

# 控制与决策

Control and Decision

## 一种面向严重受损路网的抢修队调度算法

张国富, 涂冰花, 苏兆品, 岳峰

引用本文:

张国富, 涂冰花, 苏兆品, 等. 一种面向严重受损路网的抢修队调度算法[J]. *控制与决策*, 2021, 36(7): 1663–1671.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1582>

---

## 您可能感兴趣的其他文章

Articles you may be interested in

### 基于正态云模型的状态转移算法求解多目标柔性作业车间调度问题

State transition algorithm based on normal cloud model for solving multi-objective flexible job shop scheduling problem

*控制与决策*. 2021, 36(5): 1181–1190 <https://doi.org/10.13195/j.kzyjc.2019.1233>

### 考虑效率和公平的跨区域协同应急救援路径选择

Routing optimization of cross-regional collaborative emergency rescue considering efficiency and fairness

*控制与决策*. 2021, 36(2): 483–490 <https://doi.org/10.13195/j.kzyjc.2019.0919>

### 基于改进多目标优化算法的分布式数据中心负载调度

Multi-objective optimization of energy and performance management in distributed data centers

*控制与决策*. 2021, 36(1): 159–165 <https://doi.org/10.13195/j.kzyjc.2019.0702>

### 基于改进萤火虫算法的区域交通信号配时优化

Timing optimization of regional traffic signals based on improved firefly algorithm

*控制与决策*. 2020, 35(12): 2829–2834 <https://doi.org/10.13195/j.kzyjc.2019.1835>

### 基于不变网络模型和故障注入的分布式信息系统故障溯源方法

Fault source location algorithm for distributed information system based on invariant network and fault injection

*控制与决策*. 2020, 35(11): 2723–2732 <https://doi.org/10.13195/j.kzyjc.2019.0214>

# 一种面向严重受损路网的抢修队调度算法

张国富<sup>1,2,3†</sup>, 涂冰花<sup>1</sup>, 苏兆品<sup>1,2,3</sup>, 岳峰<sup>1,2</sup>

(1. 合肥工业大学 计算机与信息学院, 合肥 230601; 2. 合肥工业大学 工业安全与应急技术安徽省重点实验室, 合肥 230601; 3. 安全关键工业测控技术教育部工程研究中心, 合肥 230601)

**摘要:** 受损路网抢修是灾害应急响应中的一个非常重要的基础环节, 主要研究如何对道路抢修队进行有效调度, 以快速恢复受灾路网的交通能力, 为后续顺利开展应急救援工作提供有效的保证. 已有方法在路网受损严重的情形下往往难以给出有效的调度策略. 为此, 在已有工作的基础上, 简化路网模型和决策模型, 并基于动作集裁减和 Q 学习设计一种面向严重受损路网的抢修队调度算法. 在该算法中, 抢修队只能从当前可达的未修复受损路段集合中选择下一个动作, 以确保 Q 学习的连续性. 仿真实验结果表明, 在节点数和受损率都较大的严重受损路网环境中, 所提算法可以保证所有需求节点均可达, 具有更高的稳定性和可靠性, 且能够在更小的时间和修复代价内给出更优的调度方案.

**关键词:** 灾害应急响应; 受损路网抢修; 严重受损路网; 抢修队调度; Q 学习; 动作集裁减

中图分类号: TP181 文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1582

开放科学(资源服务)标识码(OSID):



引用格式: 张国富, 涂冰花, 苏兆品, 等. 一种面向严重受损路网的抢修队调度算法[J]. 控制与决策, 2021, 36(7): 1663-1671.

## An algorithm for repair crew scheduling on severely damaged road network

ZHANG Guo-fu<sup>1,2,3†</sup>, TU Bing-hua<sup>1</sup>, SU Zhao-pin<sup>1,2,3</sup>, YUE Feng<sup>1,2</sup>

(1. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China; 2. Anhui Province Key Laboratory of Industry Safety and Emergency Technology, Hefei University of Technology, Hefei 230601, China; 3. Engineering Research Center of Safety Critical Industrial Measurement and Control Technology of Ministry of Education, Hefei 230601, China)

**Abstract:** Repairing the damaged road network is one of the most important and basic parts of disaster emergency response. It mainly deals with how to effectively dispatch the repair crew to quickly restore the traffic capacity of the damaged road network and provide an effective guarantee for the smooth implementation of the subsequent emergency rescue. However, when the road network is severely damaged, the existing algorithms often fail to find a feasible solution. Therefore, this paper first simplifies models of damaged road network and decision-making on the basis of the existing work. Then, an improved algorithm for repair crew scheduling on severely damaged road network is developed according to Q-learning and action set reduction. Particularly, in the proposed algorithm, the repair crew can only choose the next action from the set of current damaged road sections which are unrepaired but reachable, ensuring the continuity of Q-learning. Finally, simulation results show that the proposed algorithm can ensure that all demand nodes are reachable, has higher stability and reliability, and can obtain better scheduling schemes at lower time and repair cost, even if the road network has been seriously damaged with a great number of damaged nodes and a big damage rate.

**Keywords:** disaster emergency response; damaged road network repairs; severely damaged road network; repair crew scheduling; Q-learning; action set reduction

## 0 引言

近年来,我国自然灾害多发频发,给国民经济与社会发展造成巨大的损害.当灾害发生后,作为灾害

应急响应的核心环节,如何将救灾物资及时输送到各需求点已成为当前研究的一个热点.研究主要涉及如何在灾害发生时迅速有效地利用智能决策理论

收稿日期: 2019-11-13; 修回日期: 2020-03-19.

基金项目: 国家自然科学基金项目(61573125); 中国工程院战略咨询重点项目(2020-XZ-3); 教育部人文社会科学研究青年基金项目(19YJC870021, 18YJC870025); 中央高校基本科研业务费专项资金项目(PA2019GDQT0008, PA2019GDPK0072).

†通讯作者. E-mail: zgf@hfut.edu.cn.

和计算机辅助工具,高效合理地将各储备点的救灾物资分发到各需求点,以最大程度地减少灾害带来的损失<sup>[1]</sup>. 现有研究大都关注于救灾物资的分配和调度<sup>[2-4]</sup>,并过于理想化假设运输路网是完好无损的,显然只适合于某些特定的灾害场景.

