

**An Alternating Direction Method of Multiplier Based Problem Decomposition Scheme for
Iteratively Improving Primal and Dual Solution Quality in Vehicle Routing Problem**
基于交替方向乘子法的大规模问题分解框架：以车辆路径问题为例

Yu Yao

School of Traffic and Transportation, Beijing Jiaotong University
Beijing, 100044, China
Email: yaoyu1@bjtu.edu.cn

Xiaoning Zhu

School of Traffic and Transportation, Beijing Jiaotong University
Beijing, 100044, China
Email: xnzhu@bjtu.edu.cn

Hongyu Dong

(Corresponding author)
Logistics R&D Department, Beijing Jingdong Zhenshi Information Technology Co., Ltd.
Beijing, 100176, China
Email: donghongyu@jd.com

Shengnan Wu

Logistics R&D Department, Beijing Jingdong Zhenshi Information Technology Co., Ltd.
Beijing, 100176, China
Email: wushengnan1@jd.com

Hailong Wu

Logistics R&D Department, Beijing Jingdong Zhenshi Information Technology Co., Ltd.
Beijing, 100176, China
Email: wuhailong@jd.com

Lu (Carol) Tong

School of Electronic and Information Engineering, Beihang University, Beijing 100091, China
National Engineering Laboratory for Comprehensive Transportation Big Data Application Technology,
Beijing 100091, China

Xuesong Zhou

(Corresponding author)
School of Sustainable Engineering and the Built Environment, Arizona State University
Tempe, AZ, 85281, USA
Email: xzhou74@asu.edu

1 ABSTRACT

2 摘要

3 城市物流需要应对许多挑战，例如复杂的交通条件、时间敏感的顾客需求等。本文考虑了城市
4 物流中具有可变行驶时间和时间窗的车辆路径问题，目标是使包括运输成本、车辆等待成本和
5 车辆固定使用成本在内的总广义成本最小。我们采用高维时空网络流模型来建立具有多约束的
6 车辆路径问题。在求解 VRP 时，一个棘手的问题是如何通过迭代有效地提高原问题解和对偶
7 解的质量，以及如何打破同源车辆产生的对称性问题。因此，需要处理好模型中的耦合约束（例
8 如运输服务网络各层之间的一致性约束），来寻找问题的最优解或近似最优解。现有的研究中，
9 交替方向乘子法（ADMM）被广泛应用于凸规划领域，是一种增广拉格朗日松弛（ALR）和块
10 坐标下降法的结合。本文首次将 ADMM 引入到 VRP 问题（一个特殊的整数线性规划问题）中，
11 并提出了 ADMM 中二次惩罚项的线性化方法。更进一步，我们提出了一个具有计算可靠性的
12 对偶分解框架，以迭代的方式来改进原问题解和对偶问题解的质量。这个框架可以应用在更多
13 的交通问题中，从本质上讲，如果子问题是最短路径问题或类似的以二元变量作为决策变量的
14 问题都可以嵌入到这套求解框架中。算例部分，我们使用经典的 SOLOMON 数据集和一个基
15 于京东物流真实场景的算例来检验我们提出的方法。

16
17
18 *Keywords:* Urban logistics, Vehicle routing problem with time windows, Alternating direction method
19 of multipliers, State-space-time network

1. 引言

本文着眼于如何构建一个具有理论体系支撑的、可靠性的算法框架，以解决动态交通网络中广泛应用的调度和路由问题。同时，随着电子商务产业和乘客出行服务的快速发展，车辆路径、调度以及车辆共享也成为了在城市客货运运输中需要重点关注的问题。在这些问题上，VRP问题的求解算法是提供“门到门”货物配送(Savelsbergh和Woensel, 2016)和定制化公共交通服务(Tong et al., 2017)的重要组成部分。

1.1 VRP 研究综述

与城市间的长途运输相比，城市物流的效率在很大程度上依赖于实时的交通条件，客户对时间窗口和交货地点的要求也越来越高。这些复杂的交通状况和日益复杂的需求需要更加优化的操作决策，以便更有效地管理物流过程。在城市物流的背景下，VRP问题需要系统地考虑时变的交通条件、客户需求的时间窗和车辆承载能力。

在本研究中，我们考虑具有时间窗的车辆路径问题（VRPTW），该问题包括服务时间窗和车辆承载能力约束。VRPTW在文献中被广泛研究（Desrochers等人，1992；Kallehauge, 2008）。VRPTW的典型求解方法可以分为两类：启发式和精确求解方法。启发式方法包含有savings 算法（Clarke和Wright, 1964）、matching based算法（Altinkemer和B.Gavish, 1991）、sweep based算法（Gillet和Miller, 1974；Renaud和Boctor, 2002）、cluster first, route second算法（Fisher和Jaikumar, 1981）还有像一些元启发式算法，包含Tabu search（Taillard, 1993），自适应大邻域搜索（ALNS）算法（Goel和Gruhn等人，2005；Ghilas等人，2016），蚁群算法（Reimann等人，2004）等，这些方法虽然可以有效地生成高质量的解，但是通常无法评价解的质量。另一方面，精确或近似的优化方法，例如branch and cut、分支定价以及拉格朗日分解方法可以提供一个标准来评估所获得的解，并可以平衡求解效率和求解质量之间的关系。

1.2 问题分解研究综述

问题分解是解决大规模问题的一种方法，其核心是将原始问题分解为较小的子问题，并分别并行或串行地求解每个子问题。有许多经典的分解方法，如Dantzig-Wolfe(Dantzig and Wolfe, 1960)、Benders(Benders, 1962)、Column Generation(Ford 和 Fulkerson, 1958)、Lagrangian relaxation(Held and Karp, 1970)和 branch and Price(Nemhauser 等, 1991)。感兴趣的读者可以找到许多关于各种优化领域的研究和教科书，例如，Boyd (2004)、Lasdon (2002)、Bertsekas (1999) 第 6 章、Wolsey (1993) 和 Bradley (1977) 的第 12 章，涉及到了线性、非线性和整数规划等多个方面。

这些分解方法已广泛用于求解 VRP 问题。在以下经典文献中，Branch-and-bound (Christofides et al., 1981), Branch-and-cut (Laporte et al., 1985), Lagrangian relaxation (Fisher, 1997) and Branch-and-cut-and-price (Fukasawa et al., 2006)被用来解决车辆路径问题。近年来，可变旅行时间的 VRP 问题引起了人们的广泛关注。Dabia et al. (2013)采用了分枝定价框架分解问题，以集划分问题作为主问题，以资源约束下的最短路径问题作为价格子问题。Mahmoudi and Zhou (2016)为了在多层时空状态网中嵌入了车辆容量、时间窗约束和接取送达约束，列举了可能的运输状态，构造了高维多商品流公式，然后利用拉格朗日松弛法求解。将原始模型分解成一系

列最短路径子问题。正如 Barnhart et al. (1998)和 Niu et al. (2018)提到的, 在处理由多个同源车辆组成的多车辆路径问题时, 研究者需要充分认识到解的对称性问题, 并设计有效的破除对称性技术。

在本研究中, 我们以对偶分解为基础, 提出一个基于交替方向乘子法 (ADMM) 的可靠求解框架。关于 ADMM 的经典文献可以追溯到 Glowinski(1975), 随后其的收敛性分析和许多相关的理论也逐渐完善(Ruszczyski, 1989; Eckstein 和 Bertsekas, 1992)。Boyd et al. (2011)对 ADMM 进行了综述, 使得其在许多学科中的使用进一步普及, 特别是在大数据和机器学习, 以及分布式优化领域 (Nedic 和 Ozdaglar, 2009)。Boyd et al. (2011)也提及了一些 ADMM 在非凸统计学习问题中的应用, 例如 regressor selection and factor model fitting。在随机多阶段混合整数规划背景下, Boland et al. (2018)采用了基于 ADMM 的求解过程来处 progressive hedging model (增广拉格朗日的随机形式)。文章指出该方法在混合整数规划中的缺点——难以保证其收敛性。针对这一问题, 他们提出了一种计算下界的 FW-PH 算法, 并证明了该下界收敛于拉格朗日下界。

就我们所知, 目前很少有研究专将 ADMM 应用于在车辆路径问题, 该问题本质上是运输网络中的组合优化问题。我们也希望我们提出的基于 ADMM 的求解框架能够进一步应用于更多的交通优化问题中。

1.3 Motivation and potential contributions

我们将 ADMM 作为一种改进的对偶分解算法引入到 VRP 问题中, 希望对以下方面有更多的启示。首先, 与元启发式方法相比, 一个理想的算法不仅要获得好的可行解, 而且要建立一个可靠的下界算法来衡量解的质量。其次, 在 VRP 的经典 B&P 求解框架中, 每个车辆有多个备选路线, 通常需要通过调用线性规划 (LP) 求解器和复杂分支策略的来解决主问题。相比之下, 我们提出的 ADMM 框架在一次迭代中每个车辆只有一条路径, 这提供了相对简单的算法实现结构, 特别是对于具有大量备选路径要管理的问题中。本质上, 这个 ADMM 框架旨在迭代地提高原问题的可行性和对偶问题的可行性, 并且我们希望通过进一步扩展此分解框架, 以处理 VRP 之外的更多的运输问题 (例如各优化层之间的有一致性约束的问题)。

在提出的基于 ADMM 的求解框架中, 我们进行了一系列的模型重构步骤, 即 (a) 对偶和增广化, (b) 分解和 (c) 线性化。具体来说, 本文解决了以下挑战。

- (1) 如果通过 D-W 分解或分枝定界算法, 将 LP 松弛后的问题(集划分问题)作为用于 VRP 的有限主问题, 则通常需要复杂的分支技巧来得到可行的整数解。另一方面, 使用拉格朗日松弛法求解 VRP 的局限性在于其固有的解对称性问题。我们提出了一种改进的对偶分解方法, 通过增加增广项可以有效地打破解的对称性, 并快速生成良好的可行可行解。
- (2) ADMM 在凸规划领域有着广泛的应用, 但是对于整数规划问题的应用鲜有研究。例如 VRP 问题, 该类问题包含了一组整数决策变量, 同时还有复杂的约束条件。为了克服这一困难, 我们使用超维网络 (状态时空网络) 模型来简化约束, 并进一步使动态规划算法有效地求解对偶问题。
- (3) 将 ADMM 应用于 VRP 的另一个挑战是如何将对偶问题中包含的二次项线性化。在既有的研究中, 有学者用 Frank-Wolf 方法或一阶 Taylor 展开(Nishi 等人, 2005)进行线性化, 然而这些方法本身仍有较大的计算成本。在本文中我们清楚地表明, 基于 Gauss-Seidel 形式 block coordinate descent method, VRP-ADMM 模型中使用的二次罚项对于每个 “x-update”都是可

分的，如果只涉及二元决策值，则可以简化为更简单的线性函数形式。这种等价关系的保持了问题的特性并有较高的计算效率。

- (4) 单独应用 ADMM 方法时，只能得到上界可行解，缺乏评价解质量的信息。本文提出了一种同时计算上界和下界值的解方法。通过迭代求解拉格朗日对偶问题和增广拉格朗日模型这两个密切相关的问题，可以相应地获得并减小全局最优性差距。

本文的其余部分结构如下。在第2节中，基于状态时间空间网络来构造VRPTW的模型。在第3节中，给出了基于ADMM的分解框架和求解过程。第4节使用实例验证了ADMM可以破除对称性以及具有可靠性(计算上)。在第5节中，我们讨论了ADMM的收敛性和潜在的可扩展性。在第6节中，将所提出的模型和解决方案框架应用于所罗门数据集和基于真实数据的案例。最后，我们在第7节给出我们的结论。

