

受灾路网抢修队动态调度问题的建模与求解方法

张国富^{1,2,3}, 沈宇峰¹, 宋晓晓¹, 苏兆品^{1,2,3}

(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: 地震、受灾路网抢修作为灾后应急响应中的一个基础环节, 主要研究如何制定道路抢修队的修复方案, 从而快速打通生命救援线路, 确保救援队伍、装备、物资等及时输送到灾区各个需求点。然而, 已有研究大多专注于静态受灾路网, 难以适应地震、洪水等重特大自然灾害下的复杂应急场景。构建一种动态受灾路网模型, 模拟应急场景的动态恶化, 并基于 Markov 决策过程构建抢修队的动态决策模型, 设计相应的状态空间、动作空间和回报函数。最后, 提出一种基于改进 Q 学习 (IQL) 的动态调度 (IQLDS) 算法, 以适应当前的路网状态, 快速得到较优的修复策略。实验结果表明, 与蚁群优化算法、IQL 算法相比, IQLDS 算法在大规模、高受损率路网环境中的目标函数值降低了约 50%, 能够在精确感知路网环境变化后及时调整学习策略, 并充分利用历史经验获得较优的修复方案。
关键词: 灾后应急响应; 路网修复; 抢修队动态调度; Q 学习; 最优动作集更新

开放科学(资源服务)标志码(OSID):



中文引用格式: 张国富, 沈宇峰, 宋晓晓, 等. 受灾路网抢修队动态调度问题的建模与求解方法[J]. 计算机工程, 2023, 49(6): 300-313.

英文引用格式: ZHANG G F, SHEN Y F, SONG X X, et al. Method for modeling and solving the dynamic scheduling problem of the repair crew for restoring damaged road network[J]. Computer Engineering, 2023, 49(6): 300-313.

Method for Modeling and Solving the Dynamic Scheduling Problem of the Repair Crew for Restoring Damaged Road Network

ZHANG Guofu^{1,2,3}, SHEN Yufeng¹, SONG Xiaoxiao¹, SU Zhaopin^{1,2,3}

(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] Restoring damaged road network is a basic part of post-disaster emergency response, and studies mainly focus on how to develop a restoration scheme for the repair crew to quickly open up a rescue route and ensure that the rescue teams, equipment, and relief supplies can be transported to the disaster area in a timely manner. However, most existing works focus on a static road network, which is difficult to adapt to the complex emergency scenarios in extraordinarily serious natural disasters, such as earthquakes and floods. A dynamic disaster road network model is constructed to simulate the dynamic deterioration of emergency scenarios. Based on the Markov decision process, a dynamic decision model for the emergency repair team is constructed, and a corresponding state space, action space, and return function are designed. Finally, an Improved Q-Learning (IQL) for Dynamic Scheduling (IQLDS) algorithm is proposed to adapt to the current state of the road network and quickly make better repair strategies. Experimental results show that, compared with Ant Colony Optimization (ACO) algorithm and the IQL algorithm, the value of the objective function obtained by our method for a large-scale and high-damage road network is reduced by approximately 50%. Furthermore, our method can adjust the learning strategy in a timely manner after accurately perceiving the changes in the road network and make full use of historical experience to obtain a better repair plan. These improvements may help emergency management departments facilitate a schedule plan for the repair crew for restoring a damaged road network more efficiently and effectively.

基金项目: 教育部人文社会科学研究青年基金(19YJC870021); 中央高校基本科研业务费专项资金(PA2021GDSK0073, PA2021GDSK0074); 安徽省重点研究与开发计划项目(202104d07020001, 202004d07020011); 广东省类脑智能计算重点实验室开放课题(GBL202117)。

作者简介: 张国富(1979—), 男, 教授、博士, 主研方向为智慧应急; 沈宇峰、宋晓晓, 硕士研究生; 苏兆品, 副教授、博士。

收稿日期: 2022-05-16 **修回日期:** 2022-06-30 **E-mail:** zgfh@hfut.edu.cn

[Key words] post-disaster emergency response; road network restoration; dynamic scheduling of the repair crew; Q-Learning (QL); optimal action set updating
DOI: 10.19678/j.issn.1000-3428.0064702

0 概述

近年来,以“大智移云”为特征的新一代信息技术的快速发展,给应急管理的信息、智能化和科学化提供了新的机遇和挑战。因此,“十四五”国家应急体系规划中明确提出,要利用大数据、人工智能、机器学习等新一代信息技术提高我国重特大自然灾害风险感知、监测预警、协同救援和应急处置的能力,为推进应急管理体系和能力现代化建设提供科技支撑,实现大国智慧应急。其中,对受灾路网进行合理的修复以快速恢复路网功能,打通救援生命线,保障救援队伍、装备和应急物资能够及时输送到各个需求点是灾后开展应急处置和救援的基本前提,对迅速转移安置灾民、最大限度降低灾害损失具有重要的现实意义。

传统灾害应急智能决策研究大多聚焦于疏散^[1-2]、救灾物资的分配和调度^[3-4]等,这些应急环节能够顺利实施的重要前提是具有一个高效稳定、安全可靠的路网交通系统。因此,近年来,在灾害应急智能决策研究中,修复受灾路网以快速打通救援通道越来越受到重视。其核心问题是在时间紧迫、人力资源有限和各个需求点受灾程度不同的情况下,确定修复哪些受损路段以及按照什么顺序修复,可以迅速打通所有需求点,使抢修效率、运输效率和社会效用最大化,避免因应急响应滞后而可能发生的社会群体性事件。

为此,YAN等^[5-6]构建了一种整数多时空网络流模型,将该模型划分成若干子问题后利用CPLEX工具最小化修复路段带来的短期运营成本。为了能够将蚁群优化(Ant Colony Optimization,ACO)算法应用于求解时空网络流模型,YAN等^[7]改进了ACO算法的探索机制,并结合阈值接受技术以最小化网络流模型的修复时间。基于该网络流模型,陈钢铁等^[8]设计了一种启发式算法来同时最小化路段修复时间和应急物资运输时间。为了提高策略的求解速度,LIBERATORE等^[9]将受损路网进行简化,去掉中间不必要的过渡节点,并以出救点和需求点之间、各需求点之间的最短路径为边构造一个无向图,提出一种考虑抢修成本和运输时间的多准则层次模型,并根据目标的偏好顺序使用分层法进行优化。AKSU等^[10]采用启发式方法优化路网的连通性和路障清理时间。SUZAN等^[11]采用拉格朗日松弛与次梯度算法最小化储备站与各需求点间的累计加权距离,以此保证灾后物资运输的效率。LI等^[12]从交通系统优化角度出发,建立路网抢修与应急物资配送双层优化模型,在最大化抢修绩效的同时确保物资分配的公平性。不过,上述方法大都关注修复哪些受损路

段可以最大程度地恢复路网的连通性,而没有考虑修复时间顺序,因此并不能指导道路抢修队的实际修复活动。

为确定哪些路段需要恢复以及这些路段的修复时间顺序,DUQUE等^[13]针对乡村受灾公路网的修复问题,提出一种基于贪婪策略和动态规划的混合算法,从而最大化路网中需求点的可达性。AKBARI等^[14]将受灾路网修复问题看作弧路由问题,设计一种基于启发式和整数规划的混合算法以最小化时间开销和最大化抢修绩效。KIM等^[15]采用ACO算法优化灾区的总损失和路网修复时间。SHIN等^[16]基于ACO算法缩短路网修复后的运输总时间。SANCI等^[17]提出一种样本平均近似法求解抢修队的总路程开销和人员调度成本。MORENO等^[18-19]考虑各个需求点不可达的累计加权时间,并使其尽可能小。李兆隆等^[20]基于遗传算法优化应急恢复阶段和全面恢复阶段的路网恢复速度和累计损失。苏兆晶等^[21]认为受损路网抢修是一个典型的时序决策问题,提出一种基于Q学习(Q-Learning,QL)的道路抢修队调度算法。张国富等^[22]通过对道路抢修队的动作集进行裁减,提出一种改进的Q学习策略(Improved QL,IQL)。

考虑到多抢修队下的应急场景,LI等^[12]构建了相应的非线性规划模型,并基于改性型细菌群体趋化性算法进行求解。宣洪伟等^[23]将抢修过程分解为有核合作博弈和空核合作博弈两个子问题。AJAM等^[24]假定了抢修队从不同储备点出发的情况,利用启发式算法来优化抢修队到达关键位置的总时间开销。

需要指出的是,虽然上述已有工作在其构建的应用场景中具有一定的合理性,但只适用于静态的路网环境。而在实际的应急场景中,尤其是地震、洪水等重特大自然灾害发生后,往往伴随着余震、泥石流、山体滑坡等次生灾害和衍生灾害,此时原来的非需求点可能会变成新的需求点,原来的可通行路段有可能变成新的受损路段,甚至修复好的路段又重新成为受损路段。已有方法在处理上述路网状态进一步恶化的情况时,往往只能终止求解过程,根据变化后的路网重新进行求解,导致决策效率极低。

本文在总结和分析已有工作的基础上,针对受灾路网状态的动态变化,构建一种动态路网模型,模拟应急场景的动态恶化。同时,基于Markov决策过程构建抢修队的动态决策模型,提出一种基于Q学习和最优动作集更新策略的受灾路网抢修队动态调度算法。最后通过对比实验验证所提方法的有效性。

1 动态路网模型

为了适应灾后路网中一些非需求点可能会转换成潜在的需求点,本文考虑如图1所示的路网无向图 $G = \{V, E\}$ 。其中: V 为路网中所有节点的集合,包括物资储备站“0”、需求点集合 $\vec{V}$ 和非需求点集合 $\tilde{V}$, 即 $V = \{0\} \cup \vec{V} \cup \tilde{V}$; E 为所有路段的集合,包括可通行路段集合 $\bar{E}$ 和受损路段集合 $\tilde{E}$, 即 $E = \bar{E} \cup \tilde{E}$ 。

