

# 带有时间窗口的富网络配送问题研究及仿真

张贝金 周晓根 明洁 姚春龙 张贵军

(浙江工业大学信息工程学院 杭州 310023)

**摘要** 通过建立 GIS 富网络路网属性模型,并组合  $N$  阶最短近邻自适应聚类算法和遗传算法,来解决不确定车辆数目、较大规模网点和多层级交通网络的带时间窗口的联合配送问题。首先,为了解决传统带有时间窗口车辆线路调度模型中配送网点规模小(不超过 20 个网点)的问题,以及在建模时将各网点抽象为图的顶点的缺陷,建立基于实际道路数据的网络数据集,采用 GIS 技术精确计算各网点之间的距离,并建立距离 OD 矩阵;然后,为了降低对较大规模网点配送算法设计的复杂度,采用  $N$  阶最短近邻自适应算法确定聚类簇数,再通过聚类数划分配送网点。其次,为了确定配送车辆的种类、车辆数目以及时间窗口的限制,利用遗传算法对配送线路进行优化。最后,通过 2 个实例验证了所提方法的有效性。

**关键词** 富网络模型,聚类算法,遗传算法,OD 矩阵,路径规划

**中图分类号** TP274 **文献标识码** A

## Vehicle Routing Problem with Time Window and its Application Based on Rich Road Network Weights

ZHANG Bei-jin ZHOU Xiao-gen MING Jie YAO Chun-long ZHANG Gui-jun

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

**Abstract** To solve vehicle routing problem with indeterminate number of vehicles, large-scale outlets and multi-level transportation network, we established the road property mode based on rich network and GIS, mixed  $N$ -order nearest-neighbor adaptive clustering algorithm and genetic algorithm. Firstly, to solve the problem of small-scale outlets (less than 20 outlets) in traditional vehicle routing problem and the defect that abstracting the outlets into graph's vertices during model establishing, we established network dataset based on actual road data, utilized GIS to exactly calculate the distance between outlets and built OD matrix of distance. Secondly, to reduce the complexity of designing large-scale outlets optimization algorithm, this paper used  $N$ -order nearest-neighbor adaptive clustering algorithm to determine the number of clusters, then divided the distribution outlets by clusters. Subsequently, to determine the kind and number of distribution vehicles and restriction of time window, we used genetic algorithm to optimize the distribution path. Finally, two examples verified the effectiveness of the proposed method.

**Keywords** Rich network model, Clustering algorithm, Genetic algorithm, OD matrix, Vehicle routing

国际经济全球化趋势不断加强以及计算机技术的飞速发展,许多行业得到了兴起,物流配送就是其中一个并得到了多人的重视<sup>[1]</sup>。配送线路费用在整个物流系统中占据着较大的成本,也是物流系统的关键环节,它是伴随着市场孕育而成的市场化结果。随着客户对服务质量的不断提高,物流配送线路问题成为企业在今后市场激烈竞争中需要考虑的重要因素。

国内外科研工作者热衷于 VRPTW (vehicle routing problem with time windows) 问题的研究,主要因为它是物流配送和交通运输的核心问题,只有解决了配送车辆线路调度问题才能使配送更有效合理。1981 年,Lenstra 学者和 Kan 学者对 VRPTW 问题属于 NP 难组合优化问题这个命题进行

了证明<sup>[2]</sup>。1988 年, Solomon 对 VRPTW 问题做了较详细的综述<sup>[3]</sup>。VRPTW 问题的目标函数可以描述为车辆从仓库(单仓库)到各分散客户点之间的行驶总距离最短和线路总花费最小。线路设计原则是,每个客户点只能被一辆车访问,同时必须在规定的天窗之内。所有的线路的起点和终点必须是同一个配送中心,并且在每一条配送线路上的客户点的货运需求总量不能超过线路上的车辆的载重量。VRPTW 问题实用性较强,尤其在 3PL (third-party logistic) 中。对于上百个客户点的配送问题,通过可伸缩方法在限制的时间内能得到较好的结果。VRPTW 问题的解法丰富,比较常见的可以分为精确算法和启发式算法。采用精确算法解决 VRPTW 问题时,时间复杂度过高。近几年,虽然一些学者在研究

