Multi-USV Target Search Algorithm Based on Markov Prediction Model in Ocean Environment

Juniie Xu

School of Mechatronics and Automation, Shanghai University, Shanghai, China Shuxjj97@163.com

Honghao Cao

School of Mechatronics and Automation, Shanghai University, Shanghai, China

Abstract—Unmanned boat swarms are a modern type of marine intelligent equipment with broad applications in ocean observation, maritime patrol, and search and rescue. This paper proposes a cooperative target search strategy for USVs based on the ant colony optimization algorithm to optimize the search trajectory in the marine environment, thereby improving search efficiency and task robustness. The proposed algorithm is simulated in various search scenarios and compared with other heuristic algorithms to verify its effectiveness and feasibility.

Keywords - Markov decision, ant colony optimization, target search, Multi-USV

I. INTRODUCE

As a maritime application, target search tasks require searchers to locate targets using specified search strategies and algorithms in search areas full of uncertainty. These missions date back to World War II[1], where the need for effective search strategies for military search and rescue missions sparked an interest in target search problems. With the development of unmanned systems, the current target search task typically involves using corresponding sensor equipment carried by the unmanned system in the search area, with a specific search strategy and algorithm based on predicted target movement models. The aim is to optimize the search trajectory to minimize search time or maximize the cumulative probability of finding the target. This is a typical NP-hard problem[2] that can be solved using approximate optimization algorithms or heuristic algorithms such as Bayesian optimization, cross-entropy optimization, particle swarm optimization, and ant colony optimization[3-6].

The proposed cooperative target search strategy for USV swarms in this paper refers to the search for targets by the USV swarms in an unknown area, under the constraints of the marine environment. By improving the update and search strategies of the ant colony algorithm, the search path of the unmanned vehicle cluster is optimized to discover the target, thereby enhancing search efficiency and task robustness. Therefore, the main contributions of this paper are:

1) By using the Markov decision process (MDP) along with the target movement probability model and the sensor model of the unmanned ship, it is possible to predict the potential position of the target in an uncertain environment. The predicted target probability map is then updated based on historical information and target probability map. Rui Song*

Artificial intelligence research institute, Shanghai University, Shanghai, China

*Corresponding author: song rui@shu.edu.cn

2) Establish a model of ocean currents to study its impact on the movement and energy consumption of unmanned vehicles. At the same time, based on the ant colony optimization algorithm, its search strategy and heuristic influence are improved to obtain the optimal path for multi-USV search targets. This search algorithm is called MK-ACO.

II. THEORETICAL MODEL

A. Raster probability map

In the target search problem, the search space is often discretized into a grid map[7] of fixed size to simplify the map. The discreteness and limited precision of the raster map make it easy to visualize. Additionally, different types of information such as obstacle information and target existence probability can be stored in each grid. This allows the robot to perceive, plan, and navigate in the environment.

The grid probability map is a method for representing the probability distribution of random variables in a discrete space, and in the context of the target search problem, it can be used to represent the probability of the target's existence in each grid area. The entire search area, denoted by Ω , is discretized into a grid of size $\omega_x * \omega_y$, and the position of the target in the grid map at time t is denoted by ν_t . Hence, the grid probability map at time t can be represented by $b(\nu_t)$. The grid probability map is typically initialized based on prior knowledge, such as the last known position of the target before its signal was lost. The grid probability map is centered on this position and follows a two-dimensional normal distribution. The target search task is then carried out based on the initialized grid probability map $b(\nu_0)[8]$, as shown in Figure 1. Since the target is assumed to be within the search area, the initial grid probability map satisfies $\sum_{\nu_0 \in \Omega} b(\nu_0) = 1$.

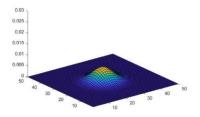


Figure 1 Initializing the grid probability map

B. Target movement model

In the target search task, the motion of the target has certain randomness and uncertainty, which belongs to the random walk model. The model can use the Markov decision process to predict the motion behavior of the target[9-11], and use $P(v_t|v_{t-1})$ to describe the target in the unit time step the probability that from the v_{t-1} position moves to the v_t position. As shown in Figure 2, (a) represents the moving direction of the target predicted by the Markov decision process when it is at each grid position. (b) represents the movement probability model of the target predicted by the Markov decision process. Here, it is assumed that the target can only move in the positive direction, which can be expressed by the following formula:

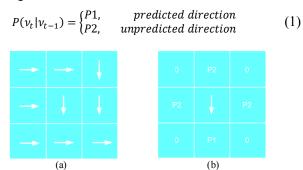


Figure 2 Target movement model

C. Sensor model

In the target search task, the sensor is mainly used to detect the target status information and obstacle information, and help the USVs to realize the search and tracking of the target. In each time step, the observation results of each USV are independent, and the observation result of the u-th USV at time t can be expressed as z_u^t , which is based on the current position of the USV and the target acquired. There are only two types of output of the observation results. When the USV finds the target, $z_u^t = D_t$, otherwise $z_u^t = \overline{D}_t$ means that the USV has not found the target. The relationship between the two is $P(z_u^t = \overline{D}_t | v_t, s_u^t) = 1 - P(z_u^t = D_t | v_t, s_u^t)$. This paper uses an ideal sensor model[12], that is, when the USV and the target are in the same position, it means that the target is found, otherwise the target is not found, which can be expressed by the following formula:

$$P(z_u^t = D_t | v_t, s_u^t) = \begin{cases} 1, & v_t = s_u^t \\ 0, & v_t \neq s_u^t \end{cases}$$
 (2)

D. Motion model of unmanned vehicle

In the grid map, the motion of the USV can be represented by discrete moving steps. In this paper, USVs will move with an eight-connected movement strategy, in addition to the horizontal and vertical directions, USVs can also move to the adjacent grid area along the diagonal. In this case, the action of USVs at each time step is selected as 8 different moving directions {E, NE, N, NW, W, SW, S, SE}. The position of the USV at time t can be represented by s_u^t , and at the end of the target search task, the complete search path of USVs is represented by $s_{1:D}^{0:t}$.

E. Prediction and update of raster probability map

The grid probability map is a probability representation of the possible location of the target in the search space, which is used to guide the search process of USVs. USVs performs priority search according to the area where the probability distribution is concentrated, so as to improve the search efficiency. The grid probability map will be dynamically updated during the search process through the sensor model and Bayesian filtering, so that the search algorithm can adapt to the movement of the target and the change of the environment[8]. The prediction and update of the raster probability map can be expressed by the following formula:

$$\bar{b}(\nu_t) = \sum_{\nu_{t-1} \in \Omega} P(\nu_t | \nu_{t-1}) \tilde{b}(\nu_{t-1})$$
(3)

$$\tilde{b}(v_t) = \prod_{u=1:U} P(z_u^t = \overline{D}_t | v_t, s_u^t) \, \bar{b}(v_t) \tag{4}$$

Among them, $\bar{b}(\nu_t)$ is the probability map predicted at time t, which is obtained from the target probability map and target movement model at the previous moment. $\tilde{b}(\nu_t)$ is the target probability map updated according to the observation results of the USV swarm at time t.

III. Objective Function and Constraint Setting

A. Objective function

In the target search problem, the objective function to evaluate the search trajectory is often chosen to be the maximization of the cumulative probability of finding the target or the minimization of the search time. In this paper, the minimum search time is selected as the optimization objective function. Given the time discretization, the evaluation of the search trajectory is not solely based on the time step required to find the target, but is also related to the target discovery probability at each time step. A higher target discovery probability indicates a smaller minimum search time. Thus, the search trajectory time can be expressed using the following formula:

$$TIME(s_{1:H}^{1:N}) = \sum_{t=1}^{N} P(\overline{D}_{1:H}^{1:t}|s_{1:H}^{0:t})$$
 (5)

Where $P(\overline{D}_{1:U}^{1:t}|s_{1:U}^{0:t})$ represents the probability that the unmanned boat swarm has not yet found the target until time step t. This is related to the location of the unmanned boat cluster and the target. If no target is found at the location of the unmanned boat cluster at time t, then $P(\overline{D}_{1:U}^{1:t}|s_{1:U}^{0:t}) = 1 - p_t$, p_t represents the probability of finding the target for the first time at time t. According to Bayesian theory, the probability that the target has not been detected at time t is related to the predicted probability map and the probability that the sensor model is not detected, which can be expressed by formula (6).

$$\bar{p}_t = \sum_{\nu_t \in \Omega} P(\bar{D}_t | \nu_t, s_u^t) \tilde{b}(\nu_t) = \sum_{\nu_t \in \Omega} \tilde{b}(\nu_t) \tag{6}$$

From $\bar{p}_t = 1 - p_t$:

$$P(\bar{D}_{1:U}^{1:t}|s_{1:U}^{0:t}) = \sum_{\nu_t \in \Omega} \tilde{b}(\nu_t)$$
 (7)

