·计算机与自动化技术·

引用格式:刘逗逗,王文发,许淳. 带约束k-means结合贪心法求解多车型软时间窗集油车辆调度问题[J]. 延安大学学报(自然科学版),2026,45(3):113-119. [LIU D D,WANG W F,XU C. Constrained k-means combined with greedy algorithm for solving the multi-vehicle type soft time window oil collection vehicle routing problem[J]. Journal of Yan'an University(Natural Science Edition),2026,45(3):113-119.] **DOI**:10.13876/J.cnki.ydnse.250149

带约束k-means结合贪心法求解多车型软时间窗集油车辆调度问题

刘逗逗[1],王文发[2*],许　淳[2]

(1. 延安大学西安创新学院 商学院,陕西 西安 710100;2. 延安大学 数学与计算机科学学院,陕西 延安 716000)

摘　要:针对车辆路径问题(VRP)的现有算法在解决采油厂带时间窗多车型多费用集油车辆调度时存在的不足,提出一种带约束k-means结合贪心法的分阶段多目标混合算法。综合考虑构成集油运输费用的各项指标,建立混合整数规划模型;应用贪心法结合车辆载重约束将油井集进行划分,得到聚类的初始子簇和集油所需车辆数;定义多因素加权组合距离,应用带约束k-means算法,通过"约束优先+费用最小化"的双目标迭代,将油井集划分为子簇;在油井子簇内,采用贪心法优化车辆集油的行驶路径;通过簇划分、路径优化及全局调优的两个阶段和三层策略,在满足约束的前提下,使集油车辆数和总费用最小化。该算法在某采油厂集油业务中的应用表明,日集油井次下降约18.28%,显著提高了生产效率和经济效益。

关键词:聚类分析;k-means算法;贪心法;车辆路径问题(VRP);整数规划模型

中图分类号:TP301　　**文献标志码**:A　　**文章编号**:1004-602X(2026)03-0113-07

采油厂集油业务是将采油厂各分散油井产出的原油通过各种方式汇集并输送到集输站(如集油点、联合站和接转站等)的过程,是石油企业油气集输的核心环节。高原山区油田的油井分散在沟谷山梁间,集油主要采用车辆运输方式,其业务技术核心是集油车辆的调度与路径优化。针对车辆调度与路径优化问题,国内外学者已提出了多种改进策略和算法[1-8]。聚类作为无监督学习的核心算法之一,以其基于相似度/距离度量将数据集样本进行划分的本质,在很多领域被广泛研究和应用[9]。k-means算法采用"聚类分组、路径优化"的两阶段策略,成为求解VRP问题的重要方法[10-14]。其中,带约束k-means算法通过降低问题规模、优化初始解,在满足约束条件的前提下,实现所有车辆的总行驶成本最小化[15-18]。采油厂集油车辆的调度与路径优化问题,由于构成运输成本的因素非常多,现有关于VRP问题的算法不完全适用该问题的求解,同时,现有算法也未考虑不同载重行驶费用的变化,且完成集油任务需要的车辆数大多给出预估值,未精准确定。本文针对多车型多费用软时间窗不确定车辆数非满载集油车辆的调度和路径优化问题,提出一种带约束k-means结合贪心法的分阶段多目标混合算法。该方法综合考虑构成集油运输费用的各项指标,建立多目标决策模型,通过"约束优

收稿日期:2025-12-13

基金项目:国家自然科学基金项目(72461032);陕西省自然科学基础研究计划项目(2023-JC-QN-0744)

作者简介:刘逗逗(1990—),女,讲师,主要从事算法设计分析与软件技术研究。

*通信作者　E-mail:yadxmyy@163.com

先+费用最小化”的双目标迭代,在满足约束的前提下,通过簇划分、路径优化、全局调优的两阶段三层策略,使集油车辆数和总费用最小化。

1 问题描述

采油厂车辆运输集油问题描述如下:采油厂拥有若干口油井、一个集油点和一个车场,车场有不同车型的车辆若干,当天部分油井需要车辆运输方式集油。设计完成当天集油任务的车辆调度方案,确定各车型发车数及各车辆集油的行驶路线,使总成本最少。