2. 问题描述及模型构建

2.1 VRPTW 问题描述

本文考虑的 VRPTW 问题，目标是找到一组广义系统成本最小的路线，包括运输成本、车辆等待成本和车辆固定使用成本。给定一个由 (N, M) 表示的物理运输网络，其中 N 是节点集合， M 是有向弧的集合。我们使用 $TT(i, j, t)$ 来表示在时刻 t 出发在弧 $(i, j) \in M$ 上的行驶时间。假定行驶时间可以根据交通条件计算得到。

在我们考虑的网络中有两种不同类型的节点，子集 P 表示的客户节点，由配送中心 o 表示的客户节点，即 $N = \{o\} \cup P$ 。每个客户 $p \in P$ 有相应的需求的体积和重量以及服务时间窗 $[e_p, l_p]$ ，其中 e_p 是最早服务时间， l_p 是最晚服务时间。 P 中的所有客户节点只能被一辆车访问一次。

配送中心 o 是所有车辆出发的起点和终点，同时车辆也可以回到配送中心重新装卸货。也就是说，在送完一批包裹之后，车辆可以返回配送中心装载下一批需要配送的货物。每个车辆也有相应的服务时段 $[e_v, l_v]$ ，其中 e_v 是最早离开配送中心的时间， l_v 是最晚返回配送中心的时间。

2.2 状态时空网络

VRP 问题的一种普遍使用的建模方法是直接基于客户节点的物理网络，使用累积时间和累积负载变量 (Cordeau, 2006) 进行建模，这些变量被用在时间窗和车辆能力约束中。另一个广泛使用的建模方式是通过时间扩展网络结构 (Boland 等人, 2017) 建立模型，这种建模方式可以在网络构建时就保证时间窗约束。在典型的列生成框架中，定价子问题通常还调用 multi-label 动态规划算法 (Boland 等人, 2006) 以考虑上述约束和其他约束，例如拾取和交付约束。一些可行性、状态空间松弛、支配准则被开发以减少搜索空间 (例如 EILON 等人, 1971)。

在对偶分解框架中，理想的模型应该具有数量较少的复杂约束需要被对偶化，并且松弛的问题应该使用高效的算法来解决。因此，我们采用 Mahmoudi 和 Zhou (2016) 提出的高维多商品流建模框架，构造了一个具有空间维度、时间维度和累积服务状态维度的基于时间索引和状态索引网络。用 $G = (E, A)$ 表示高维网络，点 $(i, t, w) \in E$ 有物理网中的节点 $i \in N$ 扩展，弧 $(i, j, t, s, w, w') \in A$ 表示从顶点 (i, t, w) 到顶点 (j, s, w') 的有向时空状态路径。其中， t 表示规划时间维度中的均匀离散时间间隔 (例如，1 分钟)，值得注意的是，这个维度自然地嵌入了时间窗

1 约束。此外， w 表示车辆的“累积服务状态”，用于记录服务了的客户并提供相应的状态和体
 2 积信息，以满足车辆容量约束。返回配送中心 o 后，车辆的累计服务状态被重置。图 1 展示了
 3 一个状态空间时间路径的简单示例，该路径对应于节点序列为 $(O,0,w_0)-(1,3,w_1)$
 4 $-(O,6,w_0)-(2,10,w_2)-(3,12,w_6)$ 。时间窗口和容量约束是通过时间和状态维度嵌入了网络。

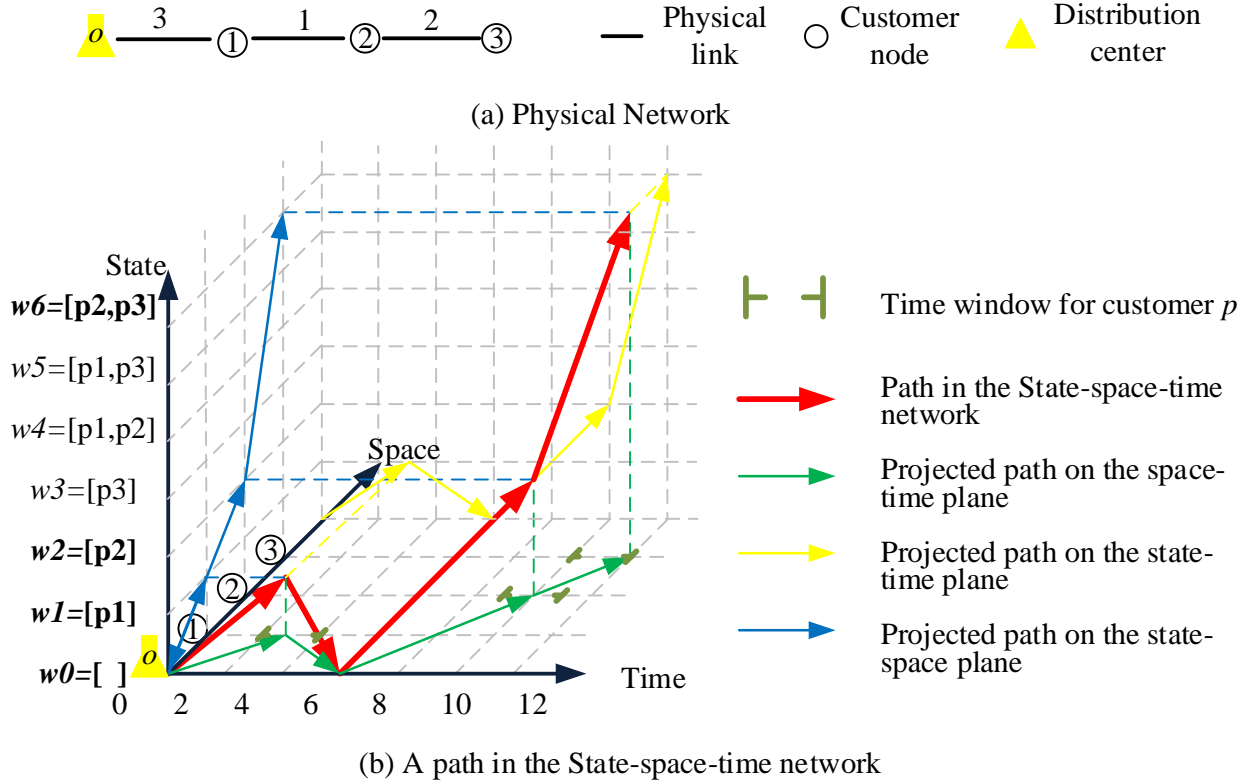


Fig. 1. A simple example of state-space-time path adopted from Mahmoudi and Zhou (2016)

2.3 VRPTW 模型

表 1 和表 2 描述了用于网络构建和模型建立所需要的集合、索引、变量和参数。

Table 1

Sets, indexes and variables used for model formulation.

Symbol	Definition
V	车的集合
P	客户集合
W	状态集合
A_v	时空状态网络中弧的集合
$\Psi_{p,v}$	V 车对 p 客户的服务弧集合
v	车的索引
p	客户的索引
w	状态索引
$(i, t, w), (j, s, w')$	时空状态网络中的节点
(i, j, t, s, w, w')	时空状态网络中弧 indicating that one travels from node i at time t with cumulative service state w to the node j at time s with state w'

a	(i, j, t, s, w, w') 的简写
$x_{i,j,t,s,w,w'}^v$	$= 1$ if arc (i, j, t, s, w, w') is used by vehicle v ; $= 0$ otherwise

Table 2

Parameters used for network construction and model formulation

Symbol	Definition
$[e_p, l_p]$	客户 p 的时间窗
$[e_v, l_v]$	车辆 v 的工作时段
o_v	车辆 v 的起点
d_v	车辆 v 的终点
w_0	初始状态
$c_{i,j,t,s,w,w'}$	弧 (i, j, t, s, w, w') 的费用
$TT(i, j, t)$	时刻出发在 link (i, j) 上的旅行时间

VRPTW 模型如下。

Objective Function:

$$\min Z = \sum_{v \in V} \sum_{(i,j,t,s,w,w') \in A_v} c_{i,j,t,s,w,w'} x_{i,j,t,s,w,w'}^v \quad (1)$$

该模型的目标函数是使所有选定的弧的总成本最小，网络中的弧可以分为以下三类：运输弧、等待弧和装卸弧。运输弧 (i, j, t, s, w, w') 表示基于给定的旅行时间 $TT(i, j, t)$, $s = t + TT(i, j, t)$ ，一辆车从节点 I 移动到节点 J 。等待弧 $(i, i, t, t+1, w, w)$ 表示从时间 t 到 $t+1$ 的节点 i 上的等待活动，即车辆位置和承载状态在一个时间间隔内保持不变。装卸弧 (i, j, t, s, w, w') 表示车辆在配送中心装载和在乘客点配送过程，只有这个弧会有状态的变化。

流平衡约束:

在我们考虑的 VRP 问题中，时间窗、车辆能力（体积、重量）能力约束已经被提前嵌入到高维网络中。因此，对每辆车而言，我们只需要确保所有选择的弧都能够遵循流量平衡约束，构成从起点到终点的可行路径。

起点的流平衡约束：

$$\sum_{(i,j,t,s,w,w') \in A_v} x_{i,j,t,s,w,w'}^v = 1 \quad i = o_v, t = e_v, w = w_0, \quad \forall v \in V \quad (2)$$

终点的流平衡约束：

$$\sum_{(i,j,t,s,w,w') \in A_v} x_{i,j,t,s,w,w'}^v = 1 \quad j = d_v, s = l_v, \quad \forall v \in V \quad (3)$$

中间点的流平衡约束：

$$\sum_{(j,s,w'')} x_{i,j,t,s,w,w''}^v - \sum_{(j',s',w')} x_{j',i,t,s',w'}^v = 0 \quad (i, t, w) \notin \{(o_v, e_v, w_0), (d_v, l_v, w)\}, \quad \forall v \in V \quad (4)$$

约束(2)和(3)确保每个车辆在研究时段开始时（即 e_v 时刻）以初始状态 w_0 离开原点 o_v ，并在研究时段结束时（即 l_v 时刻）到达目的地 d_v 。约束(4)保证其他中间节点上的流量平衡。

需求约束：

$$\sum_{v \in V} \sum_{(i,j,t,s,w,w') \in \Psi_{p,v}} x_{i,j,t,s,w,w'}^v = 1 \quad \forall p \in P \quad (5)$$

约束 (5) 确保每一个客户正好被服务一次。

$$x_{i,j,t,s,w,w'}^v \in \{0, 1\} \quad \forall (i,j,t,s,w,w') \in A_v, \forall v \in V \quad (6)$$

约束 (6) 定义了车辆对于每个弧的选择是个 0-1 变量。

3. Dualization and augmentation, decomposition and linearization techniques for applying ADMM in VRP

在本节中，首先，我们针对连续凸问题，简要回顾了 ADMM 的背景知识。其次，对于具有离散二元决策变量的车辆路径问题，我们提出了包括三个步骤的模型重构方法，即(a)对偶和增广，(b)分解和(c)线性化。具体来说，这种对偶和增广过程可以松弛模型中的复杂约束，但是二次项的引入也使得问题变得不可分解；因此，我们利用 ADMM 中的迭代原理对模型进行分解，得到一系列非线性子问题，并使用线性化技术将子问题简化为更简单的线性公式。每个子问题可以通高效的动态规划算法和搜索区域缩减技术求解。为了解决维数灾难的问题，在解决相似的列生成框架中的定价子问题时，一些学者也使用了具有状态空间松弛和 dominance criteria 的多标签动态规划算法来求解 (Boland 等人, 2006; Eilon 等人, 1974)。