需要指出的是,不管决策者制定的救灾物资分配和调度方案有多么合理可行,没有高效、稳定、可靠的路网交通系统的保障,一切都是空谈. 因此,在灾害应急响应中,修复受损路网以快速打通生命通道是开展灾后救援工作的一个非常重要的基础环节<sup>[5-6]</sup>.

## 1 相关工作

受损路网修复问题主要研究如何规划道路抢修队伍的调度,即按什么顺序修复哪些受损路段可以使得路网的修复总耗时最小、运输效率最大,这对应急救援的实施和灾民的快速安全疏散具有重要的现实意义. 因此,灾后受损路网修复问题越来越受到各国政府和学者的高度重视.

Chen<sup>[7]</sup>将应急抢修的后勤保障调度构建为一个多物资的网络流问题,并设计了一种启发式算法,可以在规定的运营时间内降低后勤保障的短期运营成本. Nurre等<sup>[8]</sup>针对在极端破坏下的基础设施系统恢复服务问题,提出了一种新的启发式调度规则用来分配工作组将节点和弧建立到网络中,以最大限度地提高网络中的累积加权流量. 但是,这种路网修复是以增加新的节点和边为代价,耗时和成本均较大.

Yan等<sup>[9]</sup>基于应急抢修和救灾物资分配的双层时空网络,将道路抢修和物资分配构建为一个多目标、混合整数、多物资的网络流问题,并引入蚁群优化搜索最优抢修、物资分配路线和时间表. 陈钢铁等<sup>[10]</sup>建立了一个多目标混合整数规划模型以最小化紧急道路抢修和应急物资配送时间,并采用启发式算法进行求解. 李双琳<sup>[11]</sup>建立了上层以抢修绩效最大为目标,下层以应急物资配送时间最短以及公平性最大为目标的双层规划模型,并提出了一种混合遗传算法进行求解. Liberatore等<sup>[12]</sup>构建了一个将运输基础设施要素恢复运作与人道主义后勤援助分配相结合的模型,并用于海地大地震后的公路恢复计划,以提高援助分配的效果. 许蓓<sup>[13]</sup>考虑震后恶劣天气对抢修工作效率的影响和各需求点的不同紧迫度,设计了一种基于蚁群混合搜索的启发式算法. 邱慧<sup>[14]</sup>将灾后公路网修复分为应急修复期和全面修复期,在时间和设备资源的约束下,构建了一种基于最大连通子图和公路网加权效率的两阶段模型. 上述工作均将受损路网抢修和救灾物资分配或调度捆绑在一起,将受损

路网抢修问题看的过于单一化. 值得注意的是,在地震、洪水等特大自然灾害发生后,受损路网抢修本身往往就是一个具有暂时性的复杂系统工程. 在有限的抢修条件下,不可能也没有必要修复所有的受损路段,在决策修复哪些受损路段时,既要考虑抢修的时效性,又要尽可能地追求受损路网的运输效率. 同时,受损路网抢修还具有鲜明的不确定性,一些次生灾害也会影响受损路网的稳定性和可靠性<sup>[15-16]</sup>.

为此, Tuzun等<sup>[17]</sup>专门研究灾后道路障碍清理规划问题,提出了一种基于动态路径的数学模型,并设计了相应的启发式方法,以最大化整个路网的连通性和最小化路障清理时间. Akbari等<sup>[18]</sup>基于弧路由问题来规划灾后道路的障碍清除,设计了基于整数规划和启发式的混合算法以最小化路网重新连通的时间开销或在给定的时间约束内最大化路网重新连通后的总收益. 不过,上述方法均是着眼于路网本身,构建的路网大都过于理想化,仅考虑修复路网中的哪些路段可以实现目标的最优化,没有考虑这些受损路段是否可达以及受损路段的修复顺序对应急救援的影响.

Duque等<sup>[19]</sup>从道路抢修队(包含人员、设备和原材料等)自身的视角出发,针对灾后道路抢修队的调度和路线规划问题,基于动态规划和贪婪策略优化路网中需求点的可达性. Kim等<sup>[20]</sup>考虑额外损失和可变损失率等灾害因素下的道路抢修队调度问题,以尽量减少因受灾点不可达而造成的总损失. 上述工作虽然考虑了抢修队的可达路线,并能够给出道路抢修队的修复策略集,但仍然理想地假设修复一个路段总是比不修复有益,而且路网中任意两节点间不存在连续受损的路段,无法处理像地震、洪水等重大自然灾害引起的路网严重受损的情形. 我国幅员辽阔,地理气候条件复杂,地震、洪水灾害发生非常频繁,是世界上地震、洪水灾害损失最严重的少数国家之一. 因此,苏兆品等<sup>[21]</sup>引入马尔科夫决策过程来模拟抢修队的修复活动,并基于Q学习算法求解抢修队的最优调度策略,能够在一定程度上提高运输效率和修复效率. 但是,在其所提的方法中,抢修队的动作集定义为所有受损路段的集合,当路网受损严重时,抢修队需要面临巨大的决策负担和计算开销,而且往往会从动作集中选择一个根本不可达的受损路段,导致学习停滞和算法失效.

基于上述背景,本文针对严重受损路网环境,首先简化路网模型和决策模型,并基于动作集裁减和Q学习提出一种改进的抢修队调度算法,旨在降低抢修队的决策负担,提高算法在严重受损路网环境中的稳

定性和可靠性.

## 2 问题描述

根据灾后路网特点及抢修队修复过程,本节给出一个面向严重受损路网的抢修队调度模型.与惯例一样<sup>[21]</sup>,假设储备点处有且仅有一个抢修队,抢修队每次修复一个路段一定能成功修复,即修复的路段在救援黄金72小时内不会再次受损.此外,将真实路网中的所有过渡节点(即非需求点)根据最短路径算法从路网中去掉,从而使路网模型的规模大幅减小.据此,受损路网可描述为图1所示的无向图  $G = (N \cup \{0\}, E)$ .

