
让信号灯“学会适应”：深度强化学习提升城市路网突发扰动韧性

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/41566.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

让信号灯“学会适应”：深度强化学习提升城市路网突发扰动韧性。英国帝国理工学院智能交通研究团队提出一种面向路网韧性的深度强化学习信号控制方法。在基准路网仿真中，该方法相较传统信号控制最高降低79.7%的累计网络性能损失；在小型路网Sioux Falls重度扰动场景下，达到新均衡所需的行为调整迭代最多减少45.7%。

从缓解拥堵转向促进系统稳定

城市道路网络支撑着通勤、应急救援、货运运输和日常经济活动。一旦交通事故、洪涝、极端天气等事件使关键路段或交叉口的通行能力骤降，影响往往不会局限在事发地点：排队会向上游扩散，瓶颈位置可能发生转移，出行者也会根据拥堵状况调整路径和出发时间。此时，交通管理的目标不能只看某一时刻的拥堵是否下降，还要同时回答两个问题：扰动造成的总体损失有多大，路网需要经历多少轮行为调整才能形成新的稳定状态？

增加道路冗余或强化基础设施能够提升抗扰能力，但通常投入高、建设周期长。相比之下，交通信号控制具有成本较低、响应迅速和便于动态调整等优势。然而，固定配时、基于容量的控制和等饱和控制主要依据预设规则分配绿灯时间，难以主动学习扰动后交通状态-信号决策-出行者行为之间持续变化的反馈关系。发表于《Engineering》的该项研究据此提出一个核心问题：能否让信号灯以降低累计损失并加快恢复为目标，在扰动过程中自主学习控制策略？

让信号智能体在动态交通环境中持续学习

研究团队构建了基于深度Q网络（deep Q-network，DQN）的信号控制框架，将交通信号视为能够与环境交互的学习智能体。智能体以链路流量和链路容量作为状态信息，并通过调整红绿灯时间比例采取行动。为刻画扰动后交通系统的真实演化逻辑，研究进一步将DQN与日内/逐日动态交通分配模型耦合：日内模型描述车辆传播、排队形成与消散以及拥堵扩散；逐日模型描述出行者如何依据前一轮出行体验更新感知成本，并调整路径和出发时间。由此，控制策略的训练不再面对静态交通快照，而是在信号控制与出行行为相互作用的连续调整过程中进行。

该框架的关键设计是把相对面积指数（relative area index，RAI）作为韧性导向的奖励。RAI衡量扰动后网络性能相对于正常基线的累计偏离：数值越小，意味着累计性能损失越低和/或恢复越快。将RAI嵌入奖励函数后，智能体学习的目标不再局限于降低短期出行成本，而是同时压缩性能下降的幅度和持续时间。