3.1 Formulation of ADMM

本质上，ADMM 是增广拉格朗日松弛（参见附录）和块坐标下降方法（block coordinate descent methods）的集成。对于一个具有可分离目标函数和线性等式约束的问题 (7)（例如 VRP 问题中的 x 车和 y 车）。

$$\min f(x) + g(y) \quad (7)$$

$$\text{subject to } Ax + By - c = 0$$

其中 $x \in R^n, y \in R^m, A \in R^{p \times n}, B \in R^{p \times m}, c \in R^p, f(x)$ 和 $g(y)$ 是凸函数。该问题的增广拉格朗日形式如 (8) 所示，对应的对偶问题如 (9) 所示。

$$L(x, y, \lambda, \rho) = f(x) + g(y) + \lambda^T (Ax + By - c) + \frac{\rho}{2} \|Ax + By - c\|_2^2 \quad (8)$$

$$\max_{\lambda} \inf_{x,y} L(x, y, \lambda, \rho) \quad (9)$$

变量 x 、 y 和乘子 λ 可以按照 (10) 中的迭代方式依此更新。

$$\begin{aligned} x^{k+1} &:= \argmin_x L(x, y^k, \lambda^k, \rho) \\ y^{k+1} &:= \argmin_y L(x^{k+1}, y, \lambda^k, \rho) \end{aligned} \quad (10)$$

$$\lambda^{k+1} := \lambda^k + \rho (Ax^{k+1} + By^{k+1} - c)$$

需要注意的是，这里我们更新 λ 的步长是 ρ ，这与拉格朗日松弛方法中的步长 α^k 不同（参见附录，等式 (A4)）。

在有不等式约束的情况下，我们可以引入松弛变量 s 来将不等式约束转换为等式约束。换言之，用 $Ax + By - c = s, s \leq 0$ 代替 $Ax + By - c \leq 0$ ，从而得到 Eq. (11)。

$$L(x, y, \lambda, \rho) = f(x) + g(y) + \lambda^T (Ax + By - c - s) + (\rho/2) \|Ax + By - c - s\|_2^2 \quad (11)$$

ADMM 模型的取得最优的充要条件是 Primal feasibility 和 Dual feasibility (Eckstein 和

1 Bertsekas,1992),如(12)所示。

Primal feasibility: $Ax + By - c = 0$

Dual feasibility: $\nabla f(x^*) + A^T \lambda^* = 0$

$\nabla g(y^*) + B^T \lambda^* = 0$

(12)

2 相应的原始和对偶残差可以用(13) 计算。

Primal residuals: $r^{k+1} = Ax^{k+1} + By^{k+1} - c$

Dual residuals: $s^{k+1} = \rho A^T B(y^{k+1} - y^k)$

(13)

3 利用 ρ 作为更新对偶变量的步长, $(x^{k+1}, y^{k+1}, \lambda^{k+1})$ 始终满足的第二条 Dual feasibility。在凸规
4 化问题中, 当 $k \rightarrow \infty$ 时, 原残差 r^{k+1} 和对偶残差 s^{k+1} 将收敛到零, Primal feasibility 和第一条 Dual
5 feasibility 也将满足。

7 3.2 Reformulation of the VRP model

8 在所提出的VRP模型中, 不同车辆之间的耦合约束 (5) 是唯一的一组复杂约束。利用拉格朗
9 日松弛, 我们将方程 (5) 对偶化为目标函数 (14)。进一步地, 添加二次惩罚项以将所提出的
10 模型转换为增广形式 L' , 如(15)所示,

11

$$\min L = \sum_{v \in V} \sum_{a \in A_v} c_a x_a^v + \sum_{p \in P} \lambda_p \left(\sum_{v \in V} \sum_{a \in \Psi_{p,v}} x_a^v - 1 \right) \quad (14)$$

$$\min L' = \sum_{v \in V} \sum_{a \in A_v} c_a x_a^v + \sum_{p \in P} \lambda_p \left(\sum_{v \in V} \sum_{a \in \Psi_{p,v}} x_a^v - 1 \right) + \frac{\rho}{2} \sum_{p \in P} \left(\sum_{v \in V} \sum_{a \in \Psi_{p,v}} x_a^v - 1 \right)^2 \quad (15)$$

12 同时满足约束 (2), (3), (4) 和 (6)。为了表示清晰, 我们将 (i, j, t, s, w, w') 简写为 a , 并定义
13 一个新的变量 μ_p^v 来表示: 除了车辆 v 之外, 客户 p 被其他车辆服务的总次数, 如 (16) 所示。

14

$$\mu_p^v = \sum_{v' \in V/\{v\}} \sum_{a \in \Psi_{p,v'}} x_a^{v'} \quad \forall p \in P \quad (16)$$

15 显然, 拉格朗日问题可以分解为每个车辆 v 的子问题 L_v , 如公式 (17)。

16

$$L_v = \sum_{a \in A_v} c_a x_a^v + \sum_{p \in P} \sum_{a \in \Psi_{p,v}} \lambda_p x_a^v \quad (17)$$

17 同样, 利用 ADMM 内部的块坐标下降方法中, 增广拉格朗日问题也可以分解为一系列子问题 L'_v ,
18 如式 (18),

$$L'_v = \sum_{a \in A_v} c_a x_a^v + \sum_{p \in P} \sum_{a \in \Psi_{p,v}} \lambda_p x_a^v + \frac{\rho}{2} \sum_{p \in P} \left(\sum_{a \in \Psi_{p,v}} x_a^v + \mu_p^v - 1 \right)^2 \quad (18)$$

19 可以看到在子问题 L'_v 中包含着二次项, 一般来说, 二次项会使得计算复杂化。然而, 下面的证
20 明表明, 基于决策变量的二元性 (只有 0, 1 两个取值), 子问题中的二次项可以被分离、化简
21 和重组为线性化的目标函数。

22 **命题:** 每个子问题可归结为具有线性目标函数的动态规划问题:

1

$$L'_v = \sum_{a \in A_v} \hat{c}_a^v x_a^v \quad (19)$$

$$\hat{c}_a^v = \begin{cases} c_a + \lambda_p + \rho \mu_p^v - \frac{\rho}{2} & a \in \Psi_{p,v} \\ c_a & otherwise \end{cases} \quad (20)$$

2 **证明:**

3 让我们从一个简单的二次形式 $(x + b)^2$ 开始, 其中 x 是 0-1 二元变量, b 是常数。显然, 由于二
4 元变量的性质, $x^2 = x$; 因此 $(x^2 + 2xb + b^2)$ 可以进一步简化为 $(x + 2xb + b^2) = (2b + 1)x +$
5 b^2 , 这是一个关于 x 的线性函数。同理, 在我们的问题中, 我们车辆 v 和其他车辆分开考虑,
6 则二次项可以拆分为 3 部分, 如等式 (21) 所示:

7

$$\left(\sum_{a \in \Psi_{p,v}} x_a^v + \mu_p^v - 1 \right)^2 = \left(\sum_{a \in \Psi_{p,v}} x_a^v \right)^2 + 2 \sum_{a \in \Psi_{p,v}} x_a^v (\mu_p^v - 1) + (\mu_p^v - 1)^2 \quad (21)$$

8 其中 $\{x_a^v\}$ 是子问题 L'_v 的决策变量, μ_p^v 可以由方程 (16) 计算。可以容易地得出 $\left(\sum_{a \in \Psi_{p,v}} x_a^v\right)^2$ 是
9 唯一的二次项。由于在我们的高维网络结构中, 每个客户最多可以由一个车辆服务一次, 所以
10 $\sum_{a \in \Psi_{p,v}} x_a^v = \{0,1\}$, 0-1 变量的平方时它本身。

$$\left(\sum_{a \in \Psi_{p,v}} x_a^v \right)^2 = \sum_{a \in \Psi_{p,v}} x_a^v \quad (22)$$

11 然后, 我们可以将每个车辆的子问题的目标函数重新组合为等式(23)中的线性形式,

$$L'_v = \sum_{a \in A_v} c_a x_a^v + \sum_{p \in P} \sum_{a \in \Psi_{p,v}} \left[\lambda_p x_a^v + \frac{\rho}{2} x_a^v (2\mu_p^v - 1) \right] + Q = \sum_{a \in A_v} \hat{c}_a^v x_a^v + Q \quad (23)$$

12 其中, \hat{c}_a^v 是车辆每次迭代的组合参数, Q 是与变量无关的项。

13 3.3 基于 ADMM 的求解算法流程

14 算法流程如下

15 **//Step 1:** 初始化

16 初始化迭代次数 $k = 0$;

17 初始化拉格朗日乘子 λ_p^0 and 二次项参数 ρ^0 ;

18 初始化上界 $\{X_{UB}^0\}$ 和下界解 $\{X_{LB}^0\}$;

19 初始化下界 $LB^* = -\infty$, 上界 $UB^* = +\infty$.

20 **//Step 2:** 对每辆车依次求增广拉格朗日函数值

21 **Step 2.1** //对每辆车的子问题调用正向动态规划算法

22 **For each vehicle** $v \in V$

23 Update the arc cost $\hat{c}_a^v \in \Psi_{p,v}$ by Eq. (20); 更新路径的广义费用

24 Find the state-space-time shortest path for vehicle v (solve L'_v) by calling the forward dynamic
25 programming algorithm; 调用 DP 算法求解 L'_v 子问题

26 **End for**

1 **Step 2.2** // 更新乘子和惩罚项参数

2 $\lambda_p^{k+1} := \lambda_p^k + \rho^k \left(\sum_{v \in V} \sum_{a \in \Psi_{p,v}} x_a^v - 1 \right), \forall p \in P;$

3 Update the quadratic penalty parameter by:

4
$$\rho^{k+1} := \begin{cases} \rho^k + \beta & \text{if } \sum_{p \in P} \left(\sum_{v \in V} \sum_{a \in \Psi_{p,v}} x_a^{v^k} - 1 \right)^2 \geq \gamma \sum_{p \in P} \left(\sum_{v \in V} \sum_{a \in \Psi_{p,v}} x_a^{v^{k-1}} - 1 \right)^2, \\ \rho^k & \text{otherwise} \end{cases},$$

5 where $1 \leq \beta \leq 10$ and $0.25 \leq \gamma \leq 0.5$ can be chosen.

6 **//Step 3:** 生成上界可行解并更新上界值

7 **//Step 3.1:** 生成上界可行解

8 从 2.1 中复制人-车配对的结果

9 **For** each customer $p \in P$ **do**

10 If the customer is served by more than one vehicle, then designate one of the vehicles for

11 him/her.