B. Fuel constraints under the influence of ocean currents

In the ocean environment, the speed and direction of the unmanned boat are often changed by the influence of ocean currents, which leads to the deviation of the trajectory of the unmanned boat. However, in the time-discretized target search algorithm, it is required that the unmanned boat swarm maintains the same velocity within a single time step and arrives at the next search area at the same time for search. Assuming that the distance of the unmanned boats to the nearby 8 areas is the same, it is required that the unmanned boat swarms go to the next grid area at the same time at the speed of v_0 . Because the influence of ocean currents at the location of each unmanned boat is different, the impact on it is also different. In order to reach the next grid area at the same time, the actual speed of each unmanned boat is also different. As shown in Figure 3, \vec{v}_c represents the ocean current speed, \vec{v}_0 represents the speed specified by the USV, and its direction angle corresponds to the action selection one by one, which can be expressed as ψ = {0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°}. Therefore, the real speed of the unmanned boat can be obtained by the formula (8), where the angle between the ocean current speed direction and the specified speed direction $\theta = 2\pi - (\psi - \omega)$.

$$Vr = \sqrt{V_0^2 + V_c^2 - 2 * V_0 * V_c * \cos \theta}$$
 (8)

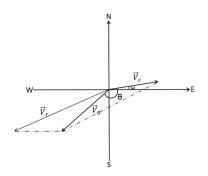


Figure 3. The actual speed of the unmanned boat under the influence of ocean currents

Usually, there are many factors that affect the fuel consumption of unmanned boats, including speed, hull involvement, sea conditions, and engine efficiency, etc., which is a complex nonlinear relationship. In order to facilitate the calculation of fuel consumption, this paper only considers the influence of speed on energy consumption, and ignores other factors, and regards them as a linear relationship, which can be expressed by formula (9). If the energy consumed by the unmanned boat moving forward at v_0 speed in each time step is E_0 , then the actual energy consumption of the unmanned boat can be expressed by formula (10).

$$Q = k * v^3 \tag{9}$$

$$Q = \left(\frac{v_r}{v_0}\right)^3 * E_0 \tag{10}$$

Since the energy state of the unmanned vehicle swarm is different before the start of the target search task, the fuel constraints of the unmanned vehicle need to be considered to improve the robustness of the target search task. Before the start of the target search task, the available fuel of each unmanned boat can be represented by $E = \{E_1, E_2, ..., E_u\}$, considering the fuel constraint is that all the available fuel of the unmanned boat cluster needs to be the target is successfully searched before exhaustion, expressed by formula (11).

$$\exists E_i \in E, E_i > 0 \tag{11}$$

C. Obstacle Avoidance Constraints

In the process of target search, it is necessary to ensure the safety of surface unmanned boats, mainly including the risk of collision between unmanned boats and obstacles and the risk of collision between unmanned boat groups. Therefore, the introduction of obstacle avoidance constraints can effectively avoid collisions, help surface unmanned vehicles plan safer paths, and improve search efficiency. The obstacle expansion method is a commonly used method in path planning, which can solve the collision avoidance problem of mobile robots in complex environments, as shown in Figure 4. By expanding the obstacle, the actual distance between the unmanned vehicle and the obstacle increases, even if there is a sensor error or a control trajectory error, the risk of collision can be effectively reduced.

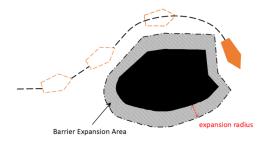


Figure 4. Schematic diagram of obstacle expansion

In the grid map, in order to achieve obstacle expansion, the obstacle expansion table (Obs_Table) is introduced, that is, information about obstacles is stored in each grid. After introducing the obstacle expansion table, it is necessary to turn the grid area around the obstacle into an obstacle buffer zone, and try to avoid the passage of unmanned surface vehicles, as shown in Figure 5. For the expansion between adjacent obstacles, the corresponding values in the grid can be accumulated, as shown in Figure 6, to obtain a complete obstacle expansion table. It can be seen from the figure that the higher the value in the expansion table, the denser the obstacles in the area. In the search strategy of this algorithm, all areas where the value of the expansion table is not 0 are unmanned boats. Therefore, the obstacle avoidance constraint in the target search task can be expressed by formula (12):

