Optimal travel strategy model study

Yiming He¹*, Yufan Li¹ and Dalin Huang¹

¹ School of Mechatronical Engineering, Beijing Institute of Technology, Beijing, 102488, China
* Corresponding author’s e-mail: 1120191946@bit.edu.cn

Abstract. In this paper, we propose an appropriate algorithm to determine the best strategy of an individual in the Through the Desert game, and give the generally applicable models for single-player and multi-player games under known and unknown weather conditions respectively. We integrate ant colony algorithm, branch-and-bound method and game theory to build and solve the optimization model.

1. Problem formulation

Question B of the 2020 China Undergraduate Mathematical Contest in Modeling (CUMCM) is a combination of graph theory and game theory, which is recapitulated as follows.

The rules of the game of crossing the desert are that the player, with a map, uses initial funds to buy a certain amount of water and food, starts from the starting point and walks in the desert, where different weather is encountered in the diagram and can obtain funds in the mines and replenish resources in the villages, with the goal of reaching the endpoint in a specified time and keeping as much money as possible.

We need to give the best decision for the player under different game settings.

2. Model assumptions

a). The player only considers the map, weather conditions, weight limit, daily consumption and final payoff when deciding on a course of action, independent of the remaining factors.
b). It is assumed that when multiple players participate in the game, each player is "rational", i.e., they do not lose their own revenue in order to reduce the revenue of others, and each player seeks to maximize their personal benefits.

3. Notations

Table 1. Symbol Definition Instruction

| Symbol      | Description                                    |
|-------------|------------------------------------------------|
| $P$         | Remaining funds at the endpoint                |
| $P_b$       | Initial funds                                  |
| $l_{total}$ | Total income of funds                          |
| $C_{total}$ | Total consumption of funds                     |
| $L_m$       | Maximum load                                   |
| $ddl$       | Deadline date                                  |
| Action      | Action strategy                                |
| $Z$         | Synthetic fuzzy decision function              |
4. Problem analysis
The main task of the authors is to determine the travel route and build an optimization model by operation planning under different conditions, giving the optimal strategy for the general case.

Problem 1: Solve the optimal action strategy for a single player under the condition that the weather is all known. We use the greedy algorithm[1] to divide the whole process into three stages: departure, resupply and return to find the local optimal solutions: a simplified ant colony algorithm is applied to solve the optimal solutions of the departure and return stages; an improved branch-and-bound method is applied to exhaust all possible cases to find the optimal solutions of the resupply stage. The local optimal solutions of the three parts are combined to obtain the overall optimal solution.

Problem 2: Solve the optimal action strategy for a single player under the condition that only the weather of the day is known. We build an optimization model for integrated decision making based on the expectation coefficients of different players for the game and introduce the game completion function on the basis of the optimization model of Problem 1.

Problem 3: Solve the optimal action strategy for multiple players who jointly participate in the game individually: we build a game-theoretic model of mixed strategies from the perspective of the most stable gain, and interpret this game from the players' perspective in a game-theoretic perspective.

5. Model building and solving

5.1 Problem 1
This problem needs to build an optimization model to make the most money when reaching the endpoint. In this problem, since the weather condition is known, in order to avoid the waste of returning the materials at half price at the end of the journey, we calculate in advance to make the purchased materials meet the number of materials needed to reach the end of the journey, thus the process of accumulating funds in this problem is only for the revenue generated by mining. Accordingly, we establish an optimization model[2]:

\[
\begin{align*}
\text{max} & \quad P = P_b + l_{\text{total}} - C_{\text{total}} \\
\text{s.t.} & \quad l_{\text{total}} = \text{income} \times n \\
& \quad C_{\text{total}} = \sum_{i=1}^{\text{final}} m \times (\epsilon p_w C_w + \epsilon_f C_f) \\
& \quad m_w \sum_{a_k} \Delta m C_w + m_f \sum_{a_k} \Delta m C_f \leq L_m \\
& \quad \text{final} \leq \text{ddl}
\end{align*}
\]

where \( n \) is the number of mining days; \( \text{final} \) is the number of days to reach the endpoint; \( m = 1,2,3 \) represents the coefficient of material consumption for different action scenarios, \( m = 1 \) for staying and consuming basic materials, \( m = 2 \) for action and consuming twice the basic materials, and \( m = 3 \) for mining and consuming three times the basic materials; \( \epsilon = 1, 2 \) represents the coefficient of material purchase for different periods, \( \epsilon = 1 \) represents the purchase of supplies at the starting point at the original price and \( \epsilon = 2 \) represents the purchase of supplies in the village at twice the base price; \( \alpha_k \) represents the date of replenishment of water and food supplies and \( \Delta \) represents the number of days between two adjacent replenishments.