1.1 涉及因素

1)集油点、车场和所有油井及其之间的道路与长度构成的网;2)车型及其数量、各车型关于行驶长度及载重的油耗率;3)车型车辆的固定费用和行驶费用;4)油井储油罐容量;5)当天需要集油油井储油罐的存油量。

1.2 问题假定

1)单车场、单集油点;2)所有车型车辆的运载能力能够保证当天集油任务的完成;3)每个油井储油罐,一定存在一种车型,其最大载重不小于该油井储油罐的存油量。

1.3 问题约束

1) 每个油井的集油任务都有完成时间限制,且由一辆车完成。

2) 每辆车完成集油任务总时间:行驶时间+油井集油时间+集油点卸油时间。

3) 集油车辆的行驶路线均为:从车场出发到分散分布的油井集油,然后到集油点卸油,最后返回车场。

2 集油车辆调度与路径优化问题的数学模型

2.1 符号定义

表1 符号定义说明

符号	定义说明
LT	$LT=\{1,2,\cdots,M\}$为车场车辆车型集合
NV	$NV=\{u^m \mid m=1,2,\cdots,M\}$为各种车型车辆数
NU	$NU=\{1,2,\cdots,u^m \mid m=1,2,\cdots,M\}$为各种车型的车辆集合
TL	$TL=\{q^m \mid m=1,2,\cdots,M\}$为各种车型单车载重上限,其中$M$车型载重最大
SS	$SS=\{0,1,\cdots,N-1,N\}$为当天需要集油的油井集合、车场和集油点,其中,0表示车场,N表示集油点
OG	$OG=\{g_i \mid i=0,1,\cdots,N-1\}$为当天需要集油油井的存油量,$g_0=0$
LGC	$LGC=\left\{\Delta_i^m=\begin{cases}1, g_i \leqslant q^m,\\ 0, \text{否则},\end{cases} \mid m=1,2,\cdots,M, i=1,2,\cdots,N-1\right\}$为集油车辆的车型$m$与油井$i$集油任务的相容系数
SLG	$SLG=\left\{\Delta_i=\sum_{m=1}^{M}\Delta_i^m \mid i=1,\cdots,N-1\right\}$为与油井$i$集油相容的所有车型之和
LD	$LD=\{w_{iu}^m \mid i=0,1,\cdots,N-1;\ m=1,\cdots,M;\ u=1,\cdots,u^m;\ w_{0u}^m=0\}$为第$u$辆$m$型车按油井编号顺序完成油井$i$的集油决策后的总载重
OD	$OD=\{d_{ij} \mid i,j=0,1,\cdots,N\}$为油井、车场、集油点之间的距离,如果两个对象$i$和$j$之间不连通,$d_{ij}=inf$(充分大的数);$d_{ii}=0(i=0,1,\cdots,N)$
FC	$FC=\{c_0^m \mid m=1,2,\cdots,M\}$为各种车型车辆的固定费用
UC	$UC=\{c_1^m(w) \mid m=1,2,\cdots,M\}$为各种车型车辆载重$w$时的单位距离行驶费用
TC	$TC=\{c_{ij}^m(w) \mid m=1,2,\cdots,M;\ i,j=1,\cdots,N-1\}$为各种车型车辆载重$w$时从油井$i$到油井$j$的行驶费用
PTC	$PTC=\{c_{0j}^m=c_0^m+c_1^m(0)\times d_{0j} \mid m=1,2,\cdots,M;\ j=1,\cdots,N-1\}$为各种车型车辆从车场到油井$j$的费用,包括固定费用和空载行驶费用
$OPTC$	$OPTC=\{c_{N0}^m=c_1^m(0)\times d_{N0} \mid m=1,2,\cdots,M\}$为各种车型车辆从集油点到车场空载行驶费用
ST	$ST=\{t_i \mid i=0,1,\cdots,N\}$为车辆到达油井$i$的时间,$t_0=0$
GT	$GT=\{gt_i \mid i=0,1,\cdots,N\}$为车辆在油井$i$的集油时间,$gt_N$表示在集油点的卸油时间,$gt_0=0$
TT	$TT=\{tt_{ij} \mid i,j=0,1,\cdots,N\}$为车辆从油井$i$到油井$j$的行驶时间
ELT	$ELT=\{[ET_j,LT_j] \mid j=1,\cdots,N-1\}$为油井的集油时间窗