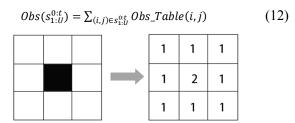


Figure 5. Obstacle expansion method

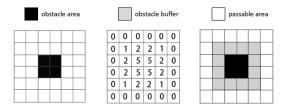


Figure 6. Obstacle expansion table

In summary, the objective function to be optimized and the constraints to be satisfied in this paper are:

$$Min \ TIME(s_{1:U}^{1:N}) = \sum_{t=1}^{N} \sum_{\nu_t \in \Omega} \tilde{b}(\nu_t)$$

$$\exists \ E_i \in E, E_i > 0$$

$$Obs(s_1^{0:t}) = 0$$
(13)

D. Improvement of Ant Colony Optimization Algorithm

Ant colony optimization algorithm is a heuristic optimization algorithm based on the foraging behavior of ants in nature. It guides the ant population to find the optimal path in the environment by simulating the mechanism of ants releasing pheromone. Multi-USV cooperative search can improve the traditional ant colony optimization into distributed search, and the information sharing among UUV clusters can support UUV to receive search information and environmental information in real time, so as to dynamically adjust its own search strategy. The following is an introduction to the ant colony optimization search strategy based on the collaborative improvement of unmanned vehicles.

1) Pheromone Update Strategy

s.t.

In the ant colony algorithm, pheromone is a positive feedback mechanism, which is the main mechanism for the ant population to make decisions and cooperate with each other in the search process. Before the algorithm starts, the pheromone needs to be initialized first, usually the pheromone concentration in all grids is set to a small positive number to ensure the randomness and global search ability of the search process. The update of pheromone is mainly divided into two parts, enhancement and volatilization.

At the end of each iteration, the pheromone concentrations in all grids will volatilize to a certain extent, ensuring the randomness and dynamics of the search process. Suppose there is a volatilization coefficient $\rho \in [0,1]$, the volatilization formula of pheromone concentration is shown in (16), where $\Delta \tau$ represents the enhanced part of pheromone, and at the same time, the larger the volatilization coefficient, the more volatilized pheromone Faster, the stronger the randomness of the search process, and the lower the convergence speed; the smaller the volatility coefficient, the stronger the certainty of the search process, but it may fall into a local optimum.

$$\tau_t = (1 - \rho)\tau_{t-1} + \Delta\tau \tag{14}$$

In the target search problem, the strengthening of pheromone is related to the optimal solution in the current iteration, and the pheromone in the grid corresponding to the optimal path route_ib obtained in the current iteration will be strengthened, and the size of the strengthening is the search time corresponding to the path solution The reciprocal of TIME ib,

the smaller the search time is, the greater the pheromone concentration of the corresponding search path will be, as shown in formula (17):

$$\Delta \tau = \Delta \tau (route_ib) + \frac{1}{TIME_ib}$$
 (15)

2) Heuristic information

In this paper, in order to better guide the USV swarm to search for targets, a heuristic formula based on search space is designed. As shown in Figure 7, the heuristic information is related to the states in all the grids in the direction of the USV's current action selection. The heuristic information of a movable direction is related to the probability of the existence of targets in all grids in this direction, which can be expressed by formula (21), where i represents the current position of the USV, are presents one of the movable directions of the USV, and j represents all grids in this movable direction. Considering the influence of distance in the formula, the farther the distance of the grid, the smaller the impact of the target existence probability on the heuristic of the USV.

$$\eta(a, i, t) = \sum_{j \in straightline(a, i)} \frac{1}{distance(i, j)} * \bar{b}(v_t = j)$$
 (16)

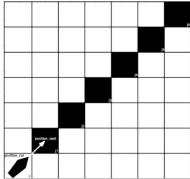


Figure 7. Heuristic information

The state transition rules of the ant population can be described by formula (17), where a represents the action selection direction of the USV, and α and β represent the weight of pheromone concentration and heuristic information.

$$p(a,t,u) = \frac{(\tau[a,s_{u}^{t},u])^{\alpha}(\eta(a,s_{u}^{t},t))^{\beta}}{\sum_{a=1:8} (\tau[a,s_{u}^{t},u])^{\alpha}(\eta(a,s_{u}^{t},t))^{\beta}}$$
(17)

IV. Simulation Result

All experiments will be performed in the same size search space, which is rasterized ($\omega_x = 50$, $\omega_y = 50$). First, verify the update strategy of the target grid probability map under the Markov prediction model through multiple scenarios under different ocean currents and terrain environments; then compare the target search algorithm experiments based on the Markov prediction model under multiple constraints verify.

