

# 传统优化方法与机器学习融合的研究进展

朱红红<sup>1</sup>, 刘雅<sup>2</sup>, 杨希望<sup>3</sup>

西安交通大学 管理学院, 陕西 西安 710049

**摘要:** 由于组合优化问题规模的不断扩大, 传统组合优化算法面临着巨大的计算压力, 很难拓展到大规模问题。近年来机器学习和传统组合优化方法的融合在处理组合优化问题上表现出良好的性能, 具有较强的泛化能力和快速的求解速度, 为解决组合优化问题提供了新的方向和思路。因此, 通过查阅近些年传统优化方法与机器学习融合的相关研究, 首先对二者结合的理论研究进行了总结, 然后以传统优化方法为分类目标, 对机器学习与组合优化方法融合的方式、逻辑、切入点等进行归纳和综述, 并对其优缺点和异同点进行了对比, 最后指出了该方向现存的一些问题和难点, 旨在为未来该方向的进一步探讨与研究提供启发和参考。

**关键词:** 组合优化, 优化方法, 机器学习, 强化学习, 神经网络

## 1、概述

在组合优化问题中, 由于可行解数量会随问题规模呈指数型增长, 无法在多项式时间内穷举得到问题的最优解<sup>[1]</sup>。为此学者们对组合优化问题的求解算法进行了大量的研究, 经典方法包括分支定界法、列生成法、分支定价法、分支切割法等, 但其中也存在一些问题, 如利用上述方法求解组合优化问题的时间复杂度是指数量级的, 本质上讲这些问题都是 NP 完全 (NP-Complete) 或 NP 难 (NP-Hard)。随着实际应用中问题规模的不断扩大, 实时性求解的要求越来越高, 传统组合优化算法面临着很大的计算压力, 很难拓展到大规模问题, 同时也很难实现问题的在线求解<sup>[1]</sup>。因此, 学者们开始尝试使用数据驱动和机器学习的方法来处理组合优化问题。

在组合优化领域, 机器学习因其自适应、数据驱动等特点而受到大量学者关注。目前已有部分研究利用机器学习或深度强化学习方法来解决一些经典的组合优化问题, 取得了较好的效果, 尤其在计算时间方面有着明显优势。还有少部分研究关注于将机器学习和传统优化方法的融合, 并提出了一些两者融合的方法论, 这类方法在对一些经典的组合优化问题求解上表现出良好的性能, 因此获得了大量研究人员的关注。

机器学习与传统优化方法的融合是近年来较新的研究方向, 目前对该类方法进行系统的研究和综述的文献较少。鉴于此, 通过查阅近些年机器学习和传统优化方法融合的相关研究, 首先对二者结合的理论研究进行了总结, 然后以分支定界法、列生成法、分支定价法、分支切割法这四大经典优化方法为分类目标, 分别对机器学习与之融合的方式、逻辑、切入点等进行归纳和综述, 并对其优缺点和异同点进行对比, 最后指出了该方向现存的一些问题和难点, 为未来该方向的进一步探讨与研究提供启发和参考。

## 2、相关基础知识

本节给出了优化方法和机器学习的基本介绍。

### 2.1 优化方法

组合优化是从离散空间集合的所有组合中找出一种最优方案以满足特定的目标。经典的求解方法主要有以下几种。

**分支定界法:** 是一种求解整数规划和混合整数规划问题的常用算法, 基本思想是在约束最优化问题的整个可行解空间内搜索且不断分割为小型子集 (称为分支), 并对每个子集计算其下界或上界 (称为定界), 对凡是超出界限的子集进行剪支, 达到缩小搜索范围的目的。

**列生成法:** 是一种用于求解大规模线性规划问题的高效算法, 其将原始问题分为主问题和子问题。主问题不需要显示全部变量, 只使用一组能产生初始可行解的变量构成模型; 子

问题的作用是生成新列从而改进当前主问题。每次迭代时，首先求解主问题并将其对偶解输入给子问题，然后求解子问题并将其解加入主问题，即为主问题生成了一个新变量，而后开始新的迭代。当无法生成更多列时停止迭代。

分支定价法：是将分支定界法和列生成法动态整合，其中列生成算法用于获得节点的下界（以最小化问题为例）。分支定价算法其他步骤与分支定界算法相同。

分支切割法：是将分支定界法和割平面法动态整合，在分支定界法的基础上，使用割平面法来进一步缩小搜索空间。割平面法的思路是：找到能割掉原可行域的部分区域而没有割掉最优解的约束条件，将其增加至模型中。

