Cooperative Target Searching and Tracking via UCT with Probability Distribution Model

Ruoxi Qin, Tian Wang*, Haotian Jiang,
Qianhong Yan, Weikang Wang
School of Automation Science and Electrical Engineering,
Beihang University, Beijing, China
E-Mail: ruoxiqin@yahoo.com, wangtian@buaa.edu.cn,
jianghaotian@buaa.edu.cn, yanqianhong@yeah.net,
weikangthebest@gmail.com

Hichem Snoussi
Institute Charles Delaunay-LM2S-UMR STMR 6279 CNRS,
University of Technology of Troyes, France
hichem.snoussi@utt.fr

Abstract

As Unmanned Aerial Vehicle's (UAV) battery life and stability develop, multiple UAVs are having more and more applications in the uninterrupted patrol and security. Thus UAV's searching, tracking and trajectory planning become important issues. This paper proposes an online distributed algorithm used in UAV's tracking and searching, with the consideration of UAV's practical need to recharge under limited power. We propose a Quantum Probability Model to describe the partially observable target positions, and we use Upper Confidence Tree (UCT) algorithm to find out the best searching and tracking route based on this model. We also introduce the Teammate Learning Model to handle the nonstationary problems in distributed reinforcement learning.

***Index Terms*—UCT Planning, UAV Tracking, Quantum Probability Distribution**

I. Introduction

At present, UAV searching is having more and more applications in civil and military[1] fields. UAVs can access the place too dangerous or inaccessible for human beings to go. Moreover, UAVs can patrol or track automatically and continuously, which can greatly reduce humans workload. Many researches have been done in this field, but there are also many important practical issues which were neglected.

The task of UAV patrolling can be divided as searching and tracking. Current literatures about searching missions concern about the exploration of irregular topography[2][3] and searching in an area whose dimension is known[4]. But they only handled searching fixed target without considering tracking. In [5], Yu *et al*. used dynamic occupancy grid to implement UAV tracking in urban areas. Although tracking and searching are combined in the paper, it did not consider the energy of UAV. Since UAV's duration is limited at present, the contradiction between searching and charging is a crucial problem.

In tracking missions, current literatures have proposed many tracking algorithms of a single UAV[6][7][8][9][10][11][12]. But since a single UAV has limited field of vision, cooperative tracking with multiple UAVs is becoming a more popular field in research[13][14][15]. However, only few documents about cooperative tracking consider the charging problem of UAV in a searching mission on a broad area. Besides, some articles concerned about the possibility to succeed in a tracking process. But losing targets is inevitable for UAVs in complex practical conditions, so the ability to search again after losing targets is necessary for UAV to patrol automatically.

Based on the truth above, we propose a tracking and searching algorithm to help the UAV fleet to overcome all these listed problems.Sperciﬁcally, our strengths over others' solutions are as follows:

1) Take the power of UAV into consideration. This is not just an easy task like ordering the UAV flies back when it used up its power. Instead, it should move back to recharge when there is no emergency so that it will have enough power in important circumstance.

2) Keep tracking with temporally losing target. We design an innovative quantum probability model to help estimate the location of the target. Thus our algorithm can guide the UAV fleet no matter whether they observe the target, which means losing target tentatively would not terminate our program. This situation occurs frequently in reality but many researches before assumes they always know the exactly location of the target.

3) On-line distributed cooperative planning algorithm. Our algorithm could guide the fleet working together to search and track the target. It is more efficient than doing this task by only one aircraft. Also, our solution avoids the curse of dimensionality, so we can deploy as many UAVs as we want for the task.

The subsequent sections are organized as follows. Section II describes and specifies the patrol and trace problem. Section III demonstrates our planning algorithm by dividing them into three parts: section III-A and section III-B illustrate two important components of our algorithm. Followed on with section III-C explaining the main algorithm and how it use the components above. After that, section IV and section V demonstrate the experiment which tests our algorithm and analyzes the experiment result.

978-1-5090-4165-7/16/$31.00 ©2016 IEEE

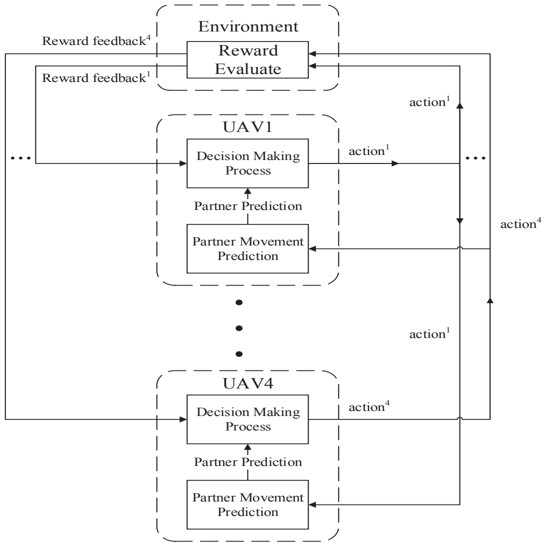

Fig. 1: The structure of UAV moving algorithm.