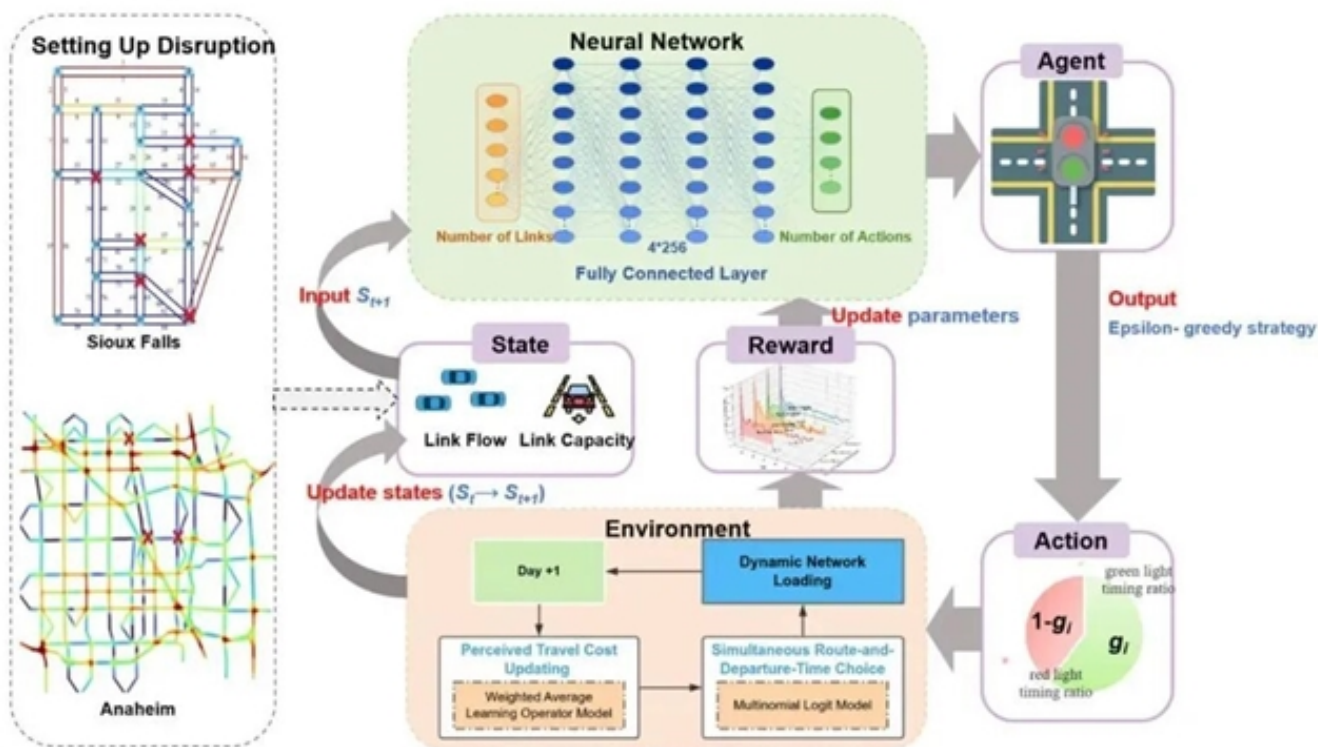


图1.基于DQN的韧性导向交通信号控制框架。智能体依据链路流量和容量调整红绿灯时间比例，并在日内/逐日动态交通分配环境中通过RAI奖励更新策略。图片来源：Qiannian Xiang et al.

两类基准路网、三级扰动与三种传统策略对比

研究在Sioux Falls和Anaheim两个常用基准道路网络上开展验证，将选定链路的通行能力分别降低25%、50%和75%，对应轻度、中度和重度扰动，并设置三叉、四叉、五叉和六叉路口场景。所提出方法与等饱和控制、基于容量的控制和固定配时控制进行比较。评估不仅关注最终出行成本，还系统考察RAI、达到新均衡所需的迭代次数以及关键链路上的成本分布。文中的逐日迭代表示一次出行者学习与路网状态更新循环，是理论模型中的行为调整周期，并不等同于现实世界的一个日历日。因此，下文以迭代次数而非现实天数描述收敛速度。

Sioux Falls：扰动越严重，韧性导向控制的价值越突出

在Sioux Falls网络的所有交叉口类型和扰动等级下，DQN方法均取得最低RAI。轻度扰动下，三叉、四叉和五叉路口场景中的RAI分别为0.0699、0.0536和0.0622；与固定配时相比，最高降幅达到79.7%。在五叉路口中度扰动场景中，该方法经过28次模型化行为调整迭代后达到均衡，RAI为0.3547，较等饱和控制、基于容量的控制和固定配时分别降低23.1%、69.4%和71.0%。