## 2.2 机器学习

机器学习是指从有限的观测数据中学习出具有一般性的规律，并利用这些规律对未知数据进行预测的方法<sup>[2]</sup>。机器学习主要模型有逻辑斯蒂回归、感知器、支持向量机、随机森林、人工神经网络等。基于学习方式，机器学习可以划分为以下三类。

1、监督学习：是从标记的训练数据中学习出一个映射函数的机器学习任务。训练数据由若干输入对象及其输出值组成。监督学习算法是分析这些训练数据，并习得一个输入与输出之间的映射函数，其可以用于推断或预测其他实例的目标输出。寻找这样的映射函数称为学习，并通过一系列函数的优化问题来解决。

2、无监督学习：是指从无标签的数据中学习出一些有用的模式。无监督学习算法一般直接从原始数据中学习，不借助与任何人工给出的标签或者反馈等指导信息<sup>[2]</sup>。

3、强化学习：用于描述和解决智能体（Agent）在与环境的交互过程中通过学习策略以达成奖励最大化或实现特定目标的问题。强化学习问题不需要给出“正确”策略作为监督信息，只需要给出策略的奖励或回报，并通过调整策略来取得最大化的期望回报<sup>[2]</sup>。强化学习有两个可以进行交互的对象：智能体和环境，其基本要素包括：状态、动作、策略、状态转移函数、奖励。

## 3、机器学习与传统优化方法融合：理论研究

近年来学者们开始将机器学习应用于组合优化领域，在机器学习与传统优化算法的融合方面进行了相关研究，部分学者对其进行了回顾与综述，总结了二者结合的角度、方式方法和基本原理，并提供了一些方法论介绍。

Bengio 等<sup>[3]</sup>调查了机器学习和运筹学两个专业领域关于利用机器学习来解决组合优化问题的研究，旨在进一步推动机器学习和组合优化结合，并详细介绍了相关方法论。首先，从组合优化领域出发，机器学习方法有两种方式来改进组合优化算法：一是利用快速的机器学习近似取代繁重的计算；二是利用机器学习找出最佳的启发式。基于这两种改进方式，在利用机器学习求解组合优化问题时有两种学习策略的方法：（1）模仿背景：对于给定问题首先利用已有方法获得大量训练数据，即其输入和相应的解，使用机器学习方法习得专家策略，再将此策略应用于其他数据或问题。这种方法的局限性较大，如训练数据的目标输出不是最优解那么利用机器学习获得的解也会受影响，而且在训练过程中，由于错误的累积，机器学习模型性能可能不稳定。（2）经验背景：用策略对应的奖励作为信号，目标是最大化总奖励值，通过不断试错的方法最终获得最大奖励值对应的策略。强化学习的训练时间一般都较长，但与学习既定专家策略相比，更灵活，有可能更优于任何专家策略。但如果训练不充分，学习过程可能会陷入较差的解决方案。

鉴于此，学者们发现可以将监督学习和强化学习结合使用，即先用监督学习习得若干专家策略，然后再用强化学习对这些策略进行选择，这样可以避免部分二者的缺陷。Bengio 等认为习得的策略可以从三方面与传统优化算法整合（表格 1）。

表格 1

类别	方法	优点	缺点
直接求解	直接使用机器学习习得的策略获得问题的解	整体求解耗时较少	1、不能保证最优化，甚至可行性； 2、只适用于特定问题，泛化能力差
学习参数等信息	利用机器学习习得传统优化算法的部分信息，如参数	改进算法性能，加速求解	只适用于特定问题，泛化能力差
嵌入传统优化算法框架	机器学习作为传统优化算法的一部分，迭代使用	可应用于不同的优化方法中	算法每次迭代都会调用机器学习模型，可能会影响计算时间

与 Bengio 等将机器学习用于组合优化领域的观点不同，Gambella 等<sup>[4]</sup>也考虑了机器学习与组合优化的结合，但 Gambella 等是将机器学习模型的底层设计视为数学优化问题，分别对机器学习的主要模型进行了数学优化方面的总结和综述。Gambella 等认为机器学习和运筹优化之间的关系可以从三个维度划分：将机器学习应用于管理科学问题；利用机器学习求解优化问题；将机器学习底层设计表示为优化问题。Gambella 等从回归模型、分类、聚类、深度学习、线性降维、对抗学习入手，分别讨论了各类机器学习模型底层设计的数学优化公式和意义。