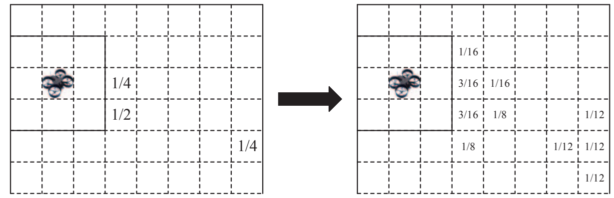

Fig. 2: Probability distribution transition if no target found. The probability that the target in one position is shown in the grid.

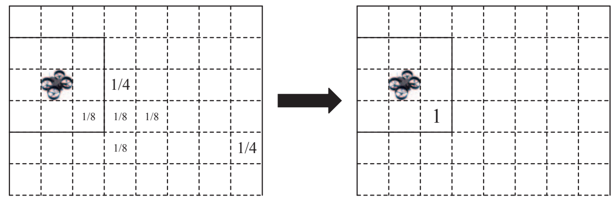

Fig. 3: Probability distribution transition if target is found within the UAV's horizon. The probability that the target in one position is shown in the grid.

II. Problem Specification

UAV automatic searching and route planning problem is a Partial Observed Markov Decision Process (POMDP). The goal of the fleet of UAVs is to find the moving target as soon as possible and tail after it. Each UAV can only monitor a finite visual field and it cannot know the target's location if none of them finds the target. But they can share the information with each other immediately. Every UAV possesses limited electric power and it can recharge back to base. We assume that UAV can fly in the direction as it planned without consideration of disturbance.

III. Cooperative Searching and Tracking Strategy

Our algorithm can be divided into three parts: Quantum Probability Module (QPM), Partner Movement Prediction (PMP) and Decision Making Procedure (DMP), as shown in Fig. 1. The QPM calculates the probability of target being in specific position. The PMP model provides the estimate of the UAV partner's next move. These two independent procedures are called by the main Decision Making Procedure (DMP), which employs Upper Confidence Tree (UCT) algorithm to decide the moving direction. Thus the decision is made based on the knowledge of target's location (or moving history) and the movement of UAV partners.

A. Quantum Probability Model

Since the exact positions of the target may not be known, we imitate the methods in quantum mechanics, transforming the exact positions of target into a probability distribution dispersed in the whole patrolling region, and establish our Quantum Probability Module of the position of target. This model can describe the probability of the appearance of the target in the searching region, which can give guidance to UAV. $P(t_i, p)$ means the probability that target appear at place p at time t_i. $P(t_i, p)$ changes according to Probability Transfer Rule (PTR) as time goes by, and PTR is determined by UAVs observation $O(i)$ and the prior knowledge of the movements of the target. In this paper, we assume target's movement are uniform random, which means the probability that target moves *North, South, West, East* or *Park* (stay at current position) is equal. This can be expressed as follows:

$$P(t_{i+1}, p_{s,j}) = \frac{P(t_i, p|O_i)}{N_s(t_i, p)} \tag{1}$$

In the equation, $p_{s,j}$ means the next position the target can reach at t_i, and $N_s(t_i, p)$ represents the number of $p_{s,j}$.

Thus the PTR can be illustrated as follow:

(1) If all UAVs do not find the target(as shown in Fig. 2), then update the probability in every position p according to Equ.1.

(2)If any UAV finds the target(Fig. 3), the probability distribution of the target's positions will collapse into a certain state [16]: The possibility of the targets position is 1,

$$P(t_{i+1}, p_{s,j}) = \begin{cases} 1 & p_{s,j} \text{ is the target position} \\ 0 & \text{else.} \end{cases} \tag{2}$$

B. Partner Movement Prediction

To avoid curse of dimensionality invoked by the increase of UAVs number, we choose distributed algorithm to let each UAV make decisions independently. In order to solve the Nonstationary problem[17][18] caused by the interactions between UAVs decisions, we introduce Partner Movement Prediction module (PMP) to help UAV predict partners behaviors when applying UCT in decision-making phase and let UAV cooperate in hunting.