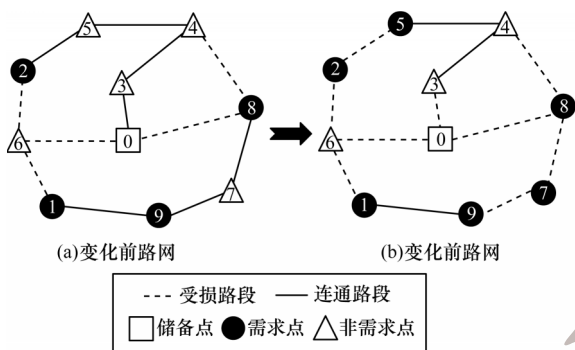


图1 动态受损路网示意图

Fig.1 Schematic diagram of dynamic damaged road network

$\forall i \in V$ 有一个权重 $I_i \in \mathbb{R}^+$, 表示节点 i 的灾情严重程度。当 $i \in \{0\} \cup \tilde{V}$, $I_i = 0$ 时, I_i 越大, 则节点 i 的受损情况越严重。为了保证灾后物资能够快速运往需求点,修复工作完成后各个需求点与储备点之间的路径长度不宜过大,每个节点到储备站“0”都有一个最大可接受距离 D_i , 表示路网修复后节点 i 到储备站“0”之间的最优路径长度 D_{i0} 不能高于 D_i , 否则会影响后续应急响应工作的运输效率。此外,用 $\theta_i \in \{0, 1\}$ 表示节点 i 的可达状态: 当 $\theta_i = 1$ 时,从物资储备站“0”到节点 i 是可达的,即节点 i 的救援线路已被打通; 当 $\theta_i = 0$ 时,表示节点 i 处于不可达的孤立状态。

$\forall e_{ij} \in E$ 有一个通行距离 $l_{ij} \in \mathbb{R}^+$ 。 $\delta_{ij} \in \{0, 1\}$ 表示路段 e_{ij} 的连通状态: 当 $\delta_{ij} = 1$ 时,路段 e_{ij} 当前是可通行的; 当 $\delta_{ij} = 0$ 时,表示 e_{ij} 为一个受损路段。因此,有 $\tilde{E} = \{e_{ij} | \delta_{ij} = 0\}$ 。假设每个受损路段都因为灾害的影响而导致无法正常使用(包括路面坍塌、山石堵塞、洪水淹没等),且只有在抢修队对其进行了恢复作业(如修路、搭桥、清障等)后才能允许通行。因此,对于每个受损路段,有一个恢复作业所对应需要的时间开销 $t_{ij} \in \mathbb{R}^+$ 。

道路抢修队从储备站“0”出发修复受损路段,假设所有需求点都连通时抢修队走过的路段有序集(包含可通行路段和受损路段)为 H 。抢修工作的首要任务就是要让救援物资与队伍尽快达到需求点,这需要需求点的连通时间与物资储备点到需求点之间的运输路径都尽可能短。

因此,本文首先希望在策略 H 下运输效率越大

越好,即每个需求点到储备站“0”的最短路径长度要尽可能小:

$$\min f_1(H) = \sum_{i \in \vec{V}} (I_i \times D_{i0}) \quad (1)$$

其中: D_{i0} 为根据策略修复受灾路网后的需求点 i 到储备站“0”的最短路径长度,可由 Dijkstra 算法直接得到,需要满足 $D_{i0} \leq D_i$ 。 D_{i0} 越小,说明从储备站“0”到需求点 i 的物资运输距离越短,运输效率越高。

其次,假设道路抢修队的通行速度为 $v \in \mathbb{R}^+$, 本文希望在策略 H 下修复效率越大越好,每个不可达的需求点 i 要尽可能早地被打通,即当每个不可达的需求点与储备站“0”连通时,抢修队的累积时间开销要尽可能小,其表达式如式(2)所示:

$$\min f_2(H) = \sum_{i \in \vec{V} \wedge \theta_i = 0} (I_i \times C_{i0}) \quad (2)$$

其中, C_{i0} 为不可达的需求点 i 与储备站“0”连通时抢修队的累积时间开销,包含通行时间和修复时间:

$$C_{i0} = \sum_{e_{ij} \in H_{i0}} \left[\frac{l_{ij}}{v} + (1 - \delta_{ij}) \times t_{ij} \right] \quad (3)$$

其中, H_{i0} 为抢修队从储备站“0”到 i 所经历的路段有序子集。

基于上述考虑,动态受灾路网的修复问题可描述为如式(4)所示的约束动态优化问题:

$$\begin{aligned} \min f(H) &= (1 - \lambda) f_1(H) + \lambda f_2(H) \\ D_{i0} &\leq D_i, \forall i \in \vec{V} \end{aligned} \quad (4)$$

其中: $\lambda \in (0, 1)$ 为加权权重,表示运输效率和修复效率之间的偏好。

需要注意的是,在式(4)中, $\vec{V}$ 、 $\tilde{V}$ 、 $\bar{E}$ 、 $\tilde{E}$ 等可能会随着时间的推移而发生变化。例如,在图1中,初始路网环境的需求节点包含 $\vec{V} = \{1, 2, 8, 9\}$, 受损路段 $\tilde{E} = \{e_{06}, e_{08}, e_{16}, e_{26}, e_{34}, e_{48}\}$ 。在某一时间点,由于次生灾害的发生,路网环境发生变化,其中需求点集合为 $\vec{V} = \{1, 2, 5, 7, 8, 9\}$, 受损路段集合为 $\tilde{E} = \{e_{03}, e_{06}, e_{08}, e_{16}, e_{25}, e_{26}, e_{34}, e_{48}, e_{78}, e_{79}\}$ 。

2 抢修队动态决策模型

受灾路网修复是一个序贯决策,而 Markov 决策过程是序贯决策的一种经典表达形式,通常包括状态空间、动作空间和回报函数3个基本要素。本文将抢修队看作一个智能体(agent)。在动态路网环境下,agent 需要适应当前最新的路网状态,从而能够选择较优的修复策略。因此,状态空间 S 、动作空间 A 和回报函数 R 的设计应该能够精确感知路网环境的变化。

2.1 状态空间

在已有工作中^[21-22],把 agent 的状态表示为可达需求节点集合与已修复受损路段集合的一个二元组合。这种描述方法聚焦的是 agent 已执行的修复动作和产生的效果,虽然可以直观地表达抢修队的修复绩效,但是很难适应路网的动态变化。例如,在

图1(a)中,假设agent从储备站“0”出发,沿着规划好的抢修路线前进,先后修复了受损路段 e_{06} 和 e_{26} ,此时agent的状态应为 $\langle \{2\}, \{e_{06}, e_{26}\} \rangle$ 。假设在某一时间点发生了次生灾害,路网环境的变化表现在图1(b)中,原来的非需求节点5和节点7变为新的需求节点,原来的连通路段 e_{03}, e_{25}, e_{78} 和 e_{79} 变为新的受损路段。在新一轮的探索中,如果agent重复变化前的修复规划,修复受损路段 e_{06} 和 e_{26} 到达节点2,此时agent状态还是 $\langle \{2\}, \{e_{06}, e_{26}\} \rangle$,因此会直接覆盖历史学习经验(相当于抹掉了历史学习经验),更无法感知路网环境的变化,不能根据变化后的路网环境自主地适应决策空间的变化,因此无法有效刺激抢修队的决策。

为解决上述问题,本文专注于路网中有哪些节点和路段可以响应,设计了一个三元组 $s = \langle \bar{V}, \bar{E}, k \rangle$ 来表示agent的状态。其中: $\bar{V} = \{i | \theta_i = 0\} \subseteq \vec{V}$ 表示不可达的需求点集合; $\bar{E}$ 为受损路段集合; k 为agent所在的节点序号;状态空间可表示为 $S = \{s\}_{2^{|\bar{V}|} \times 2^{|\bar{E}|} \times |\vec{V}|}$ 。

据此,在上述例子中,当变化前agent到达节点2时所处的状态为 $\langle \{1, 8, 9\}, \{e_{08}, e_{16}, e_{34}, e_{48}\}, 2 \rangle$,路网发生变化后agent在节点2的状态则更新为 $\langle \{1, 5, 7, 8, 9\}, \{e_{03}, e_{08}, e_{16}, e_{25}, e_{34}, e_{48}, e_{78}, e_{79}\}, 2 \rangle$ 。这种状态表示方法不仅保留了变化前的历史学习经验,而且能够准确地感知路网环境的细微变化给策略空间带来的巨大变化,有利于刺激agent在变化后的策略空间中利用历史经验进行探索。

2.2 动作空间

在受灾路网修复过程中,agent所能执行的动作与其当前所处的状态密切相关。当处于状态 $s = \langle \bar{V}, \bar{E}, k \rangle$ 时,agent需要考虑从 k 出发,选择哪条受损路段进行修复。为了提高agent执行抢修动作的效率,避免agent选择无法执行的修复动作,和文献[22]一样,本文将agent的动作集定义为agent可到达的受损路段集合:

$$A = \{e_{ij} | \delta_{ij} = 0 \wedge \theta_{i,j} = 1\} \quad (5)$$

即受损路段 e_{ij} 连接的2个节点 i, j 中至少有一个是可达的。决策模型每次转移到新状态 s 时,都会根据路网中各边的通行情况,更新相应的动作空间。

2.3 回报函数

通常,在包含大量需求点的路网环境下,agent探索的目标是尽可能优先打通邻近且受损严重的需求节点。即使没有需求节点可被打通,agent也会在尽可能短的时间内打通中间的非需求节点,以期加速需求点的连通。基于这个考虑,假设agent在状态为 $s_1 = \langle \bar{V}_1, \bar{E}_1, k_1 \rangle$ 时,从动作空间中选择一条受损路段 e_{ij} ,进行修复后到达下一个状态 $s_2 = \langle \bar{V}_2, \bar{E}_2, k_2 \rangle$,可能的情况有:

1)有新的需求点被打通,即 $V_2 \subset V_1$ 。例如,在图1(a)中,当agent修复路段 e_{16} 后,需求点1和需求点9均可达,显然这个动作的执行给路网的修复带来了极大的绩效,需要给予强烈的正向反馈。因此,本文把这种情形下动作的即时奖励值设置为:

$$R = \sum_{i \in \bar{V}_1 - \bar{V}_2} \left[(1 - \lambda) \frac{I_i}{D_{i0}} + \lambda \frac{I_i}{C_{i0}} \right] \quad (6)$$

即新打通需求节点在最短路径长度和累积时间开销上带来的回报。

2)没有新的需求点被打通,即 $V_2 = V_1$ 。例如,在图1(a)中,agent修复路段 e_{06} 后,并没有新的需求点能够被打通,但是到达节点6后对打通需求点1和需求点2是非常有利的。因此,本文把这种情形下动作的即时奖励值设置为:

$$R = \frac{1}{\frac{l_{ij}}{v} + t_{ij}} \quad (7)$$

即agent在受损路段 e_{ij} 上耗费的时间开销越小,带来的回报越大。

与已有工作不同的是,在上述回报函数中,本文以新打通的需求点的最短路径长度和累积时间开销作为奖励值的计算,而没有考虑前期其他已经可达的需求点,进一步降低了计算开销和存储代价。这是因为在强化学习中,通常只需根据目标给予agent合理的奖励刺激,agent即会按照奖励的趋势自主学习,并朝着目标积极探索,而没有必要完全按照目标函数来设计奖励值。

3 抢修队动态调度算法

本节基于Q学习对抢修队动态调度问题进行求解。在Q学习中,问题求解的策略通常用一个Q值进行衡量。其中,Q值反映的是在状态 s 下采取动作 e_{ij} 能够获得的总期望收益,并储存在以 $\langle s, e_{ij} \rangle$ 为索引的Q值表中。Q值越大,表示对应动作选择策略所带来的效益越多。由于重大灾害事故大多具有突发性和偶然性,难以对各个 $\langle s, e_{ij} \rangle$ 下的期望收益做出评估。因此,本文将Q值表中的Q值初始化为0,在每轮迭代中学习经验,使Q值不断趋于实际收益。Q值的迭代表达式^[21-22]如式(8)所示:

$$Q \leftarrow (1 - \alpha) Q(s_1, e_{ij}) + \alpha [R + \gamma \max Q(s_2, e_{ij})] \quad (8)$$

其中: $\alpha \in (0, 1)$ 为学习速率,用于调控前期训练产生的奖励对当前Q值更新的影响, α 越大,保留之前训练的效果就越少; $e_{ij} \in A$ 是agent从动作空间中选择的动作; R 为agent选择动作 e_{ij} 可获得的即时奖励,可根据式(6)或式(7)计算得到; γ 是折扣因子,表示未来回报与当前回报之间的差异, γ 越大,表示越重视历史经验; $\max Q(s_2, e_{ij})$ 是在状态 s_1 下选择动作 e_{ij} 进入下一个状态 s_2 时能得到的最大Q值。

3.1 最优动作更新策略

基于 Q 值的计算, Q 学习在每个状态上根据动作空间,会形成一张 Q 值表。根据贪心策略,状态 s 下的agent会有较大的概率从最优动作集中随机选取一条边作为执行动作。其中,最优动作集为状态 s 在 Q 表中最大 Q 值所对应的动作集合。然而,在传统的 Q 学习中,初始 Q 值表中的值均是从0开始, Q 值的更新缓慢。在没有足够周期训练下, Q 值不能准确反映相应 $\langle s, e_{ij} \rangle$ 下的 Q 值,可能会导致状态 s 下收益更高的动作 e_{ij} 无法更新到 s 所对应的最优动作集中。其次,如果路网规模和受损路段数较大,使得每个状态所对应的动作数目巨多,这种传统方法就会导致 Q 值表的更新变得更加迟缓。而且,由于本文的路网环境是动态恶化的,仅依靠变化前的历史经验很难在有限的训练周期内学习到较好的策略。

为解决上述问题,本文提出一种最优动作集更新策略。在每个训练周期结束后,基于学习到的修复策略 H ,首先根据式(4)计算修复策略 H 对应的目标函数值,然后与历史最佳目标值进行对比,如果优于历史目标值,就说明在该周期找到了一个更好的修复策略,根据修复策略 H 中每个状态动作对 $\langle s, e_{ij} \rangle$,将动作 e_{ij} 放入 s 状态下的最优动作集中,并直接以新环境下的目标函数值来驱使agent学习和适应新的环境。这样,在下一个学习周期中,agent就会有一定的概率选择到当前对目标函数值最有利的历史经验,从而提升agent的学习效率。

3.2 IQLDS算法框架

本文基于改进 Q 学习的抢修队动态调度(Improved Q-Learning for Dynamic Scheduling, IQLDS)算法的基本流程如图2所示,具体描述如下:

步骤1 对路网模型、决策模型和 Q 学习的相关参数进行初始化。

步骤2 判断路网环境是否发生变化。如果是,则更新路网环境。

步骤3 将agent放置在储备站“0”并设置初始状态 $s_1 = \langle \bar{V}_1, \bar{E}, 0 \rangle$ 和初始动作集 A 。

步骤4 agent根据 $\varepsilon \in (0, 1)$ 贪心策略从 A 中选取动作 e_{ij} ,即按照大概率 $(1-\varepsilon)$ 从 Q 值表的最优动作集中随机选择一个动作,按照小概率 ε 放弃历史学习经验,从 A 中随机选取一个动作。

步骤5 执行选取的动作 e_{ij} ,从状态 s_1 过渡到状态 $s_2 = \langle \bar{V}_2, \bar{E}_2, k_2 \rangle$,根据式(6)~式(8)计算对应的 Q 值并更新 Q 值表。

步骤6 如果当前状态 s_2 已达到最终状态(即 $\bar{V}_2 = \emptyset$),则终止本轮学习,根据式(4)计算调度策略 H 的目标函数值,如果优于历史最佳目标值,则基于 H 更新 Q 值表中的最优动作集。如果 H 不是最终状态(即 $\bar{V}_2 \neq \emptyset$),则更新agent的动作集 A ,转步骤4继续学习。

续学习。

步骤7 如果已达最大训练周期数,则终止训练并输出当前最佳目标函数值对应的最优调度策略 H ,否则回到步骤2继续训练。

每次训练都会得到相应的受损路段序列作为最终的策略,抢修队可以以最终修复策略中的路段次序为参照进行实际的修复工作。

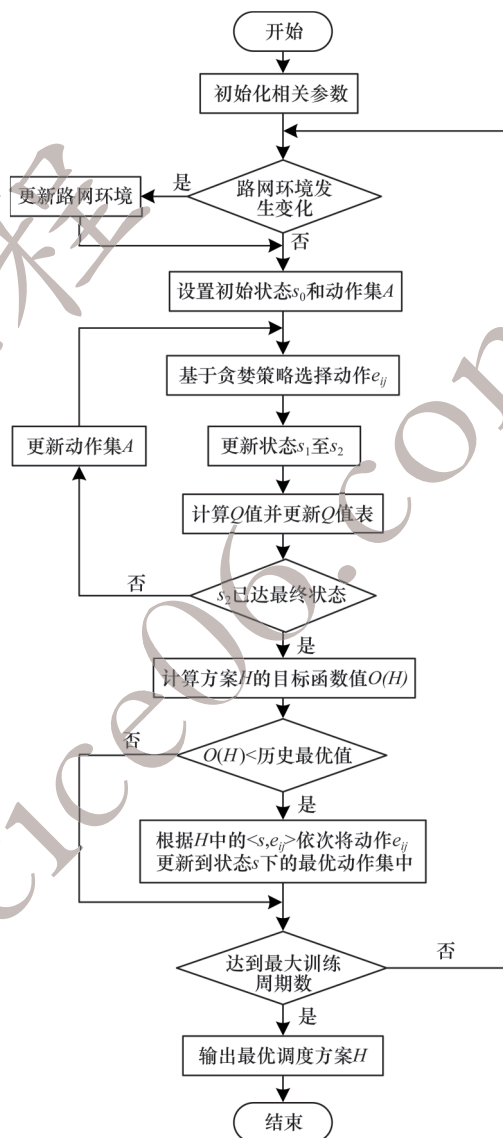


图2 IQLDS算法的流程

Fig.2 Procedure of IQLDS algorithm

4 实验结果与分析

为验证本文所提模型和IQLDS算法的有效性,在本节中,本文首先给出实验环境和参数设置,然后对比分析本文所设计的最优动作集更新策略的优劣,最后将本文设计的IQLDS算法与已有的ACO算法^[15-16]和IQL算法^[22]进行深入地对比分析。IQL算法与本文IQLDS算法均是基于 Q 学习的算法,而ACO是群智能优化算法。IQL算法和IQLDS算法的训练周期与ACO算法的迭代次数不具有相容性。

因此,为了对比的公平性,本文以静态 ACO 算法的结果为参照物,以 IQL 与 IQLDS 算法来模拟动态场景的优化。在实验中,使用 ACO 算法优化最后变化的静态路网,采用 IQL 与 IQLDS 算法优化动态路网,以 ACO 算法的结果来反映 IQL 算法与 IQLDS 算法在动态环境下的训练程度。