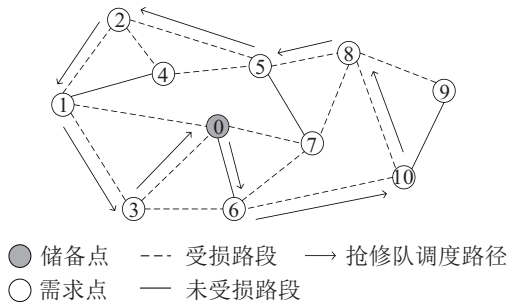


图1 受损路网示意图

在无向图  $G$  中,0代表储备点,  $N = \{1, \dots, m\}$  表示需求点集合,  $m \in \mathbf{N}$ . 抢修队需要从储备点0出发修复受损路段,从而让  $N$  中的每个需求点都可达,即与储备点0连通.这是因为,在灾害应急响应中,从社会心理学角度考虑,为了避免可能的群体性恐慌事件,每个需求点都应该得到必要的响应.对于  $\forall k \in N$  都有一个应急响应的紧迫度  $\theta_k \in \mathbf{Z}^+$ ,  $\theta_k$  越大,表示需求点  $k$  的受灾程度越严重,需要优先考虑.此外,每个需求点  $k$  还有一个可达标志位  $\delta(0, k) \in \{0, 1\}$ .  $\delta(0, k) = 1$  表示节点  $k$  与储备点0是连通的,即已经打通节点  $k$  与储备点0的路径;如果  $\delta(0, k) = 0$ ,则表示节点  $k$  目前不可达.另外,每个需求点  $k$  与储备点0之间有一个最短路径长度  $d(0, k) \in \mathbf{Z}^+$ ,如果  $\delta(0, k) = 0$ ,则  $d(0, k) = \infty$ .

$E$  为无向图  $G$  中所有边的集合,  $\forall e(i, j) \in E$  表示从起点  $i \in N \cup \{0\}$  到其相邻点  $j \in N \cup \{0\}$  的一条边(即路段),其物理距离为  $l(i, j)$ . 每个  $e(i, j)$  还具有一个修复标志位  $\beta(i, j) \in \{0, 1\}$ ,  $\beta(i, j) = 0$  表示  $e(i, j)$  是受损路段且还未修复,  $\beta(i, j) = 1$  表示  $e(i, j)$  是畅通的.由此可知,  $\delta(0, i/j) = 1$  表示路段  $e(i, j)$  的两个端点  $i, j$  至少有一个与储备点0是连通的,即抢修队可以到达路段  $e(i, j)$  的一端;如果  $\delta(0, i/j) = 0$ ,则表示路段  $e(i, j)$  目前不可达.

抢修队从储备点0出发进行路网修复,其平均

行驶速度为  $v \in \mathbf{Z}^+$ ;修复任意路段  $e(i, j)$  的耗时为  $t(i, j) \in \mathbf{Z}_0^+$ ,如果  $e(i, j)$  是未损路段,则  $t(i, j) = 0$ .

基于上述描述,抢修队调度问题指的是,抢修队应该选择哪些受损路段以及按照什么调度顺序对这些受损路段进行修复(如图1中的箭头实线所示),可以使受损路网的修复效率最大.不失一般性,假设所有需求点打通时抢修队的调度路径为  $P$ .显然,不同的  $P$  对应的修复效率也不同.在受损路网修复中,一般重点考虑调度策略的路径长度和修复耗时两个方面.首先,每个需求点  $k$  到储备点0的最短路径长度  $d(0, k)$  要尽可能地小,从而让救援物资可以更快地被输送到每个需求点,即

$$\min f_1(P) = \sum_{k=1}^m [\theta_k d(0, k)]. \quad (1)$$

其次,每个需求点被打通的总耗时(包括打通这个需求点时抢修队的累积行驶时间和修复路段时间)要尽可能地少,从而让每个需求点能尽快与储备点连通,及时得到响应.单个需求点  $k$  被打通的耗时可按照下式计算:

$$t_k = \sum_{e(i,j) \in P_k} \left\{ \frac{l(i,j)}{v} + [1 - \beta(i,j)]t(i,j) \right\},$$

其中  $P_k \subseteq P$  为抢修队从储备点0出发到打通需求点  $k$  时当前所走过的路径.据此,所有需求点被打通时的累积总耗时为

$$\min f_2(P) = \sum_{k=1}^m (\theta_k t_k). \quad (2)$$

综上所述,抢修队调度策略的目标函数<sup>[21]</sup>可设计为

$$\min f(P) = w f_1(P) + (1 - w) f_2(P), \quad (3)$$

其中权重  $w \in (0, 1)$  用来控制在两个衡量指标上的偏好.

## 3 决策模型

抢修队的调度问题显然是一个典型的序贯决策,因此本文沿用文献[21]的Markov决策模型,将抢修队看成是一个智能体(agent),并对agent的动作集进行必要的裁减,以提高agent的决策效率.

agent的动作集  $A = \{e(i, j) | \beta(i, j) = 0 \wedge \delta(0, i/j) = 1\}$ ,即当前可达的未修复受损路段集合,而不是文献[21]采用的所有受损路段的集合.这是因为,首先,如果采用所有受损路段,在agent选择下一个动作(即选择要修复哪一条受损路段)时,很可能选择的受损路段根本不连通,也就是说agent无法到达,从而造成agent停滞不前,导致决策的毫无目的性.其次,

选择当前可达的受损路段, agent 的动作集规模会大大减少, 从而可以在很大程度上减轻 agent 的决策负担, 而且可以增强 agent 每一个选择的动作对打通需求点的贡献, 因为要想继续修复路网, 只有这些可达受损路段是最佳选择, 目的性更强. 同时, 从常理上分析, 让 agent 从动作集中选择一个可达的就近受损路段也符合应急响应的时效性要求.

每当 agent 修复一个受损路段后, 路网主要发生两个变化, 首先是有新的需求点被打通, 其次是因为有新的可行路段产生, 路网中可通行部分的结构发生了变化, 当前已经打通的一些需求点到储备点 0 的最短路径长度可能会变得更小. 为了记录上述状态的变化, agent 的状态集  $S = \{s_0, s_1, s_2, \dots, s^*\}$ , 其中:  $\forall s_h \in S$  为一个二元组  $s_h = \langle K, D \rangle$ , 表示第  $h \in \mathbf{N}$  步的状态,  $K = \{k | \delta(0, k) = 1\}$  表示当前所有已经打通的需求点集合,  $D = \{d(0, k) | \delta(0, k) = 1\}$  表示这些连通的需求点对应的最短路径长度集合;  $s^*$  为最终状态, 表示所有  $m$  个需求点都与储备点 0 连通, 即  $|K| = m$ .

agent 从动作集  $A$  中选择一个动作  $e(i, j)$  执行, 从状态  $s = \langle K, D \rangle$  过渡到状态  $s' = \langle K', D' \rangle$ , 会得到相应的奖励, 以刺激 agent 自主学习. 根据式(3)的抢修队调度策略目标函数可以知道, 每个需求点  $k$  到储备点 0 的最短路径长度  $d(0, k)$  要尽可能地小, 同时  $k$  被打通的总耗时要尽可能地少, 就以式(3)的目标作为 agent 探索的内在诱因来设计奖励函数.