本文受国家自然科学基金(61075062),浙江省自然科学基金(LY13F030008),浙江省科技厅公益项目(2014C33088),浙江省重中之重学科开放基金(20120811),杭州市产学研合作资助项目(20131631E31),浙江省大学生“新苗计划”(2014R403090)资助。

张贝金(1990—),男,硕士,主要研究方向为优化调度、智能交通;周晓根(1987—),男,硕士,主要研究方向为智能优化;明洁(1990—),女,硕士,主要研究方向为优化调度、智能交通;姚春龙(1988—),男,硕士,主要研究方向为优化调度、智能交通;张贵军(1974—),男,博士,教授,主要研究方向为优化调度、智能交通, E-mail: zgj@zjut.edu.cn (通信作者)。

VRPTW 问题中用到了精确算法,但是启发式算法能在可行的时间复杂度之内优化问题,大多数学者还是着力研究启发式算法来解决 VRPTW 问题。Thangiah 以服务成本最小为目标,提出一种改进的遗传算法 Gideon<sup>[4]</sup>。Blanton 和 Hwang 对遗传算法中的交叉、变异算子做出了改进<sup>[5,6]</sup>。Ozdemir 提出一种基于有向无环图的进化算法<sup>[7]</sup>。Tan 采用了变长编码的混乱遗传算法(mGA)<sup>[8]</sup>。Baker 提出了一种隐含编码方式的遗传算法<sup>[9]</sup>。戚铭尧等人提出了一种同时考虑顾客的时间和空间邻近性的路径改进方法<sup>[10]</sup>。总的来说,VRPTW 问题比 VRP 问题更难。因此,采用启发式算法解决 VRPTW 问题更好,并且采用启发式算法来求解 VRPTW 问题已获得了一些效果不错的成果。但是,不管用哪种方法求解 VRPTW 问题都很少考虑到配送过程中的道路行车规则以及配送客户点数较大的情况,相应的仿真也没能贴近实际配送方案。

对于上面描述中存在的两方面不足,本文引入了 GIS 富网络模型和  $N$  阶近邻自适应分类算法。在配送线路设计中,一个首要难点就是对实际路网模型的抽象,城市街道纵横交错,道路等级不一而同,引入 GIS 富网络属性模型也就是为了解决实际道路抽象问题,GIS 具有较好的空间数据处理能力,可以有效地处理实际道路复杂性问题<sup>[11-13]</sup>。由于 VRPTW 问题是 NP 难问题,网点数的规模问题直接影响算法的运算效率,引入  $N$  阶近邻自适应算法<sup>[14]</sup>能较好地解决这个问题, $N$  阶近邻自适应算法能够将网点进行合理的划分,为配送算法的高效设计提供了保障。

## 1 问题的描述与建模

### 1.1 分类模型

本文所研究的配送问题处理方法可以描述为:建立 GIS 富网络模型,考虑到车辆在不同道路上的行车速度、行车规则,同时对全部客户点采用  $N$  阶近邻算法来确定聚类数量( $K$  值),并通过  $K$ -means 分类,将配送网点大而化小,再为每一类的配送建立调度模型。

从整体优化层面考虑,并结合聚类性质,将单个聚类形成的单簇视为概率分布中的单峰值,聚类中心也就是单个网点中心。同理,当聚类形成  $K$  个簇时,可以看成多峰概率分布,求解  $K$  值也就转化成求解多峰峰值。 $K$  个簇的概率密度函数为:

$$f(x) = \sum_{i=1}^K p_i f(x) \quad (1)$$

其中,

$$\sum_{i=1}^K p_i = 1$$

并且  $p_i > 0$ 。从式(1)中可知,概率密度函数由  $K$  个簇的概率密度叠加形成。 $K$  通常是未知的,这取决于样本数据。采用  $N$  阶近邻计算最短邻阶距离时,样本数据的稠密度会影响算法效率,稀疏数据计算量往往更大。