12 If the customer is not served by any vehicle, then assign a backup vehicle for him/her.

13 **End for**

14 **// Step 3.2:** Compute UB^k

15 Compute UB^k by substituting solution $\{X_{UB}^k\}$ in Eq. (1);

16 $UB^* = \min\{UB^*, UB^k\}.$

17 **// Step 4:** 生成下界可行解并更新下界值

18 **// Step 4.1:** Generate lower bound solution $\{X_{LB}^k\}$ by solving the pure Lagrangian dual problem

19 **For** each vehicle $v \in V$

20 Solve sub-problem (17) by calling the forward dynamic programming algorithm, where the

21 Lagrangian multipliers $\{\lambda_p^k\}$ are adopted from step 2; note that sub-problem (17) does not

22 contain any quadratic penalty terms related to ρ , as it is a sub-problem of pure Lagrangian

23 dual problem;

24 **End for**

25 **// Step 4.2:** Compute LB^k

26 Initialize $LB^k=0$;

27 Compute LB^k by substituting solution $\{X_{LB}^k\}$ in Eq. (14);

28 $LB^* = \max\{LB^*, LB^k\}.$

29 **//Step 5:** 评估解的质量和终止条件测试

30 Compute the relative gap between LB^* and UB^* by:

31 Gap = $\frac{UB^* - LB^*}{UB^*} \times 100\%$;

32 If the solution satisfies $r^{k+1} = 0$ and $s^{k+1} = 0$, or k reaches the maximum iteration number,

33 terminate the algorithm and output the best lower bound LB^* and best upper bound UB^* ; otherwise,

34 $k:=k+1$ and go back to Step 2.

35 值得注意的是，当解决第 m 个单独车辆的子问题时，在这个内部迭代步骤中，仅优化第 m 个

1 车辆的路线决策变量。如图 2 所示，所有其他车辆的路线暂时固定。

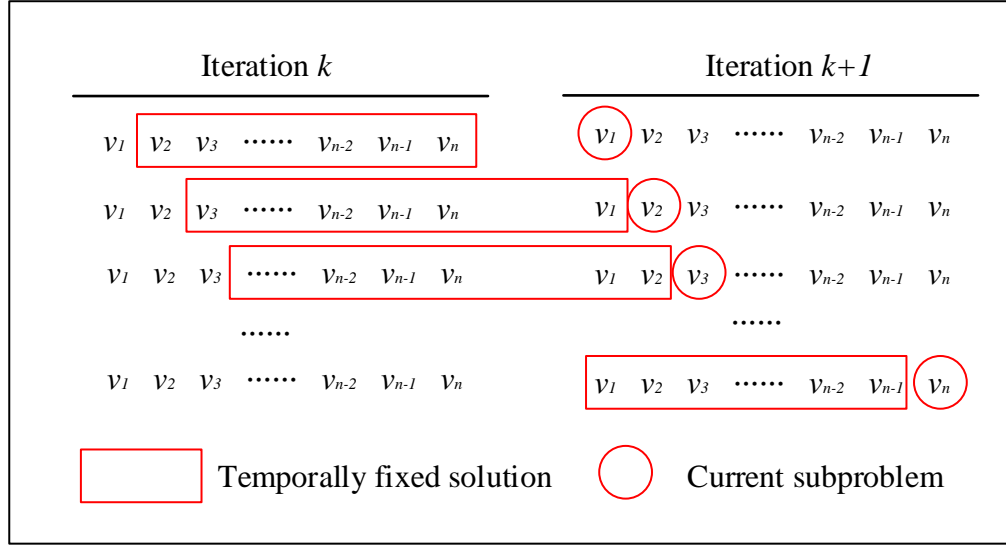


Fig. 2. Iterative pattern in the ADMM based framework for VRP.

4. Illustration example for breaking symmetry and improving the primal and dual solution

为了证明 ADMM 的计算可靠性和可破除对称性，我们用一个简单的例子来比较 ADMM 技术和标准的拉格朗日松弛方法。考虑一个如图 3 所示的物理网络，其中节点 1 和节点 2 代表两个客户，并且假设每条弧的成本是常数，并已在相应的弧旁标出。每个车辆从原始节点 o 开始，最后返回目的地 d ，设置容量为 1。其他参数和初始值包括： $\lambda_p^0 = -6$; $\rho = 2$; $\alpha^0 = 2$, $\alpha^{k+1} = 0.5\alpha^k$, 以及 $\pi_{p,v}^k$ 表示车辆 v 在第 k 次迭代中为顾客 p 服务的价格。

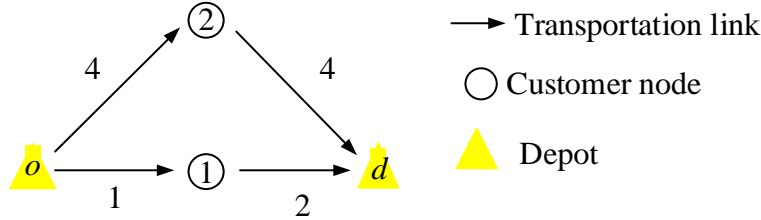


Fig. 3. A simple physical network

在标准的拉格朗日松弛方法中，原问题同差被分解为两条针对单个车辆的最短路径问题。在迭代过程中，通过调整乘数 λ_p 来保证松弛约束，其中 λ_p 可以看作是顾客 p 的服务利润，反之亦然。然而，由于两个子问题是相同的，所以两辆车会走相同的路线。只有当我们能找到一组最小成本路径，恰好能够精确地覆盖每个客户一次时，才能找到可行解也就是最优解决 (Kohl 和 Medsen, 1997)，否则拉格朗日松弛只能提供原始问题的下界。

我们在图 4 中画出松弛问题的函数，其中在底面中的四个点表示问题的 4 个解，具体含义在表 4 中给出。该图显示了三个轴分别代表 v_1 , v_2 和 L 。为了清晰起见，函数图像是以连续的形式绘制。在表 5 中演示了详细的优化过程，其中我们可以看到，在迭代 5 中获得了两条具有相同成本的路径。

Table 4

Details of the corner points in the graph of solution geometry.

The points in GSG	Assignment results		Feasibility in the primal problem	Times of service	
	v_1	v_2		p_1	p_2
(1, 1)	p_1	p_1	Infeasible	2	0

(1, 2)	p_1	p_2	Feasible	1	1
(2, 1)	p_2	p_1	Feasible	1	1
(2, 2)	p_2	p_2	Infeasible	0	2

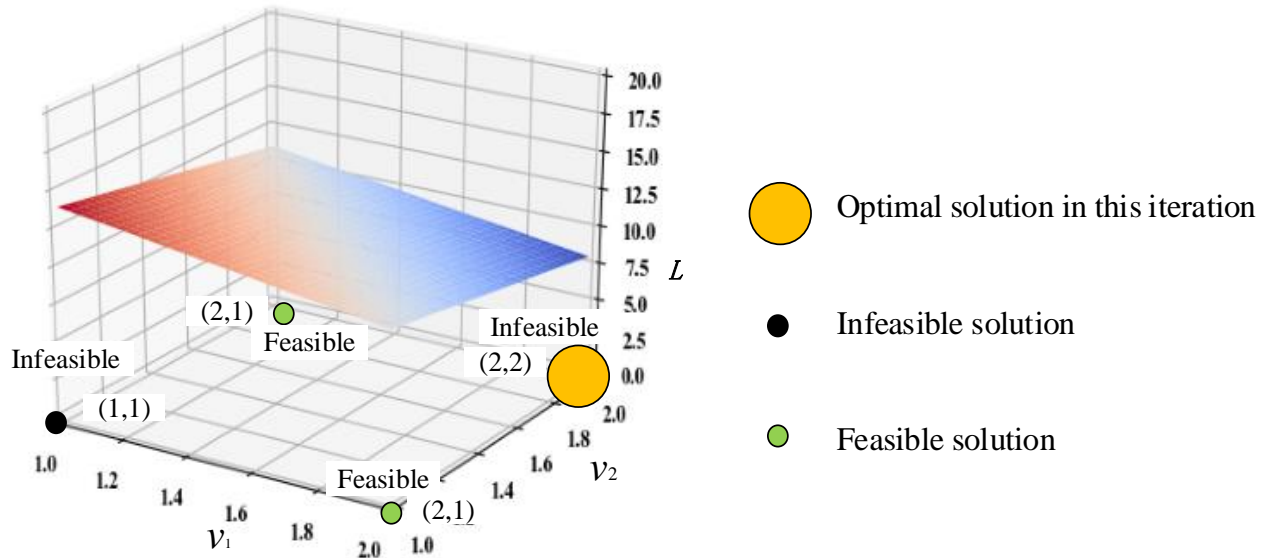


Fig.4. An example of the graph of solution geometry

Table 5

Detailed optimization process using Lagrangian relaxation method.

Details of computation **Geometry of sub-problem** **Graphical solution**

● Optimal solution of the dual problem at each iteration
 ● Feasible solution
 ● Infeasible solution

[] Service price
 → Path of Vehicle 1
 → Path of Vehicle 2

Iteration 0:

$$\lambda_1^0 = \lambda_2^0 = -6$$

$$\pi_{1,1}^0 = \pi_{1,2}^0 = -\lambda_1^0$$

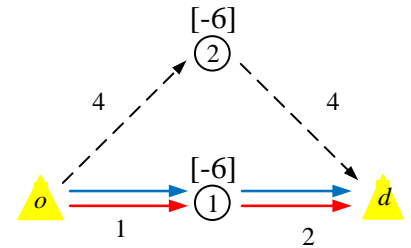
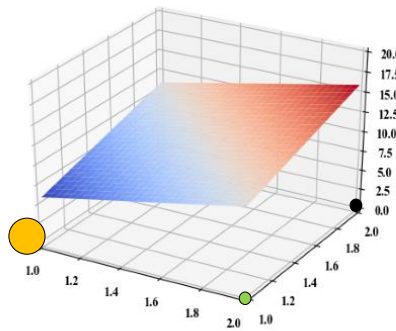
$$\pi_{2,1}^0 = \pi_{2,2}^0 = -\lambda_2^0$$

The shortest path: [o-1-d]

$$L_1 = 1 + 2 - \pi_{1,1}^0 = -3$$

$$L_2 = 1 + 2 - \pi_{1,2}^0 = -3$$

$$L = 6$$



Iteration 1:

$$\alpha^0 = 4$$

$$\lambda_1^1 = -2, \lambda_2^1 = -10$$

$$\pi_{1,1}^1 = \pi_{1,2}^1 = -\lambda_1^1$$

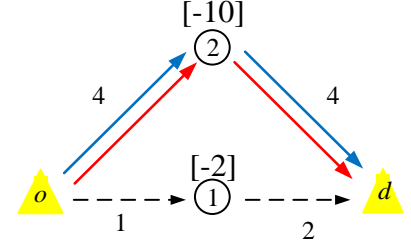
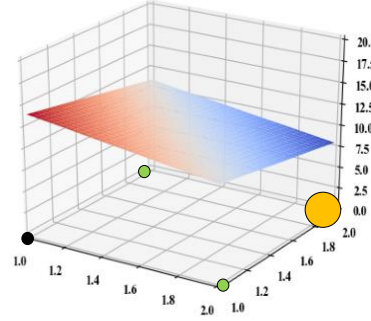
$$\pi_{2,1}^1 = \pi_{2,2}^1 = -\lambda_2^1$$

The shortest path: [o-2-d]

$$L_1 = 4 + 4 - \pi_{2,1}^1 = -2$$

$$L_2 = 4 + 4 - \pi_{2,2}^1 = -2$$

$$L = 8$$

**Iteration 2:**

$$\alpha^1 = 2$$

$$\lambda_1^2 = -4, \lambda_2^2 = -8$$

$$\pi_{1,1}^2 = \pi_{1,2}^2 = -\lambda_1^2$$

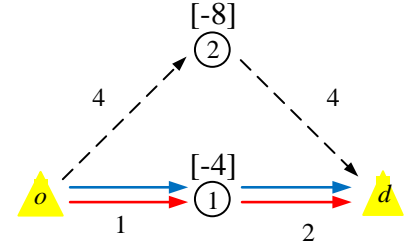
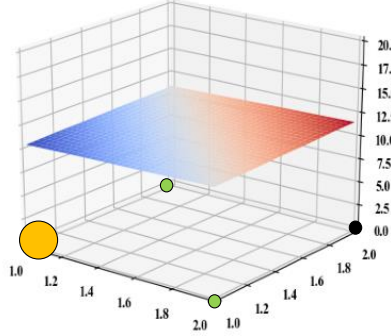
$$\pi_{2,1}^2 = \pi_{2,2}^2 = -\lambda_2^2$$

