

基于动态规划模型的穿越沙漠路程优化问题

付芷睿 李雪纯 李兴睿

(武汉大学, 湖北 武汉 430000)

摘要 在“穿越沙漠”的游戏中, 玩家需要规划路线在规定时间内到达终点, 并保留尽可能多的资金。利用动态规划等数学工具可以针对不同的关卡设置寻找玩家的最优路线。

第一问中, 只有一名玩家且该玩家已知全程天气。首先以到达终点时剩余资金最大为目标函数, 以游戏规则为约束条件, 以玩家的策略为决策变量, 建立单目标规划模型。接着将全路程分为三个阶段: 从起点到矿山、在矿山与村庄活动、从矿山到终点。在一、三阶段利用 Dijkstra 算法求出最短路径, 对于第二阶段设计循环算法求解。

第二问中, 只有一名玩家且仅知道当天天气。首先对时间离散化, 将本问转化为多阶段决策问题, 而后建立动态规划模型。求解过程分为两步:

(1) 借助贪心算法的思想, 构造符合约束条件的玩家策略;

(2) 将采用该策略求解的路径与第一问中的算法求解的路径进行比较, 优化该玩家策略, 使其逼近最优解。

第三问中, 推广到多名玩家与有顺序游戏模式。为简化问题, 仅考虑玩家 A 的最优策略。此时 A 需要综合考虑其前面的 $i-1$ 位玩家的策略以及后 $n-1$ 位玩家的策略给出最终策略。

关键词 Dijkstra 算法 单目标规划模型 多阶段决策 动态规划模型

中图分类号: U491.2

文献标识码: A

文章编号: 1007-0745(2021)01-0052-04

1 模型假设

1. 假设玩家都是理性的, 追求的目标仅是到达终点时获取的总利益最大;

2. 假设区域可以抽象成一个点;

3. 假设以下的初始值是不变的: 负重上限恒为 1200kg, 初始资金恒为 10000 元, 每箱水的质量为 3kg, 食物的质量为 2kg, 每箱水的基准价格为 5 元/箱, 食物的基准价格为 10 元/箱。

2 模型准备

2.1 地图的要素提取及简化

首先, 对于形状不规则的地图进行抽象处理, 提取出起点、终点、村庄和矿山四个要素, 其余点均视为普通点。

2.2 最短路径

为了使得到达终点时剩余资金最大, 玩家应尽量减少行走过程中的消耗, 因此玩家在起点出发后, 会立即以最短路径前往村庄、矿山或终点其中之一。

2.3 定义消耗上限与收益下限

消耗上限: 假设全程 30 天均为沙暴天气, 计算出玩家的物资消耗量, 称作消耗上限。

收益下限: 假设挖矿时均遇到沙暴天气, 计算出此时的挖矿收益, 称作收益下限。

在只知道当天天气的情况下, 玩家应对此后的天气进行最坏的打算, 计算出消耗上限与收益下限, 权衡下一步的行动。

2.4 停留

初步分析可知, 停留的条件有两种:

(1) 沙暴时必须在原地, 不能移动;

(2) 到达矿山后, 由于挖矿的消耗为基础消耗量的三倍, 并且高温天气与沙暴天气相对于晴朗天气的基础消耗量都显著提高, 所以为了挖矿期间的消耗尽可能少, 可能会根据需要在沙暴天或高温天不挖矿, 通过简单演算, 我们规定每一关卡除沙暴天以外, 其余总的停留时间不超过两天。

3 基于单目标规划模型的最优策略求解

3.1 单目标规划模型的建立

3.1.1 决策变量的确定

(1) 玩家在第 i 天的位置 x_i , $i=0, \dots, N$, $x_i \in A$ 。其中, N 表示游戏的截止日期, A 表示地图上所有区域的序号的集合。特别地, 当玩家在第 n 天 ($1 \leq n \leq N$) 到达终点后, 我们约定玩家位置始终位于终点;

(2) 玩家在第 i 天 ($0 \leq i \leq N$) 是否位于村庄, 对此引入 0-1 变量 v_i ;

(3) 玩家在第 i 天 ($0 \leq i \leq N$) 是否挖矿, 对此引入 0-1 变量 u_i ;