假设给定有  $N$  点的样本数据集  $S = \{x_i | x_i \in R^n, i=1, 2, \dots, N\}$ ,令  $d_1$  为样本点  $x_i$  的一阶最近邻接距离, $d_2$  为样本点  $x_i$  的二阶最近邻接距离,依此类推,可以得到样本点  $x_i$  的  $N-1$  阶最近邻接距离,从小到大的排列顺序有:

$$d_1 < d_2 < \dots < d_{N-1}, \forall i=1, 2, \dots, N-1 \quad (2)$$

进一步计算  $j$  阶近邻接距离的平局值  $\bar{d}_j$  和方差  $\bar{d}_j^2$ ,

得:

$$\bar{d}_j = \sum_i^N d_i^j / N, \forall j=1, 2, \dots, N-1 \quad (3)$$

$$\bar{d}_j^2 = \sum_i^N (d_i^j)^2 / N, \forall j=1, 2, \dots, N-1 \quad (4)$$

为了表示样本点的稀疏度,再计算出样本方差:

$$\sigma_{N,m}^2 = \frac{1}{m} \sum (\bar{d}_j^2 - (\bar{d}_j)^2), \forall j, m=1, 2, \dots, N-1 \quad (5)$$

引入样本点  $x_i$  的排列顺序  $y_i$  和对应的有相同概率密度函数  $f(\tilde{y}_i)$  的  $\tilde{y}_i$ ,并满足条件  $y_1 \leq y_2 \leq \dots \leq y_{N-1}, \tilde{y}_i \geq 0$ 。由次序统计的相关原理可以得到概率密度的  $K$  阶统计:

$$f_{N-1,K(y)} = \frac{(N-1)!}{[(K-1)! (N-K-1)!]} F(y)^{K-1} \times (1-F(y))^{N-K-1} f(y) \quad (6)$$

其中, $F(y)$  为累积概率分布函数,可表示:

$$F(y) = \int_0^y f(t) dt \quad (7)$$

式(3)和式(4)可以转化为:

$$\bar{d}_j \equiv \overline{y^j} = \int_0^\infty y^j f_{N-1,K(y)} dy \quad (8)$$

$$\bar{d}_j^2 \equiv \overline{y^{2j}} = \int_0^\infty y^{2j} f_{N-1,K(y)} dy \quad (9)$$

由次序学理论可以知道,已知给定样本数据集集聚为单簇时,计算方差为:

$$\sigma_{N,m}^2 = \alpha(N)(m+1)^2 + \beta(N)(m+1) \quad (10)$$

假定通过手段得到系数  $\alpha(N)$  和  $\beta(N)$ ,又可知给定  $m$  值,将式(8)和式(9)带入式(5),可以得到  $\sigma_{N,m}^2$  的值:

$$\sigma_{N,m}^2 = \frac{1}{m} \sum_{K=1}^m \left[ \int_0^\infty y^2 f_{N-1,K(y)} dy - \left( \int_0^\infty y f_{N-1,K(y)} dy \right)^2 \right] \quad (11)$$

把式(6)和式(7)都代入式(11)得:

$$\sigma_{N,m}^2 = \frac{1}{m} \sum_{K=1}^m \left\{ \frac{(N-1)!}{(K-1)! (N-K-1)!} \int_0^\infty y^2 \left( \int_0^y f(t) dt \right)^{K-1} (1 - \int_0^y f(t) dt)^{N-K-1} f(y) dy - \left[ \frac{(N-1)!}{(K-1)! (N-K-1)!} \times \int_0^\infty y \left( \int_0^y f(t) dt \right)^{K-1} (1 - \int_0^y f(t) dt)^{N-K-1} f(y) dy \right]^2 \right\} \quad (12)$$