Mazyavkina 等<sup>[5]</sup>研究了将强化学习用于求解组合优化问题上的进展。Mazyavkina 等首先介绍了强化学习的组成部分、主要网络模型及使用方法，然后通过五大经典组合优化问题（旅行商问题、最大割问题、装箱问题、最小顶点覆盖问题、最大独立集问题）展示了如何利用强化学习来求解这些问题。Mazyavkina 等将强化学习求解组合优化问题的方式分为两类：一是强化学习直接求解组合优化问题；二是强化学习与传统优化方法融合来改进原算法。在每类经典组合优化问题中，Mazyavkina 等总结了使用强化学习与传统优化算法融合的方法来求解该类问题的逻辑要点。

## 4、机器学习与传统优化方法融合：方法研究

将机器学习直接用于求解组合优化问题虽然速度较快但其解无法保证精确性甚至可行性。学者们开始关注将机器学习与传统优化算法相融合的方法。一个突出的结合点是将机器学习嵌入传统优化算法的框架中，用传统优化算法来保证解的可行性，而机器学习作为优化算法中的一部分可以加快迭代或克服一些传统优化算法中的难点。

### 4.1 分支定界法

分支定界法是一种搜索与迭代的方法，其在每一步都要选择搜索方向即选择子节点进行分支，因此搜索方向及分支策略对于缩减搜索空间和减少整体求解时间至关重要。Applegate 等<sup>[6]</sup>提出了强分支策略，就搜索树中的节点数而言，强分支是目前较为有效的分支策略。然而，这种效率的实现是以巨大的计算时间为代价的。

由于完全实现强分支的计算要求太高，Yu 等<sup>[7]</sup>提出了广义强分支策略，使用极端梯度增强技术来训练模型，以预测一组候选变量的排名。广义强分支策略是将强分支策略中选择单个变量进行分支扩展到选择一组变量进行分支，并使用强分支分数来确定要分支的候选变量集。选择变量集进行分支，不仅可以在节点上提供更紧的界限，还有可能缩小搜索树范围。为了在搜索树的当前节点上获得候选变量集的强分支分数，Yu 等训练了一个机器学习模型来模仿强分支策略以做出强分支分数预测。然后，Yu 等提出了广义强分支策略模型来指导变量集合分支。Yu 等分别在集合覆盖问题、装箱问题上进行了实验。研究表明，在变量集上的分支优于在单个变量上的分支，习得的模型可以对候选变量进行有效选择，并且该策略在搜索节点数和搜索时间方面都优于求解器 CPLEX 的默认分支策略。

Nair 等<sup>[8]</sup>则是在分支定界法中引入深度神经网络进行分支选取来提升算法性能。其主要思想是通过神经网络模仿学习强分支策略和专家策略，从而在未知实例中做出分支预测。该方法可以保留高质量的解并且大幅减少计算时长。Nair 等首先在训练实例集上使用专

家策略获得训练数据，并对节点分支的信息进行特征提取。然后在训练过程中，使用深度神经网络技术对专家策略进行学习。实验研究表明，该方法在测试数据集上有明显的加速效果，在相同的运行时间内可以获得较高质量的解。

Alvarez 等<sup>[9]</sup>提出了在使用分支定界法求解混合整数线性规划问题时变量分支的一种新颖的通用方法，其使用监督学习模型习得的函数来提供对已有强分支策略的快速近似，克服了强分支决策导致的巨大计算开销。为了有效地模拟强分支决策，Alvarez 等提出了一种两阶段的方法来习得分支策略，此分支策略可以在分支定界算法中用作强分支的近似。第一阶段需要解决一组具有强分支的训练问题，以生成一组分支决策，每个分支决策都记录在训练数据集中，然后机器学习算法将使用该数据集学习模拟强分支决策的函数。第二阶段，将习得的分支策略引入分支定界算法中，并在一组测试数据集上评估这种分支方法的效率。Alvarez 等在随机生成的算例和基准算例上分别进行了实验，比较了该分支策略与其他分支策略在不同问题上的求解性能。研究表明，该方法习得的分支策略求解速度较快，与其他方法相比有一定优势。