4.1 实验环境与参数设置

为全面评估算法的性能,本文设计了 3 种不同规模的受灾路网,如表 1 所示,节点数和路段数逐渐增加,以考察算法对问题规模的鲁棒性。此外,经过一定训练周期后路网会发生变化,以模拟次生灾害带来的影响。次生灾害发生前后需求点和受损路段的受损程度 $(|\vec{v}|/|V|,|\vec{E}|/|E|)$ 如表 2 所示。本文考虑了 9 种不同的变化,分别对应需求点数 $|\vec{v}|$ 增加、受损路段数 $|\vec{E}|$ 增加、 $|\vec{v}|$ 和 $|\vec{E}|$ 同时增加这 3 种情形,每种情形又分别考虑了 3 种不同的路网受损情况。本文根据表 1 的路网规模、表 2 的需求点占比 $|\vec{v}|/|V|$ 、受损路段占比 $|\vec{E}|/|E|$ 共模拟了 27 种动态路网,每种动态路网分别生成 10 个不同的测试实例,每个对比算法总共运行 270 个测试实例。每个算法对每个测试实例均在 Intel® Xeon® CPU 2.10 GHz、64 GB 内存、Windows Server 2012 操作系统的个人计算机上独立运行 30 次。为了比较不同算法的性能,本文利用式(4)对 30 次不同结果计算相应的目标函数并进行统计分析。

表 1 不同路网规模下的参数设置

Table 1 Parameter settings under different scales road network

规模	训练周期数	发生变化的周期数
$ V =40, E =60$	1 000	400
$ V =60, E =90$	1 500	600
$ V =80, E =120$	2 000	800

表 2 路网变化前后的受损程度

Table 2 Damage degree before and after the changes of the road network

变化情形	变化前	变化后
$ \vec{v} $ 增加	(0.30,0.40)	(0.35,0.40)
	(0.40,0.50)	(0.45,0.50)
	(0.50,0.60)	(0.55,0.60)
$ \vec{E} $ 增加	(0.30,0.40)	(0.30,0.50)
	(0.40,0.50)	(0.40,0.60)
	(0.50,0.60)	(0.50,0.70)
$ \vec{v} $ 和 $ \vec{E} $ 同时增加	(0.30,0.40)	(0.35,0.40)
	(0.40,0.50)	(0.45,0.60)
	(0.50,0.60)	(0.55,0.70)

此外,式(4)中目标函数的加权权重 $\lambda=0.9$ 。在 IQL 算法和 IQLDS 算法中,探索率 $\varepsilon=0.1$,折扣因子

$\gamma=0.2$,学习速率 $\alpha=0.4$,最大训练周期数和发生变化的周期如表 1 所示。在 ACO 算法中,蚂蚁数为 80,总迭代次数为 300,信息素挥发率为 0.5,信息素调控因子为 0.003,路程调控因子为 0.02。

4.2 加权权重 λ 对修复策略的影响

本文将运输效率与修复效率的加权和最小化作为调度算法的优化目标,如式(4)所示,其中 $\lambda\in(0,1)$ 为加权权重,用来调控修复策略下的偏好。

为了反映 λ 对修复策略的影响,本文将不同 λ 取值下的 IQLDS 算法对规模为 $(|V|=80,|E|=120)$,受损程度为 $(|\vec{v}|/|V|=0.5,|\vec{E}|/|E|=0.6)$ 的路网求解,并分别对修复时间和运输距离进行数值分析(均值和标准差),所得结果如表 3 所示。其中,当 $\lambda=0.1$ 时,最终方案所需要的修复时间最少,但是相应的运输距离最长。随着 λ 取值的提升,运输距离逐步缩小,修复时间也逐步增加。这是因为 λ 越大,运输效率所占的权重越大,在求解过程中会更加注重于优化方案的运输距离; λ 越小,则修复效率所占的权重越大,agent 会倾向选择时间开销更小的策略。

表 3 不同加权权重 λ 对修复策略的影响

Table 3 Influence of different weighting weights λ on repair strategy

加权权重 λ	运输距离	修复时间
0.1	1 022.75±77.86	2 148.0±139.37
0.3	991.50±47.10	2 550.8±79.43
0.5	926.20±22.58	2 731.9±89.04
0.7	903.10±16.92	2 924.8±98.99
0.9	892.90±12.99	3 127.6±99.40

4.3 最优动作集更新策略对 IQLDS 算法性能的影响

如上文所述,传统 Q 学习在选择动作时大概率会从 Q 值表中选择 Q 值最大的最优动作,但是 Q 值表的更新较为缓慢,在路网规模较大时,Q 学习的探索效率降低。为此,本文基于 IQLDS 算法提出一种最优动作集更新策略,让当前最佳目标函数值对应的动作来引导 agent 快速适应路网的变化。因此,在第 1 个实验中,本文基于 IQLDS 算法来评估最优动作集更新策略对 IQLDS 算法性能的影响。此外,为了验证算法求解复杂问题的鲁棒性,本文基于 $|\vec{v}|$ 和 $|\vec{E}|$ 同时增加这种最复杂的情形来进行测试。

表 4 给出了引入最优动作集更新策略的 IQLDS 算法在不同规模下的路网受损程度 $(|\vec{v}|/|V|,|\vec{E}|/|E|)$ 从(0.30,0.40)变为(0.35,0.50)后获得的目标函数值均值和标准差,其中加粗数字表示该组数据最小值。由表 4 可知,引入最优动作集更新策略的 IQLDS 算法在 30 个不同测试实例上获得的目标函数值均显著优于没有引入更新策略的 IQLDS 算法。具体来说,目标函数均值在 3 种路网规模上分别降低了约 25%、25% 和 28%。

表4 IQLDS算法在不同规模路网下的目标函数值(均值和标准差)对比

Table 4 Comparison of the target function values (mean and standard deviation) of the IQLDS algorithm under different scale road networks

路网规模	$ V =40, E =60$		$ V =60, E =90$		$ V =80, E =120$	
	有	无	有	无	有	无
实例1	1 771.77±124.01	2 260.35±273.34	4 951.46±459.17	7 074.72±1 434.76	12 842.63±1 198.56	17 877.98±3 079.52
实例2	1 732.88±120.05	2 592.76±567.30	7 579.46±559.42	11 531.13±1 622.89	11 583.75±1 362.03	16 133.32±2 721.38
实例3	1 683.71±166.37	2 289.41±494.11	3 965.27±309.55	5 171.64±853.36	12 544.22±1 485.72	17 772.43±2 184.21
实例4	2 507.97±260.90	3 324.29±438.47	3 612.22±406.56	5 014.14±1 229.79	11 338.04±1 019.42	15 092.86±1 788.85
实例5	2 269.84±184.03	2 891.16±362.95	6 248.78±368.18	7 871.02±823.03	9 708.69±1 118.73	15 647.58±3 571.38
实例6	2 406.54±195.84	2 838.18±345.05	3 843.59±450.82	5 066.29±575.68	9 315.63±842.03	11 590.87±1 869.04
实例7	1 944.64±196.84	2 698.73±483.44	4 954.95±386.91	6 406.80±948.90	13 041.96±1 156.23	17 587.08±2 465.95
实例8	3 616.05±290.61	4 797.18±991.27	4 891.38±433.80	6 290.11±987.84	11 338.5±832.70	15 601.87± 2682.11
实例9	1 517.08±21.11	1 633.66±127.76	5 370.41±487.93	7 320.09±944.43	1 4361.67±1 201.81	19 220.75±2 846.58
实例10	2 856.06±156.17	5 329.08±487.03	10 153.67±585.51	12 864.96±1 382.29	9 524.72±1 242.68	13 770.56±1 938.26

表5给出了有无最优动作集更新策略在不同路网受损程度($|V|/|V|, |E|/|E|$)变化情形下获得的目标函数值对比,表中加粗数字表示该组数据最小值。由表5可知,有更新策略的算法在30个不同测试实例上的目标函数值显著优于无更新策略的算法,其目标函数均值在3种路网受损程度变化情形下分别降低了约24%、25%和27%。上述实验结果表明,在相同受损程度变化情形下,采用最优动作集更新策略可以提高Q学习的探索性能,且路网规模越大,性能

提升越明显。此外,在相同路网规模下,随着路网受损程度的增加,最优动作集更新策略也显著提升了Q学习的求解性能。这是因为随着路网规模和受损程度的增加,agent的状态集与动作集规模呈指数增长,而传统Q学习很难迅速适应路网环境的动态变化。相反的,最优动作集更新策略可以充分利用历史学习经验,用当前对目标函数值最有利的动作来指导agent快速适应路网环境的变化,从而在一定程度上提升Q学习的学习效率。

表5 IQLDS算法在规模为 $|V|=80, |E|=120$ 路网下的目标函数值(均值和标准差)对比

Table 5 Comparison of the target function values (mean and standard deviation) of the IQLDS algorithm under a road network with a scale of $|V|=80, |E|=120$

路网的受损程度变化	(0.30, 0, 40)→(0.35, 0.50)		(0.40, 0, 50)→(0.45, 0.60)		(0.50, 0, 60)→(0.55, 0.70)	
	有	无	有	无	有	无
实例1	7 340.92±973.39	9 597.40±1 429.98	13 981.81±1 587.34	18 782.62±4 477.23	24 175.51±1 316.70	31 693.94±4 228.08
实例2	8 922.77±1 153.41	12 286.44±2 817.36	18 003.12±2 021.76	25 096.36±3 254.43	24 531.87±2 094.08	36 829.54±6 134.28
实例3	10 961.76±1 217.01	14 089.41±2 150.08	18 362.87±1 362.25	23 607.73±3 458.73	28 200.33±1 335.54	37 163.27±6 413.32
实例4	6 328.50±555.46	8 619.20±1 060.48	12 701.04±996.36	16 198.39±1 760.02	19 958.95±1 078.82	25 427.80±2 538.59
实例5	10 784.66±806.74	13 838.33±1 768.48	18 040.79±1 251.61	22 382.98±1 992.24	28 214.02±1 513.42	41 465.97±6 122.94
实例6	10 652.73±700.40	13 440.94±1 525.62	15 584.72±789.76	20 865.36±2 198.73	24 725.23±1 160.69	31 604.01±3 401.86
实例7	11 084.92±1 196.07	14 848.45±1 985.94	16 515.31±1 423.54	22 963.58±2 855.63	22 808.69±1 308.86	27 495.41±4 032.59
实例8	5 692.94±471.05	7 538.19±972.39	15 705.61±1 334.52	22 718.41±2 860.50	20 177.27±2 000.21	33 446.40±5 326.77
实例9	8 321.95±970.94	11 320.91±1 999.97	16 708.33±1 179.05	21 792.21±2 272.26	27 512.73±1 183.53	35 826.73±2 987.91
实例10	8 883.34±619.04	11 397.77±1 410.65	17 173.63±1 059.72	23 042.98±2 874.05	24 604.59±1 587.00	31 776.12±3 386.71