从上式可以看出,要求得方差  $\sigma_{N,m}^2$ ,需要确定概率密度函数  $f(y)$ 。假设样本满足(0,1)均匀分布,则  $f(y)=1$ ,代入式(12)得:

$$\sigma_{N,m}^2 = \frac{1}{m} \sum_{K=1}^m \left\{ \frac{(N-1)!}{(K-1)! (N-K-1)!} \int_0^1 y^{K+1} (1-y)^{N-K-1} dy - \left[ \frac{(N-1)!}{(K-1)! (N-K-1)!} \int_0^1 y^{K+1} (1-y)^{N-K-1} dy \right]^2 \right\} \quad (13)$$

利用 Beta 函数化简式(13),得:

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx = \frac{(p-1)! (q-1)!}{(p+q-1)!} \quad (14)$$

$$\sigma_{N,m}^2 = \frac{1}{N^2 (N+1)} \left( \frac{N}{m} \sum_{K=1}^m K - \frac{1}{m} \sum_{K=1}^m K^2 \right) \quad (15)$$

式(15)进一步简化为:

$$\sigma_{N,m}^2 = -\frac{1}{3N^2 (N+1)} (m+1)^2 + \frac{3N+1}{6N^2 (N+1)} (m+1) \quad (16)$$

将式(16)和式(10)相比较:

$$\alpha(N) = -\frac{1}{3N^2(N+1)} \quad (17)$$

$$\beta(N) = \frac{3N+1}{6N^2(N+1)} \quad (18)$$

上述公式得出建立的样本数据分布满足(0,1)均匀分布,但是在其它分布的条件下,推导方式非常困难,故只能用数值分析计算做无限的接近期望值的近似解。

$$\left| \frac{\sigma_{N,m}^2}{m+2} + \frac{\sigma_{N,m-1}^2}{m} - 2 \frac{\sigma_{N,m}^2}{m+1} \right| < \epsilon \left| \frac{\sigma_{N,m}^2}{m+1} \right| \quad (19)$$

其中,敏感参数  $\epsilon \in (0,1)$ , 变量  $m \geq 2$ 。式(19)中,给定样本数据集中,其多簇情况可以看成是单簇情况的叠加。

将  $\sigma_{N,m}^2$  做  $\frac{\sigma_{N,m}^2}{(m+1)}$  处理,带入式(10)可得:

$$\frac{\sigma_{N,m}^2}{m+1} = \alpha(N)(m+1) + \beta(N) \quad (20)$$

式(20)满足线性关系,通过判断邻阶样本点斜率的变化可以确定阶跃的发生。假定邻阶样本点斜率发生较大变化,可以认定为发生一次阶跃,即当前存在两个簇,使得计算机在处理邻阶样本之间的关系时发生跳跃式变化。迭代全部样点,统计发生的所有样本的阶跃个体  $m^*$ ,即为样本数据聚类簇  $K$  值。同时,记录所有的阶跃点  $m^*$  可知簇中包含的样本数目。

对目标节点的划分采用一种数据样本分布算法,算法依照数据分布动态选取聚类中心。假设样本数据集  $U = \{x_i | i = 1, 2, \dots, n\}$ ,  $C_j (j = 1, 2, \dots, K)$  为聚类的  $K$  个类。

$$c_j = \frac{1}{n_j} \sum_{x \in C_j} x, j = 1, 2, \dots, K \quad (21)$$

其中,式(21)为各个类别的聚集中心,其中  $n_j$  为各类数据集中的样本个数。已知样本数据带有  $m$  个属性特征,那么样本数据  $x$  可以用  $x = (x_1, x_2, \dots, x_m)$  来表示。同理,样本数据  $y = (y_1, y_2, \dots, y_m)$ , 样本数据  $x$  和样本数据  $y$  两点之间的欧氏距离可表示为:

$$dist[x, y] = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_m - y_m)^2} \quad (22)$$

