基于超启发式算法的CVRP问题研究

陈一杭，张韬

（南京邮电大学计算机学院 ，江苏 南京 210046）

摘要：为了更有效地求解带容量车辆路径问题，避免陷入局部最优的情况，提出一种基于神经网络学习的超启发式算法。设计了算法的高层启发式策略，基于学习机制，使用强化学习中的深度Q神经网络算法构造该算法的选择策略，对底层算子的性能进行奖惩评价；利用元启发式算法作为底层算子，为优质解建立序列池，从而引导算法更有效地搜索解空间。对带容量车辆路径问题的标准算例进行计算，并和其他算法对比，统计分析了最优值、误差率和平均值，实验结果表明了所提算法在该问题求解上的有效性和稳定性，总体求解效果均优于对比算法。

关键词：超启发式算法；车辆路径优化；禁忌搜索

中图分类号：TP301.6 文献标志码：A

1. 引言(相关工作回顾)

车辆路径问题(Vehicle Routing Problem，VRP)是1956年Dantzig和Ramse[[[1]](#endnote-1)]首次提出的，带容量约束的多点配送车辆路径优化问题（Capacitated Vehicle Routing Problem，CVRP）是车辆路径问题的一种经典类型。CVRP问题已经被证明是NP难题，随着问题规模的不断增大，求解的复杂性也随之提高。因为精确算法无法满足要求，所以学术界常常使用启发式算法等近似算法求解。

Akhand等人[[[2]](#endnote-2)]针对CVRP问题对多种路线优化方法进行了比较，发现使用粒子群优化算法进行路线优化能够得到更好的结果。Ahmed等人[[[3]](#endnote-3)]提出基于双层局部搜索的粒子群优化算法以提高求解精度，该算法能在较短优化时间内得到较优求解。Stanley等人[[[4]](#endnote-4)]设计了用二进制编码的方法通过遗传算法求解 CVRP 问题，为求解 CVRP 问题提供了新的解决方案。传统的某一启发式算法针对某些案例求解性能良好，但针对其他案例时往往表现不佳，已不再具有优势。因而需要对传统的启发式算法进行局部改进，或将不同的元启发式算法作为底层算子组成超启发式算法（Hyper-heuristic Algorithms, HHA）。

超启发算法最早由Cowling[[[5]](#endnote-5)]提出，将超启发算法阐述为“寻找启发式算法的启发式算法”。Burke[[[6]](#endnote-6)]对超启发算法进行了更精确的描述：“超启发算法提供了一种高层启发式方法，通过管理或操纵一系列低层启发式算法，用于求解各种组合优化问题”。超启发式算法研究的主要任务在于选择策略、接收策略和底层算子三部分。根据超启发算法搜索空间的性质，可以将超启发算法分为选择式超启发算法（Heuristic Selection）和生成式超启发算法（Heuristic Generation）。选择式超启发算法根据反馈机制来源不同，可以将超启发算法分为不学习（No Learning）、离线学习（Offline Learning）和在线学习（Online Learning）三种类型，在线学习机制又可以分为元启发、基于强化学习和基于函数选择3种选择方法[[[7]](#endnote-7)]。Chen等人[[[8]](#endnote-8)]研究了基于蚁群的超启发式算法，采用蚁群算法来管理和操纵底层算子以获得新的启发式算法，每只蚂蚁均构造一个新的启发式算法。王万良等人[[[9]](#endnote-9)]将蛙跳算法作为选择策略用于启发式算法，具有更快的收敛性和稳定性，对跨问题领域求解效果也比较好。蔡延光等人[[[10]](#endnote-10)]提出求解CVRP问题的带分裂机制的帝国竞争算法，解决了算法“早熟”问题。

本文提出了一种将DQN（Deep Q Network）神经网络作为选择策略的超启发式算法，用于对CVRP问题进行求解，对公共标准算例进行计算，结果显示本文算法相对于来说，无论是结果的准确性还是运算时间的长短都表现出了较好的性能。

1. 问题描述
   1. 数学模型

CVRP是从同一车场出发的多辆车辆访问多个客户进行运送货物的问题。已知出发车场的位置以及客户的位置，客户的数量以及需求量，车辆的数量以及最大负荷量，为车辆组织适当的行车路线，有序地访问所有客户并且只访问一次，在满足所有设定的约束条件下，力争实现所设定的目标（车辆所走的路径最短）。将现实问题转化为数学模型，可以将CVRP描述为辆车从车场出发访问位客户进行运送货物并返回车场的问题。具体详情见表1。

表1 符号与变量声明

**Table 1 The declaration of symbols and variable**

|  |  |
| --- | --- |
| 符号与变量 | 含义 |
|  | 车辆数量 |
|  | 客户的数量 |
|  | 客户的需求量 |
|  | 客户到客户的运输距离 |
|  | 车辆 的最大负荷量 |
|  | 当车辆服务客户后立刻服务客户为1，其他为0 |
| Z | 总的运输长度 |