4.4 算法有效性分析

在第2个实验中,本文基于27种不同动态路网和270个测试实例对比分析ACO算法、IQL算法和本文IQLDS算法的性能。其中,IQLDS算法和IQL

算法在优化过程中需要面对路网的变化,而ACO算法则优化变化后的静态路网。

图3~图5分别给出了3种算法在不同路网规模下不同测试实例的目标函数值(均值和标准差)。从

图3~图5中可以看出,在27种不同的动态受损路网下,IQLDS算法得到的目标函数值均显著优于IQL算法和ACO算法。特别是随着路网规模的增大,IQL算法和ACO算法所得的解波动很大,与IQLDS算法差距越来越明显。与ACO算法和IQL算法相比,IQLDS算法在大规模、高受损率路网环境中的目标函数值降低了约50%。上述实验结果表明,IQLDS算法可以很好地适应路网环境的动态恶化,所得修复策略明显优于IQL算法和ACO算法。这是因为ACO算法中动作的选择只根据当前节点位置的信息素浓度大小,导致信息素浓度的更新只考虑当前节点的连通情况,而不考虑其他需求点的连

通性。随着路网规模的增加,ACO算法的探索能力变差,很容易陷入局部最优。对于IQL算法,在路网发生变化后,只能根据路网更新后的信息重新优化,不能有效利用前面周期的优化结果。而本文IQLDS算法考虑受灾路网的总体连通情况,将当前路网的孤立需求点、未被修复的受损路段以及当前节点位置作为agent的状态,可以精准地感知路网环境的变化。同时基于最优动作集更新策略,充分利用历史经验,可以在剩余的有限训练周期内迅速适应路网环境的变化,从而得到更好的修复策略。图6~图8分别给出了3种算法在不同路网规模下不同测试实例所耗费的运行时间(均值和标准误差)。

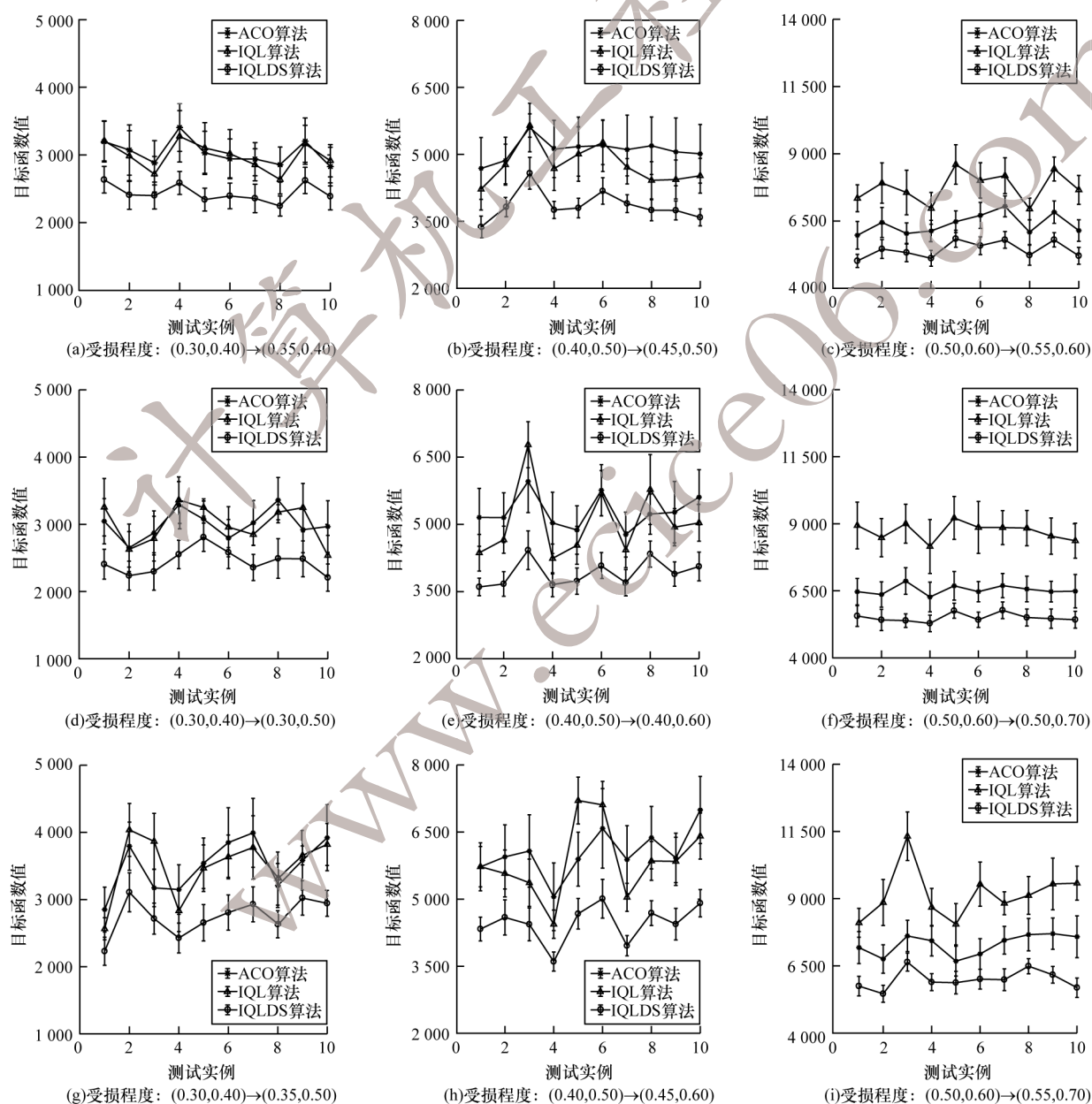
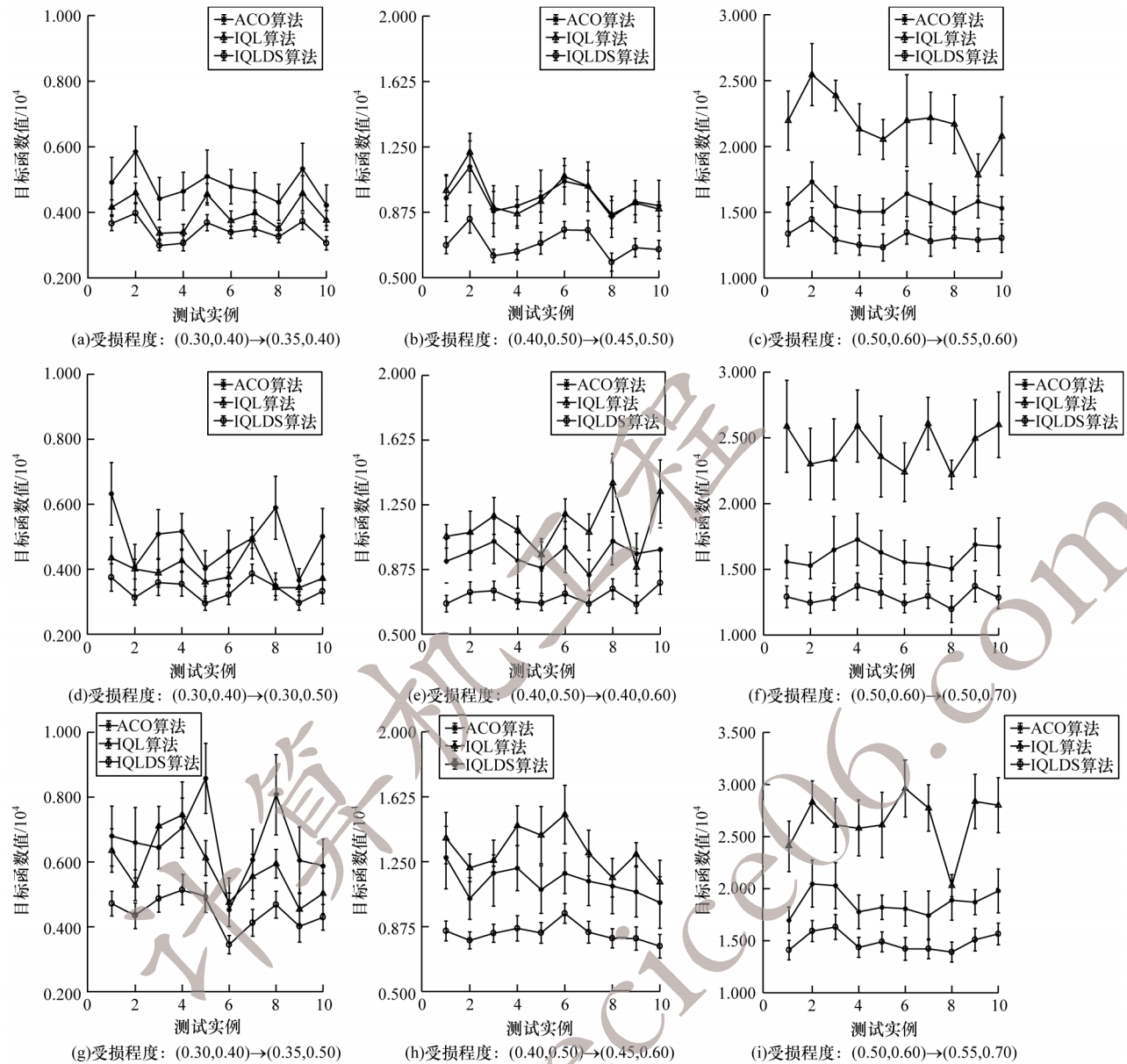
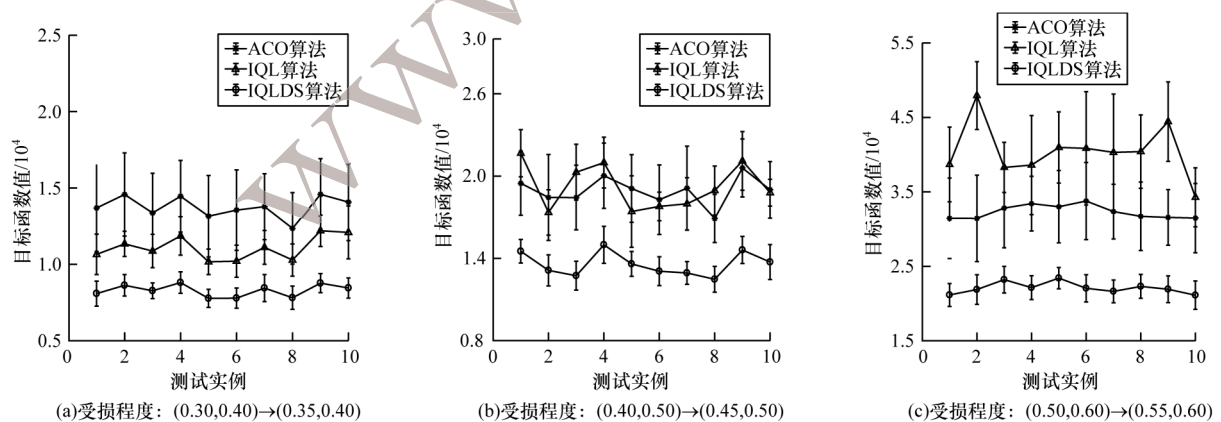