$limit$  为允许加入集合的最大空间距离:

$$limit = \frac{\max_{1 \leq I \leq N, 1 \leq J \leq N} (dist[I, J]) - \min_{1 \leq i \leq N, 1 \leq j \leq N} (dist[i, j])}{K} \quad (23)$$

样本数据  $x$  以及样本数据集  $V$  两者间的最小距离值可以表示为:

$$dist[x, V] = \min(dist[x, v]), v \in V \quad (24)$$

样本数据  $x$  与样本数据集  $V$  之间的最大距离为:

$$Dist[x, V] = \max(dist[x, v]), v \in V \quad (25)$$

## 1.2 配送模型

假设配送中心往  $L$  个网点配货,各个网点需求量可表示为  $g_i (i = 1, 2, \dots, L)$ ,同时,要求车辆到达网点的时段和在网点停留的时间分别为  $[et_i, lt_i]$  和  $ut_i (i = 1, 2, \dots, L)$ ; 设送货车辆提前到达网点的每小时等待消耗为  $e_i$ , 超出时段每小时延迟耗费  $f_i (i = 1, 2, \dots, L)$ ; 配送中心与网点、网点与网点之间的最短距离、行车速度以及车辆每公里耗费分别为  $d_{ij}$ ,  $v_{ij}$  和  $\omega_{ij} r_{ij} (i, j = 0, 1, 2, \dots, L)$ , 如  $i = 0$  时,则  $i$  表示为配送中心,  $\omega_{ij}$  为道路状况权重; 配送车辆类型可以分为  $q_0$  类, 其中,第  $q$  类型号的车辆共有  $p_0$  辆,同时,可知  $q$  类型车辆载重量为  $v_{qp} (p = 1, 2, \dots, p_0)$ , 并且,每车每次配送线路最大长度不

超过  $D_q$ ; 配货驾驶员出车补助和加班费用别为  $s/h$  和  $es/h$ ; 驾驶员在出车途中到中午  $HL$  时刻(比如 12:00 点)和  $HD$  时刻(比如 18:00 点)时安排  $HT_{min}$  就餐时间(比如 30min), 车辆当天返回配送中心。

配送线路费用由 4 部分组成,可以概括为:车辆使用消耗、驾驶员工资、车辆提前到达网点的等待耗费以及晚点到达网点的延迟耗费。车辆使用消耗又包括行车油耗、车辆新旧损耗和车辆维修,相关人员可由以往经营情况核算出每类车辆每公里需耗费的费,这个统计值为常量,每类车辆耗费也呈线性关系。在驾驶员费用中,如果正常工作时间在 8 小时以内,则以工资发放;一旦超出 8 小时,那么超出部分以加班费用计算。每个网点对配送中心送过来的货物设定了接待时间,当送货车辆提前或者晚点到达都得受到相应的惩罚,即需付出不同的费用代价,提前到达时,需支付等待费用;晚点到达时,需支付延迟费用。

令车辆  $qp$  (第  $q$  类车的第  $p$  辆)从客户  $j$  到达客户  $i$  时刻为  $t_i$ , 则  $t_i = t_j + ut_j + d_{ij}/v_{ij}$ , 其中  $j$  为  $i$  的前一个客户点,若  $t_j < HL$  且  $t_i \geq HL$  或  $t_j < HD$  且  $t_i \geq HD$ , 则需要考虑驾驶员的就餐时间。对于  $t_j < HL$  且  $t_i \geq HL$  的情况,有:

$$t_i = \begin{cases} t_j + ut_j + d_{ij}/v_{ij} + HT, & t_j + ut_j < HL \\ d_{ij}/v_{ij} + HL + HT, & HL \leq t_j + ut_j \leq (HL + HT) \\ t_j + ut_j + d_{ij}/v_{ij}, & t_j + ut_j > (HL + HT) \end{cases} \quad (26)$$