The shortest path: [o-1-d]

$$L_1 = 1 + 2 - \pi_{1,1}^2 = -1$$

$$L_2 = 1 + 2 - \pi_{1,2}^2 = -1$$

$$L = 10$$

**Iteration 3:**

$$\alpha^2 = 1$$

$$\lambda_1^3 = -3, \lambda_2^3 = -9$$

$$\pi_{1,1}^3 = \pi_{1,2}^3 = -\lambda_1^3$$

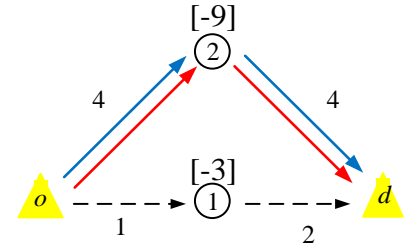
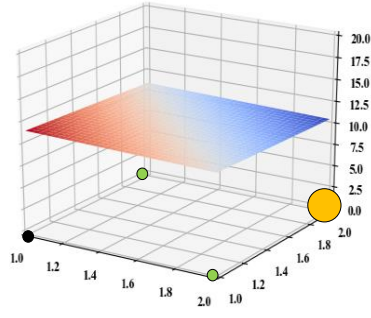
$$\pi_{2,1}^3 = \pi_{2,2}^3 = -\lambda_2^3$$

The shortest path: [o-2-d]

$$L_1 = 4 + 4 - \pi_{2,1}^3 = -1$$

$$L_2 = 4 + 4 - \pi_{2,2}^3 = -1$$

$$L = 10$$

**Iteration 4:**

$$\alpha^3 = 0.5$$

$$\lambda_1^4 = -3.5, \lambda_2^4 = -8.5$$

$$\pi_{1,1}^4 = \pi_{1,2}^4 = -\lambda_1^4$$

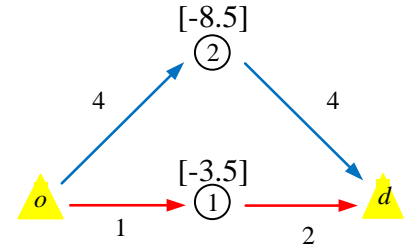
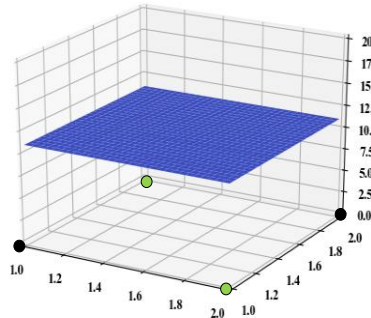
$$\pi_{2,1}^4 = \pi_{2,2}^4 = -\lambda_2^4$$

The shortest paths: [o-1-d], [o-2-d]

$$L_1 = 1 + 2 - \pi_{1,1}^4 = -0.5,$$

$$L_2 = 4 + 4 - \pi_{2,2}^4 = -0.5$$

$$L = 11$$

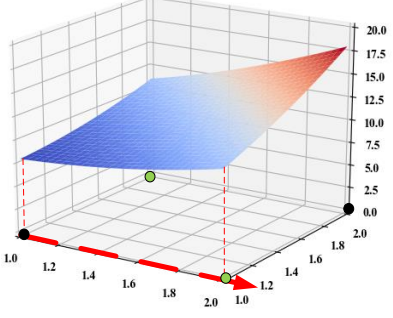
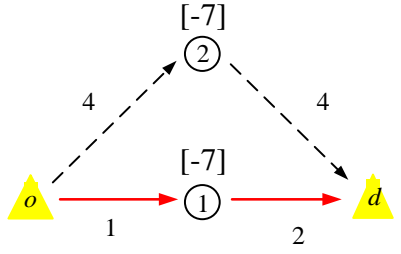
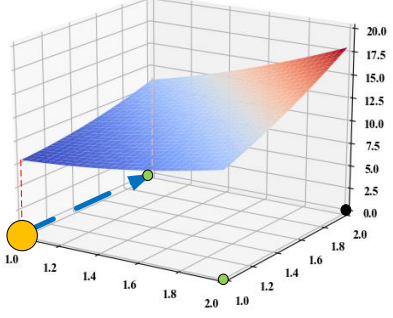
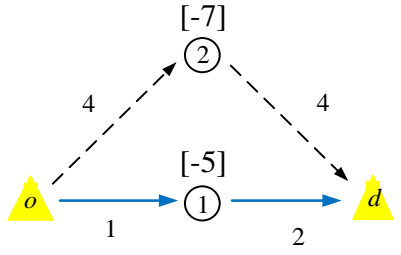
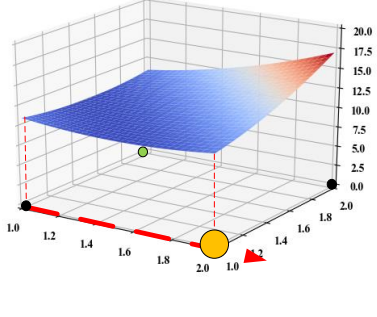
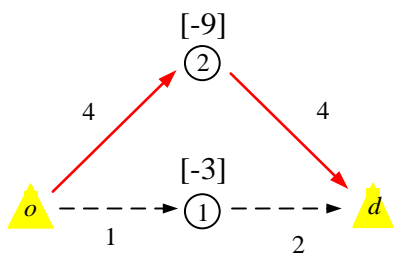


- 1 与拉格朗日松弛法相比，ADMM 的对偶问题具有附加的二次惩罚项，可以区分同源子问
- 2 题的组合成本，以减轻问题的对称性。具体而言，如等式 (20) 所示，每辆车服务顾客 p 的价
- 3 钱不仅取决于拉格朗日乘子 λ_p ，还取决于惩罚参数 ρ 。也就是说，如果车辆已经由其他车辆服
- 4 务，则会有一个阻力尽量避免当前车辆去服务该乘客。有趣的是，当 ρ 的值被设置为无穷大时，
- 5 就相当于一个与基于原问题解的启发式算法：车辆将找到完全可行的解决方案。详细的求解过

程列于表 6 中。可以看出，在 ADMM 的第二次迭代中得到了最优解，与拉格朗日松弛算法相比，它具有更快的收敛性。

Table 6

Detailed optimization process of ADMM approach.

Details of computation	Graph of sub-problems	Graphical Solution
<p> ● Optimal solution of the dual problem at each iteration ● Feasible solution ● Infeasible solution --> Search direction V_1 --> Search direction V_2 → Service price → Path of Vehicle 1 → Path of Vehicle 2 </p>		
Iteration 0: $\lambda_1^0 = \lambda_2^0 = 6$ $\rho = 2$ Step 1: $\pi_{1,1}^0 = -\lambda_1^0 + \frac{\rho}{2} = 7$ $\pi_{2,1}^0 = -\lambda_2^0 + \frac{\rho}{2} = 7$ The shortest path: [o-1-d] $L'_1 = 1 + 2 - \pi_{1,1}^0 = -4$		
Step 2: $\pi_{1,2}^0 = -\lambda_1^0 - \frac{\rho}{2} = 5,$ $\pi_{2,2}^0 = -\lambda_2^0 + \frac{\rho}{2} = 7$ The shortest path: [o-1-d] $L'_2 = 1 + 2 - \pi_{1,2}^0 = -2$ $L' = 10$		
Iteration 1: $\lambda_1^1 = 4, \lambda_2^1 = 8$ $\rho = 2$ Step 1: $\pi_{1,1}^1 = -\lambda_1^1 - \frac{\rho}{2} = 3,$ $\pi_{2,1}^1 = \lambda_2^1 + \frac{\rho}{2} = 9$ The shortest path: [o-2-d] $L'_1 = 4 + 4 - \pi_{2,1}^1 = -1$		

Step 2:

$$\pi_{1,2}^1 = -\lambda_1^2 + \frac{\rho}{2} = 5,$$

$$\pi_{2,2}^1 = -\lambda_2^2 - \frac{\rho}{2} = 7$$

The shortest path: [o-1-d]

$$L'_2 = 1 + 2 - \pi_{1,2}^1 = -2$$

$$L' = 11$$

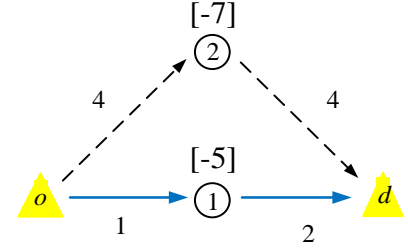
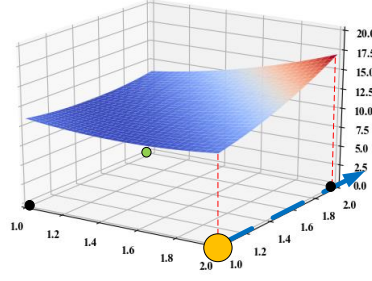
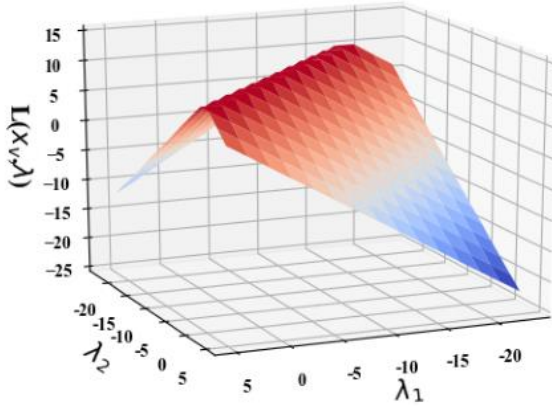
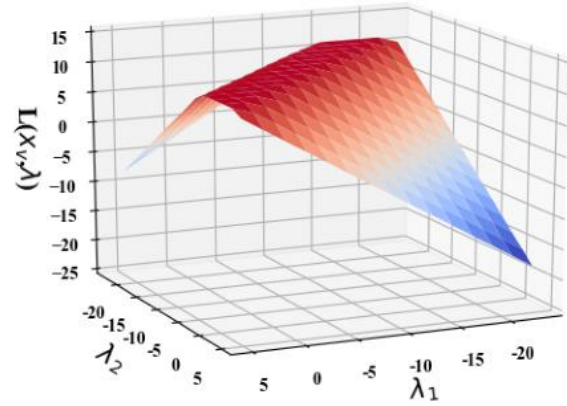