图3 不同算法在规模为 $|V|=40$, $|E|=60$ 路网下的目标函数值(均值和标准差)对比

Fig.3 Comparison of objective function values (mean and standard deviation) of different algorithms under road networks with a scale of $|V|=40$ and $|E|=60$

图4 不同算法在 $|V|=60, |E|=90$ 路网下的目标函数值(均值和标准差)对比Fig.4 Comparison of objective function values (mean and standard deviation) of different algorithms under road networks with a scale of $|V|=60$ and $|E|=90$ 图5 不同算法在规模为 $|V|=80, |E|=120$ 路网下的目标函数值(均值和标准差)对比Fig.5 Comparison of objective function values (mean and standard deviation) of different algorithms under road networks with a scale of $|V|=80$ and $|E|=120$

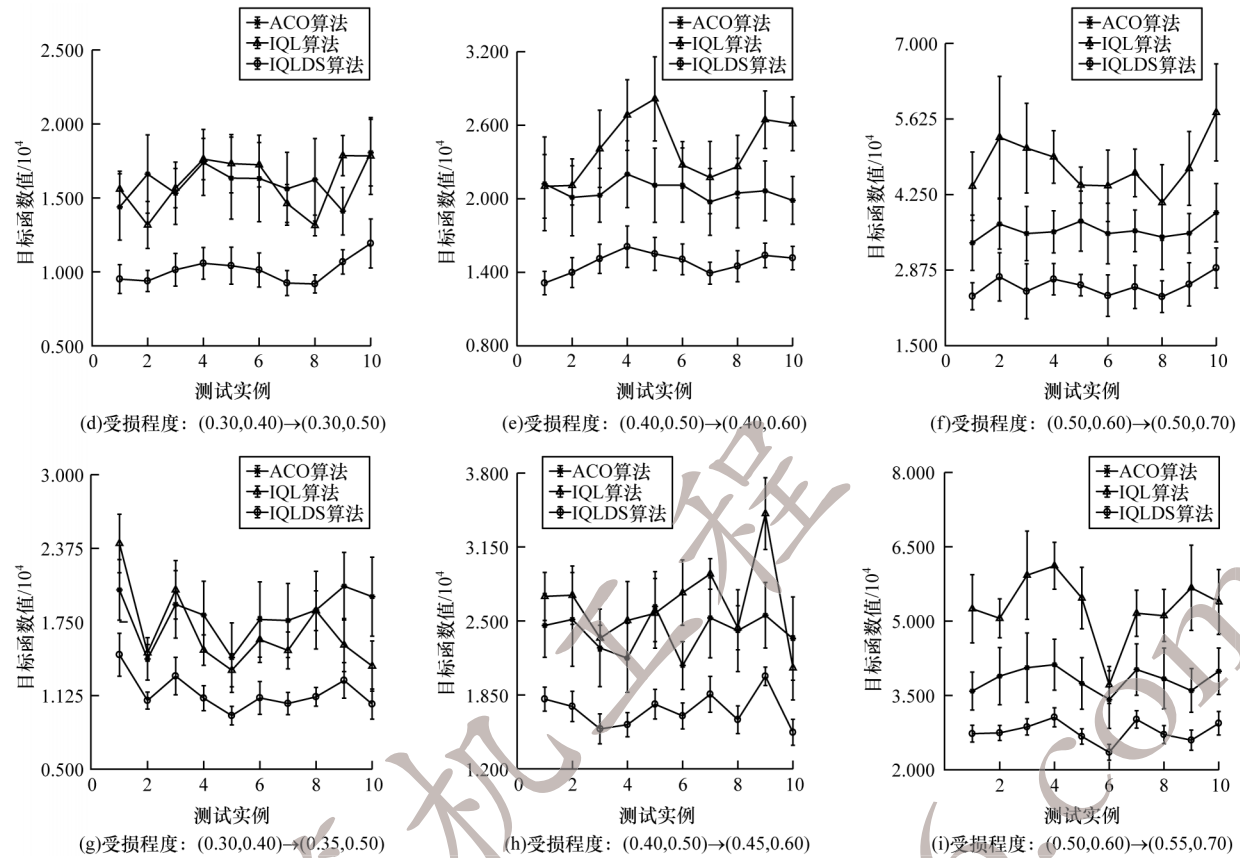


图 5 (续)

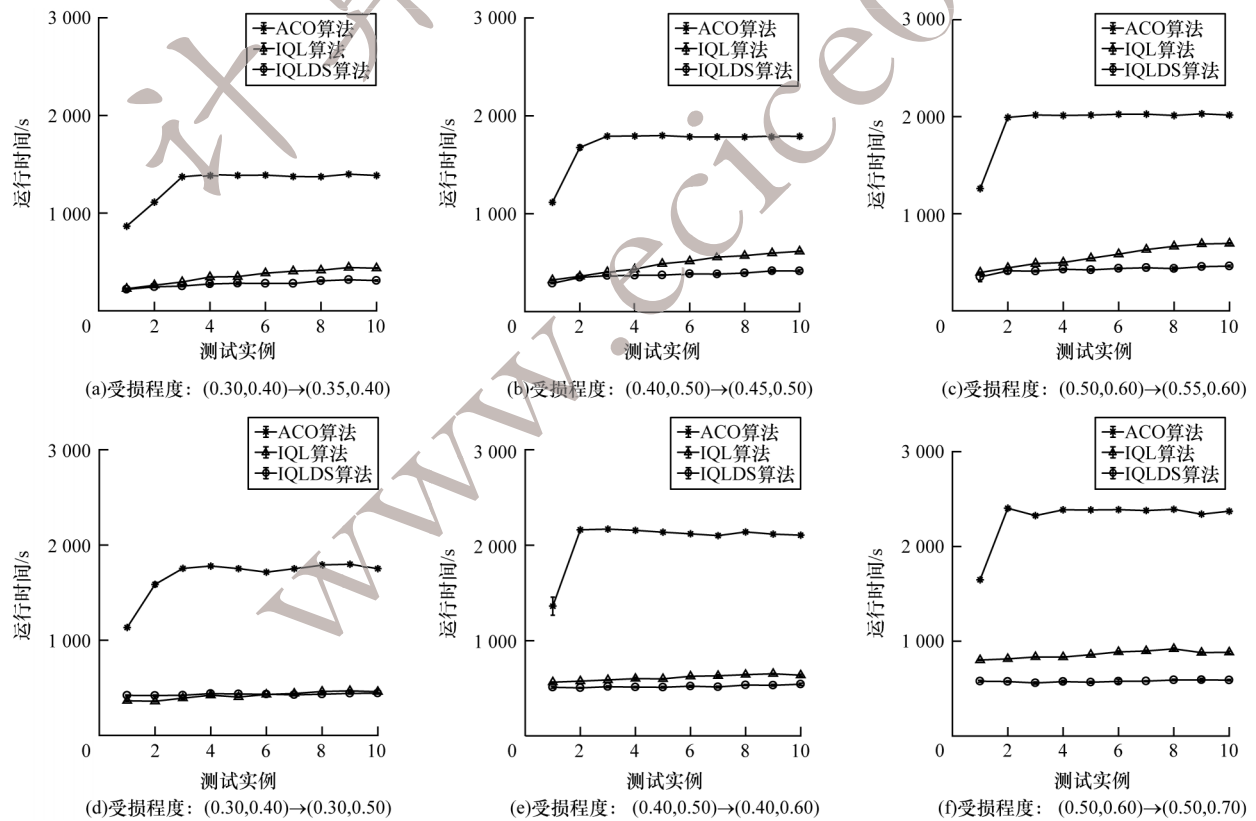


图 6 不同算法在规模为 $|V|=40, |E|=60$ 路网下的运行时间(均值和标准误差)对比

Fig.6 Comparison of running times (mean and standard error) of different algorithms under road networks with a scale of $|V|=40$ and $|E|=60$