首先, 从最短路径长度角度考虑, 对于  $\forall k \in K'$ , 如果  $k \notin K$ , 即  $k$  是一个新打通的需求点, 则应该获得相应较大的奖励, 以诱导 agent 积极打通所有需求节点, 且对应的最短路径长度越小, 奖励应该越大. 此外,  $k$  的紧迫度  $\theta_k$  越大, 奖励也应该越大, 从而诱导 agent 优先考虑受灾严重的需求点. 基于此, 将 agent 通过  $k$  可获得的奖励设计为  $\frac{\theta_k}{d_{s'}(0, k)}$ . 同时, 如果  $k \in K$ , 即需求点  $k$  是前一个状态的历史需求节点, 则需要判断  $k$  的最短路径长度是否变小, 如果没有, 则 agent 获得奖励 0; 如果  $d_{s'}(0, k) < d_s(0, k)$ , 则找到了更好的最短路径, 理应获得些许额外的奖励, 但不应该太大, 因为 agent 的首要目标是尽可能地让新的需求点尽快地打通, 而不是优化已有连通节点的最短路径. 因此, 将 agent 通过  $k$  可获得的奖励设计为前后状态的奖励差值, 即  $\frac{\theta_k}{d_{s'}(0, k)} - \frac{\theta_k}{d_s(0, k)}$ . 基于上述分析, agent 在状态  $s'$  通过  $k$  在最短路径长度上可获得的奖励为

$$r_k^d = \begin{cases} \frac{\theta_k}{d_{s'}(0, k)}, & k \notin K; \\ \frac{\theta_k}{d_{s'}(0, k)} - \frac{\theta_k}{d_s(0, k)}, & k \in K, d_{s'}(0, k) < d_s(0, k); \\ 0, & k \in K, d_{s'}(0, k) = d_s(0, k). \end{cases}$$

根据上式, agent 在状态  $s'$  通过整个修复后的路网在最短路径长度上可获得的奖励为

$$R_d = \sum_{k \in K'} r_k^d. \quad (4)$$

其次, 从抢修耗时角度考虑, 当 agent 从状态  $s$  到达状态  $s'$  时, 当前所耗费的修复累积时间为

$$t_{s'} = \sum_{e(i, j) \in P_{s'}} \left\{ \frac{l(i, j)}{v} + [1 - \beta(i, j)]t(i, j) \right\},$$

其中  $P_{s'}$  为 agent 执行一系列动作到达状态  $s'$  时所走的路径(即经历的路段序列集合). 显然, 很容易得到

$$t_{s'} = t_s + \sum_{e(i, j) \in P_{s'} - P_s} \left\{ \frac{l(i, j)}{v} + [1 - \beta(i, j)]t(i, j) \right\}.$$

对于  $\forall k \in K'$ , 同样需要根据前面最短路径长度的 3 种情况讨论 agent 在状态  $s'$  时通过抢修时间获得的奖励. 同理, 如果  $k \notin K$ , agent 通过  $k$  可获得的奖励设计为  $\frac{\theta_k}{t_{s'}}$ . 如果  $k \in K$ , 而  $k$  的最短路径长度没有变小, 则 agent 获得奖励 0; 如果  $d_{s'}(0, k) < d_s(0, k)$ , 则将 agent 通过  $k$  可获得的奖励设计为前后状态的奖励差值, 即  $\frac{\theta_k}{t_{s'}} - \frac{\theta_k}{t_s}$ . 因此, agent 在状态  $s'$  通过  $k$  在抢修耗时上可获得的奖励为

$$r_k^t = \begin{cases} \frac{\theta_k}{t_{s'}}, & k \notin K; \\ \frac{\theta_k}{t_{s'}} - \frac{\theta_k}{t_s}, & k \in K, d_{s'}(0, k) < d_s(0, k); \\ 0, & k \in K, d_{s'}(0, k) = d_s(0, k). \end{cases}$$

根据上式, agent 在状态  $s'$  通过整个路网在抢修耗时上可获得的奖励为

$$R_t = \sum_{k \in K'} r_k^t. \quad (5)$$

综合考虑上述两因素, agent 在状态  $s'$  通过整个路网可获得的即时奖励为

$$R = wR_d + (1 - w)R_t, \quad (6)$$

其中权重  $w$  与式(3)一致. 需要说明的是, 这里的奖励函数并没有完全按照目标函数(3)来设计, 这是因为: 首先, 两者的设计目标是一致的, 都是为了追求路网修复后每个需求点到储备点的最短路径长度最小, 同时每个需求点被打通的累计修复时间尽可能地少, 且紧迫度越大的需求点越要优先考虑; 其次, 采用紧迫

度相除可以降低奖励函数取值范围,减少奖励值表的存储空间;第三,对于机器学习而言,只要按照要实现的目标给予agent合理的奖励刺激,agent就会按照奖励的趋势自主学习,朝着目标积极探索。

### 4 抢修队调度算法

本节基于Q学习对抢修队调度问题进行求解,与文献[21]不同的是,在本文的改进算法中,根据前面提出的路网模型和决策模型,在每次状态更新后,为了强化学习的正反馈机制,会对agent的动作集进行更新,只保留当前可达的未修复受损路段,目的性更强,从而保证agent的每个动作都是可执行的,同时减少了对Q值表的遍历代价,让agent的决策空间和选择压力逐步变小,提高Q学习的训练效率。

在Q学习中,Q值的迭代公式为

$$Q \leftarrow (1 - \alpha)Q(s, a) + \alpha[R + \gamma \max Q(s', a)]. \quad (7)$$