Wehenkel 等<sup>[10]</sup>提出的方法是在优化过程中使用在线学习算法来学习一个近似于强分支的函数，达到减少计算成本的目的。强分支近似是用一种简单的机器学习技术，即线性回归来完成的，然后使用习得的线性回归函数来预测给定变量的强分支分数。当需要进行分支时，从当前节点的变量集中选择一组候选变量，并计算每个变量的分支分数，然后选择得分最大的变量进行分支。在这个过程中，Wehenkel 等将学习分支和可靠性分支的思想结合起来考虑，如果候选变量的强分支分数近似值不可靠，则使用真实的强分支分数，反之，则使用习得的近似函数来生成分数。由于学习是在线进行的，用于训练函数的数据是在求解过程中生成的，因此不需要任何初始化阶段，也不会浪费任何计算时间，而且在线学习采用了梯度下降算法，允许近似函数在优化过程中更新。Wehenkel 等在随机算例和基准算例上分别进行了实验，比较了该分支策略与其他分支策略在不同问题上的求解性能，实验结果表明，该方法在时间和节点选择方面都优于其他分支策略。

在分支定界法的整个过程中如果使用单一的搜索策略则忽略了搜索树的动态结构。基于此，He 等<sup>[11]</sup>提出了利用机器学习来习得一种适应不同问题类型和不同解决阶段的搜索和分支策略，充分考虑了搜索树的动态结构和特征。这种分支策略是通过模仿学习来习得的，使用了 DAgger 算法，其思想是在每一次迭代中利用当前的策略收集数据，然后利用所有的数据集训练下一次的策略。He 等首先利用 DAgger 迭代模仿学习算法获得训练数据集并习得节点选择和分支的策略，即在对问题的不断迭代求解过程中，根据迄今收集到的所有数据，利用机器学习的回归模型和分类模型不断学习新策略，通过运行新习得的策略继续收集其他数据，直到获得最优策略。然后将其嵌入到分支定界算法框架中，用于求解其它问题实例。

## 4.2 列生成法

对于大多数列生成求解的问题，子问题很可能是 NP-Hard 问题，为解决它而开发的方法非常有限，一般消耗计算时间较大。

为了加速列生成算法，针对定价问题的解决，Morabit 等<sup>[11]</sup>提出一种基于机器学习的启发式定价算法。该方法利用训练数据集来学习分类器，此分类器将选择出有较大机会成为最优解中的部分边，以减小定价问题的网络规模从而减少定价问题的计算时间，而不会显著增加迭代次数。在机器学习中，这是一个二元分类问题，如果生成的列被添加至主问题，那么此列所涉及的每条边标签编码为 1，否则为 0。为了收集足够多的训练数据，Morabit 等在几个问题实例上，使用专家启发式算法经过迭代获得近似最优解，然后在每条边上收集特征和标签的元组对，构成训练数据集。在训练阶段，该算法试图最小化已知目标值与模型预测值之间的差异。训练阶段完成后，学习的模型将嵌入到列生成算法中。该算法首先提取边的特征并获得模型预测，然后使用预测值构建简化网络。在获得主问题初始解以及对偶值后开始迭代，在每次迭代中，对当前活动网络进行定价问题求解，并生成一组新的列，直到无法生成更多列时停止。Morabit 等将该方法应用于公交车辆和乘务员调度问题和带时间窗的车辆路径问题，结果表明计算时间分别减少了 27% 和 40%。

同样在机器学习与列生成算法融合的方法方面,与上述提出的侧重于定价问题的加速方法不同,Morabit 等<sup>[13]</sup>中的策略是针对在求解主问题时花费大量计算时间而设计的。这种方法基于机器学习训练出一个列选择器,该选择器从大量生成的列中选择要在每次迭代时添加到主问题的列,目标是添加最有可能改进当前解决方案的列。这种列的选择被视为一个二元分类问题,将每个生成的列分类为选择或不选(1或0)。通过解决多个问题实例来收集数据,从中提取列的特征并分配标签。所用模型基于图形神经网络,该问题使用二部图表示,其中使用了具有非线性激活函数的两层前馈神经网络,输出为列分类的概率,使用了具有归一化函数的三层前馈神经网络。作者对公交车辆和乘务员调度问题和带时间窗的车辆路径问题进行了实验,计算结果表明所需的总计算时间平均减少了20%至30%。