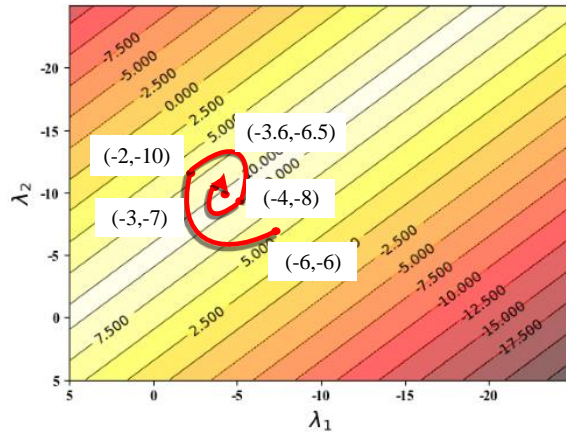
图 5 进一步显示了拉格朗日对偶问题和增广拉格朗日对偶问题的函数和等高线。拉格朗日对偶函数的最优区域会一条较窄的边（图 5 (a)）；在增广拉格朗日对偶问题（见图 5 (b)）中，在二次惩罚使得，表面更光滑，并且最优搜索区域也被扩大了。收敛轨迹在图 5 (c) 和 5 (d) 所示的等高线图画出。可以看出，即使在乘子变化很小时，拉格朗日对偶问题的解也会穿过最优区域并来回振荡，。图 5 (d) 所示的 ADMM 求解过程有更可靠的收敛性，乘子的振荡问题被大大地减少，并且以较少的迭代次数达到最优解。



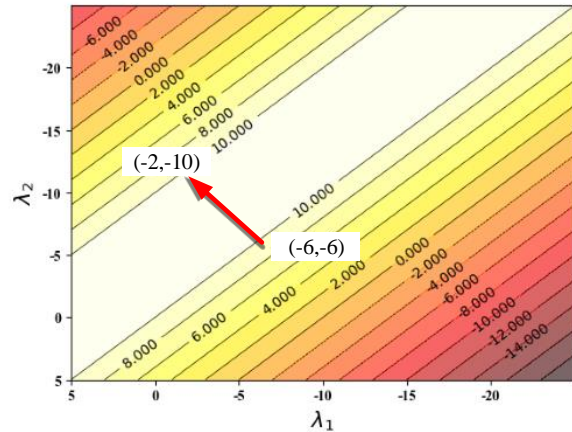
(a) surface of Lagrangian problem



(b) surface of augmented Lagrangian problem



(c) contour and solution process of LR



(d) contour and solution process of ADMM

Fig. 4. The surface, contour and empirical convergence process for dual problem of Lagrangian and ADMM.

5. 讨论

5.1 收敛性讨论

ADMM 在凸问题中的收敛性已在很多研究中被讨论过 (Gabay, 1983; Eckstein 和 Bertsekas, 1992), Boyd 等人 (2011) 给出了一个通用的理论结果, 证明了残差和目标收敛性可以在以下两个假设下得到证明。

假设 1: f and $g: R^n \rightarrow R \cup \{+\infty\}$ are closed, proper, and convex.

假设 2: The unaugmented (original) Lagrangian L has a saddle point.

既有的研究认识到, 当 ADMM 应用于非凸问题时, 并不能保证其收敛性。在我们的研究中, 由于 VRP 问题是整数规划, 所以这个问题显然是非凸的。换句话说, 虽然我们可以通过动态规划算法在每个步骤中取得精确的最优解, 但是依然不能保证 ADMM 的收敛性, 这意味着取不同的参数值可能导致不同的局部最优解 (而且通常不是全局最优解)。因此, 在非凸问题中, ADMM 被认为是一种基于 primal and dual 求解框架或改进型次梯度方法的启发式算法。为了解决这个问题, 最近的研究 (Gade 等人, 2016; 和 Boland 等人, 2018) 提出了 Progressive Hedging Algorithm (PHA) 下界的有效方法。虽然这些研究是针对随机多阶段混合整数规划的, 但它们为我们进一步研究运输优化问题也提供了重要的理论参考。在本研究中, 我们将 ADMM 迭代过程中得到的对偶价格带入到拉格朗日问题中, 来计算目标函数的下界。这样, 可以在 ADMM 的每次迭代中同时得到上界和下界。

总之, ADMM 本质上不是解决非凸问题的精确算法。未来的研究可以进一步采用 Branch and Price 方法进一步减少 GAP。另一方面, 我们提出的方法的主要优点可以突出如下: (a) 对偶乘子可以提供全系统的价格信息; (b) 二次项可以减少对偶问题的解求解过程中的, 以更好地达到原始可行性。 (c) 利用拉个朗日方法可以算出问题的下界, 以使得 ADMM 与其他启发式方法相比, 可以更好衡量解的质量。

5.2 ADMM 的拓展应用

本文提出的基于 ADMM 的框架还可以应用于具有更多类别的交通优化模型, 这些约束可以分为以下几类。

(1) Consensus constraints

Consensus 的通常指不同个体或决策者的决策。具体而言, 在我们的 VRPTW 模型中, 需求约束是车辆间的 Consensus 约束。这种约束在其他交通问题中也有广泛的应用, 例如, 铁路时刻表问题中的车头时距约束和交通分配问题中的道路通行能力约束。

(2) Consistency constraints

Consistency 是指在不同的决策层或空间和时间层之间应该确保一致性。在运输系统优化中, 基于 ADMM 的求解框架可以坚实的理论基础上提供可靠的算法, 用于迭代地优化不同层中的决策变量。例如, 在列车运输组织优化问题中, 使用跨层的一致性约束来整合三种类型的决策: (a) 旅客选择行为, (b) 运输计划和时间表以及 (c) 固定设备和移动设备的能力利用。

(3) Non-anticipative constraints

在随机规划中，在第一静态阶段和第二动态阶段之间需要满足 Non-anticipative 约束 (Boland 等人, 2018)。考虑一个基于场景的随机规划问题，它可以被重新表示为对应的确定形式 (24)，

$$\min \sum_{s \in S} p_s (cx_s + q_s y_s) \quad (24)$$

subject to, $x_s = z, \forall (x_s, y_s) \in K_s, s \in S$

其中 S 是场景集，约束 $x_s = z$ 是 Non-anticipative 约束。具有非预期约束松弛的增广拉格朗日函数为 (25)，

$$L_s^\rho = cx_s + q_s y_s + \omega_s (x_s - z) + (\rho/2) \|x_s - z\|_2^2 \quad (25)$$

可以利用 ADMM 将它分解为每个场景的独立子问题。

(4) Pair-based consensus constraint

在分部优化问题中，可以在相邻的 agents 之间建立 consensus constraint (Ling and Ribeiro, 2014; Tsianos et al., 2012)。考虑一个可分离的问题，它将 multi-agent 网络中的局部成本函数之和最小化为 (26)，

$$\min f(x) = \frac{1}{n} \sum_{i=1}^n f_i(x) \quad (26)$$

subject to, $x_i = z_{ij}, x_j = z_{ij}, \forall (i, j) \in \mathcal{A}$

where \mathcal{A} is a set of agent pairs, the variable z_{ij} is introduced to enforce the consensus between agents i and j . Using the definitions of the \mathcal{A} , the constraints of (26) can further be reformulated as (27).

subject to, $Ax + Bz = 0 \quad (27)$

原问题的增广拉格朗日函数可表示为等式 (28)

$$L(x, z, \lambda, \rho) = f(x) + g(z) + \lambda^T (Ax + Bz - c) + (\rho/2) \|Ax + Bz - c\|_2^2 \quad (28)$$

Then the problem can be solved by using the iterative principle of ADMM.

此外，我们用一个具体的例子来说明 ADMM 在一类具有一致性约束的 VRP 模型中的应用。首先，我们通过变量分解技术 (Fisher 等人, 1997; Niu 等人, 2018) 重新构造了之前提出的 VRP 模型。assignment 与 routing 的一致性约束，能更好地帮助我们利用拉格朗日乘子来协调多乘客多车辆的分配任务。具体地，我们引入一个附加的分配决策变量 y_p^v ，如果顾客 p 由车辆 v 服务，则 $y_p^v = 1$ ，否则 $y_p^v = 0$ ，并且用约束(29)、(30)和 (31)来替换约束 (5)。

Request assignment constraint:

$$\sum_{v \in V} y_p^v = 1 \quad \forall p \in P \quad (29)$$

Consistency constraint between assignment and routing:

$$\sum_{(i,j,t,s,w,w') \in \Psi_{p,v}} x_{i,j,t,s,w,w'}^v = y_p^v \quad \forall p \in P, v \in V \quad (30)$$

Binary definitional constraint:

$$y_p^v \in \{0, 1\} \quad (31)$$

约束 (29) 确保每个客户被访问一次。约束 (30) 是车辆路径决策和分配决策之间的一致性约束。约束 (31) 是 0-1 变量定义。在所提出的 ADMM 分解框架中，约束(29)被对偶到目标函数中，然后松弛模型可以被分解为两个子问题：(a)一个广义分配子问题来求解车辆分配，(b)

一个最短路径问题用于详细描述车辆路线的。这两个不同决策层次的子问题交替迭求解，如图 6 所示。这种架跨不同决策层的 ADMM 分解框也提供了将高效的启发式算法嵌入到求解过程的机会，例如 savings 算法 (Clarke and Wright, 1964) 和 sweep based 算法 (Gillet and Miller, 1974)。

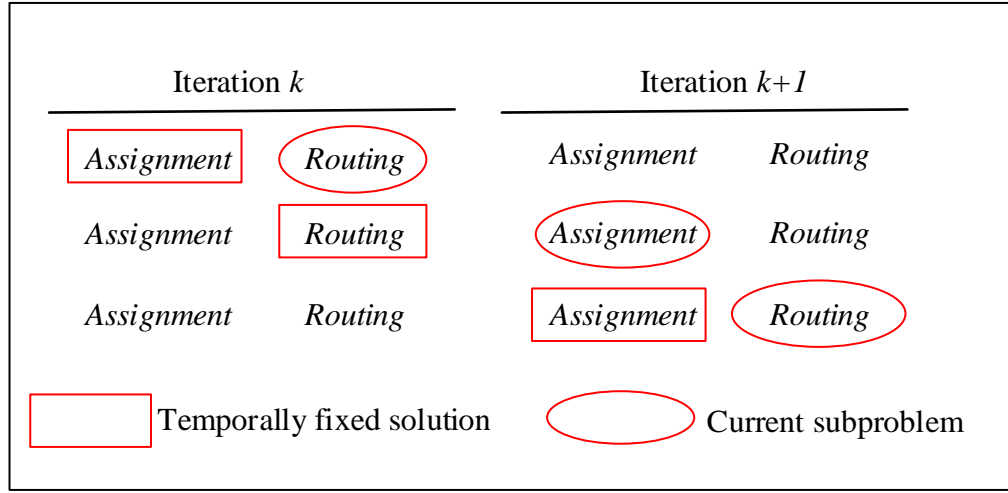


Fig. 6. Iterative pattern in the ADMM based framework for assignment and routing problem.