Figure4 illustrates Partner Movement Prediction model. Every UAV observes and records its partners behaviors a and its current state s (including its battery level and the probability distribution of surrounding positions) in each time step, and counts the times of visiting this action or state as $N(a^*, s)$

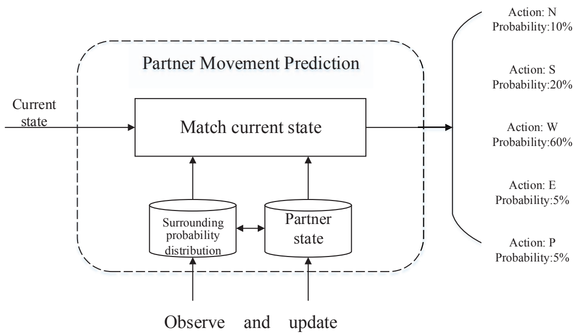

Fig. 4: The partner movement prediction model.

and $N(s)$ separately. With more record, action frequency can be regarded as probability of next movement under the same state s, so we can predict partner movement a^* according to $P(a^*, s)$ in which s matches current state in UCT simulation phase:

$$P(a^*, s) = \frac{N(a^*, s)}{N(s)} \tag{3}$$

The definition of "state" of partners influence the prediction accuracy. A suitable definition can make the prediction become more precise and can save computation resources. In this paper, we choose the maximal sum of probability in one direction for the partner UAV as the "state". We denote the decision a of move direction this partner made under state s at time t, and update PMP module according to

$$\begin{aligned} N_t(s) &= N_{t-1}(s) + 1 \\ N_t(a, s) &= N_{t-1}(a, s) + 1. \end{aligned} \tag{4}$$

At the beginning of a patrolling and tracking mission, we initialize probabilities of moving in each directions equal because each UAV knows nothing about its partners.

C. Decision Making Algorithm

We use Upper Confidence Tree (UCT) algorithm[19] to decide the move directions of UAV. The tree policy and default policy of our decision making algorithm are UCB1 [19] and uniform random selection respectively.

In the rollout procedure, we use QPM mentioned in section III-A to conduct simulation. The depth of the simulation is controlled by D, and the times of simulation (number of trajectories) are limited by T. The greater T and D, the more consumption of computation resource and the longer time taken to make decisions. In the simulation process, we apply QPM to decide whether UAV finds the target. For every position p in all UAVs visual fields, the probability that the target appears at p is $P(t_i, p)$. If the target appears, UAV would apply Probability Transfer Rule (2) to update probability distribution; if no target appears, UAV applies Probability Transfer Rule (1) to update probability distribution.

Each UAV can get reward from every simulation, which is defined as follows:

$$R(t_i, O_i) = \begin{cases} 1 & \text{if find target} \\ -1.5 & \text{in risky state} \\ 0 & \text{else} \end{cases} \tag{5}$$

where the risky state means that UAVs remaining power is only enough for it fly back to the charging base.

The whole distributed planning algorithm is shown in algorithm1.

Algorithm 1 Cooperative searching and tracking strategy

Observe UpdateQPM(QPM, observation o)
procedure UCTPLAN(root state s_0,QPM,PMP,D,T)
  for $\tau \leftarrow 1$ to T **do**
    TREEPOLICY($s_0, QPM, PMP, d = D$)
  end for
  return $a = \arg\max_a Q(s_0, a_i)$
end procedure
procedure TREEPOLICY(state s,QPM,PMP,depth d)
  if $d \leq 0$ **then**
    return $q = 0$
  end if
  if $s \in searchTree$ **then**
    return $q =$DEFAULTPOLICY(s,QPM,PMP,$d-1$)
  end if
  selfAction $a^* \leftarrow$ UCB1 selection
  partnerActions $a^i \leftarrow PMP(s)$
  next state s',reward r, simulating observation $o \leftarrow$ Simulation(a^*, a^i, QPM)
  UpdateQPM(QPM, o)
  $q = r + \gamma \times$TREEPOLICY(s', QPM,PMP,$d-1$)
  BACKPROPAGATION(s,a^*,q)
  return q
end procedure
procedure DEFAULTPOLICY(state s,QPM,PMP,depth d)
  if $d \leq 0$ **then**
    return $q = 0$
  end if
  selfAction $a^* \leftarrow$ uniformly at random
  partnerActions $a^i \leftarrow PMP$
  next state s',reward r, simulating observation $o \leftarrow$ Simulation(a^*, a^i, QPM)
  UpdateQPM(QPM, o)
  $q = r + \gamma \times$DEFAULTPOLICY(s', QPM,PMP,$d-1$)
  return q
end procedure

IV. EXPERIMENT OBJECTIVE AND STRUCTURE

A. Experiment Objective

As illustrated in the last section, there are two key parameters which could influence algorithm performance dramatically: the amount of trajectory(T) and depth(D) of searching tree. Thus we conduct this experiment to explore the property of the optimized parameter tuning.

