. As can be seen, the PB approach is very sensitive to the probability variation since the probability is the only criteria for the robot to make searching decisions. It makes sense that the probability variation has no effect on the RS approach at all. The GT and UG approach are much more robust than the PB approach, where the GT beats UG in both mean searching time and variance. This indicates that the GT is more robust to handle environmental uncertainty compared to UG, although we have to pay the penalty of more computational time for GT.

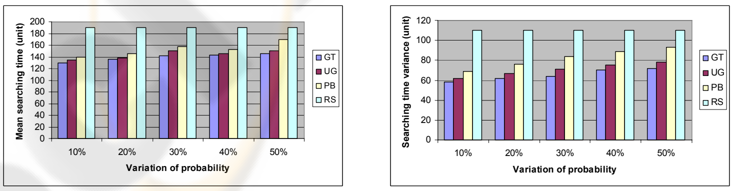


**Fig. 2.** Searching time *(mean and variance)* with different variations in probability of target distribution with a 10-room configuration.

## 5 Conclusions

Utility greedy and game-theory based strategic searching approaches are proposed in this paper for a cooperative multi-robot searching task. Comparing to other heuristic searching strategies, the simulation results demonstrated that the proposed two approaches are more efficient and robust. In addition, the game theory has guaranteed better worst-case performance and be more robust to handle the environmental uncertainty. Another advantage of using game theory based approach is that the explicit communication between the robots can be reduced significantly due to their mutual rationality. Therefore, it can be applied to some emergency scenarios, such as USAR, where the RF communication tends to attenuate or even broken.

Our preliminary simulation only contains two homogeneous robots and one target in the searching task. The proposed algorithm can easily be extended to the heterogeneous robots with different moving speeds and local sensing capabilities by setting up different travel and covering time for each robot. Our future research will focus on multi-target searching by a large scale robot ream.

## References

1. Murphy, R.R.: Biomimetic search for urbane search and rescue. IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 3, pp. 2073-2078 (2000).
2. Bourgault, F, Furukawa, T., and Durrant-Whyte, H.F.: Coordinated decentralized search for a lost target in a bayesian world. IEEE/RSJ International Conference on Intelligent Robots and Systems (2003).
3. Lau, H, Huang, S., and Dissanayake, G. :Optimal search for multiple targets in a built environment. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'05), Edmonton, Alberta, Canada (2005).
4. L′opez-Ortiz1, A., and Schuierer, S. : Online parallel heuristics and robot searching under the competitive framework. 8th Scandinavian Workshop on Algorithm Theory, Turku, Finland, July 3-5 (2002).
5. Kao, M., Reif, J. H., and Tate, S. R.:Searching in an unknown environment: An optimal randomized algorithm for the cow-path problem. In Proc. 4th ACM-SIAM Sympos. Discrete Algorithms, pages 441–447 (1993).
6. Baeza-Yates, R., Culberson, J., and Rawlins, R.: Searching in the plane. Informationand Computation, 106:234–252 (1993).
7. Skrzypczyk, K.: Game theory based target following by a team of robots. Fourth International Workshop on Robot Motion and Control, June 17-20 (2004).
8. LaValle, S., and Hutchinson, S.: Path selection and coordination for multiple robots via Nash equilibria. Proc. IEEE Int. Conf. Robot and Automation, pp. 1847-1852 (1994).
9. Hespanha, J. P., Prandini, M., and Sastry, S.:Probabilistic pursuit-evasion games: a one-step Nash approach. In proc. of the 39th conf. on decision and control, vol. 3, (2000).