(4) 玩家在第 i 天 ($0 \leq i \leq N$) 购买的物资数量, a_i 表示水的数量, b_i 表示食物的数量。特别地, $i=0$ 时表示第 0 天玩家在起点处购买的物资数量, 并且需要注意玩家不位于村庄时无法购买物资, 因此 a_i 和 b_i 均取值为 0。

综合上述变量, 决策变量可统一表示为向量 $X_i=(x_i, v_i, u_i, a_i, b_i)$ 。

a_i, b_i), ($0 \leq i \leq N$)。

3.1.2 目标函数的确定

设玩家在第 n 天 ($0 \leq i \leq N$) 抵达终点, 为使抵达终点时剩余的资金 y_n ($1 \leq n \leq N$) 尽可能多, 目标函数为: $\text{Max } y_n$ 。其中:

$$y_n = 10000 + \sum_{j=0}^n [w_j S_0 - (5a_j + 10b_j)]$$

S_0 表示挖矿的基础收益, a_i 和 b_i 分别表示玩家在第 i 天 ($0 \leq i \leq N$) 购买的水和食物的数量, w_i 为第 i 天 ($0 \leq i \leq N$) 的天气状况:

$$w_i = \begin{cases} 1, & \text{第 } i \text{ 天为晴朗天气} \\ 2, & \text{第 } i \text{ 天为高温天气} \\ 3, & \text{第 } i \text{ 天为沙暴天气} \end{cases}$$

3.1.3 约束条件

条件一: 任意第 i 天包内都有剩余物资且总负重 m_i 不超过负重上限 1200kg;

条件二: 第 i 天 ($0 \leq i \leq N$) 购买物资所耗费的资金 S_i 应当小于等于当天剩余资金 y_i ;

条件三: 当天气状况为沙暴时, 玩家必须在原地停留一天, 即第 $i+1$ 天和第 i 天位置相同 ($0 \leq i \leq N$);

条件四: 玩家每天只能从目前区域到达与之相邻的区域, 或在目前区域停留一天;

条件五: 当且仅当第 i 天 ($0 \leq i \leq N$) 玩家位于起点或村庄时, 才可以购买物资。即玩家不位于起点或村庄时, 购买物资数量为 0。

综上, 建立单目标规划模型即可。

3.2 求解模型得出一般策略

3.2.1 求解一般策略的思维过程

由于直接求解上述优化问题时间和电脑储存空间开销较大, 因此我们首先根据不同类型的地图所含要素进行分类讨论, 将全程路线分为挖矿与不挖矿两种路线。

路线一: 不经过村庄或矿山。通过选择最短路径, 计算出玩家最快到达终点的日期 n , 可得最终剩余资金为:

$$y_n = 10000 - \sum_{j=0}^n (5a_j + 10b_j)$$

路线二: 经过村庄和矿山。当玩家采取挖矿的路线时, 我们再将全程分为从起点到矿山、挖矿过程和从矿山到终点三个阶段处理, 并且给出每个阶段玩家应采取的策略, 最后给出整体算法。

阶段 1: 起点 - 矿山

1. 起点物资量

由于村庄中物资的价格为基准价格的 2 倍, 因此为了

节约资金, 玩家应在起点处购买尽可能多的物资。

2. 到达矿山的路径及日期

为了减少行走的消耗, 玩家应当尽快前往矿山, 仅在沙暴天气停留。通过选择最短路径, 可以得到玩家最快到达矿山的路径及日期。

阶段 2: 矿山挖矿, 中途去村庄补给

1. 挖矿天数

根据阶段 1 和阶段 2 计算的到达与离开矿山的日期, 可得玩家挖矿的天数。

2. 去村庄补给的日期

在物资即将不足但是恰好能维持玩家抵达村庄时, 离开矿山前往村庄。根据往返矿山和村庄之间需要花费的时间, 计算出最多可以前往村庄的次数。

3. 购买补给的量

应满足补给后的物资足够玩家维持到下一次补给。

阶段 3: 矿山 - 终点

离开矿山的日期。为了使收益最大, 玩家应当用尽可能多的天数挖矿, 通过选择最短路径并且考虑食物和水的约束可以得到玩家最晚离开矿山的日期。

3.2.2 求解一般过程的整体算法

Step1: 采用 Dijkstra 算法计算从起点到矿山的日期;

Step2: 根据食物和水的余量的初步计算从矿山离开到终点的日期范围, 而后采用 Dijkstra 算法计算从矿山离开到终点的确切日期;