上述两个研究都使用的是监督学习与列生成算法融合,而Chi等<sup>[14]</sup>提出了强化学习与列生成算法的融合(RLCG),旨在通过这种融合作出更好的列选择决策,从而加快算法的收敛速度。考虑到迭代中所选择的列会影响算法的后续迭代,Chi等将列生成视为一个顺序决策问题。在RLCG算法中,强化学习智能体与列生成循环融合,使用图神经网络学习动作价值函数(Q函数),基于从当前主问题和子问题中提取的信息主动选择要添加到下一次迭代中的列。首先,Chi等采用机器学习中的课程学习方法,按照增加难度的顺序提供实例,并将列生成表示为马尔可夫决策过程。然后将图神经网络用作Q函数近似器,使用具有经验回放的深度Q函数学习算法对强化学习代理进行训练。为评估RLCG算法的性能,Chi等分别在切材问题、带有时间窗的车辆路径问题上进行实验,并将结果与Morabit等<sup>[13]</sup>方法的求解结果进行比较,结果表明,RLCG在迭代次数和时间方面都显示出优越的性能和更好的收敛性。

Firat等<sup>[15]</sup>将列生成方法用于解决构建最优分类决策树的问题。决策树是一种基本的分类与回归方法,是解决分类任务的主要方法之一。构造最优分类决策树是根据一组标记的目标类对实例中的数据进行分类并选择最优决策树的问题。决策树的元素包含节点和边,节点会根据实例的特征作出分类判断,边表示进一步决策分类的方向,叶子节点表示最终的分类结果。构建决策树的本质是找出决策树中根节点到叶子节点的决策路径。Firat等提出了一种整数线性规划模型来构建分类决策树,该模型定义了与决策路径相关的决策变量,其中决策路径是从根节点到某个叶子节点的一系列决策拆分。该模型使用列生成方法进行求解,可以在较短计算时间内获得高质量的解。

#### 4.3 分支定价法

在分支定价算法中,定价问题的计算耗时占比较大。而定价问题本质上是重复的,连续定价问题的解是有内在联系的,即当前新增的列会影响主问题的求解。机器学习为利用从已解决的定价问题中获得的信息来加速定价问题的当前迭代的研究提供了帮助。

Quesnel等<sup>[16]</sup>提出了一种结合机器学习的分支定价算法用以解决人员排班问题。人员排班问题是合理安排所有配对以形成一个周期性的工作排班表,其中配对为车次序列、空车、中转、成员及其休息时间段的组合。该问题中,由于子问题需要获得所有成员的所有可行配对,这会大幅增加整体算法的计算时间,因此,Quesnel等提出了基于机器学习的人员排班问题的部分定价方案,其中每个成员的子问题只包含可能在最优解或近最优解中的配对,目的是减少子问题的搜索空间。在该算法中,机器学习模型用来对每个成员的配对进行降序排序,然后选择排名靠前的配对进入算法迭代。Quesnel等搭建的机器学习模型是一个深度神经网络模型,从监督学习训练框架中训练得出。该神经网络是一个前馈全连接深度神经网络,所有隐藏层的神经元数量都小于或等于前一层的神经元数量,除输入输出层的神经元外,所有神经元都是非线性激活单元,输出层由单个归一化单元组成。Quesnel等在大型实例上测试了该算法的性能,该算法较传统的分支定价算法速度快。

Furian等<sup>[17]</sup>提出了一种基于机器学习的分支定价法用以解决具有时间窗的车辆路径问题。在这种方法中,Furian等引入了两种用途的机器学习模型,一是预测决策变量(两个节点间的边是否在最优解中)的值,这些预测被整合到分支定价算法的节点选择策略中;二是预测节点分支分数,这些预测被整合到分支定价算法的变量选择策略中。Furian等设计了两种随机森林模型并以二者结合的方式来预测决策变量的值,即做出节点选择。在变量选

择方面, Furian 等提出了一种基于分数预测进行分支的方法, 该法基于在训练阶段习得的回归森林模型, 为每个候选变量预测一个分数, 然后选择得分最高的变量。计算实验表明, 这种方法在搜索过程中处理的节点数量以及运行时间方面都优于标准算法, 而且对于实例数据的大小变化是鲁棒的。

Vaclavik 等<sup>[18]</sup> 提出了一种方法来加速分支定价算法的计算时间, 并展示了如何利用从前期求解定价问题中获得的知识来减少后续求解定价问题的解空间。这种方法基于在线机器学习算法, 使用一个简单快速的回归模型, 在保持分支定价算法精确性的同时, 预测了定价问题当前迭代的紧上界, 然后, 上界被用来更有效地删减定价问题的搜索空间, 从而减少选择检查的状态。Vaclavik 等采用的核心思想是基于连续定价问题之间的联系。在该方法中, 每一次列生成迭代都训练回归模型, 然后回归模型将从主问题中获得的对偶值转换为随后求解的定价问题的目标函数的上界的预测, 上界被用作定价问题的约束, 进而加速求解定价问题。求解定价问题之后, 算法会创建一个新的数据点, 并更新训练数据集, 然后重新训练回归模型。如果找到新列, 则列生成算法将继续后续迭代。这种改进的分支定价算法应用于护士排班问题、时分复用问题, 取得了高质量的解, 并且显著地减少了整体求解时间。这种方法可以用于使用分支定价算法求解的任意组合优化问题。对于更复杂的问题来说, 在合理的时间框架内很难计算出一个严格的上界, 而这种方法利用机器学习可以快速的获得一个较紧的上界, 加速整体进程。