$t_j < HD$  且  $t_i \geq HD$  的情况与式(26)类似,这里不再做描述。线路  $(i, j)$  表示配送中心与网点或者网点与网点之间的最短路径,当  $x_{ijqp} = 1$  时,则说明车辆  $qp$  路过  $(i, j)$  线路,当  $x_{ijqp} = 0$  时,则说明车辆  $qp$  行车并未路过  $(i, j)$  段线路; 当  $y_{iqp} = 1$  时,表示车辆  $qp$  需要给网点  $i$  配货。相反,当  $y_{iqp} = 0$  时,表示车辆  $qp$  不需给网点  $i$  配货; 令  $wt_{qp}$  表示驾驶员工作时间在 8 小时之内,可表示为  $wt_{qp} = \min(t_{00} - t_0, 8)$ , 其中  $t_0$  是发车时刻,  $t_0 = et_i - dt_i - d_{0i}/v_{0i}$  ( $i$  是第一个客户点,  $dt_i$  为到达第一个客户点的等待时间)或  $t_0 = et_i + yt_i - d_{0i}/v_{0i}$  ( $i$  为第一个客户点,  $yt_i$  为达到第一个客户点的延迟时间),  $t_{00}$  为车辆返回仓库时间点;  $ewt_{qp}$  表示驾驶员的加班时间,可表示为  $ewt_{qp} = \max(t_{00} - t_0 - 8, 0)$ ; 每条线路客户点配送量之和要小于线路车载量,可表示为:

$$\sum_{i=0}^L \sum_{j=0}^L y_{ijqp} q_i \leq v_{qp} \quad (27)$$

$A$  为惩罚因子,是一个足够大的正整数,  $Z$  为总配送费用。假定各项费用呈线性变化,则有目标函数:

$$\begin{aligned} \min Z = & \sum_{i=0}^L \sum_{j=0}^L \sum_{q=1}^{q_0} \sum_{p=1}^{p_0} d_{ij} \omega_{ij} r_{ij} x_{ijqp} + \sum_{i=0}^L \sum_{q=1}^{q_0} \sum_{p=1}^{p_0} y_{iqp} (\omega t_{qp} \cdot \\ & s + ewt_{qp} \cdot es) + \sum_{i=1}^L e_i \cdot \max(et_i - t_i, 0) + \sum_{i=1}^L f_i \cdot \\ & \max(t_i - lt_i, 0) + A \sum_{q=1}^{q_0} \sum_{p=1}^{p_0} \max(\sum_{i=0}^L \sum_{j=0}^L d_{ij} x_{ijqp} - \\ & D_p, 0) \end{aligned} \quad (28)$$

上述式子共由 5 项组成,第一项为配送类型车辆自身耗费,第二项为出车驾驶员一趟费用,第三项为配货车辆提前到达网点时的等待费用,第四项为配货车辆晚点到达网点的延迟费用,如果是硬时间窗口问题,可以取  $f_i$  为一个足够大的正数,第五项为对配货车辆的一些限制,包括最大载重量,一次行驶最长距离。在线路中,任何一辆车超过限制的最大行车长度都会引起上述目标函数值急剧增大,也就保证了搜索

最优路径时的可行有效解必须在可行域之内。

## 2 算法设计

### 2.1 分类算法设计

根据上面的理论分析,算法的设计分为3个部分:第1部分为 $K$ 值的确定,即通过 $N$ 阶近邻算法确定;第2部分为划分算法的设计,对目标节点聚类;第3部分为配送算法,目标节点数目较大,即这部分的算法应该采用启发式算法。

根据 $N$ 阶近邻理论的设计,算法设计步骤分为:最短邻近距离求解、对邻近距离排序、样本方差求解和样本数值计算统计。合理、正确、高效地求出节约点 $m^*$ 为算法的关键部分。 $N$ 阶最短近邻算法的流程图如图1所示。