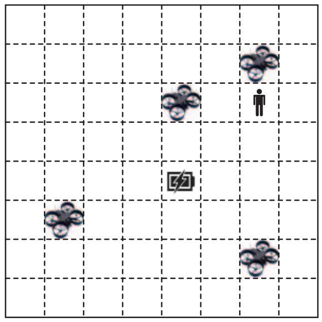

Fig. 5: 8×8 grid world with 4 UAVs, 1 target and 1 base.

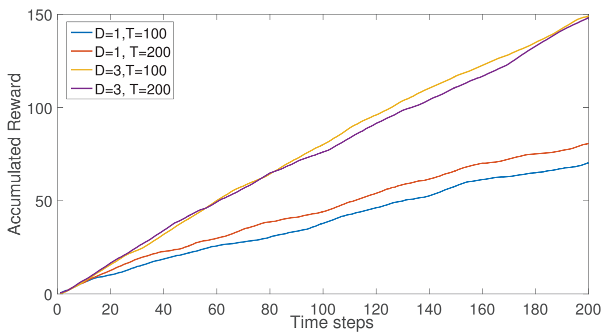

Fig. 6: Experiment result with different T and D.

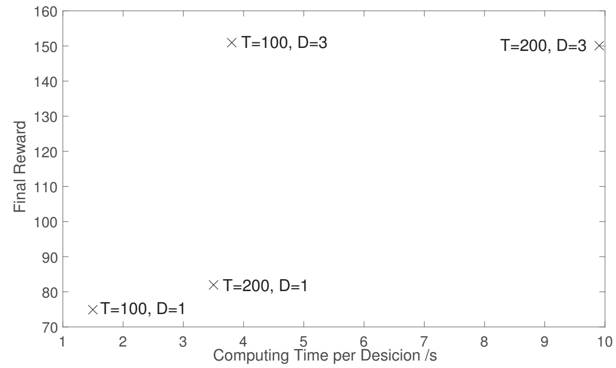

Fig. 7: Computing time comparison[1].

B. Experiment Environment

Fig. 5 shows the searching and tracking experiment. 4 UAVs is responsible to patrol and track 1 target in an 8×8 grid world. Each UAV and the target can choose to move *North, South, West, East* or *Park* (stay at current position) at each time step. Every aircraft can observe their surrounding grids (9 grides in total). The UAV consumes 1 unit of power in each time step. An UAV would crush if it runs out of power. There is a fixed power station near the middle of the map. When UAV flies to the station, it can be full recharged in within 1 time step.

*1) Dynamics.*At the beginning of experiment, 4 UAVs powers locate at the 4 corners in the map with full power. The target is first randomly generated in the map. In each time step, the target choose a direction to move at uniformly random, and then each UAV independently decide their moving directions. These two processes execute in turn. The experiment ends when $steptime = 200$.

2) Objective reward. Different rewards of different states are set as function(5).

3) Observation and prior-knowledge. Every UAV can get its partners location and power information and would share the position of target if found. Of course, they dont know its partners current or future decision, but they can observe and record past moves of their partners.

V. Experiment Results

We test the algorithm performance of every combination of $T = \{100, 200\}$ and $D = \{1, 3\}$ for 30 times in the experiment. The result is shown in Fig. 6,in which the curve indicates how the accumulated reward that UAV group gains varies as time goes by. As can be seen, greater depth D improve the performance remarkably, and the best result reaches 75% of successful tracking ratio. However, greater trajectory T does not contribute a lot to performance improvement, even though it consumes more computing time.

To get better results with limited computation resource, Fig. 7 draws the distribution of the accumulated reward of different (T, D) combinations according to computing consumption. From the figure, we can draw the following conclusion: when computing resource is limited, we should increase the depth D of searching tree in priority in order to achieve better performance.

Besides, no aircraft crash happens during the experiments, which proves that our algorithm successfully manages the power of each UAV.

VI. Conclusion and Future Work

This paper discusses cooperative ground target searching and tracking by multiple UAVs with consideration of limited power. The proposed quantum probability model effectively helps the UCT algorithm to decide moving directions. And the partner learning model makes the fleet searches and tracks cooperatively. The experiment finds out the optimized parameter setting of our algorithm.