以具有二维装载约束的车辆路径问题为背景, Zhang 等<sup>[19]</sup> 开发了一种分支定价算法, 并嵌入机器学习模型来求解, 提出了一种新的基于监督学习模型的列生成机制。由于车辆路径问题和二维装箱问题的集成带来的复杂性, 该领域的精确算法研究非常有限。该方法可以有效的减轻二维装箱问题带来的计算负担, 以便在合理的时间预算内实现精确性。分支定价算法使用分支定界树, 其中每个节点处考虑的子问题的线性松弛通过列生成来解决, 需要解决定价问题以找到成本为负的列。定价问题通常作为一个具有资源约束的基本最短路径问题来处理, 属于 NP-Hard 问题, 它通常通过标记算法来求解。然而, 当涉及到具有二维装载约束的车辆路径问题时, 定价问题在结构上发生了变化。因为必须通过解决二维装载问题来检查所有定价列, 以确保装载的可行性。已有研究大部分使用的是精确可行性检查, 其对所有定价列全部执行可行性检查, 消耗了大量时间。为了加快检查速度, Zhang 等提出的策略是在进行可行性检查之前, 利用机器学习模型对定价列进行预测分类, 然后根据分类结果选择性地检查, 目的是减少可行性检查程序的使用频率。此预测任务可以表示为一个二元分类任务, 应用一个二元分类器作出可行性预测, 目的是阻止尽可能多的不可行列、并允许尽可能多的可行列进入主问题。为了习得可行性预测器, Zhang 等使用前馈神经网络作为机器学习模型, 选择小批量梯度下降算法作为训练算法来解决此问题。实验结果表明这种新算法的性能优于未结合机器学习模型的算法, 解决方案得到了显著改进。

上述研究都是将机器学习与分支定价算法融合, 均使用了监督学习。在此基础上, 专家学者可以考虑从两个方向进行深入研究, 一种是为更复杂的情况开发机器学习模型; 另一个是探索将预测器插入其他算法的方法。

#### 4.4 分支切割法

分支切割法是将分支定界法和割平面法相结合, 每个分支都包含作为附加约束的割平面。在迭代地向模型中添加割的过程中, 所添加割可以通过消除部分可行区域来收紧上下界, 同时保留最优解。该方法的重点在于如何设置或选择有效割, 从而加速收敛到最优解。已有文献大部分研究的是设计若干种有效割。Gomory 等<sup>[20]</sup> 提出了一种割平面方法(称为 Gomory 割), 目前已广泛应用于现代求解器中, 最常见的是作为分支切割方法的子程序使用。Gomory 等的研究表明, 在每一步添加适当 Gomory 割的割平面方法能在有限次数的迭代中终止。然而, 在每次迭代中, 存在若干个可供选择的割, 全部使用会大大增加计算时间。因此, 解决方案的效率和质量在很大程度上取决于生成的割平面的顺序。

基于 Gomory 割, Tang 等<sup>[21]</sup> 开发了一种基于强化学习的智能自适应选择切割平面的方法, 并将其与分支切割方法结合使用, 以有效地解决整数规划问题。强化学习算法通过反复尝试求解随机生成的训练实例池来学习, 而无需访问任何已求解的实例。该方法将割平面选

择问题作为强化学习问题来表述，引入了一个有效的马尔可夫决策过程来解决线性规划选择割平面的顺序问题，分别指定了状态空间、动作空间、奖励和状态转移函数。为了鼓励积极添加割平面，每次迭代中的即时奖励设置为连续迭代过程中线性规划解之间的差距，尽可能快地接近整数最优解。由于 Gomory 割的顺序不可知，Tang 等使用注意力网络对每个割和可行区域进行打分，然后通过归一化函数输出概率。同时使用长短期记忆网络嵌入每个约束以便能够对不同大小的实例使用相同的策略网络。为了检查强化学习选择的割平面的质量，Tang 等分别在装箱问题、生产计划问题、背包问题、最大割问题上进行实验，并从割的有效性、解的质量等方面评估了该方法的性能。结果表明，此方法选择割平面所需时间较少，提高了整体算法的效率。