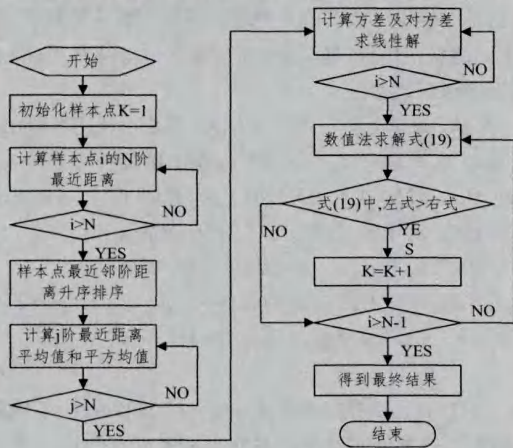


图1  $N$ 阶最短近邻算法流程图

采用 $K$ -means算法时,聚类簇具有在样本数据分布密集地形成区域的特点,因此,在选择初始聚类中心的时候往往先采用这些区域的距离平均值。为了找到与样本数据在空间坐标分布上相似,同时能够表示 $K$ 个聚类的样本数据集合,输入采用 $N$ 阶近邻自适应算法获得的 $K$ 值,再输入具有 $n$ 个样本的数据集合,经过运算,得到 $K$ 个数据中心,并且确定以每个数据中心为簇的成员数据的坐标值。算法步骤如下:

Step1 使用式(22)获取样本数据点之间的距离值的矩阵为 $D=(d_{ij})$ ,其中 $d_{ij}=dist[i,j](1 \leq i \leq n, 1 \leq j \leq n)$ ;

Step2 搜索样本数据点之间的距离矩阵 $D$ ,并寻找样本数据之间距离的最大以及最小值,再利用式(23)计算 $limit$ 值的大小;

Step3 搜索样本数据点之间的距离矩阵 $D$ ,在矩阵 $D$ 中找到最小值,并提取最小值相应的2个样本数据数据 $a, b$ ;然后,把样本数据 $a, b$ 添加到集合 $C_j(j=1, 2, \dots, k), C_j=\{a, b\}$ ;同时,把获得的样本数据 $a, b$ 从集合 $U$ 中去掉,并更新距离矩阵 $D$ ;

Step4 使用式(4)在集合 $U$ 中扫描距离集合 $C_j$ 最近的样本数据 $t$ ,如果 $Dist[t, C_j]$ 距离小于 $limit$ 阈值,那么需把样本数据 $t$ 添加到集合 $C_j$ 中,同时,把样本数据 $t$ 从集合 $U$ 中去掉,并更新距离矩阵 $D$ ,重复Step5,否则停止;

Step5 如果 $i < k, i = i + 1$ ,则需跳转至步骤Step3、Step4,直至 $k$ 个集合数完成;

Step6 取集合 $C_j(1 \leq j \leq K)$ 中样本数据的算术平均值作为样本数据中心 $c_j$ ,并获取 $c_j$ 的坐标值以及已完成的 $k$ 个数据中心的坐标值。

### 2.2 配送算法设计

以上步骤可以将配送客户点分类成 $k$ 类,因此每个类中的客户点的数量平均减缩为分类之前的 $1/k$ ,对于每一类,本文采用遗传算法来进一步解决配送问题。算法设计步骤如下:

Step1 根据类中客户点数目产生初始种群进行遗传编码;

Step2 计算种群的适应度函数;

Step3 用轮盘赌选择方式选取父代染色体,并对其进行删减、复制等染色体操作;

Step4 采用交叉概率对选取的父代染色体进行交叉操作,交叉操作方式以一定的规则进行,并检查交叉后的染色体是否满足约束条件;