Based on this model and the known weather conditions and parameters, we use the greedy algorithm to divide this game into three parts, find the local optimal solutions separately, and finally merge the local optimal solutions of the sub-problems into the solution of the original problem.

1. Departure phase: Since the player needs to keep as much money as possible at the end of the
game, the player obviously needs to go to the mine to replenish the money.

2. Replenishment phase: here the replenishment material refers to both the financial gain from mining and the food and water replenished by going to the village halfway.

3. Return phase: return from the mine or village to the end of the game within the time limit.

---

**5.1.1 Simplified ant colony algorithm.** For the departure and return phases, we design a simplified ant colony algorithm[3]: we stipulate that if there is no pheromone left by other ants in the travel route of ant $k$, then ant $k$ can continue to travel, and conversely, if there is already pheromone left by other ants in its forward route, then the ant dies immediately. As time passes, the route of the first ant to reach the end is recorded as the optimal route.

1) Initialization parameters: In the departure and return phases, the starting point and the endpoint are set as the starting point and the mine, and the mine and the endpoint, respectively. $m$ ants start from the starting point at the same time. They search for the solution of the shortest path in parallel with the same speed, and communicate with each other through the medium of pheromones, and use the number of pheromones on the edge connected to each node as the trial basis for searching the next node until they find the feasible solution of the optimal route.

2) The survival probability of artificial ant $k$ at time $t$ after it is transferred from region $i$ to region $j$ is

$$p_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}(t)\eta_{ij}(t)}{\sum_{j \in S}\tau_{ij}(t)\eta_{ij}(t)}, & j \in S \\ 0, & j \notin S \end{cases} \quad (2)$$

where $S$ is the set of feasible points, parameter $\alpha$ is the relative importance of trajectory; $\beta$ is the relative importance of visibility; $\tau_{ij}(t)$ is the trajectory intensity of arc $(i,j)$ at time $t$; $\eta_{ij}(t)$ is the visibility of arc $(i,j)$ at time $t$, reflecting the degree of expectation from region $i$ to region $j$, defined by the 0-1 variable $\eta_{ij}(t)$ to determine the survival or non-survival of ant $k$.

$$\eta_{ij}(t) = \begin{cases} 1, & \tau_j = 0 \\ 0, & \tau_j > 0 \end{cases} \quad (3)$$

If $\tau_j > 0$, it is proved that ants have already reached region $j$ before ant $k$, so $\eta_{ij} = 0$, which means that $p_{ij}(t) = 0$, i.e., the survival probability of worker ant $k$ to region $j$ at time $t$ is 0.

3) Under the condition that the speed of the ant colony is the same, the optimal route is recorded at the end of the line for the first ants to arrive.
5.1.2 **Improved branch-and-bound method.** We branch the actions done by players on that day by specifying that\[4]:

\[
\text{Action} = \begin{cases} 
1, & \text{Return phase} \\
2, & \text{Mine} \\
3, & \text{Stay in the mine} \\
4, & \text{Go to the village for resupply}
\end{cases}
\] (4)

In this regard, we have the following instructions.

1. when \text{Action}=1, that means end the resupply phase, enter the return phase, depart to the endpoint.
2. when \text{Action}=3, we allow players to stay in the mine without mining under certain conditions, considering the limitation of their own supplies and the high consumption of mining in sandstorm weather, which does not conflict with the third point of the prerequisite principle.
3. when \text{Action}=4, we specify that the player must return to the mine after resupplying from the village.
4. when \text{Action}=1 or 4, the game rules are also observed during the journey, i.e., stopping at the same place in case of the sandstorm.

5.2 **Problem 2**

This question builds on the optimization model developed in Problem 1. Since the weather conditions are unknown, we need to make a choice about the ambiguity of future weather conditions and thus determine the number of supplies to purchase at the starting point. In addition, the unpredictability of future weather conditions inevitably leads to the possibility of failure of our game, so we need to weigh the relationship between the successful completion of the game and the amount of money retained at the end of the game.

Therefore, we set the expectation coefficient \(\gamma (\gamma \in [0,1])\), the magnitude of which represents the player's expectation of completing the game, the closer \(\gamma\) is to 1, the more the player wants to score high (i.e., the more money he or she will have with him or her at the end of the game), and the more he or she is willing to take the high risk of failure; the closer \(\gamma\) is to 0, the more the player wants to complete the game, and the more he or she is willing to finish with a lower score. and are willing to
complete the game with a lower score. We give the integrated fuzzy decision model.