Step3: 遍历玩家离开矿山到达村庄进行补给的日期, 对于多种可能的路线, 分别计算出终点时玩家的资金剩余量, 求出最大资金剩余量所在路线, 则该路线即为所求玩家最佳策略;

step4: 综合前四步, 得到全过程的策略。

4 基于动态规划的最优决策求解

4.1 多阶段决策过程的动态规划模型

本问题中, 玩家不知道全程的天气, 此时不便利用第一问中分 3 个阶段的方法, 而是应该每天根据当前的状态进行下一步行动的决策, 因此我们建立动态规划模型如下:

划分阶段: 以天为单位, 假设游戏的截止日期为第 n 天, 则全程分为 $n-1$ 个阶段。

状态变量: $X_i = (x_i, v_i, u_i, a_i, b_i)$, ($0 \leq i \leq N$), 其中 i 表示第 i 天, x_i 表示玩家所在位置, v_i 表示玩家是否在村庄, u_i 表示玩家是否挖矿, a_i 表示购买水的数量, b_i 表示购买食物的数量^[1]。

决策变量: 当一个阶段的状态确定后, 玩家需要做出选择从而演变到下一阶段的某个状态, 这种选择手段称为决策, 描述决策的变量称为决策变量。设玩家在位置 x_i ($0 \leq i \leq N$) 的决策变量为 $f_i(x_i)$ 。

状态转移方程: 在确定性过程中, 一旦某阶段的状态和决策已知, 下阶段的状态便完全确定。用状态转移方

程表示这种演变规律,即通过玩家目前位置 x_i ($0 \leq i \leq N$) 和决策变量 $f_i(x_i)$ 来确定下一步的位置 x_{i+1} 的方程, 设为 $x_{i+1} = T_i(x_i, f_i)$ 。

指标函数: 是衡量过程优劣的数量指标。由于玩家的目标是使剩余的资金最大, 因此选择每一天的净收益作为指标函数, 净收益即玩家当天可能挖矿的收益与消耗资金之差, 记为 $V_{i,n}$ ($0 \leq i \leq N$)。

最优值函数: 在玩家位置 x_i ($0 \leq i \leq N$) 给定时指标函数 $V_{i,n}$ 对策略的最优值称为最优值函数, 记为 $g_i(x_i)$, 即:

$$g_i(x_i) = \text{opt} V_{i,n}(x_i)$$

在上述动态规划模型的基础上增加第一问中的约束条件, 即为第二问的总模型。

4.2 求解模型得出一般策略

通过该动态规划模型, 我们希望求解出全过程的最优策略 $S^* = \{f_1(x_1)^*, f_2(x_2)^*, \dots, f_n(x_n)^*\}$, 即玩家在每一天的决策 $f_i(x_i)^*$ 的集合。但是, 由于该动态规划模型的时间复杂度是 $O(n^3)$, 直接采用逆序求解法的消耗巨大, 无法在满意的时间范围内求得最优解。因此我们采用拆分的思想, 将求解过程拆分成两步^[2]:

(1) 首先利用贪心算法的思想制定出某一特定的策略 S_0 , 作为初值;

(2) 接着通过与最优策略的比较, 不断修改该策略 S_0 , 得到最终的一般策略 S^* , S^* 就是上述动态规划问题的近似最优解。

4.2.1 基于贪心算法的特定策略 S_1

(1) 当玩家位于起点时, 假如玩家经过矿山的时间超过天数上限, 则直接前往终点; 否则, 时间充足, 可以经过矿山。

(2) 当玩家在第 i 天位于村庄时, 我们利用贪心算法, 使得玩家始终做出利益最大化的选择。即, 先计算从当前日期到最后一天的消耗上限, 若此时包内水的数量或者包内食物的数量小于该数值, 那么玩家进行补给, 否则不补给。

(3) 当玩家位于矿山时, 利用贪心算法, 使得玩家始终做出利益最大化的选择。即, 先将水和食物的消耗上限按照村庄的价格折算成资金, 再与挖矿收益进行比较。若 $S_0 \leq 2(3P^* + 2Q^*)$, 那么玩家不挖矿, 否则挖矿。