（续表1）

符号	定义说明
GTC	$GTC=\left\{T_j=\begin{cases}0, & ET_j\leqslant t_j\leqslant LT_j,\\ t_j-ET_j, & t_j\leqslant ET_j,\\ t_j-LT_j, & LT_j\leqslant t_j,\end{cases}\middle| j=1,\cdots,N-1\right\}$为油井$j$的集油时间相容系数
SC	$SC=\left\{sc_j\middle| j=0,1,\cdots,N\right\}$为早到油井$j$的单位时间惩罚成本
DC	$DC=\left\{dc_j\middle| j=0,1,\cdots,N\right\}$为晚到油井$j$的单位时间损失成本

2.2　决策变量定义

$$y_{iu}^{m}=\begin{cases}1, \text{油井}i\text{由}m\text{型车的第}u\text{辆集油},\\ 0, \text{否则},\end{cases}$$
$$i=1,2,\cdots,N-1,\ m=1,2,\cdots,M,$$
$$u=1,2,\cdots,u^{m};$$

$$x_{iju}^{m}=\begin{cases}1, \text{第}u\text{辆}m\text{型车油井}i\text{集油后到油井}j\text{集油},\\ 0, \text{否则},\end{cases}$$
$$i,j=1,2,\cdots,N-1,\ m=1,2,\cdots,M,$$
$$u=1,2,\cdots,u^{m}。$$

2.3　数学模型

集油车辆到达油井j的时间t_j通过公式(1)递推计算：

$$\begin{cases}t_0=0,\\ t_j=\sum\limits_{i=0}^{N}\sum\limits_{m=1}^{M}\sum\limits_{u=1}^{u^m}(t_i+tt_{ij}+gt_i)\times x_{iju}^{m},\\ \quad j=1,2,\cdots,N-1。\end{cases}\tag{1}$$

集油车辆到达油井$j(j=1,2,\cdots,N-1)$开始集油时间t_j，需满足油井j的集油时间窗约束，即$ET_j\leqslant t_j\leqslant LT_j$。否则，集油时间早于$ET_i$，存油量不足，导致集油成本增加，产生惩罚成本；集油时间晚于LT_i，可能出现油井储油罐溢出事故，造成原油流失浪费和环境污染，产生损失成本。

第u辆m型车按油井编号顺序完成油井i的集油决策后的总载重通过式(2)计算：

$$\begin{cases}w_{0u}^{m}=0,\\ w_{iu}^{m}=\sum\limits_{j=1}^{i}\Delta_j^{m}\times g_j\times y_{ju}^{m}, i=1,2,\cdots,N-1,\\ m=1,\cdots,M,\ u=1,\cdots,u^{m}。\end{cases}\tag{2}$$

综上所述，带时间窗的多车型多费用不确定车辆数非满载集油问题的数学模型如下：

$$\min Z=\sum_{i=0}^{N}\sum_{j=0}^{N}\sum_{m=1}^{M}\sum_{u=1}^{u^m}c_{ij}^{m}\left(w_{iu}^{m}\right)\times x_{iju}^{m}-\sum_{j=1}^{N-1}T_j\times sc_j+\sum_{j=1}^{N-1}T_j\times dc_j;\tag{3}$$

$$\text{s.t.}\sum_{m=1}^{M}\sum_{u=1}^{u^m}\Delta_i^{m}\times y_{iu}^{m}=1,\ i=1,2,\cdots,N-1;\tag{4}$$

$$\sum_{i=0}^{N-1}x_{iju}^{m}=y_{ju}^{m},\ j=1,2,\cdots,N-1,\ \forall m,u;\tag{5}$$

$$\sum_{j=0}^{N-1}x_{iju}^{m}=y_{iu}^{m},\ i=1,2,\cdots,N-1,\ \forall m,u;\tag{6}$$

$$\sum_{i=1}^{N-1}\Delta_i^{m}\times g_i\times y_{iu}^{m}=w_{N-1u}^{m}\leqslant q^{m},\ \forall m,u;\tag{7}$$

$$\sum_{j=1}^{N-1}x_{0ju}^{m}=1,\ \forall m,u;\tag{8}$$

$$\sum_{i=0}^{N-1}x_{ihu}^{m}=\sum_{j=0}^{N-1}x_{hju}^{m}=1,\ h=1,\cdots,N-1,\forall m,u;\tag{9}$$