约束条件为：

（1）

（2）

（3）

（4）

目标函数为：

（5）

式（1）表示所有车辆都从一个车场出发，完成所担负的任务后全部返回车场，每一辆车走出的路径都要形成一个哈密尔顿回环；式（2）和式（3）表示车辆要对所有的客户进行服务；并且只能服务一次；式（4）表示每辆车服务客户时，自身的最大负载量不能低于被服务客户的需求量；式（5）是CVRP模型中的路径总长度计算方法，即所有车辆完成配送任务后的总路径长度，令总路径长度最小。

* 1. 算子介绍

本文提出了一种将DQN神经网络作为选择策略的超启发式算法，并根据CVRP问题设计了启发式优化算子、局部优化算子、变异算子三大类。所采用的具体算子如下：

LLH1：SA算子。利用退火算法快速优化路径。

LLH2：GA-SA算子。利用改进的遗传退火算法快速优化路径。

LLH3：TABU算子。利用改进的禁忌算法快速优化路径。

LLH4：or-opt算子。在一条路径上，客户数量为N(N>2)，随机选取两个相邻的客户点，将这两个客户点倒序插入该路径的其他位置，且只接受改进解。

LLH5：exchange算子。在一条路径上，随机选取两个客户，交换这两个客户的位置，且只接受改进解。

LLH6：2-opt算子。在一条路径上，随机选取两个客户，逆序排列这两个客户之间的所有客户，且只接受改进解。

LLH7：interchange算子。选择两条路径，随机从两条路径中各选取一个客户点进行位置交换，且只接受改进解。

LLH8：insert算子。选择一条路径，选取一个客户并将该客户插入该路径的任意其他位置，且只接受改进解。

LLH9：shift算子。选择两条路径，从第一条路径中选取一个客户点，将它插入到第二条路径任意位置，且只接受改进解。

LLH10：2-opt\*。和之前的2-opt算子一样，只不过此算子接受所有解。

LLH11：insert\*。和之前的insert算子一样，只不过此算子接受所有解。

LLH12：shift\*。和之前的shift算子一样，只不过此算子接受所有解。

其中LLH1～LLH3是启发式优化算子，LLH4～LLH9是局部优化算子，LLH10～LLH12是变异算子。

1. 基于DQN的选择策略

本文设计了基于强化学习的DQN来选择底层算子，是在线学习机制中将Q-Learning与神经网络相结合的选择方法[[[11]](#endnote-11)]。算法框架如图1所示。



图1 基于DQN算法的选择策略框架

Fig.1 Framework of hyper-heuristic algorithm in DQN algorithm

* 1. DQN参数设计

DQN算法主要基于Q-Learning算法的思想，利用当前状态，下一步行动，下一步状态，延用值的计算方法，建立损失函数，构建神经网络结构，起到预测下一步底层算子性能表现的作用。由于传统的单一神经网络会导致对动作评价出现较大偏差，本文引入估值网络与目标网络结合的双重神经网络结构，利用估值网络所对应的值来修正目标网络所对应的值，从而达到更精确预测下一步行动性能表现的目的。利用DQN算法作为底层算子的选择策略，根据当前解的状态（或者解的发展趋势），挑选下一步底层算子。

（1）设计

表示事物当前的状态。设计目的：尽可能体现对事物产生的影响以及能够为挑选之后的起到预测选择的作用。由于本文将该算法用于求解CVRP时，难以对值提前划分确切范围，设计当前代的适应度值为，上一代的适应度值为，值完全取决于各对解的适应度值的影响，所以设计的表达式为：

（6）

由于本文在算子选择中采用了三种不同类型的算子，鉴于三类算子的优化性能不同，为了更好地体现算子对当前状态的影响，所以添加值的辅助判别机制：利用常数值作为的基数，通过对不同算子所得出的值选取不同的常数加权，可以达到平衡算子被选择概率以及使结果更快收敛的效果。因此的表达式更新如下：

（7）

其中：

（8）

(2)设计

的值为LLH算子的编号。

在考虑值时，已经对算子种类进行区分，因此变异算子的立即回报值的判断方法，不再区别于其他种类的算子。因此定义如下：

（9）

* 1. 算法流程

步骤一 初始化参数。利用聚类划分的思想生成初始解集，设为最优解，为最佳适应度，设LLH算子的数目为，取值为整数，从中随机抽取解，初始化，，(随机挑选一个范围中的数)。

步骤二 经验池、序列池存储。操作上步Action后，产生的个体为,适应度值为，根据适应度值，计算，此时变为，计算两者偏差，利用式（7）计算值。设代表经验池，将上述值存入，则,代表经验池中数据组数。

步骤三 解的接受保留。如果,则说明此时解的适应度值更好，则保存解，令，；如果，则采用模拟退火判别，若随机值，则同样保留该解，同时更新状态。反之，则舍去该解，此时，。