(4) 当玩家位于其他位置时, 由于约束条件过于复杂, 无法确定应该前往终点、矿山还是村庄。此时我们首先计算出所有可能的移动路线的消耗上限, 当包内水的数量 $\sum_{j=0}^i a_j - n_{\text{水}}^{(j)}$ 和食物的数量 $\sum_{j=0}^i b_j - n_{\text{食物}}^{(j)}$ 均大于消耗上限时, 在这些路线中, 我们约定玩家等可能选择其中一条。

(5) 当玩家位于终点时, 游戏结束。

4.2.2 通过比较 Z^* 对策略 Z_1 进行修正

1. 定义两条路径的差别 ε_{ij} 。在该问题中, 定义向量 $\varepsilon_{ij} = (x_1 - y_1, x_2 - y_2, \dots, x_n - y_n)$, ε_{ij} 中的元素为两条路径 x_i 和路径 y_j 中每个位置的差, 用于衡量两条路线的相似程度。

其中, ε_{ij} 的零位越多, 两条路径越相似。

2. 根据 ε_{0k} 修正一般策略 Z_1 的算法。

Step1: 编写程序模拟玩家采用一般策略 Z_1 后走的路径 K_1 ;

Step2: 采用第一问的算法计算玩家采用的策略 Z_0 以及对应的路径 K_0 ;

Step3: 根据 ε_{0k} 修正策略 Z_1 , 得到策略 Z_2 , 当向量 ε_{0k} 中有超过两位元素不为零时, 继续修正策略;

Step4: 直到向量 ε_{0k} 中只有小于等于两位元素不为零时, 可以认为两条路径足够接近, 此时迭代得到的 Z_k 即为所求 Z^* 。

5 多玩家参与情况下的最优策略求解

现在, 将第一问和第二问的情况分别推广到存在 n 名玩家的情况下。由于不同玩家处于同一区域时, 他们的消耗量会增加为 $2k$ 倍而挖矿的收益变为 $1/k$, 此时会出现多人博弈的问题, 每个玩家的策略会受到环境和他人决策的共同影响。

由于无顺序选择模式较为简单, 我们假设游戏模式为有顺序模式: n 个玩家同一天从起点出发, 但选择策略有先后顺序, 即抽到的编号越小, 越先选择策略。

为了简化模型, 我们只考虑玩家 A 应选择的策略。

5.1 第一小问: 已知全部天气状况

首先, 根据第一问, 确定单名玩家的最佳策略 S^* , A 的策略选择流程为:

(1) n 名玩家在游戏开始前分别获得其编号, 按编号从小到大先后进入游戏并确定策略, 即抽到第 i 号的玩家是第 i 个确定策略的玩家;

(2) 玩家 A 根据第二问中的一般策略推测前 $i-1$ 名玩家的策略的集合 P_i , 并根据 P_i 推测出自己的策略;

(3) 玩家 A 根据第二问中的一般策略推测后 $n-i$ 名玩家的策略的集合 Q_i , 并根据 P_i 和 Q_i 以及自己的策略得出自己的最终收益。

5.2 第二小问: 不知道天气状况

首先根据第二问, 确定单名玩家的最佳策略 S^* , A 的选择流程为:

(1) n 名玩家在游戏开始前分别获得其编号, 按编号从小到大先后进入游戏并确定策略, 即抽到第 i 号的玩家是第 i 个确定策略的玩家;

(2) 玩家 A 根据第二问中的一般策略推测前 $i-1$ 名玩家的策略的集合 P_i , 并根据 P_i 推测出自己的策略;

(3) 玩家 A 根据第二问中的一般策略推测后 $n-i$ 名玩

家的策略的集合 Q_i , 并根据 P_i 和 Q_i 以及自己的策略得出自己的最终收益。

6 模型评价

6.1 模型优点

(1) 综合考虑各种因素, 建立路径决策的单目标规划模型, 模型可迁移性强;

(2) 建立动态规划模型, 详细考虑玩家在各个情况下的决策, 模型完备;

(3) 采用拆分的思想对模型进行求解, 求解迅速且结果较准确。

6.2 模型缺点

(1) 第三问求解多名玩家的游戏时, 简化了模型, 并未考虑多者博弈的情况;

(2) 算法普适性有待提高, 对于不同的关卡, 运行程

序得到的结果需要进行一定步数的手动调整, 结果才能更准确。

6.3 模型推广

(1) 可以尝试手动调整得到更复杂的算法, 允许算法运行更长时间得到更优的结果;

(2) 可以尝试采用强化学习的方法学习得到的游戏策略与本文中算法求得的策略进行对比, 并叠加强化学习的状态转移等要素改进本文算法, 得到更好的结果。

参考文献:

- [1] 韦化, 龙丹丽, 黎静华. 求解大规模机组组合问题的策略迭代近似动态规划 [J]. 中国电机工程学报, 2014, 34(25): 4420-4429.
- [2] 司守奎, 孙兆亮. 数学建模算法与应用 [M] (第 2 版). 北京: 国防工业出版社, 2016.