$$\sum_{i=0}^{N-1}x_{iNu}^{m}=1,\ \forall m,u;\tag{10}$$

$$x_{N0u}^{m}=1,\ \forall m,u;\tag{11}$$

$$x_{iju}^{m}=1\text{或}0, i=0\text{时}, j\neq N, i=N\text{时}, j=0,$$
$$i,j=0,1,2,\cdots,N,\forall m,u;\tag{12}$$

$$y_{iu}^{m}=1\text{或}0, i=1,2,\cdots,N-1,\forall m,u。\tag{13}$$

目标函数(3)综合考虑集油车辆的固定费用、行驶费用、时间惩罚成本。约束(4)保证每一口油井只有一种相容车型的一辆车进行集油，约束(5)、(6)表示两决策变量之间的约束关系，约束(7)表示每辆车的集油总量w_{N-1u}^{m}不超过该车的载重上限，约束(8)、(9)、(10)、(11)表示每辆车从车场出发，完成集油任务，到集油点交付原油，最后回到车场，约束(12)、(13)为对相应决策变量的约束。

本文设计了带约束k-means结合贪心策略的混合求解算法，实现车辆调度和行驶路径的同步优化。

3　带约束k-means结合贪心策略的算法设计

应用贪心法结合车辆载重约束将油井集划分为互不相交的K个子集，得到聚类的K个子簇，每个子簇对应1辆载重相容集油车；定义“多因素加权组合距离”，应用改进的带约束k-means算法，通过“约束优先+费用最小化”的双目标迭代，将油井集划分为K个子簇；应用贪心法优化集油车辆在对应子簇的行驶路径；通过簇划分、簇中心更新和簇内路径优化的两个阶段三层策略，最小化集油总费用，得

到数学模型满足约束条件、满足精度的近似最优解。算法流程图见图1。

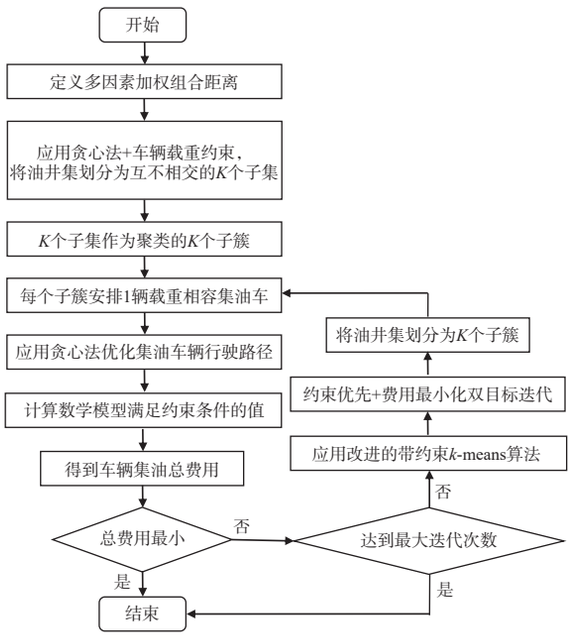


图1 带约束k-means结合贪心策略的算法流程图

3.1 基于贪心法的油井初始簇划分

当天需要集油油井$N-1$口，油井i的存油量为$g_i(i=1,\cdots,N-1)$，总存油量$G=\sum_{i=1}^{N-1}g_i$。车场车辆总数为$U=\sum_{m=1}^{M}u^m$，车辆总运力为$Q=\sum_{m=1}^{M}u^m\times q^m$。

1）定义$(N-1)\times(N-1)$阶对称矩阵$\boldsymbol{D}_{(N-1)\times(N-1)}=(d_{ij})_{(N-1)\times(N-1)}$表示$N-1$口油井相互之间的距离。

2）将每口油井$i(i=1,\cdots,N-1)$的信息定义为结点$tnode_i=(i,sg_i)$，构造以结点$tnode_i=(i,sg_i)$为根结点的树T_i，i为树T_i的编号，sg_i为树T_i上所有叶子结点存油量的和。

3）创建$N-1$棵树的森林$FT=\{T_1,T_2,\cdots,T_{N-1}\}$。

4）定义森林FT中两棵树T_n和T_m的距离$TD_{nm}=\min\left\{d_{ij}\middle|\left((i,sg_i)\in T_n\text{且是叶子结点}\right),\text{同时}\left((j,sg_j)\in T_m\text{且是叶子结点}\right)\right\}$。