\[
\max \quad Z = \gamma \frac{p}{p_{\text{max}}} + (1 - \gamma)Q
\]

\[
P = P_b + I_{\text{total}} - C_{\text{total}}
\]

\[
l_{\text{total}} = \text{income} \times n + 50\% \times (p_w W_{\text{remain}} + p_f F_{\text{remain}})
\]

\[
c_{\text{total}} = \sum_{i=1}^{\text{final}} m \times (\epsilon p_w c_{w_j} + \epsilon p_f c_{f_j})
\]

\[
m_w \sum_{\alpha_k}^{\alpha_k+\Delta} m c_{w_j} + m_f \sum_{\alpha_k}^{\alpha_k+\Delta} m c_{f_j} \leq L_m
\]

\[
f_{\text{final}} \leq d_{\text{final}}
\]

\[
Q = 0, 1
\]

In equation (5), \( Q \) is the game completion function, defined as \( Q = 0 \) for game failure and \( Q = 1 \) for game completion; under the condition of avoiding game failure as much as possible, it is necessary to buy more supplies than needed, so the process of accumulating funds needs to add 50% of the base price obtained by returning the resources at the end; \( W_{\text{remain}} \) is the amount of water left at the end, \( F_{\text{remain}} \) is the amount of food left at the end. \( F_{\text{remain}} \) is the amount of food remaining at the end of the journey; the remaining parameters are identical to equation (1).

By normalizing \( P \) and defining the game completion function in terms of 0-1 variables, we constrain the overall evaluation function to be between 0 and 1, with the corresponding \( \gamma \) determined for a particular player, so that we have to give the optimal strategy under the above constraints.

5.3 Problem 3

5.3.1 Game theoretic model of mixed strategy: Here we build a "non-cooperative" game model, i.e., players do not communicate in advance, and use the mixed strategy Nash equilibrium with the ultimate goal of maximizing their own interests, not caring whether to win among the two players through the negative-sum game, not seeking to limit each other's interests, but seeking to reduce their own losses[5].

We choose a mixed strategy and the mixed strategy must have saddle points, i.e., there exist probability vectors \( \vec{x} \) and \( \vec{y} \), such that

\[
\vec{x}^T \vec{A} \vec{y} = \max_{x} \min_{y} x^T A y = \min_{y} \max_{x} x^T A y
\]

(6)

For a single player, the probability vector of his choice of strategy is \( \vec{x} = (p_1, p_2, \ldots, p_m) \)

5.3.2 Game-theoretic model with known weather: First, the player's own winning matrix is determined based on the material consumption, and then the probability \( p_i \) of choosing each route is obtained from the mixed strategy constraints, and finally, the decision is made based on the probability to maximize the benefit stability.

\[
\sum_{i=1}^{n} p_i a_{ik} = \text{Const} (k = 1, 2, \ldots, n)
\]

\[
\sum_{i=1}^{n} p_i = 1
\]

(7)

5.3.3 Game-theoretic model with unknown weather: For the event that the weather is unknown and all \( n \) game players know the action plan and the number of resources remaining for the rest of the players at the end of the day, all players have the knowledge and all players know that other players also know this event, so this event is common knowledge. We give the variance evaluation function to solve for the probability of choosing each strategy.
\[
\begin{aligned}
\min & \quad A(p_1, p_2, \cdots, p_{m+1}) \\
\text{s. t.} & \quad X_i = p_1 a_{i1} + p_2 a_{i2} + \cdots + p_{m+1} a_{i(m+1)} (i = 1, 2, \cdots, 2(m + 1)) \\
& \quad \bar{X} = \frac{\sum_{i=1}^{2(m+1)} X_i}{2(m+1)} \quad (8)
\end{aligned}
\]

6. Conclusion

In summary, we comprehensively analyze the influence of relevant factors on the player's decision making when proceeding under different conditions respectively, and analyze the player's strategy for the Through the Desert game under different conditions from shallow to deep, and for increasingly complex problems we transform from deterministic strategies to optimal solutions with randomness and localization, in order to achieve the maximum gain at the endpoint. After analysis and verification, the model of this paper is reasonable and has some practical significance.

References

[1] Edmonds J. (1971) Matroids and the greedy algorithm. J. Mathematical Programming, 1(1):127-136.
[2] Li, Z. C., Zhou, X., Dai, Z., & Zou, X. Y., (2011) Identification of protein methylation sites by coupling improved ant colony optimization algorithm and support vector machine. J, Analytica chimica acta, 703(2), 163-171.
[3] Kenderov P S. (1984) Most of the Optimization Problems have Unique Solution. J. Comptes rendus de l'Académie bulgare des sciences: sciences mathématiques et naturelles, 37.
[4] Shih W., (1979) A Branch and Bound Method for the Multiconstraint Zero-One Knapsack Problem. J, Journal of the Operational Research Society, 30(4):369-378.
[5] Border K C., (1985) Fixed Point Theorems with Applications to Economics and Game Theory. J, Economic Journal, 96(381):255-255.