Future work involves more experiments of different scales of searching region and various parameter settings, including physical testing by real UAVs. We will also improve PMP model for more accuracy and better cooperation.

Acknowledgment

This work is partially supported by the ANR AutoFerm project and the Platform CAPSEC funded by Région Champagne-Ardenne and FEDER, the Fundamental Research Funds for the Central Universities (Grant No. YWF-14-RSC-102) and the National Natural Science Foundation of China (Grant No. U1435220, 61503017).

[1]Tested by Intel Core i7, 2.80GHz.

REFERENCES

[1] "Uav path planning based on bilateral searching new model," in *International Conference on Modelling, Identification and Control*, 2015.

[2] V. Ablavsky, "Search path optimization for uavs using stochastic sampling with abstract pattern descriptors," in *Aiaa Guidance, Navigation, and Control Conference and Exhibit*, 2003.

[3] B. B. Meng and X. Gao, "Uav path planning based on bidirectional sparse a* search algorithm," in *2010 International Conference on Intelligent Computation Technology and Automation*, 2010, pp. 1106–1109.

[4] B. Liu, S. Singh, R. L. Lewis, and S. Qin, "Optimal rewards for cooperative agents," *IEEE Transactions on Autonomous Mental Development*, vol. 6, no. 4, pp. 286–297, 2014.

[5] H. Yu, K. Meier, M. Argyle, and R. W. Beard, "Cooperative path planning for target tracking in urban environments using unmanned air and ground vehicles," *IEEE/ASME Transactions on Mechatronics*, vol. 20, pp. 541–552, 2015.

[6] S. Zhu, D. Wang, and B. L. Chang, "Ground target tracking using uav with input constraints," *Journal of Intelligent & Robotic Systems*, vol. 69, no. 1-4, pp. 417–429, 2013.

[7] Y. Sang, Z. Cai, Q. Lin, and Y. Wang, "Planning algorithm based on airborne sensor for uav to track and intercept moving target in dynamic environment," in *Guidance, Navigation and Control Conference*, 2015.

[8] S. A. P. Quintero, D. A. Copp, and J. P. Hespanha, "Robust uav coordination for target tracking using output-feedback model predictive control with moving horizon estimation," in *American Control Conference*, 2015.

[9] H. Chen, K. Chang, and C. S. Agate, "Uav path planning with tangent-plus-lyapunov vector field guidance and obstacle avoidance," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 49, no. 2, pp. 840–856, 2013.

[10] T. Yamasaki, H. Sakaida, K. Enomoto, H. Takano, and Y. Baba, "Robust trajectory-tracking method for uav guidance using proportional navigation," in *Control, Automation and Systems, 2007. ICCAS '07. International Conference on*, 2007, pp. 1404 – 1409.

[11] K. Krishnamoorthy, D. Casbeer, P. Chandler, M. Pachter, and S. Darbha, "Uav search & capture of a moving ground target under delayed information," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, 2012, pp. 3092–3097.

[12] N. Regina and M. Zanzi, "Uav guidance law for ground-based target trajectory tracking and loitering," in *IEEE Aerospace Conference*, 2011, pp. 1–9.

[13] T. Sun and M. Xin, "Multiple uav target tracking using consensus-based distributed high degree cubature information filter," 2000.

[14] L. Ma and N. Hovakimyan, "Cooperative target tracking in balanced circular formation: Multiple uavs tracking a ground vehicle," in *American Control Conference*, 2013, pp. 5386–5391.

[15] S. Ragi and E. K. P. Chong, "Uav path planning in a dynamic environment via partially observable markov decision process," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 49, no. 4, pp. 2397–2412, 2013.

[16] G. Greenstein, A. G. Zajonc, and A. Peres, "The quantum challenge: Modern research on the foundations of quantum mechanics." *American Journal of Physics*, vol. 66, no. 5, pp. 455–456, 1998.

[17] A. Gosavi, "Reinforcement learning: A tutorial survey and recent advances," *Informs Journal on Computing*, vol. 21, no. 2, pp. 178–192, 2009.

[18] C. Claus and C. Boutilier, "The dynamics of reinforcement learning in cooperative multiagent systems," in *Fifteenth National/tenth Conference on Artificial Intelligence/innovative Applications of Artificial Intelligence*, 1970, pp. 746–752.

[19] D. Silver and J. Veness, "Monte-carlo planning in large pomdps," in *Advances in Neural Information Processing Systems 23: Conference on Neural Information Processing Systems 2010. Proceedings of A Meeting Held 6-9 December 2010, Vancouver, British Columbia, Canada*, 2010, pp. 2164–2172.