其中: $\alpha \in (0, 1)$ 为学习程度,用于调控之前训练产生的奖励对当前Q值更新的影响; $a \in A$ 为agent从动作集中选择的当前动作; $R$ 为agent采取动作 $a$ 可获得的即时奖励,可根据式(6)计算; $\gamma \in [0, 1]$ 为折扣因子,表示未来回报与当前回报之间的差异; $\max Q(s', a)$ 为在状态 $s$ 下选择动作 $a$ 进入下一状态 $s'$ 时能得到的最大Q值。

据此,本文抢修队调度算法的基本流程如图2所示,具体描述如下。

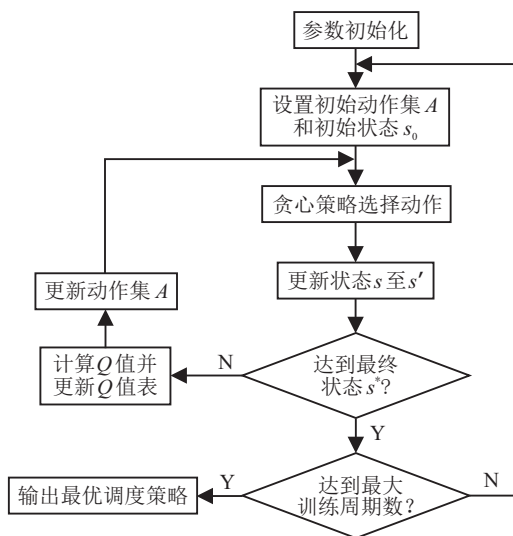


图2 基于Q学习的抢修队调度算法流程

step 1: 对路网模型、决策模型和Q学习的相关参数进行初始化;

step 2: 设置初始动作集 $A$ 和初始状态 $s_0$ ;

step 3: agent根据 $\epsilon \in (0, 1)$ 贪心策略从 $A$ 中选取动作,即按照大概率 $1 - \epsilon$ 从Q值表中选择Q值最大

的动作,按照小概率 $\epsilon$ 放弃历史学习经验,从 $A$ 中随机选取一个动作;

step 4: 执行选取的动作,从状态 $s$ 更新到状态 $s'$ ,根据式(6)和(7)计算对应的Q值并更新Q值表;

step 5: 如果当前状态 $s'$ 已达到最终状态 $s^*$ ,则结束本轮学习,否则根据当前的可达受损路段情况更新agent的动作集 $A$ ,转入step 3继续学习;

step 6: 如果已达最大训练周期数,则结束训练并根据Q值表输出agent的最优调度策略,否则回到step 2继续训练。

### 5 仿真实验与分析

#### 5.1 实验环境和评价指标

为了验证本文所提算法的有效性,随机模拟3组不同节点数{37, 50, 62}的路网环境.为了模拟严重受损路网,路段受损率分别为{80%, 90%}. $\theta_k \in (0, 10)$ , $l(i, j) \in [1, 10]$ , $t(i, j) \in [1, 10]$ ,均在所给的取值范围内随机生成. $v = 1, w = 0.9$ .为了对比的公平性,Q学习的关键参数设置与文献[21]保持一致,训练周期数在[100, 1500]之间, $\epsilon = 0.1, \alpha = 0.4, \gamma = 0.2$ .

在下面的实验中,将本文所提算法与文献[21]算法分别从算法运行时间、目标值、修复路段占比、需求点连通率这4个指标进行对比分析.每个测试实例在Intel Core i5 CPU 2.50 GHz、4 GB内存、Windows 10操作系统的个人计算机上独立运行30次,并根据30次不同结果进行统计分析。

#### 5.2 算法时效性

在本节,根据式(3)计算最优调度方案对应的目标函数值来对比分析两种算法在路网严重受损环境下的时效性。

表1给出了两种算法在24个不同测试实例(对应6种路网参数和4种不同算法参数)下的运行时间(均值 $\pm 95\%$ 置信区间,单位为s).从表1可以看出,在所有24个实例上,本文算法的运行时间都远远小于文献[21]算法,平均下来大约只有文献[21]算法耗时的1/3.这是因为,文献[21]算法的动作集是路网中的所有受损路段,在路网严重受损时,其动作集异常庞大,agent在Q值表中的遍历代价也相当大,从而带来了巨大的计算开销.与之相比,本文算法对动作集进行裁减,促使agent不再作全局盲目的探索学习,而是仅在当前可达的受损路段中进行动作的选择,因而极大减少了对Q值表的遍历代价,提高了agent的学习效率。

表1 两种算法在不同测试实例下的运行时间(均值±95%置信区间)

路网参数	算法	训练周期数			
		100	250	500	1000
节点数37,受损率80%	文献[21]算法	5.86±0.11	12.27±0.21	22.75±0.35	44.12±0.91
	本文算法	<b>1.61±0.08</b>	<b>2.77±0.05</b>	<b>5.16±0.2</b>	<b>9.85±0.34</b>
节点数37,受损率90%	文献[21]算法	5.27±0.26	12.94±0.44	24.49±0.5	46.46±2
	本文算法	<b>1.96±0.11</b>	<b>4.39±0.27</b>	<b>8.8±0.28</b>	<b>17.59±0.48</b>
节点数50,受损率80%	文献[21]算法	13.15±0.29	30.72±1.02	55.47±0.55	116.64±3.26
	本文算法	<b>2.71±0.08</b>	<b>5.69±0.12</b>	<b>10.89±0.24</b>	<b>26.12±1.77</b>
节点数50,受损率90%	文献[21]算法	15.67±0.26	37.61±0.72	94.77±8.25	118.73±2.15
	本文算法	<b>3.73±0.13</b>	<b>8.73±0.37</b>	<b>15.65±0.59</b>	<b>30.86±1.3</b>

路网参数	算法	训练周期数			
		250	500	1000	1500
节点数62,受损率80%	文献[21]算法	30.25±0.92	59.5±1.97	128.86±7.88	176.05±6.4
	本文算法	<b>9.95±0.31</b>	<b>20.55±1.07</b>	<b>38.36±1.41</b>	<b>51.7±0.76</b>
节点数62,受损率90%	文献[21]算法	37.12±1.48	72±2.83	142.44±5.27	221.9±8.14
	本文算法	<b>10.36±0.18</b>	<b>23.28±1.85</b>	<b>59.82±0.64</b>	<b>89.59±1.45</b>

表2给出了两种算法在不同测试实例下的目标函数值(均值±标准差).从表2可以看出,在所有的24个实例上,本文算法在20个实例上都获得了更好的目标值,而文献[21]算法仅在4个实例上取得了占优.特别地,文献[21]算法在节点数50、受损率80%、训练周期数100的实例下,30次运行仅仅获得了1个解,而且在受损率提高到90%时算法失效.上述实验表明,本文算法在路网严重受损时明显比文献[21]算

法更加有效,可以得到更佳的抢修队调度方案.此外,文献[21]算法的可靠性也不如本文算法,存在算法失效的风险,而本文算法在所有720次运行中均能稳定地给出调度方案.这是因为,文献[21]算法中agent在选择动作时可能选择了不可达的受损边,而本文算法中agent只会从当前所有可达的受损边里面选择下一个动作,从而确保了算法的有效性.