5）选择当前森林FT中使距离TD_{nm}最小的两棵树T_h和$T_l(h<l)$。

6）如果$sg_h+sg_l\leqslant q^M$，将两棵树T_h和T_l合并为一棵树，同时从森林FT中删除树T_h和T_l；将合并的新树命名为T_h，加入森林FT中，令$sg_h=sg_h+sg_l$，则T_h的根结点为(h,sg_h)；然后继续步骤5）、步骤6）；否则，如果$sg_h+sg_l>q^M$，按照TD_{nm}值从小到大的次序，逐次选择对应的两颗树T_h和$T_l(h<l)$，继续步骤6）。

7）直到FT中任意两颗树T_h和T_l，都有$sg_h+sg_l>q^M$，结束森林FT中树的合并操作，最终森林FT中有K颗树。

8）当前森林FT中每一棵树的叶子结点的编号构成了油井集SS的一个子集，所有子集互不相交。因此，森林FT中的K颗树将油井集SS划分为K个簇$SS=\{SS_1,SS_2,\cdots,SS_K\}$，$K$就是$k$-means聚类的簇数，$SS=\{SS_1,SS_2,\cdots,SS_K\}$是油井集的初始簇划分，$|SS_k|=sn_k(k=1,2,\cdots,K)$。

9）对$SS=\{SS_1,SS_2,\cdots,SS_K\}$的每个簇指派一辆载重相容的车辆集油，所以，$K$也是完成当天集油任务需要调度的车辆数。

10）若$SS_k(k=1,2,\cdots,K)$是树T_h的叶子结点的编号构成的油井簇，树T_h的根结点为(h,sg_h)，sg_h即为簇SS_k的总存油量；选择满足$sg_h\leqslant q^m(m=1,2,\cdots,M)$的单车载重上限最小的车型车辆，完成簇$SS_k$的集油任务。

3.2 基于油井之间路径长度、油井存油量及集油时间窗的加权距离

综合考量路径距离、集油时间窗和存油量等集油相关因素，将油井i与油井j之间的路径距离、油井j的集油时间窗$[ET_j,LT_j]$和油井j存油量g_j的加权和作为聚类算法中从油井i到油井j的距离度量。

定义1 油井i到油井j的距离为

$$df_{ij}=\alpha d_{ij}+\beta g_j+\gamma(LT_j-ET_j)\times ET_j,\quad i,j=0,1,\cdots,N-1,\tag{14}$$

其中，α、β、γ为组合系数，满足$0\leqslant\alpha,\beta,\gamma\leqslant1$，$\alpha+\beta+\gamma=1$。

3.3 初始簇中心选择

对油井集的初始簇划分$SS=\{SS_1,SS_2,\cdots,SS_K\}$的每个簇$SS_k(k=1,2,\cdots,K)$，选择簇中与车场距离最小的油井作为簇的初始中心。

$SCP_k=\{j|df_{0j}\leqslant df_{0i}\text{同时}j\in SS_k\text{且}i\in SS_k\}$是簇$SS_k$中与车场距离最小的油井子集，任取$cp_k\in SCP_k$作为簇$SS_k$的初始中心。

3.4 带约束k-means聚类算法设计

带约束k-means算法求解模型(3)主要包括“约束满足优先的油井分配”、“概率性分散的簇中心更新”和“簇内路径优化”三个步骤，三步骤循环迭代直至满足结束条件。

3.4.1　约束优先的油井分配

针对油井集的簇划分$SS=\{SS_1,SS_2,\cdots,SS_K\}$，依次将油井$i(i=1,2,\cdots,N-1)$分两步分配到簇：

1)约束优先的候选簇筛选

分配油井i时，前面$i-1$口油井的分配已完成，簇SS_k的当前存油量为$sg_k(k=1,2,\cdots,K)$。

油井i的候选簇必须满足簇存油量与集油车辆的最大载重相容的约束条件，因此，油井i的候选簇集合$CS=\{SS_k|sg_k+g_i\leqslant q^M,k=1,2,\cdots,K\}$。