步骤四 判断经验池容量。若，则进入步骤七，否则，进入步骤五。

步骤五 选择。判断此时值所属状态，如果为，则此时为局部优化算子和变异算子。设置值，若随机值，将值输入估值网络，输出值，取max()所对应的，若随机值，则根据此时值，令，此时为对应值的算子序号。

步骤六 保留全局最优解，判断算法是否结束。若，更新全局最优解，，，否则保留原有解。若，返回步骤二，否则算法结束。

步骤七 选择学习样本，并初始化神经网络。从EP中随机挑选组，作为学习样本，记为ESP。初始化估值网络和目标网络的阈值和权值。

步骤八 神经网络学习更新。估值网络中输入**（样本中第样本中的值，余同），计算后取（**）。目标值网络中输入**，计算后取。计算损失值，更新估值网络参数。

步骤九 判断代数。若代数，进入步骤五，返回主循环，反之，则程序结束。

算法流程图如图2所示。

**

开始

生成初始化解集

随机选择一组解

初始化

，，

计算种群，

适应度值

更新Pb，Fb

根据

选择

结束

从Ep中挑选，获取N组ESP

学习并更新DQN网络参数

将加入

是

是

是

是

否

是

否

否

否

否

图2 HH-DQNBM算法流程图

Fig.2 Flow chart of Particle Swarm optimization algorithm

1. 数值仿真与分析
   1. 实验设计与参数分析

实验环境 AMD Ryzen-5pro-4650U，16GB RAM，Python语言编写程序。经过反复测试，程序使用到的参数有值函数中的折扣率，初始值，迭代的最大代数为，经验池，学习挑选样本。

实验利用HH-DQNBM，对CVRP的标准算例进行求解。选取Set A的算例，并对每个算例计算20次。所有的算例可在网址：<http://vrp.atd-lab.inf.puc-rio.br/index.php/en/>下载。其中算例最优距离（BK），实验所得最优解距离解（Min），所得最优解距离解平均值（Avg），实验所得最优解与算例最优解误差（Dev%）（Dev=（Min-BK）/BK）。

* 1. 结果分析

表2为本文使用的HH-DQNBM算法及文献[]对Set A标准算例的计算结果。表中标有“\*”的数字表示当前算例最优解。

表2 CVRP SetA-标准算例测试



1. 结论

参考文献：

1. [] Dantzig G B , Ramser J H . The Truck Dispatching Problem[J]. Management Science, 1959, 6(1):80-91. [↑](#endnote-ref-1)
2. [] Akhand M A H, Peya Z J, Sultana T, *et al.* Solving capacitated vehicle routing problem using variant sweep and swarm intelligence [J]. Journal of Applied Science and Engineering,2017, 20 (4): 511-524. [↑](#endnote-ref-2)
3. [] Ahmed A , Ji S . Bilayer Local Search Enhanced Particle Swarm Optimization for the Capacitated Vehicle Routing Problem[J]. Algorithms, 2018, 11(3):31. [↑](#endnote-ref-3)
4. [] Lima S J D A , Araujo S A D . A New Binary Encoding Scheme in Genetic Algorithm for Solving the Capacitated Vehicle Routing Problem[C]// International conference on bioinspired optimization methods and their applications. 0. [↑](#endnote-ref-4)
5. [] Cowling P , Kendall G , Soubeiga E . A Hyperheuristic Approach to Scheduling a Sales Summit[C]// International Conference on the Practice and Theory of Automated Timetabling. Springer, Berlin, Heidelberg, 2000. [↑](#endnote-ref-5)
6. [] Pillay N , Banzhaf W . A Genetic Programming Approach to the Generation of Hyper-Heuristics for the Uncapacitated Examination Timetabling Problem[J]. 2007. [↑](#endnote-ref-6)
7. [] 谢毅, 侯彦娥, 陈小潘,等. 超启发算法研究进展综述[J]. 计算机工程与应用, 2017(14):1-8. [↑](#endnote-ref-7)
8. [] Chen P C , Kendall G , Berghe G V . An Ant Based Hyper-heuristic for the Travelling Tournament Problem[C]// 2007 IEEE Symposium on Computational Intelligence in Scheduling. IEEE, 2007. [↑](#endnote-ref-8)
9. [] 王万良, 徐昶, 赵燕伟,等. 基于超启发式算法的选址-路径问题研究[J]. 浙江工业大学学报, 2019. [↑](#endnote-ref-9)
10. [] [1]蔡延光,王世豪,戚远航,王福杰,林卓胜.帝国竞争算法求解CVRP[J/OL].计算机应用研究:1-7[2021-03-08].https://doi.org/10.19734/j.issn.1001-3695.2020.01.0006. [↑](#endnote-ref-10)
11. [] 张景玲, 冯勤炳, 赵燕伟,等. 基于强化学习的超启发算法求解有容量车辆路径问题[J]. 计算机集成制造系统, 2020, 026(004):1118-1129. [↑](#endnote-ref-11)