表2 两种算法在不同测试实例下的目标函数值(均值±标准差)

路网参数	算法	训练周期数			
		100	250	500	1000
节点数37,受损率80%	文献[21]算法	9 220.46±392.97	9 456.34±418.13	9 173.37±692.22	9 554.06±868.69
	本文算法	<b>8 931.68±672.65</b>	<b>8 681.28±599.46</b>	<b>8 616.76±489.88</b>	<b>8 576.39±555.49</b>
节点数37,受损率90%	文献[21]算法	<b>10 828.33±624.53</b>	11 236.65±994.81	<b>10 083.54±1 342.22</b>	11 544.89±1 680.62
	本文算法	10 972.53±944.45	<b>11 136.97±920.73</b>	10 708.86±861.11	<b>11 137.41±1 047.07</b>
节点数50,受损率80%	文献[21]算法	28 019.7	23 791.33±347.16	22 910.45±234.26	23 103.49±1148.03
	本文算法	<b>22 952.62±1 865.98</b>	<b>23 061.14±1 235.53</b>	<b>21 948.91±1 157.86</b>	<b>22 324.64±1 303.82</b>
节点数50,受损率90%	文献[21]算法	—	<b>23 256.67±1 259.32</b>	23 943.7±2 298.56	25 506.27±3 186.75
	本文算法	<b>23 010.69±2 073.48</b>	23 753.45±1 927.89	<b>23 366.04±1 493.57</b>	<b>24 124.88±2 062.72</b>

路网参数	算法	训练周期数			
		250	500	1000	1500
节点数62,受损率80%	文献[21]算法	<b>24 490.05±507.63</b>	27 626.57±2 260.67	27 117.5±1 752.88	27 660.88±2 464.65
	本文算法	25 807.53±1 922.6	<b>26 628.05±1 875.03</b>	<b>26 444.08±1 695.5</b>	<b>25 901.48±1 494.61</b>
节点数62,受损率90%	文献[21]算法	28 287.88±3 158.88	27 285.78±1 466.45	27 702.1±1 542.89	29 337.3±2 281.91
	本文算法	<b>26 358.96±1 681.36</b>	<b>26 293.34±1 695.01</b>	<b>27 684.28±1855.44</b>	<b>27 322.09±1 598.29</b>

为了进一步对比算法的稳定性和可靠性,图3统计了两种算法在不同受损率和不同节点数下所有120次运行获得的有效解占比.从图3可以看出,文献[21]算法在路网受损严重的情形下找到有效解的几率普遍较低,大约都在30%以内,例如在受损率80%和节点数50时有效解占比仅有11.67%,在受损率90%和节点数62时,有效解占比大约在30.82%左右.相反地,无论受损率和节点数如何设置,本文算法在每个实例下的每次运行全部都能给出有效解,具有显著的可靠性.

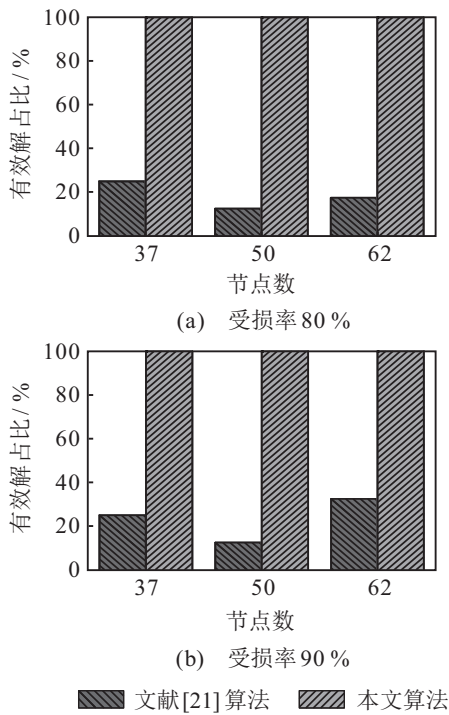


图3 两种算法在不同受损率和节点数下的有效解占比

上述实验结果表明,文献[21]算法将路网中全部受损路段作为agent的动作集,这种全搜索策略在面向严重受损路网时,agent经常会选择到根本不可达的受损路段,导致agent学习停滞,因此其算法极不稳定,经常性地失效,大大降低了其算法的可用性.而本文算法对动作集进行合理地裁减,只保留agent当前可达的受损路段,有效保证了agent学习的连续性,显著增强了算法的稳定性.

从整体上而言,在路网受损严重的环境中,特别是节点数较多时,路网中存在着大量的受损路段,这给求解抢修队调度问题带来极大的困难.与文献[21]算法相比,本文算法从微观视角出发,优先考虑agent可达的受损路段,从局部探索逐步向全局覆盖,可以更快地给出更优的调度方案,而且具有更高的稳定性和可靠性.

### 5.3 调度方案的修复效率

抢修队调度的最终目标是在尽可能短的时间内,即尽可能地修复少的路段就能迅速打通所有需求点,同时让每个需求点到储备点的最短路径长度尽可能地小.鉴于此,在本节,从两种算法获得的最优调度方案对应的修复路段占比和需求点连通率来进一步对比分析路网修复效率.

图4给出了两种算法在不同受损率和不同节点数下所获得的最优调度方案对应修复路段占比,即总共修复的路段数占受损路段总数的比例.从图4可以看出,文献[21]算法在路网受损严重情形下的修复路段占比较大,平均在86.8%左右,而本文算法的修复路段占比要比文献[21]算法小,最小占比只有67.9%,最大占比也只有81.6%,低于文献[21]算法的平均值.