2)距离优先的簇分配

对于候选簇中的每个簇$SS_k\in CS$，定义油井i与簇SS_k的距离如下。

定义2　油井i与簇SS_k的距离为

$$sf_{ik}=\min\{d_{ij}|j\in SS_k\}。\tag{15}$$

根据定义2，构造候选簇CS关于油井i的子集$CS_i=\{SS_j|sf_{ij}\leqslant sf_{il}\ ;\ \forall SS_l\in CS\}$；将油井$i$分配到候选簇子集$CS_i$的任一簇$SS_r=SS_r\cup\{i\}$，同时更新簇$SS_r$存油量$sg_r=sg_r+g_i$，直到所有油井都分配到$K$个簇为止。

3.4.2　概率性分散的簇中心更新

1)从油井集SS中随机选择一口油井作为第一个簇的聚类中心，记为cc_1。

2)设已选定了油井集SS的前h个簇中心$CC_h=\{cc_1,cc_2,\cdots,cc_h\}$。

3)对油井集SS中每个油井$i\notin CC_h$，计算油井i到每个已选簇中心的距离并取最小值，记为$dist_{ih}=\min\{d_{ij}|j\in CC_h\}$。

4)对油井集SS中所有油井$i\notin CC_h$，计算$dist_{ih}$的平均值$ad_h=\left(\sum_{i\notin CC_h}dist_{ih}\right)$。

5)选择下一个簇中心cc_{h+1}时，选中油井集SS中油井i的概率为

$$p_{ih+1}=\begin{cases}(dist_{ih})^2\times ad_h, & i\notin CC_h,\\ 0, & i\in CC_h。\end{cases}$$

按照油井i的概率p_{ih+1}，选择下一个聚类中心$cc_{h+1}\in SS-CC_h$；距离已选定簇中心越远的油井被选中的概率越大，保证聚类簇中心的分散性。

6)重复步骤3)~5)，直到K个簇中心都选定为止。

3.4.3　基于贪心策略的簇内路径优化

应用贪心法对集油车辆在簇SS_k内的行驶路径进行优化，即求“车场→簇SS_k内集油路径→集油点→车场”的最短路径，使簇SS_k的集油费用最小。

1)SU_k表示已集油油井，初始状态$SU_k=\varnothing$，$SS_k=ns_k$。

2)车辆从车场出发，首先选择SS_k中与车场距离最小的油井k_1集油，油井k_1满足$d_{0k_1}\leqslant d_{0j}(j\in SS_k)$；同时，$SS_k=SS_k-\{k_1\}$，$SU_k=SU_k\cup\{k_1\}$。

3)设SS_k中已有i口油井完成集油，然后选择SS_k中未集油且与油井k_i距离最小的油井k_{i+1}集油，油井k_{i+1}满足$d_{k_jk_{i+1}}\leqslant d_{k_ij}(j\in SS_k\ \text{and}\ j\notin SU_k)$；同时，$SS_k=SS_k-\{k_{i+1}\}$，$SU_k=SU_k\cup\{k_{i+1}\}$。

4)重复步骤3)，直到$SS_k=\varnothing$为止。

5)车辆集油路径：车场$\rightarrow k_1\rightarrow k_2\rightarrow\cdots\rightarrow k_{ns_k}\rightarrow$集油点$\rightarrow$车场。

3.4.4　带约束k-means聚类算法迭代过程与收敛判断

输入：精度ε，最大迭代次数num。

输出：集油车辆调度方案，车辆集油的行驶路径，总费用。

1)应用贪心法将待集油油井集SS初始划分为K个子簇。

2)在每个子簇中，选择与车场距离最小油井作为簇的初始质心，得到初始质心集$CC_K\subseteq SS$。

3)设迭代次数$nc=1$。

4)依据约束优先的油井分配策略，将油井集SS重新划分为K个子簇。

5)采用分散优先策略，按概率重新选择簇中心。

6)应用贪心法优化簇内车辆集油的路径。

7)计算簇内集油费用。

8)计算本次车辆调度及路径优化的集油总费用$Z^{(nc)}$。

9)若$Z^{(nc)}-Z^{(nc-1)}<\varepsilon$，转11)。

10)$nc=nc+1$，如果$nc<num$，转4)，继续迭代；否则转11)。

11)算法结束。

4　结果与分析

某采油厂地处黄土高原山区，有3个采油队，每个采油队有1个车场、1个原油中转联合站。分属3个采油队的12 000多口油井分散在沟谷山梁间，因此，主要采用车辆运输方式集油。