为增强割平面选择效率，Huang 等<sup>[22]</sup> 提出了一种基于机器学习的割平面选择方法，称为割排序。该方法在多实例学习环境中对学习过程建模，并训练出基于排名的割选择策略。与 Tang 等提出的迭代选择割平面的方法相比，Huang 等主要在两个方面有所不同。首先是提出了一个割排序规划，并对多实例学习环境中的学习问题进行了建模；其次，Tang 等人的主要目标是提高割平面选择效率，即减少添加的割平面总数，而 Huang 等的目标是通过割排序来提升算法的性能。其基本思想是学习一个可以衡量每个割平面质量的评分函数，其表示为割排序问题，通过习得的评分函数对每个生成的割平面进行评分，并选择得分最高的割平面。割排序的过程包括训练阶段和测试阶段。在训练阶段，学习过程在多实例学习背景下建模，训练数据被分组到若干个割集中，标签分配在割集层次上。训练数据由割集层次的特征和标签组成，并用其来训练一个评分函数以预测每个候选割的分数。Huang 等将评分函数参数化为神经网络，对于每个割集，模型将其特征作为输入，并输出一个二维概率分布。Huang 等选取了六种割平面，分别在集合覆盖问题、背包问题、生产计划问题、混合整数规划上进行实验。结果表明，习得的割排序策略与其他人工启发式算法相比更具竞争力，并且对不同大小和结构的问题具有泛化能力。

上述研究的实证结果表明机器学习技术用于割平面选择这一方法具有较好的泛化能力，并揭示了机器学习模块可以嵌入到分支切割算法中，以改进求解过程，尤其对于大规模组合优化问题。

## 5、总结与展望

本文调查了近些年机器学习与传统优化方法融合的相关方法和应用研究，从四大经典优化方法出发，综述了近些年机器学习与之结合的相关研究和方法，并对其结合原理及各方法的优劣异同进行了总结。

现有研究已经表现出机器学习与传统优化方法融合在解决组合优化问题方面的优势，但是相关研究还较少且处于初探阶段，其中仍然存在很多问题以及一些新出现的难点。未来可以从以下几个方面展开进一步研究，在现有机器学习与传统优化方法融合的理论知识上逐步拓展和完善。

- 1) 在机器学习与传统优化方法融合的切入点方面。从优化方法层次看，同类研究的切入点基本相同，如与分支定界法融合时，机器学习大部分解决的是分支问题，与列生成融合时则是列选择问题。下一步研究可以尝试二者结合的不同切入点，使结合点多面化。
- 2) 在研究问题方面。为了较好地说明机器学习与传统优化方法的融合及其解决问题的性能，当前研究大部分的实验针对于经典组合优化问题，这些问题大多是简单的、静态的。而实际生产生活中的问题是实时动态变化的，约束多而复杂，未来可以在二者结合的基础上对这些问题研究更好的求解方法。
- 3) 在机器学习模型方面。在当前的研究中，使用监督学习较多，这类模型是离线训练，然后在线预测，容易出现预测偏差，而强化学习模型较之有更好的效果，表现出很好的前景。这说明在模型选取方面还有较大的提升空间。