(上接第 44 页)

4.4 做好应急数据管理

数据的汇总能力在应急处理上无疑是核心, 在通过物联网获得信息、大数据对收集来的信息进行处理时, 应注意将小数据进行结构化, 大数据多种化处理。同时应在加强各区域之间的协调合作的基础上, 打破以往不同区域不同数据的治理方式, 将数据进行统一规范化管理。对数据进行汇总融合后, 再进行深度挖掘, 找到各地区之间的关联和潜关联, 在以上基础上做出决策。同时将决策进行数据收集的反向行动, 下放给各部门, 由各部门规划追踪, 再将实际调查结果反馈给区域数据中心。基层也可采取用二维码的方式直接把数据传至数据中心, 数据进行提取选择后, 下放给基层组织, 解决反复收集信息的困扰。

个人隐私问题是在信息大量收集时必须去考虑的问题, 用户隐私应通过法律和技术进行保障, 对数据做好脱敏处理。在城市突发应急事件时, 应在加紧收集数据的同时打击恶意攻击信息, 形成良好的网络环境, 保障好每位公民的利益。在应急卫生事件发生时, 也可减少公民瞒报、误报的情况发生。

4.5 应急服务平台的标准化建设

现阶段, 数字化信息服务存在着数字化程度低、数据采集的渠道相对分散、数据流通能力低、数据标准化程度低等问题^[8]。如在前段时间出现不同地区不同的健康码, 加大了居民出行的难度, 不利于社会的发展, 且现在数据的相关收集大多还是通过人力实现, 在突发事件出现时, 人力就成为了不确定因素, 因此需加强数字化进程, 以及提高相关系统的数字化标准。相关部门通过搭建标准化的公共卫生应急服务平台, 充分利用应急服务平台对公共卫生事件进行研判、处置, 有助于提高政府在处理突发事件的

效率和质量。

5 结语

智慧城市下的应急管理体系积极采取了物联网、5G 无人机、大数据等先进科学技术, 完善地区应急预案以及面对突然应急事件的处理能力, 通过地区试点、区域结合、社区搭建平台、流程标准化建设等方面进行完善。这些措施可大幅提高应急救援能力和应急水平, 搭建智慧城市应急管理服务平台, 打造更加完善的智慧城市^[9]。

参考文献:

- [1] 龙海波. 健全公共卫生应急管理体系的几点思考 [J]. 重庆理工大学学报 (社会科学), 2020, 34(06): 1-4.
- [2] 杜明芳. 疫情考验触发的“数智网+”城市智慧治理体系构建 [J]. 中国建设信息化, 2020(07): 18-21.
- [3] 李建华. 借力新基建, 提升河南新型智慧城市建设水平 [J]. 决策探索 (下), 2020(12): 19-20.
- [4] 王家喜. 创新城市应急管理新模式 [J]. 中国应急管理, 2020(07): 40-41.
- [5] 王晨, 徐同德. 公共安全应急管理信息化建设研究 [J]. 辽宁警察学院学报, 2021, 23(01): 69-75.
- [6] 邹东升, 孙彦博. 完善智慧应急体系建设提升应急管理整体效能 [J]. 重庆行政, 2021, 22(01): 38-41.
- [7] 蔡龙江. 社区应急管理系统平台设计与实现 [J]. 中国新通信, 2020, 22(16): 38-39.
- [8] 李赞, 徐炎, 卫志远, 胡欣, 张亚军, 王昊. 智慧城市公共卫生应急服务平台实践及标准化探索 [J]. 信息技术与标准化, 2020(07): 22-25, 35.
- [9] 伊长新. 智慧应急系统的研究及应用 [J]. 科技创新与应用, 2021(07): 172-174.