4.1　数据集

数据集来自3个采油厂共12 000多口油井某个月平均每天的集油数据，集油数据集的描述如表2所示。

表2 集油数据集

数据集名称	数据对象个数	属性个数	类别个数
采油一大队	5 330	7	2
采油二大队	4 120	7	2
采油三大队	3 240	7	2

4.2 聚类结果评价指标

采用分簇准确率(AC)δ_{AC}、簇精度(PR)δ_{PR}召回率(RE)δ_{RE}三种评价指标,评估集油井聚类结果的有效性,三种评价指标分别定义如下:

$$\delta_{AC}=\frac{\sum_{i=1}^{K}a_i}{n} \tag{16}$$

$$\delta_{PR}=\frac{\sum_{i=1}^{K}\left[a_i/(a_i+b_i)\right]}{K}, \tag{17}$$

$$\delta_{RE}=\frac{\sum_{i=1}^{K}\left[a_i/(a_i+c_i)\right]}{K}, \tag{18}$$

其中,n表示数据集对象个数,a_i表示正确分到i子簇的油井数,b_i表示误分到i子簇的油井数,c_i表示本该分到i子簇但未分到的油井数,K是油井聚类子簇数。

4.3 聚类结果与分析

在集油数据集上运行传统的k-means算法,实验结果如表3所示。

表3 传统k-means算法的实验结果

数据集名称	δ_{AC}	δ_{PR}	δ_{RE}
采油一大队	0.771 6	0.816 3	0.886 5
采油二大队	0.753 4	0.836 0	0.879 6
采油三大队	0.812 5	0.842 6	0.857 3

在集油数据集上运行本文的带约束k-means结合贪心算法,实验结果如表4所示。

表4 带约束k-means结合贪心算法的实验结果

数据集名称	δ_{AC}	δ_{PR}	δ_{RE}
采油一大队	0.779 2	0.846 7	0.890 6
采油二大队	0.786 5	0.885 3	0.925 6
采油三大队	0.793 4	0.853 1	0.873 2

实验结果表明,3个大队的9个指标值中,三大队因为油井分布范围大、个别油井偏远、大部分油井产量较低,导致本文的带约束k-means对三大队油井的聚类结果不太理想,AC指标值略低于传统k-means算法,其他8个指标值均大于传统k-means算法,说明本文的带约束k-means结合贪心算法的聚类效果总体优于传统k-means算法。

4.4 实际应用结果

采油厂非集输油井多为低产、间歇出油井,以前的集油车辆调度及车辆行驶路线主要依赖历史经验人工确定,导致数据的实时性和合理性差、调度方案的效率和效益不高。在本文作者参与研发的《采油厂生产统计综合管理系统》中,将本文的研究成果应用于采油厂3个采油大队的集油生产业务,推行大调度体系,实行定人、定车、定井、定时、定量、定路线的"六定"模式,由车辆管理中心统一派单。分析汇总管理系统中近两年的油井集油数据和中转联合站收油数据,与应用前的数据进行比较,忽略油井产量波动、间歇采油实施、油井检修停产、自然灾害等因素影响,比较结果表明,采油厂日均集油井次由人工调度时的约930井次,降到了约760井次,日均集油井次下降约18.28%,集油车辆发车数从120车次/天,降到了85车次/天,下降约29.17%,显著降低车辆空驶率和燃油消耗,提高了生产效率和经济效益。

5 结论

本文以总费用最小化为目标,建立了带时间窗的多车型多费用不确定车辆数非满载集油问题的数学模型,设计了带约束k-means结合贪心策略的混合求解算法,应用于《采油厂生产统计综合管理系统》的集油生产业务中,显著提升了集油调度效率,降低了生产运营成本。

参考文献:

[1]轩苗苗. 基于改进遗传算法的冷链物流配送路径优化研究[D]. 淮南:安徽理工大学,2025.

[2]丘凯. 深度强化学习在车辆路径问题中的应用研究[D]. 广州:广东财经大学,2025.

[3]李二超,张智钊. 在线动态订单需求车辆路径规划[J]. 山东大学学报(工学版),2024,54(5):62-73.