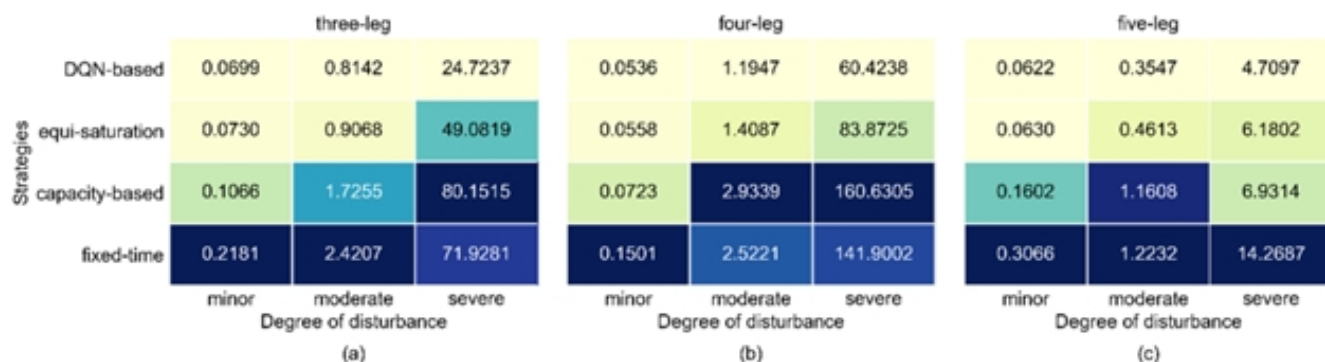


图2.Sioux Falls路网不同交叉口类型、扰动强度和控制策略下的RAI。RAI越小，表示累计性能损失越低、恢复越快。图片来源：Qiannian Xiang et al.。

重度扰动下，优势进一步显现。在三叉路口场景中，DQN方法经过38次逐日行为更新后达到新均衡；等饱和控制、基于容量的控制和固定配时分别需要49、70和59次迭代。DQN方法的RAI为24.7237，较三种传统策略分别降低49.6%、69.2%和65.6%，达到新均衡所需的迭代次数减少22.4%—45.7%。在五叉路口重度扰动场景中，该方法经过32次模型化调整迭代后达到均衡，RAI为4.7097，较三种传统策略分别降低23.8%、32.1%和67.0%，并比三种基准策略少用13—18次迭代。

链路层面结果揭示了网络韧性提升的来源：DQN控制能够抑制关键瓶颈上的出行成本累积，并减少局部拥堵向其他区域传播。例如，链路26的出行成本较基于容量的控制降低44.5%；链路29的出行成本较等饱和控制降低63.7%。

图3.Sioux Falls五叉路口重度扰动场景下的链路出行成本分布：（a）DQN控制；（b）等饱和控制；（c）基于容量的控制；（d）固定配时控制。图片来源：Qiannian Xiang et al.。

消融实验：RAI奖励兼顾少损失与快恢复

奖励函数决定了智能体究竟在学习什么。研究在五叉路口重度扰动场景中比较了三种奖励机制。以RAI为奖励时，模型取得4.7097的RAI，并在第32次迭代达到稳定；仅以总出行成本为目标时，RAI升至5.5535，第38次迭代才稳定；仅以收敛迭代次数为目标时，RAI为5.7261，第33次迭代稳定。这说明，单独追求较低的即时成本或较少的收敛迭代，都可能忽略另一维度；RAI奖励能够更有效地平衡累计损失与收敛速度。

图4.五叉路口重度扰动场景中的奖励消融实验：（a）不同策略的RAI与达到稳定所需的迭代次数；（b）不同奖励函数及固定配时控制下的总出行成本演化。横轴Day表示模型化逐日行为更新迭代，并非现实日历天数。图片来源：Qiannian Xiang et al.。

Anaheim：更大规模网络上的可扩展性验证

Anaheim网络的实验进一步检验了方法在更大规模路网上的表现。在六叉枢纽交叉口重度扰动场景中，DQN方法经过51次模型化行为调整迭代后达到新均衡，RAI为1.8660；相较等饱和控制、基于容量的控制和固定配时控制，RAI分别降低44.6%、59.4%和51.7%。这一结果表明，韧性导向的信号控制优势并非局限于单一路网或单一交叉口形式。

图5.Anaheim路网在轻度、中度和重度扰动下四种信号控制策略的收敛过程与RAI结果。横轴Day表示模型化逐日行为更新迭代，并非现实日历天数。图片来源：Qiannian Xiang et al.。

应用意义与下一步

这项研究将交通信号控制从常规的拥堵治理工具拓展为面向突发事件的路网韧性调控工具。通过把信号动作、动态交通传播、出行者迭代适应和韧性指标纳入同一框架，该方法可为灾前预案推演、数字孪生测试、事件响应和应急交通组织提供量化决策支持。相较新增道路或大规模基础设施改造，信号配时是一种更灵活的运营干预手段，有望帮助管理者在既有路网条件下减轻扰动影响并加快系统稳定。总体而言，该研究为人工智能赋能韧性交通管理提供了一条可度量、可迭代和可迁移的技术路径。（来源：EngineeringJournals微信公众号）

相关论文信息：<https://www.sciencedirect.com/science/article/pii/S2095809926003346>

作者：Wen-Long Shang 来源：《工程》

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发