6. 算例

该算法基于 Python 平台实现，并在 1.70GHz CPU，8GB 内存的个人计算机上运行。在这一节，我们应用了两组数据集检验了我们提出的模型和算法，一组是经典的所罗门数据集 (Solomon, 1987)，另一组是京东物流提供的基于实际场景的算例。本文使用的算法程序和数据集可在以下链接获取 https://github.com/YaoYuBJTU/ADMM_Python。同时，这个算法框架也可以很容易地扩展并应用于其他问题测试和数据集。例如，对大规模周期时刻表问题的相关研究可以在以下链接查看 https://www.researchgate.net/profile/Yongxiang_Zhang7。

6.1 所罗门数据集

所罗门数据集是一个经典的测试集，其突出了影响算法效果的几个因素：包括地理数据、客户数量、有时间窗的客户所占百分比以及时间窗长度和时段等。所罗门数据集中包含三类不同的场景（分别是 C、R 和 RC），其中客户的地理位置在集合 C 中是聚合的，在集合 R 中是随机的，在集合 RC 中随机和聚合结构混合。

我们分别选取了数据集 C101、R101 和 RC101，并在不同的规模下来测试我们的求解框架。问题规模从小到大分别是 25、50 和 100 个客户。车队规模设为 25 辆，每辆车容量为 200。表 7 详细地给出了 ADMM 算法得到的解，并和已知最优解进行了比较。

对于小规模算例（25 个客户），ADMM 可以在 5 次迭代内获得很好的可行解。在 50 个客户和 100 个客户情况下，分别需要 25 次和 50 次左右的迭代。还可以看出，在 C101 数据集中，我们的解非常接近于最优解，三种问题规模下，我们和最优解之间的 gap 分别是 0.26%、0.22% 和 1.77%。在 R101 数据集中，相对 gap 在 5% 以内，在 RC101 中，相对 gap 在 10% 以内。计算时间随着问题的规模而增加，但在我们最大的算例也不超过 150 秒。总的来说，结果表明我们提出的 ADMM 框架可以有效地获得较好的解。

Table 7

Comparison between ADMM solutions and the best known solutions.

Dataset	Number of customers	Best known Solution	ADMM Solution	Computing time	Gap between best known and ADMM solution
R101	25	617.1	634.9	0.34s	2.80%
	50	1044.0	1083.6	3.52s	3.65%
	100	1637.7	1724.9	48.49s	5.01%
C101	25	191.3	191.8	1.55s	0.26%
	50	362.4	363.2	12.11s	0.22%
	100	827.3	842.2	106.28s	1.77%
RC101	25	461.1	470.7	3.08s	2.04%
	50	944	1012.5	21.76s	6.71%
	100	1619.8	1761.8	143.26s	8.06%

6.2 基于真实场景的案例

我们还用了京东物流发起的全球运筹大赛中提供的数据集对我们的算法框架进行了测试。着眼于智能物流和供应链，基于京东在北京 B2B 配送的真实场景和数据，竞赛的题目是配送车辆的路由和调度问题。更多细节和比赛的数据可以从网站上获得：<https://jdata.joy.com/en/>。

如图 6 所示，我们从竞赛数据集中选取了部分客户数据的子集作为我们的测试用例，配送中心为北京市的 100 个客户提供城市配送服务。此外，问题假设如下：

- (1) 每辆车 8:00 后开始工作，00:00 前返回配送中心。
 - (2) 每辆车每天的固定使用费是 200 元，运输费是 12 元/公里，等待费是 24 元/小时。并假设车辆的数量是足够的。
 - (3) 每辆车最多能装载 12000 公斤或 12m³的货物。
 - (4) 每个客户都有指定的时间窗。
 - (5) 将各路段的行驶时间作为常数给出。
 - (6) 每个客户的服务时间设置为 30 分钟。
- 目标函数是包括运输成本、等待成本和固定车辆成本在内的总成本最少。

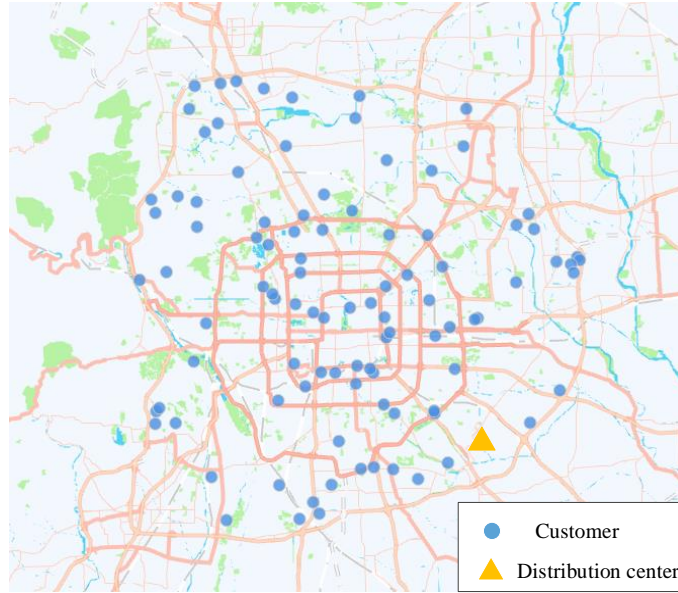


Fig. 6. The distribution center and customers in Beijing road network.

我们进行了 100 次迭代，每次迭代的计算时间约为 35s，参数 β 为 3， γ 为 0.25， ρ 为 2，初始拉格朗日乘子 λ_p 为 180。

下界和上界的收敛趋势如在图 7 所示。可以看出，在最初的 26 次迭代中，上界解得到显著改善，在第 73 次迭代中，上界解得到了稳定的较好的解，最终的 gap 为 11.79%。同样，在前 20 次迭代中，下界也不断地紧缩。

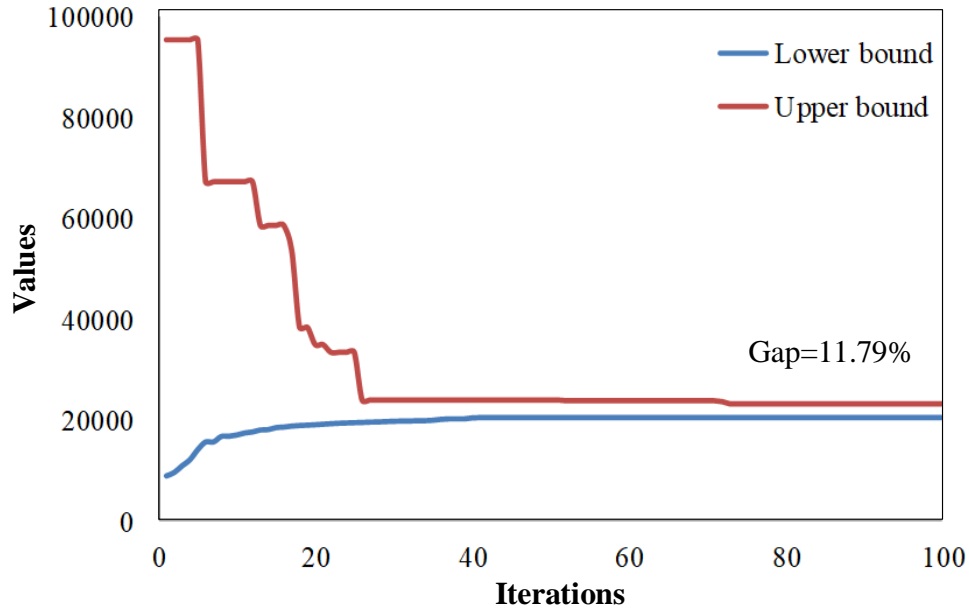


Fig. 7. Evolution curves of the upper bounds and lower bounds.

在最佳上界解中，需要 12 辆车来服务所有给定客户，总等待时间为 103 分钟，总运输距离为 1659.19 公里。优化结果的细节如表 8 所示，第二列中的粗体数字表示在该点车辆需要等待直到时间窗开始。具体的车辆线路如图 8 所示。

Table 8
The Details of the Optimized Solution.

Vehicle id	Vehicle routing solution	Waiting time (min)	Travel distance (km)
1	0-98-61-16-26-39-51- 13 -96-76-100	1	202.99
2	7-45-18-40-85-93-41-33-86-74-100	0	175.48
3	4-58-36-63-99-97- 2 -11-27-28-100	8	164.65
4	31-59-88-78-71-23-91-47-43-100	0	105.55
5	9-22-15-34-5-69-68-87-55-100	0	121.44
6	57-54-17-84-79-38-95-80-24-100	0	78.88
7	21-70-60-32-6- 82 -29-46-90-100	9	181.25
8	25-52-1-50-56-65-19-77-12-00	0	170.42
9	20-42-35-73- 53 -66-44-100	36	155.41
10	14- 37 -64-92- 94 -8-100	22	83.74
11	62-67- 48 -72- 3 -75-49-83-100	9	137.86
12	30-81- 10 -89-100	18	81.52
Total	-	103	1659.19
Average	-	8.58	138.27

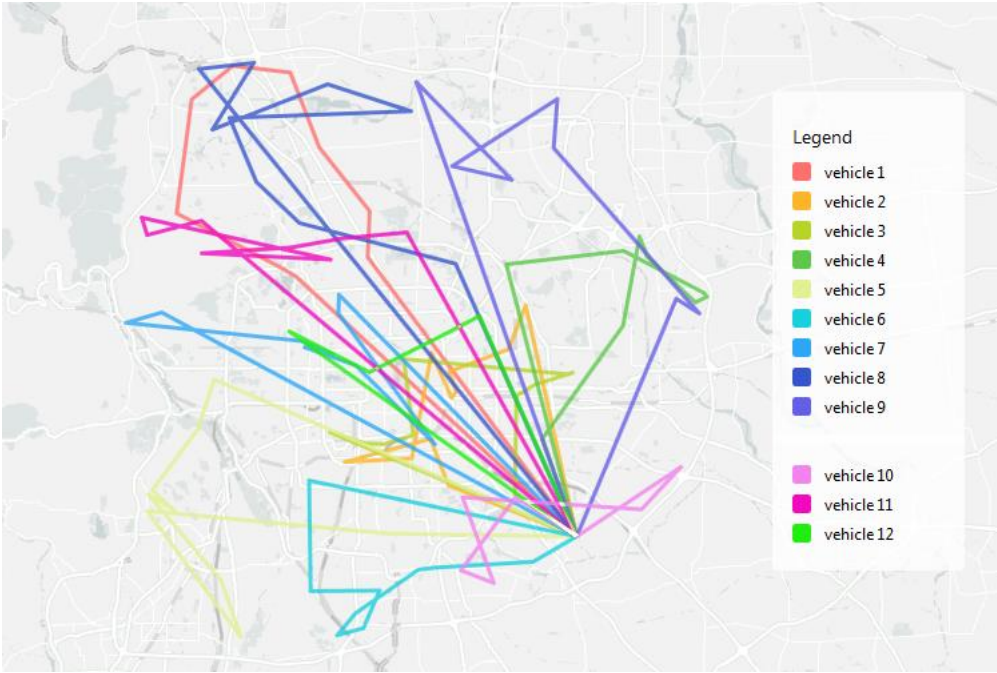


Fig. 8. The optimized routes and schedules for vehicles.

7. 结论

针对车辆路径问题中的耦合约束，我提出了一种基于 ADMM 的分解框架，可以同时迭代改进原解和对偶解的质量。并且，我们提出的算法框架可以更广泛地应用在其他带有耦合约束的运输问题中。本文中，对于 VRP 问题，我们首先利用状态-空间-时间网络构造了 VRP 的多维商品流模型，其中嵌入了时间窗和车辆容量约束，以简化对偶过程。然后，我们提出了一个基于 ADMM 可靠的分解框架，在该框架中，原始模型经过对偶和增广、分解和线性化三个步

骤分解为一系列最短路径问题，并通过动态规划算法求解。为了衡量解的质量，我们还通过求解拉格朗日松弛问题来计算下界。

为了检验所提出的 ADMM 框架的有效性，我们用两组数据集进行了测试，包括一组所罗门经典数据集和京东物流提供的基于真实场景数据。计算结果表明，基于 ADMM 的方法能够有效地获得质量较好的解，且有相对好的下界。

总的来说，我们提出的 ADMM 框架提供了一种相对简单可靠的算法结构，可以普遍用于解决许多其他交通问题。我们未来的研究将主要集中在以下两个方面。(1)将 ADMM 框架扩展到其他更复杂的运输问题，例如具有多层决策变量的问题。(2)在框架中引入分支技术，提高了解的质量，缩减上下界之间的相对差距。

Appendix A: Formulation of Lagrangian relaxation and augmented Lagrangian method

For the sake of completeness, in this appendix we review the basic framework of Lagrangian relaxation method and augmented Lagrangian method briefly.