[4]刘狄. 考虑碳排放和服务水平的多中心配送路径优化研究[D]. 重庆:重庆交通大学,2025.

[5]MOON I. K,LEE J-H,SEONG J. Vehicle routing problem with time windows considering overtime and outsourcing vehicles [J]. Expert Systems with Applications,2012,39(18):13202-13213.

[6]CETINKAYA C, KARAOGLAN S, GOKCEN H. Two-stage vehicle routing problem with arc time windows: A mixed integer programming formulation and a heuristic approach [J]. European Journal of Operational Research,

2013,230(3):539-550.
[7]GHANNADPOUR S F, NOORI S, TAVAKKOLI-MOGHADDAM R, et al. A multi objective dynamic vehicle routing problem with fuzzy time windows: Model, solution and application [J]. Applied Soft Computing, 2014, 14: 504-527.
[8]DHAHRI A, ZIDI K, GHEDIRA K. Variable neighborhood search based set covering ILP model for the vehicle routing problem with time windows [J]. Procedia Computer Science, 2014,V(29):844-854.
[9]张冬梅,李敏,徐大川,等.k-均值问题的理论与算法综述[J].中国科学:数学,2020,50(9): 1387-1404.
[10]赵雄,李琳.基于聚类的LNS算法求解异构VRP问题[J].计算机技术与发展,2023,33(9):98-104.
[11]朱珠,温召成,臧洁.电子商务快递末端配送模式优化研究[J].武汉理工大学学报(交通科学与工程版),2024,46(1):139-145.
[12]宋虹颖.考虑多车型的成品油主动配送规划方法研究[D].长春:吉林大学,2025.
[13]戴永寿,李韶光,李立刚等.基于改进k-means和遗传算法的油田特种车辆优化调度[J].计算机应用,2016,36(S1):86-89.
[14]张雪婷.基于K-Means++和遗传算法的提货车辆路径研究[D].北京:首都经济贸易大学,2017.
[15]梁学恒,杨家其,向子权.两阶段BSO-SA算法求解带单边软时间窗的多车型VRP问题[J].武汉理工大学学报(交通科学与工程版),2024,48(1):19-24.
[16]罗蒙,高超,王震.基于带约束谱聚类的启发式车辆路径规划算法优化方法[J].计算机应用,2025,45(5):1387-1394.
[17]蔡祥.带时间窗的同时取送货车辆路径问题优化模型及算法研究[D].南昌:南昌大学,2023.
[18]李敏,倪少权,周凌.基于订单邻域的成品油二次配送中带时间窗车辆路径规划问题[J].计算机集成制造系统,2015,21(8):2158-2169.

[责任编辑　毕　伟]

Constrained k-means combined with greedy algorithm for solving the multi-vehicle type soft time window oil collection vehicle routing problem

LIU Doudou[1], WANG Wenfa[2*], XU Chun[2]

(1.Business School, Xi'an Innovation College of Yan'an University, Xi'an 710100;
2.Department of Computer Science, Yan'an University, Yan'an 716000, China)

Abstract: Aiming at the deficiencies of existing Vehicle Routing Problem (VRP) algorithms in addressing the oil-gathering vehicle scheduling problem with time windows, multiple vehicle types and multi-dimensional costs in oil production plants, a phased multi-objective hybrid algorithm combining constrained k-means and the greedy algorithm is proposed. A mixed integer programming model is established by comprehensively considering all indicators that constitute the oil-gathering transportation cost. The greedy algorithm integrated with vehicle load constraints is applied to partition the set of oil wells to obtain the initial clustering sub-clusters and the number of vehicles required for oil gathering. A multi-factor weighted composite distance is defined and the constrained k-means algorithm is adopted to divide the oil well set into sub-clusters through dual-objective iteration following the principle of "constraint priority + cost minimization", the greedy algorithm is employed to optimize the driving routes of oil-gathering vehicles within each oil well sub-cluster, and based on the two-stage and three-layer strategy encompassing cluster partitioning, route optimization and global refinement, the number of oil-gathering vehicles and the total cost are minimized on the premise of satisfying all constraints. The application of this algorithm in the oil-gathering business of a certain oil production plant shows that the daily number of oil well visits has decreased by nearly 18.28%, which significantly improves production efficiency and economic benefits.

Key words: clustering analysis; k-means algorithm; greedy algorithm; VRP; integer programming model