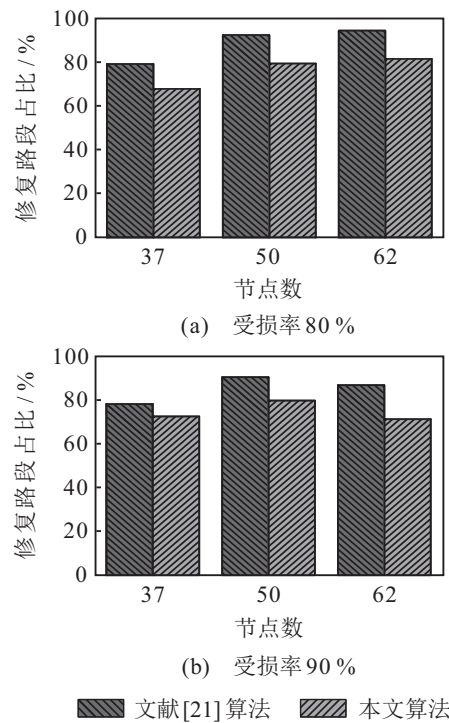


图4 两种算法在不同受损率和节点数下的修复路段占比

上述实验结果表明,本文算法给出的调度方案只需修复较少的路段就能让所有需求点尽快连通,因此本文算法所给的调度方案修复代价更小、效率更高.

为了更加直观形象地展示两种算法调度方案的修复效率,图5给出了在不同受损率和不同节点数下两种算法最优调度方案在修复过程中的需求点连通率变化情况,即每修复一条受损路段后,整个路网中与储备点连通的需求点占比.从图5可以看出,在每个测试环境下,本文算法的需求点连通率相比文献[21]算法增长趋势更快,能够在修复较少受损路段时就能让需求点连通率快速达到100%.例如,在图5(d)中,本文算法在修复了23条路段时需求点连通率就

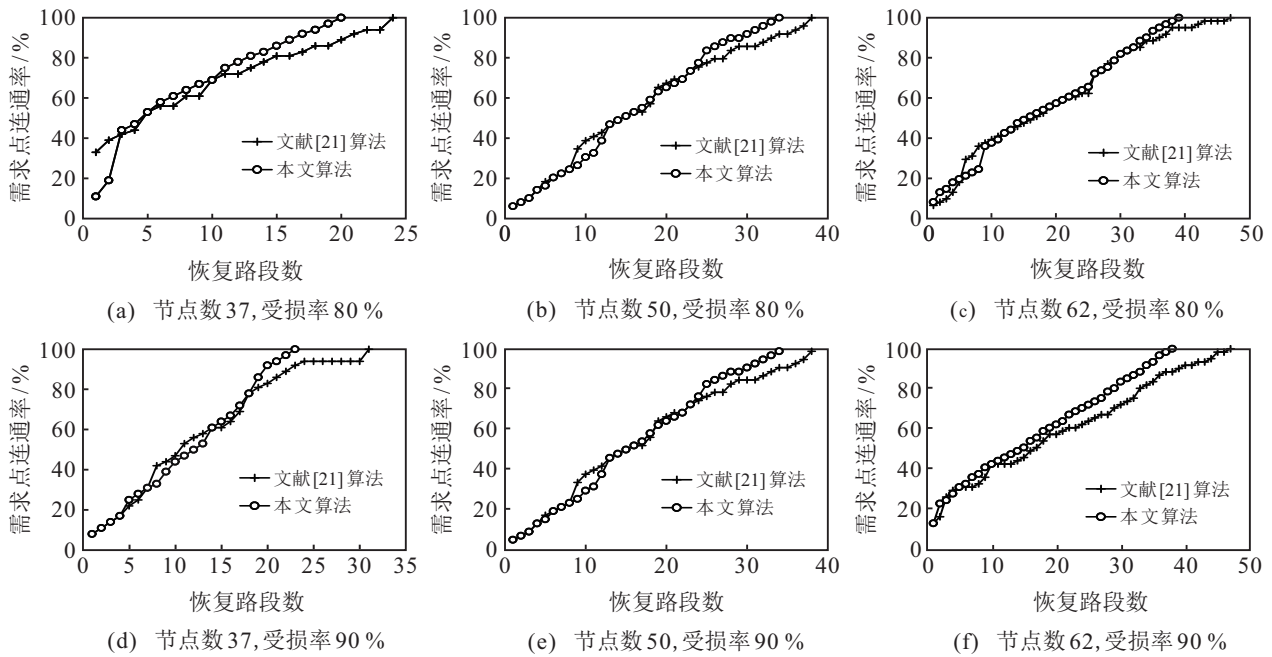


图5 两种算法在不同受损率和节点数下的需求点连通率变化

已达100%,而文献[21]算法在修复23个路段时的需求点连通率不到92%。特别地,在节点数和受损率都比较大时,此时路网受损极其严重,而本文算法越是在这种恶劣情形下修复效率越高。在图5(f)中,本文算法在修复了38条路段时需求点全部连通,而文献[21]算法此时的需求点连通率只有88.5%。此外,还可以看出,越往后期,本文算法效率越高,修复更快,这是因为在本文算法中,随着修复进程的推进,agent的动作集只专注于当前所能到达的受损路段,因而随着待打通的需求点数的减少,算法的决策空间和选择压力越来越小,目的性更强,正反馈机制更加突出。反观文献[21]算法,每次都是从整个路网的受损路段集中选择动作,而不管这些受损路段与需要打通的需求节点的关联性,决策空间和选择压力巨大,因此其算法可能要执行一系列动作,通过整个路网长期的奖励之后才能看到明显的修复效果。

## 6 结论

修复受损路网以快速打通生命通道是开展灾后救援的一个非常重要的基础环节。本文针对严重受损路网环境下已有抢修队调度算法无法保证给出有效解的问题,通过简化路网模型和决策模型,基于动作集裁减和Q学习提出了一种面向严重受损路网的抢修队调度改进算法。对比实验结果表明,在节点数和受损率都较大的严重受损路网环境中,本文算法更加稳定可靠,而且能够以更小的代价给出更优的抢修队调度方案。在今后的工作中,将考虑次生灾害的影响,构建动态变化的受损路网模型。同时,还要研究