Consider a linear equality-constrained problem expressed as (A1),

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & Ax - c = 0 \end{aligned} \quad (\text{A1})$$

where $x \in R^n$, $A \in R^{p \times n}$, and $c \in R^p$, and $f(x)$ are assumed to be convex. The Lagrangian relaxation form for problem (A1) can be written as Eq. (A2) and the Lagrangian dual problem is shown as Eq. (A3).

$$L(x, \lambda) = f(x) + \lambda^T (Ax - c) \quad (\text{A2})$$

$$\max_x \inf L(x, \lambda) \quad (\text{A3})$$

Explicitly, the dual problem can be solved through the following step (A4),

$$\begin{aligned} x^{k+1} &:= \operatorname{argmin}_x L(x, \lambda^k) \\ \lambda^{k+1} &:= \lambda^k + \alpha^{k+1} (Ax^{k+1} - c) \end{aligned} \quad (\text{A4})$$

where λ is the Lagrange multiplier for constraint $Ax - c = 0$ and α is the step size for updating λ . In the general case, after each iteration the equality constraint residual contributions $(Ax^k - c)$ are required to be collected in order to compute the Lagrangian multipliers of the next iteration. The main advantage of Lagrangian relaxation method is that each x -minimization step can be split into separate sub-problems and solved in parallel when $f(x)$ is separable.

By adding a quadratic penalty, the augmented Lagrangian function is shown as Eq. (A5),

$$L(x, \lambda, \rho) = f(x) + \lambda^T (Ax - c) + (\rho/2) \|Ax - c\|_2^2 \quad (\text{A5})$$

where ρ is the parameter of the quadratic penalty. The standard augmented Lagrangian method would minimize $L(x, \lambda, \rho)$ with respect to each sub-vector of x jointly.

参考文献

- Altinkemer, K., Gavish, B., 1991. Parallel savings based heuristics for the delivery problem. *Operations Research* 39(3), 456-469.
- Barnhart, C., Johnson, E.L., Nemhauser, G.L., Savelsbergh, M.W., Vance, P.H., 1998. Branch-and-price: Column generation for solving huge integer programs. *Operations research* 46(3), 316-329.
- Benders, J.F., 1962. Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik* 4(1), 238-252.
- Bertsekas, D.P., 1999. *Nonlinear programming*. Athena scientific Belmont.
- Boland, N., Christiansen, J., Dandurand, B., Eberhard, A., Linderoth, J., Luedtke, J., Oliveira, F., 2018. Combining Progressive Hedging with a Frank--Wolfe Method to Compute Lagrangian Dual Bounds in Stochastic Mixed-Integer Programming. *SIAM Journal on Optimization* 28(2), 1312-1336.
- Boland, N., Dethridge, J., Dumitrescu, I., 2006. Accelerated label setting algorithms for the elementary resource

- 1 constrained shortest path problem. *Operations Research Letters* 34(1), 58-68.
- 2 Boland, N., Hewitt, M., Marshall, L., Savelsbergh, M., 2017. The continuous-time service network design
- 3 problem. *Operations Research* 65(5), 1303-1321.
- 4 Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2011. Distributed optimization and statistical learning via
- 5 the alternating direction method of multipliers. *Foundations and Trends® in Machine learning* 3(1), 1-122.
- 6 Boyd, S., Vandenberghe, L., 2004. *Convex optimization*. Cambridge university press.
- 7 Bradley, S.P., Hax, A.C., Magnanti, T.L., 1977. Applied mathematical programming.
- 8 Christofides, N., Mingozzi, A., Toth, P., 1981. Exact algorithms for the vehicle routing problem, based on
- 9 spanning tree and shortest path relaxations. *Mathematical programming* 20(1), 255-282.
- 10 Clarke, G., Wright, J.W., 1964. Scheduling of vehicles from a central depot to a number of delivery points.
- 11 *Operations research* 12(4), 568-581.
- 12 Cordeau, J.-F., 2006. A branch-and-cut algorithm for the dial-a-ride problem. *Operations Research* 54(3),
- 13 573-586.
- 14 Dabia, S., Ropke, S., Van Woensel, T., De Kok, T., 2013. Branch and price for the time-dependent vehicle
- 15 routing problem with time windows. *Transportation science* 47(3), 380-396.
- 16 Dantzig, G.B., Wolfe, P., 1960. Decomposition principle for linear programs. *Operations research* 8(1),
- 17 101-111.
- 18 Desrochers, M., Desrosiers, J., Solomon, M., 1992. A new optimization algorithm for the vehicle routing
- 19 problem with time windows. *Operations research* 40(2), 342-354.
- 20 Eckstein, J., Bertsekas, D.P., 1992. On the Douglas—Rachford splitting method and the proximal point
- 21 algorithm for maximal monotone operators. *Mathematical Programming* 55(1-3), 293-318.
- 22 Eilon, S., Watson-Gandy, C., Christofides, N., de Neufville, R., 1974. Distribution management-mathematical
- 23 modelling and practical analysis. *IEEE Transactions on Systems, Man, and Cybernetics*(6), 589-589.
- 24 Fisher, M.L., Jörnsten, K.O., Madsen, O.B., 1997. Vehicle routing with time windows: Two optimization
- 25 algorithms. *Operations research* 45(3), 488-492.
- 26 Fisher, M.L., Jaikumar, R., 1981. A generalized assignment heuristic for vehicle routing. *Networks* 11(2),
- 27 109-124.
- 28 Ford Jr, L.R., Fulkerson, D.R., 1958. A suggested computation for maximal multi-commodity network flows.
- 29 *Management Science* 5(1), 97-101.
- 30 Fukasawa, R., Longo, H., Lysgaard, J., de Aragão, M.P., Reis, M., Uchoa, E., Werneck, R.F., 2006. Robust
- 31 branch-and-cut-and-price for the capacitated vehicle routing problem. *Mathematical programming* 106(3),
- 32 491-511.
- 33 Gade, D., Hackebeil, G., Ryan, S.M., Watson, J.-P., Wets, R.J.-B., Woodruff, D.L., 2016. Obtaining lower
- 34 bounds from the progressive hedging algorithm for stochastic mixed-integer programs. *Mathematical*
- 35 *Programming* 157(1), 47-67.
- 36 Ghilas, V., Demir, E., Van Woensel, T., 2016. An adaptive large neighborhood search heuristic for the pickup and
- 37 delivery problem with time windows and scheduled lines. *Computers & Operations Research* 72, 12-30.
- 38 Gillett, B.E., Miller, L.R., 1974. A heuristic algorithm for the vehicle-dispatch problem. *Operations research*
- 39 22(2), 340-349.
- 40 Glowinski, R., Marroco, A., 1975. Sur l'approximation, par éléments finis d'ordre un, et la résolution, par
- 41 pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. *Revue française d'automatique,*
- 42 *informatique, recherche opérationnelle. Analyse numérique* 9(R2), 41-76.
- 43 Goel, A., Gruhn, V., 2005. Large Neighborhood Search for rich VRP with multiple pickup and delivery
- 44 locations, *Proceedings of the 18th Mini EURO Conference on VNS (MEC-VNS)*.
- 45 Held, M., Karp, R.M., 1970. The traveling-salesman problem and minimum spanning trees. *Operations*
- 46 *Research* 18(6), 1138-1162.
- 47 Kallehauge, B., 2008. Formulations and exact algorithms for the vehicle routing problem with time windows.
- 48 *Computers & Operations Research* 35(7), 2307-2330.
- 49 Kohl, N., Madsen, O.B., 1997. An optimization algorithm for the vehicle routing problem with time windows
- 50 based on Lagrangian relaxation. *Operations research* 45(3), 395-406.
- 51 Laporte, G., Nobert, Y., Desrochers, M., 1985. Optimal routing under capacity and distance restrictions.
- 52 *Operations research* 33(5), 1050-1073.

- 1 Lasdon, L.S., 2002. *Optimization theory for large systems*. Courier Corporation.
- 2 Ling, Q., Ribeiro, A., 2014. Decentralized linearized alternating direction method of multipliers, *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, pp. 5447-5451.
- 3 Liu, L., Dessouky, M., 2018. Stochastic passenger train timetabling using a branch and bound approach. *Computers & Industrial Engineering*.
- 4 Mahmoudi, M., Zhou, X., 2016. Finding optimal solutions for vehicle routing problem with pickup and delivery services with time windows: A dynamic programming approach based on state-space-time network representations. *Transportation Research Part B: Methodological* 89, 19-42.
- 5 Nedic, A., Ozdaglar, A., 2009. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control* 54(1), 48-61.
- 6 Nemhauser, G.L., Savelsbergh, M.W.P., Sigismondi, G.C., 1991. *Constraint classification for mixed integer programming formulations*. Department of Mathematics and Computing Science, University of Technology.
- 7 Nishi, T., Ando, M., Konishi, M., 2005. Distributed route planning for multiple mobile robots using an augmented Lagrangian decomposition and coordination technique. *IEEE Transactions on Robotics* 21(6), 1191-1200.
- 8 Niu, H., Zhou, X., Tian, X., 2018. Coordinating assignment and routing decisions in transit vehicle schedules: A variable-splitting Lagrangian decomposition approach for solution symmetry breaking. *Transportation Research Part B: Methodological* 107, 70-101.
- 9 Reimann, M., Doerner, K., Hartl, R.F., 2004. D-ants: Savings based ants divide and conquer the vehicle routing problem. *Computers & Operations Research* 31(4), 563-591.
- 10 Renaud, J., Boctor, F.F., 2002. A sweep-based algorithm for the fleet size and mix vehicle routing problem. *European Journal of Operational Research* 140(3), 618-628.
- 11 Ruszczyński, A., 1989. An augmented Lagrangian decomposition method for block diagonal linear programming problems. *Operations Research Letters* 8(5), 287-294.
- 12 Savelsbergh, M., Van Woensel, T., 2016. 50th anniversary invited article—city logistics: Challenges and opportunities. *Transportation Science* 50(2), 579-590.
- 13 Solomon, M.M., 1987. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations research* 35(2), 254-265.
- 14 Taillard, É., 1993. Parallel iterative search methods for vehicle routing problems. *Networks* 23(8), 661-673.
- 15 Tong, L.C., Zhou, L., Liu, J., Zhou, X., 2017. Customized bus service design for jointly optimizing passenger-to-vehicle assignment and vehicle routing. *Transportation Research Part C: Emerging Technologies* 85, 451-475.
- 16 Tsianos, K.I., Lawlor, S., Rabbat, M.G., 2012. Consensus-based distributed optimization: Practical issues and applications in large-scale machine learning, *Communication, Control, and Computing (Allerton), 2012 50th Annual Allerton Conference on*. IEEE, pp. 1543-1550.
- 17 Wolsey, L., Rinaldi, G., 1993. Integer programming and combinatorial optimization.
- 18 Zhou, X., Tong, L., Mahmoudi, M., Zhuge, L., Yao, Y., Zhang, Y., Shang, P., Liu, J., Shi, T., 2018. Open-source VRPLite Package for Vehicle Routing with Pickup and Delivery: A Path Finding Engine for Scheduled Transportation Systems. *Urban Rail Transit* 4(2), 68-85.