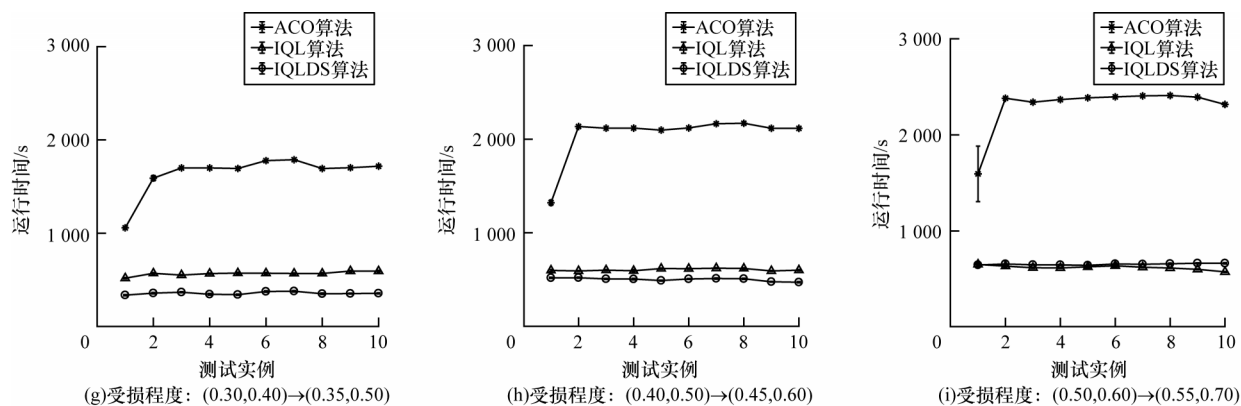
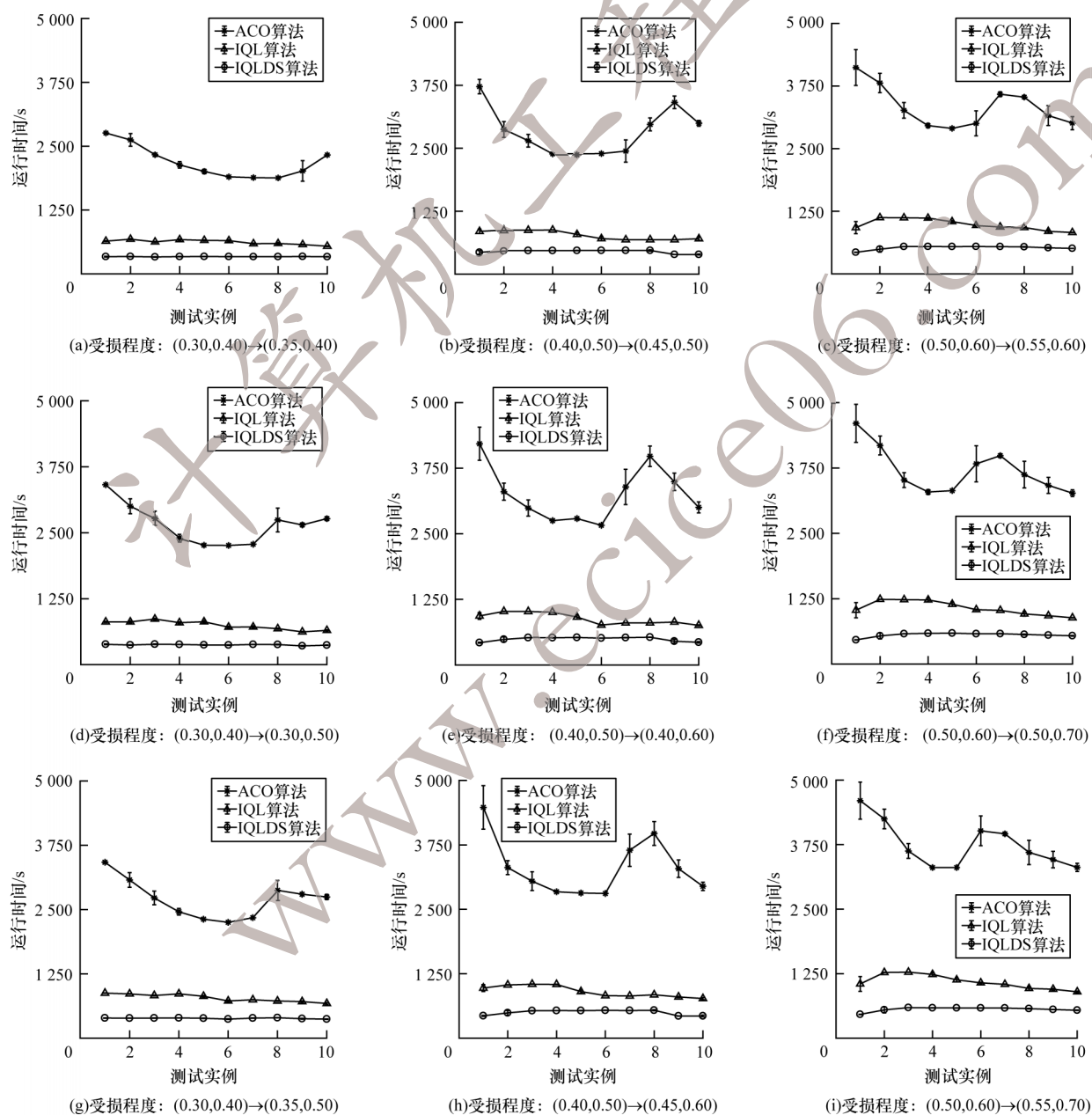
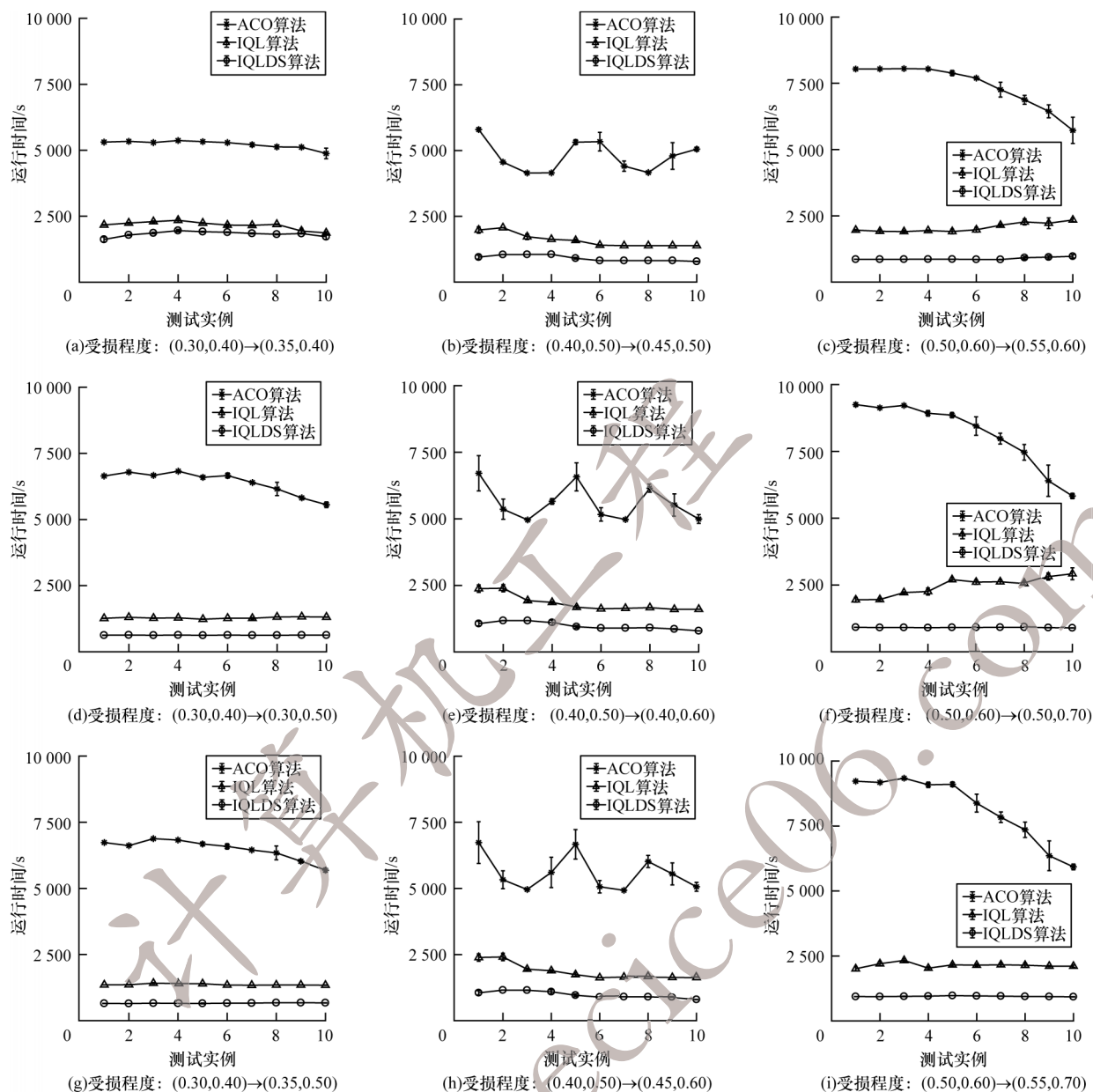


图6 (续)

图7 不同算法在 $|V|=60, |E|=90$ 路网下的运行时间(均值和标准误差)对比Fig.7 Comparison of running times (mean and standard error) of different algorithms under road networks with a scale of $|V|=60$ and $|E|=90$