多抢修队从不同储备点出发,并协同合作修复受损路网,这个时候每个agent的学习终止条件不再是探索完所有需求点,而是只需探索其周围的部分需求点,可以考虑给每个agent设置一个模糊可达度,这将是下一步的研究重点。

## 参考文献(References)

- [1] Anaya-Arenas A M, Renaud J, Ruiz A. Relief distribution networks: A systematic review[J]. *Annals of Operations Research*, 2014, 223(1): 53-79.
- [2] 刘扬, 张国富, 苏兆品, 等. 救灾物资多阶段分配与调度问题建模与求解[J]. *控制与决策*, 2019, 34(9): 2015-2022.  
(Liu Y, Zhang G F, Su Z P, et al. Modeling and solving multi-phase allocation and scheduling of emergency relief supplies[J]. *Control and Decision*, 2019, 34(9): 2015-2022.)
- [3] 张国富, 王永奇, 苏兆品, 等. 应急救援物资多目标分配与调度问题建模与求解[J]. *控制与决策*, 2017, 32(1): 86-92.  
(Zhang G F, Wang Y Q, Su Z P, et al. Modeling and solving multi-objective allocation-scheduling of emergency relief supplies[J]. *Control and Decision*, 2017, 32(1): 86-92.)
- [4] 苏兆品, 张国富, 蒋建国, 等. 基于非支配排序差异演化的应急资源多目标分配算法[J]. *自动化学报*, 2017, 43(2): 195-214.  
(Su Z P, Zhang G F, Jiang J G, et al. Multi-objective approach to emergency resource allocation using none-dominated sorting based differential evolution[J]. *Acta Automatica Sinica*, 2017, 43(2): 195-214.)

- [5] Galindo G, Batta R. Review of recent developments in OR/MS research in disaster operations management[J]. *European Journal of Operational Research*, 2013, 230(2): 201-211.
- [6] Çelik M. Network restoration and recovery in humanitarian operations: Framework, literature review, and research directions[J]. *Surveys in Operations Research and Management Science*, 2016, 21(2): 47-61.
- [7] Chen S Y. Optimal scheduling of logistical support for an emergency roadway repair work schedule[J]. *Engineering Optimization*, 2012, 44(9): 1035-1055.
- [8] Nurre S G, Cavdaroglu B, Mitchell J E, et al. Restoring infrastructure systems: An integrated network design and scheduling (INDS) problem[J]. *European Journal of Operational Research*, 2012, 223(3): 794-806.
- [9] Yan S, Shih Y L. An ant colony system-based hybrid algorithm for an emergency roadway repair time-space network flow problem[J]. *Transportmetrica*, 2012, 8(5): 361-386.
- [10] 陈钢铁, 帅斌. 震后道路抢修和应急物资配送优化调度研究[J]. *中国安全科学学报*, 2012, 22(9): 166-171. (Chen G T, Shuai B. Optimizing emergency road repair and distribution of relief supplies after earthquake[J]. *China Safety Science Journal*, 2012, 22(9): 166-171.)
- [11] 李双琳. 基于震后疏散行为和应急物资调运的道路交通应急恢复策略研究[D]. 成都: 西南交通大学, 2014. (Li S L. Emergency restoration strategies for post-earthquake road-networks based on evacuation behavior and relief distribution[D]. Chengdu: Southwest Jiaotong University, 2014.)
- [12] Liberatore F, Ortuño M T, Tirado G, et al. A hierarchical compromise model for the joint optimization of recovery operations and distribution of emergency goods in humanitarian logistics[J]. *Computers & Operations Research*, 2014, 42(3): 3-13.
- [13] 许蓓. 震后应急期道路抢修与物资运输联合调度[D]. 上海: 上海交通大学, 2015. (Xu B. Joint scheduling of post-earthquake road repair and relief distribution[D]. Shanghai: Shanghai Jiao Tong University, 2015.)
- [14] 邱慧. 灾后公路网修复序列研究[D]. 西安: 长安大学, 2016. (Qiu H. Study on the restoration sequence of highway network after disaster[D]. Xi'an: Chang'an University, 2016.)
- [15] Ahmad Hosseini S, Wadbro E. Connectivity reliability in uncertain networks with stability analysis[J]. *Expert Systems with Applications*, 2016, 57: 337-344.
- [16] Soltani-Sobh A, Heaslip K, Stevanovic A, et al. Evaluation of transportation network reliability during unexpected events with multiple uncertainties[J]. *International Journal of Disaster Risk Reduction*, 2016, 17: 128-136.
- [17] Tuzun Aksu D, Özdamar L. A mathematical model for post-disaster road restoration: Enabling accessibility and evacuation[J]. *Transportation Research Part E: Logistics and Transportation Review*, 2014, 61(1): 56-67.
- [18] Akbari V, Salman F S. Multi-vehicle synchronized arc routing problem to restore post-disaster network connectivity[J]. *European Journal of Operational Research*, 2017, 257(2): 625-640.
- [19] Duque P A M, Dolinskaya I S, Sörensen K. Network repair crew scheduling and routing for emergency relief distribution problem[J]. *European Journal of Operational Research*, 2016, 248(1): 272-285.
- [20] Kim S, Park Y, Lee K, et al. Repair crew scheduling considering variable disaster aspects[C]. *Proceedings of the IFIP International Conference on Advances in Production Management Systems*. Hamburg: Springer, 2017: 57-63.
- [21] 苏兆品, 李沫晗, 张国富, 等. 基于Q学习的受灾路网抢修队调度问题建模与求解[J]. *自动化学报*, 2020, 46(7): 1467-1478. (Su Z P, Li M H, Zhang G F, et al. Modeling and solving the repair crew scheduling for the damaged road networks based on Q-learning[J]. *Acta Automatica Sinica*, 2020, 46(7): 1467-1478.)

## 作者简介

张国富(1979—),男,教授,博士,从事智慧应急、进化算法、软件工程等研究, E-mail: zgf@hfut.edu.cn;

涂冰花(1995—),女,硕士生,从事智慧应急的研究, E-mail: tubh@mail.hfut.edu.cn;

苏兆品(1983—),女,副教授,博士,从事智慧应急、多媒体安全等研究, E-mail: szp@hfut.edu.cn;

岳峰(1981—),男,副研究员,博士,从事软件工程、多媒体安全等研究, E-mail: yuefeng@hfut.edu.cn.

(责任编辑: 齐 霖)