A. Probability Map Update Strategy Based on Markov Decision

Through three different marine environment scenarios, this experiment demonstrates the update status of the target probability map when performing search tasks in these environments, and further verifies the effect of the target search algorithm based on the Markov decision process on the target

trajectory under the influence of complex marine environments. predictive effect. The three marine environment scenarios are shown in Figure 8.

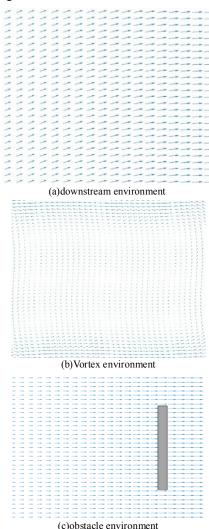
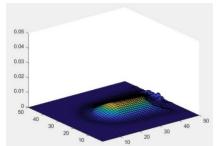
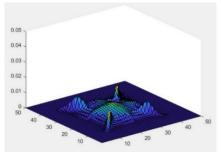


Figure 8. Description of Ocean Current and Obstacle Environment

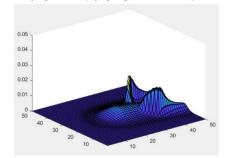
Under the influence of ocean currents and obstacles, the Markov decision process will predict the movement of the target, and the target probability map will be updated with the predicted target trajectory. Fig. 9 is the update of the target probability map at the end of the search task under various ocean currents and obstacle environments.



(a) Target probability graph update in downstream flow



(b) Target probability graph update under eddy current



(c) The target probability graph update under the obstacle Figure 9. Target probability map update status

In all target probability maps, the warmer the tone, the higher the probability of the target in the area, and on the contrary, the lower the probability of the target in the cool tone area. In the downstream environment, it can be seen from Figure 9(a) that the high-probability area will be updated along the direction of the ocean current, and the area with a high probability of target existence will attract unmanned boat groups to go, improving the possibility of searching for the target. The target existence probability in Figure 9(b) will gradually increase towards the four vortex centers, indicating that the target will be attracted by the vortex with a high probability, and the algorithm will also pay more attention to the vortex area and drive the unmanned boat swarm to the nearby vortex area for search . Figure 9(c) In the downstream case of adding obstacles, the target existence probability in the grid corresponding to the obstacle area will be 0, and at the same time, the high-probability area will disperse to both sides around the obstacle. Low probability areas where obstacles are located will be intentionally avoided.

B. Target Search under Multiple Constraints

Before the experiment begins, the number and energy state of surface unmanned boats participating in the search will be initialized. The number of unmanned boat cluster U is 3, and the energy state before the search is $E = \{50E_0, 30E_0, 40E_0\}$. Moreover, the relevant parameters of ant colony optimization algorithm were initialized. The number of ant population M was 100, the concentration of initial pheromone in each grid was set to 0.1, the volatility coefficient of pheromone was 0.1, the weight factor α and β were 1, and the maximum number of iterations NC=80.

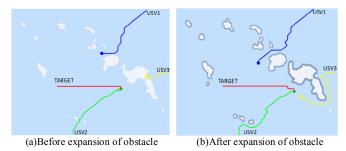


Figure 10 MKACO Search result

FIG. 10 is the simulation experiment diagram of the search task conducted by the usvs based on the MK-ACO algorithm, in which red represents the trajectory roadmap of the target, and the search trajectory of the three usvs is represented by blue, green and yellow respectively. Obstacle avoidance constraints are not considered in Figure (a), and it can be seen that part of the trajectory of the usvs is very close to the obstacle, which greatly increases the risk of the search task. Meanwhile, USV3 falls into a dead end at the groove of the obstacle, reducing the search efficiency. Figure (b) shows the effect after the obstacle expansion method is introduced into MK-ACO. When encountering obstacles, USVs makes the choice to avoid in advance, which greatly reduces the risk of the search task. At the same time, the algorithm can avoid falling into a dead end, greatly improving the search efficiency.

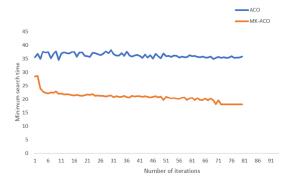


Figure 11 Target search time

Comparing the original ant colony optimization algorithm with MK-ACO in Figure 11, it can be seen that the search time of MK-ACO has been greatly improved in terms of overall performance, and the surface unmanned boat can search the target in a shorter time. In addition, it can be seen from the figure that MK-ACO algorithm can quickly converge to the optimal solution or sub-optimal solution, reflecting the advantages of the algorithm such as adaptability and strong global search ability.