图8 不同算法在规模为 $|V|=80$ 、 $|E|=120$ 路网下的运行时间(均值和标准误差)对比Fig.8 Comparison of running times (mean and standard error) of different algorithms under road network with a scale of $|V|=80$ and $|E|=120$

从图6~图8中可以看出,ACO算法的运行时间要远大于IQL算法和IQLDS算法。平均下来,ACO算法在3种路网规模上的耗时分别约为IQL算法和IQLDS算法的4倍、7倍和8倍。此外,随着路网规模及路网受损程度的增加,3种算法耗时越来越长,但ACO算法比IQL算法和IQLDS算法增加的速度明显变慢得多。这是因为,ACO算法在求解高维优化问题时搜索时间较长,收敛速度缓慢,而Q学习算法随机性小,运行时间波动较小。此外,在图7和图8中,在路网规模和受损程度均较大时,IQLDS算法要比IQL算法的运行速度更快。这是因为与IQL算法

相比,在IQLDS算法中嵌入的最优动作集更新策略直接以当前最佳目标函数值对应的动作来刺激agent的决策,加速agent的学习效率,即agent在一个周期内可以通过最优动作集的指引选择较少的动作,并达到所有需求点。

为进一步说明在灾情严重情况下的算法性能与修复效率,图9给出了在 $|V|=80$ 、 $|E|=120$ 规模、不同受损程度的路网下,3种规划方案的修复路段数与需求点连通率的变化趋势。由图9可以看出,IQLDS算法连通率的增长趋势比IQL算法和ACO算法更陡峭。尤其在 $|\vec{V}|/|V|=0.55$ 、 $|\vec{E}|/|E|=0.7$ 的高路网受损率测试实例中,如图9(c)所示,由于路网包含大量的

非需求点与连续受损路段,导致 IQL 算法和 ACO 算法的需求点可达率上升平缓,而 IQLDS 算法依旧可以通过修复较少的受损路段就可以明显提升路网的整体连通率。这说明 IQLDS 算法不仅能够很好地适应路网环境的变化,还可以在剩余的有限周期内给出更加合理的修复方案。

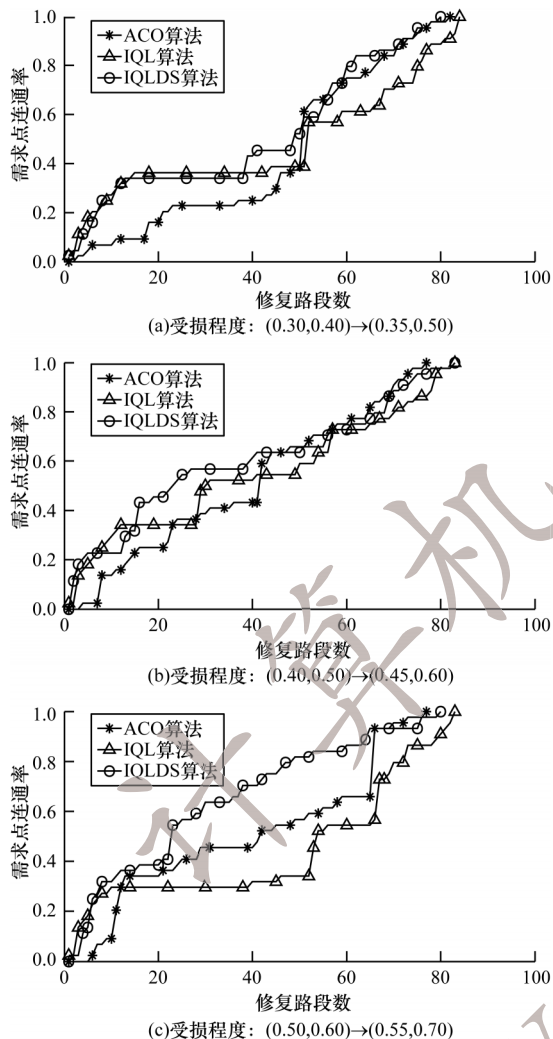


图9 3种算法的修复路段数和需求点连通率

Fig.9 Number of repair links and connectivity of demand points of three algorithms

5 结束语

在地震、洪水等重特大自然灾害发生后,往往伴随着余震、泥石流、山体滑坡等次生灾害和衍生灾害,给受灾路网修复带来极大的挑战。已有方法往往只能终止求解过程,根据变化后的路网重新搜索,导致决策效率较低。本文提出一种动态路网模型,依据动态路网和Markov决策过程构建抢修队的决策模型,并基于最优动作集更新策略提出一种抢修队动态调度算法。实验结果表明,在动态路网环境下,本文方法能够精确感知路网环境的变化,可在有限的训练周期内提供较优的抢修队调度策略,具有

很好的稳定性和可靠性。但本文在测试中发现,当面临高维受损路网时,Q学习算法会遭遇维度灾难问题而完全失效,因此下一步将基于深度Q学习求解高维受损路网抢修问题,并考虑使多个道路抢修队从不同的储备站出发修复受灾路网,通过提高多个抢修队之间的协同和合作效率,提高抢修队动态调度算法的性能。

参考文献

- [1] 汪洋,李楠,张磊.基于正六边形元胞自动机的行人疏散研究[J].计算机工程,2017,43(8):310-315.
WANG Y, LI N, ZHANG L. Research on pedestrian evacuation based on hexagonal cellular automata [J]. Computer Engineering, 2017, 43(8): 310-315. (in Chinese)
- [2] 谭峭,刘士豪,周婉,等.基于深度时空Q网络的机器人疏散人群算法[J].计算机工程,2021,47(6):305-311.
TAN M, LIU S H, ZHOU W, et al. Robot-assisted crowd evacuation algorithm based on deep spatio-temporal Q-network [J]. Computer Engineering, 2021, 47(6): 305-311. (in Chinese)
- [3] 李志,焦琴琴,周愉峰.震后应急物资供应点的多目标动态定位-分配模型[J].计算机工程,2017,43(6):281-288.
LI Z, JIAO Q Q, ZHOU Y F. Multi-objective dynamic location-allocation model for post-earthquake emergency facilities [J]. Computer Engineering, 2017, 43(6): 281-288. (in Chinese)
- [4] 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):8-16.
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): 8-16. (in Chinese)
- [5] YAN S, LIN C K, CHEN S Y. Optimal scheduling of logistical support for an emergency roadway repair work schedule [J]. Engineering Optimization, 2012, 44(9): 1035-1055.
- [6] YAN S, LIN C K, CHEN S Y. Logistical support scheduling under stochastic travel times given an emergency repair work schedule [J]. Computers & Industrial Engineering, 2014, 67: 20-35.
- [7] YAN S Y, 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.
- [8] 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. (in Chinese)
- [9] LIBERATORE F, ORTUO 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: 3-13.
- [10] AKSU D T, OZDAMAR 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.

- [11] SUZAN I, ALBERT L A. An integrated network design and scheduling problem for network recovery and emergency response[J]. *Operations Research Perspectives*, 2018, 5: 218-231.
- [12] LI S, MA Z, TEO K L. A new model for road network repair after natural disasters: integrating logistics support scheduling with repair crew scheduling and routing activities[J]. *Computers & Industrial Engineering*, 2020, 145(7): 506-512.
- [13] DUQUE P A M, DOLINSKAYA I S, SORENSEN K. Network repair crew scheduling and routing for emergency relief distribution problem [J]. *European Journal of Operational Research*, 2016, 248(1): 272-285.
- [14] 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.
- [15] KIM S, SHIN Y, LEE G M, et al. Network repair crew scheduling for short-term disasters [J]. *Applied Mathematical Modelling*, 2018, 64(11): 510-523.
- [16] SHIN Y, KIM S, MOON I. Integrated optimal scheduling of repair crew and relief vehicle after disaster [J]. *Computers & Operations Research*, 2019, 105: 237-247.
- [17] SANCI E, DASKIN M S. Integrating location and network restoration decisions in relief networks under uncertainty [J]. *European Journal of Operational Research*, 2019, 279(2): 335-350.
- [18] MORENO A, MUNARI P, ALEM D. A branch-and-Benders-cut algorithm for the crew scheduling and routing problem in road restoration [J]. *European Journal of Operational Research*, 2019, 275(1): 16-34.
- [19] MORENO A, ALEM D, GENDREAU M, et al. The heterogeneous multicrew scheduling and routing problem in road restoration [J]. *Transportation Research Part B: Methodological*, 2020, 141: 24-58.
- [20] 李兆隆, 金淳, 胡畔, 等. 基于弹复性的交通网络应急恢复阶段策略优化[J]. *系统工程理论与实践*, 2019, 39(11): 2828-2841.
- LI Z L, JIN C, HU P, et al. Resilience-based recovery strategy optimization in emergency recovery phase for transportation networks [J]. *Systems Engineering-Theory & Practice*, 2019, 39(11): 2828-2841. (in Chinese)
- [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. (in Chinese)
- [22] 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.
- ZHANG G F, TU B H, SU Z P, et al. An algorithm for repair crew scheduling on severely damaged road network [J]. *Control and Decision*, 2021, 36(7): 1663-1671. (in Chinese)
- [23] 宣洪伟, 李振东, 盛舟山, 等. 灾后运输网络中的最短路修复合作博弈[J]. *运筹学学报*, 2021, 25(3): 183-199.
- XUAN H W, LI Z D, SHENG Z S, et al. Cost allocation for multiple pairs of shortest paths recovery game on disrupted networks [J]. *Operations Research Transactions*, 2021, 25(3): 183-199. (in Chinese)
- [24] AJAM M, AKBARI V, SALMAN F. Routing multiple work teams to minimize latency in post-disaster road network restoration [J]. *European Journal of Operational Research*, 2022, 300(1): 237-254.