## 参考文献

- [1] 李凯文, 张涛, 王锐等. 基于深度强化学习的组合优化研究进展[J]. 自动化学报, 2021, 47(11):2521-2537.
- [2] 邱锡鹏. 神经网络与深度学习[M]. 北京:机械工业出版社, 2020:7-360.
- [3] BENGIO Y, LODI A, PROUVOST A. Machine learning for combinatorial optimization: a methodological tour d'horizon[J]. European Journal of Operational Research, 2021, 290(2): 405-421.
- [4] GAMBELLA C, GHADDAR B, NAOUM-SAWAYA J. Optimization problems for machine learning: a survey[J]. European Journal of Operational Research, 2020, 290(3): 807-828.
- [5] MAZYAVKINA N, SVIRIDOV S, IVANOV S, et al. Reinforcement learning for combinatorial optimization: a survey[J]. Computers & Operations Research, 2021, 134(10):105400.1-105400.15.
- [6] APPLEGATE D, BIXBY R, CHVATAL V, et al. Finding cuts in the tsp (a preliminary report) [J]. Technical Report, 1995, 89(12):1-66.
- [7] YU Y, NATASHIA B, BISTRÀ D, et al. Learning generalized strong branching for set covering, set packing, and 0-1 knapsack problem[J]. European Journal of Operational Research, 2021, 301(3): 828-840.
- [8] NAIR V, BARTUNOV S, GIMENO F, et al. Solving mixed integer programs using neural networks[J]. arXiv, 2020.
- [9] ALVAREZ A M, LOUVEAUX Q, WEHENKEL L. A machine learning-based approximation of strong branching[J]. INFORMS Journal on Computing, 2017, 29(1):185-207.
- [10] WEHENKEL L, ALVAREZ A M, LOUVEAUX Q. Online learning for strong branching approximation in branch-and-bound[J]. Department of Electrical Engineering and Computer Science, 2016, 32(1): 13813-13822.
- [11] HE H, HAL DAUMÉ, EISNER J. Learning to search in branch-and-bound algorithms[C]. NeurIPS, 2014:3293-3301.
- [12] MORABIT M, DESAULNIERS G, LODI A. Machine-learning-based arc selection for constrained shortest path problems in column generation[J]. arXiv, 2022.
- [13] MORABIT M, DESAULNIERS G, LODI A. Machine-Learning-Based Column Selection for Column Generation[J]. Transportation Science, 2021, 55(4):815-831.
- [14] CHI C, ABOUSSALAH A M, KHALIL E B, et al. A deep reinforcement learning framework for column generation[J]. arXiv, 2022.
- [15] FIRAT M, CROIGNIER G, GABOR A F, et al. Column generation based heuristic for learning classification trees[J]. Computers & Operations Research, 2020, 116(4):104866.1-104866.11.
- [16] QUESNEL F, WU A, DESAULNIERS G, et al. Deep-learning-based partial pricing in a branch-and-price algorithm for personalized crew rostering[J]. Computers & Operations Research, 2022, 138(2):105554.1-105554.15.
- [17] FURIAN N, OSULLIVAN M, WALKERC, et al. A machine learning-based branch and price algorithm for a sampled vehicle routing problem[J]. OR Spectrum, 2021, 43(3): 693-732.
- [18] VACLAVIK R, NOVAK A, SUCHA P, et al. Accelerating the branch-and-price algorithm using machine learning[J]. European Journal of Operational Research, 2018, 271(3):1055-1069.
- [19] ZHANG X Y, CHEN L, GENDREAU M, et al. Learning-based branch-and-price algorithms for the vehicle routing problem with time windows and two-dimensional loading constraints[J]. INFORMS Journal on Computing, 2022, 34(3): 1419-1437.
- [20] GOMORY R. An Algorithm for the Mixed Integer Problem[J]. Rand Corporation, 1960.
- [21] TANG Y, AGRAWAL S, FAENZA Y. Reinforcement learning for integer programming: learning to cut[C]// International Conference on Machine Learning. PMLR. Online. New York: Curran Associates, Inc, 2021: 9367-9376.
- [22] HUANG Z, WANG K, LIU F, et al. Learning to select cuts for efficient mixed-integer programming[J]. Pattern Recognition, 2021, 123:11-21.

# **Research Reviews of the Integration of Traditional Optimization Methods and Machine Learning**

ZHU Honghong<sup>1</sup>, LIU Ya<sup>2</sup>, YANG Xiwang<sup>3</sup>

School of Management, Xi'an Jiaotong University, Xi'an 710049, China

**Abstract:** Due to the continuous expansion of the scale of combinatorial optimization problems in practical applications, traditional combinatorial optimization methods suffer from huge computational burdens, and are difficult to expand to large-scale problems. In recent years, the integration of machine learning and traditional combinatorial optimization methods has shown good performance in dealing with combinatorial optimization problems, with strong generalization ability and fast solving speed, and provides a new direction and idea for solving combinatorial optimization problems. Therefore, by reviewing relevant research on the integration of traditional optimization methods and machine learning in recent years, firstly, the theoretical research on the integration of traditional optimization methods and machine learning is summarized, secondly, taking traditional optimization methods as the classification goal, the methods, logic and combinatorial points of the integration of machine learning and combinatorial optimization methods are summarized, and their advantages, disadvantages, similarities and differences are compared. Finally, some existing problems and difficulties in this direction are pointed out, aiming to provide inspiration and reference for further discussion and research in this direction in the future.

**Key words:** Combination optimization; optimization methods; machine learning; reinforcement learning, neural network