V. CONCLUSIONS

Aiming at the problem of target search in complex Marine environment, this paper proposes a target search algorithm based on Markov prediction model. The algorithm uses Markov decision process to predict the trajectory of the target under the influence of ocean currents, and optimizes the search path of the unmanned boat cluster by improving the pheromone and heuristic strategy of the ant colony optimization algorithm,

which greatly improves the success rate and efficiency of finding the target. At the same time, the algorithm takes into account the energy state of the usvs and the obstacles in the ocean, and introduces fuel constraints and obstacle avoidance constraints. The search task is performed under the guarantee of sufficient fuel and safety, which further improves the safety and search efficiency. In the future, the algorithm can be more focused on the heterogeneous and multi-target search tasks of USVs.

References

- [1] Hollinger, Geoffrey & Gordon, Geoff & Simmons, Reid. (2010). Search in the Physical World. https://www.researchgate.net/publication/241363822 Search in the Physical World.
- [2] K. E. Trummel, J. R. Weisinger, (1986) Technical Note—The Complexity of the Optimal Searcher Path Problem. Operations Research 34(2):324-327.https://doi.org/10.1287/opre.34.2.324
- [3] Pablo Lanillos, Javier Yañez-Zuluaga, José Jaime Ruz, and Eva Besada-Portas. 2013. A bayesian approach for constrained multi-agent minimum time search in uncertain dynamic domains. In Proceedings of the 15th annual conference on Genetic and evolutionary computation (GECCO '13). Association for Computing Machinery, New York, NY, USA, 391–398. https://doi.org/10.1145/2463372.2463417.
- [4] P. Lanillos, E. Besada-Portas, G. Pajares and J. J. Ruz, "Minimum time search for lost targets using cross entropy optimization," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura-Algarve, Portugal, 2012, pp. 602-609, doi: 10.1109/IROS.2012.6385510.
- [5] Phung, Manh Duong & Ha, Quang. (2020). Motion-Encoded Particle Swarm Optimization for Moving Target Search Using UAVs. https://doi.org/10.1016/j.asoc.2020.106705.
- [6] M. Dorigo, V. Maniezzo and A. Colorni, "Ant system: optimization by a colony of cooperating agents," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 26, no. 1, pp. 29-41, Feb. 1996, doi: 10.1109/3477.484436.
- [7] Sara Perez-Carabaza, Eva Besada-Portas, Jose A. Lopez-Orozco, and Jesus M. de la Cruz. 2016. A Real World Multi-UAV Evolutionary Planner for Minimum Time Target Detection. In Proceedings of the Genetic and Evolutionary Computation Conference 2016 (GECCO '16). Association for Computing Machinery, New York, NY, USA, 981–988. https://doi.org/10.1145/2908812.2908876
- [8] Sara Perez-Carabaza, Eva Besada-Portas, Jose A. Lopez-Orozco, and Jesus M. de la Cruz. 2016. A Real World Multi-UAV Evolutionary Planner for Minimum Time Target Detection. In Proceedings of the Genetic and Evolutionary Computation Conference 2016 (GECCO '16). Association for Computing Machinery, New York, NY, USA, 981–988. https://doi.org/10.1145/2908812.2908876
- [9] Pablo Lanillos, Javier Yañez-Zuluaga, José Jaime Ruz, and Eva Besada-Portas. 2013. A bayesian approach for constrained multi-agent minimum time search in uncertain dynamic domains. In Proceedings of the 15th annual conference on Genetic and evolutionary computation (GECCO '13). Association for Computing Machinery, New York, NY, USA, 391–398. https://doi.org/10.1145/2463372.2463417
- [10] 1984. The Optimal Search for a Moving Target When the Search Path Is Constrained. Oper. Res. 32, 5 (October 1984), 1107–1115. https://doi.org/10.1287/opre.32.5.1107
- [11] Raap, M., Meyer-Nieberg, S., Pickl, S. et al. Aerial Vehicle Search-Path Optimization: A Novel Method for Emergency Operations. J Optim Theory Appl 172, 965–983 (2017). https://doi.org/10.1007/s10957-016-1014-y
- [12] L. Lin and M. A. Goodrich, "UAV intelligent path planning for Wilderness Search and Rescue," 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, St. Louis, MO, USA, 2009, pp. 709-714, doi: 10.1109/IROS.2009.5354455.