Step5 采用变异概率对交叉产生的子染色体进行变异操作,变异操作以一定的规则实行,并检查变异后的染色体是否满足约束;

Step6 判断变异产生的染色是否满足终止条件,即运算是否已达到设置的迭代次数或者已经获得最优解的要求解,如满足要求则进入下一步;否则跳转至步骤Step3;

Step7 对计算结果进行解码;

Step8 选取步骤Step7解码的所有结果值,并对这些值进行筛选,获得最小值;

## 3 实例应用

根据上述模型及相应的计算方法,本文通过对杭州X牌卷烟配送实例进行分析求解,以验证方法的有效性。杭州X牌卷烟共有6个卷烟仓库,40个客户配送点,分布相对分散,且都在市区内,若动用卡车配送会增加不必要的成本,因此选择用小面包车配送。可总结出以下特点:①每个客户店的配送量少;②客户点相对比较分散;③要在一定的时间之内送达。

具体的方法可以描述为:以杭州市矢量数据为基础,通过建立网络连通性、道路等级以及设置车辆在每条道路上的行驶速度,来构建基于实际道路的杭州市路网模型,并利用GIS软件(ArcGIS)的网络分析模块,可以精确得到各仓库和各客户点之间的最短距离和最短时间(道路等级不一样,车速也不同,最短距离和最短时间不呈线性关系),同时建立最短距离和最短时间的OD矩阵,为上述算法奠定基础。

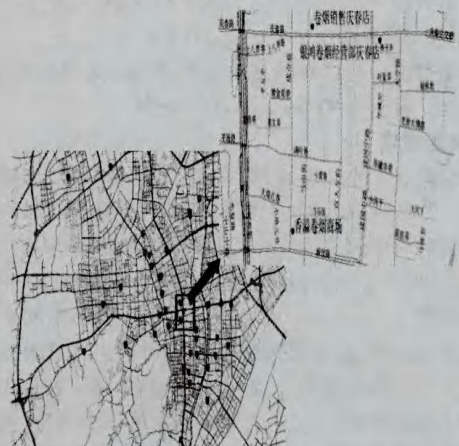


图2 杭州市卷烟配送客户点分布

图2显示了40个配送客户点的分布状况,以及市内的道路状况,一般来说,道路越宽敞,行车速度越快,相应的配送时间也是越短。

在表1中,40个客户点被分为9类,在9类中,最少的为一个客户点,最多的为8个客户点。这8个客户点可以在离它们最近的长河仓库(6个仓库中的其中一个)取货。同时可以采用遗传算法解决带有时间窗口的配送问题。

表1 分类后40个客户点的聚类中心及包含客户点数

类簇	X坐标	Y坐标	包含个数
第1类	120.093500	30.326000	2
第2类	120.197000	30.266750	4
第3类	120.154000	30.274286	7
第4类	120.176200	30.274800	5
第5类	120.143429	30.294857	7
第6类	120.114667	30.275333	3
第7类	120.173667	30.313333	3
第8类	120.090000	30.261000	1
第9类	120.174125	30.250875	8

客户点详细分类情况如图3所示。通过上述分类,本文以带有8个客户点的类为例进行求解。8个客户点加上1个配送仓库,共有9个点,编号为{0,1,2,3,4,5,6,7,8},其中0为卷烟仓库。通过GIS软件的网络分析模块分析这9个点,容易得到它们之间的最短距离OD矩阵。

从表2中可以知道配送中心到达各个网点之间的最短距离值以及各网点之间的最短距离值的情况。除此之外,从表2中还可以得到每个客户点对X牌卷烟的条数、车辆在每个客户点之间的卸货时间以及每个客户点对取货的开放时间。配送车辆可以提前到达配送地点,但是不允许迟到。若提前到达配送地点,则要等待到规定的客户取货时间才能开始卸货。时间窗中的时间计算方法为:时间窗时间/60,即0号节点的时间为:[8:00,21:00],其它几个采用同样方法计算。



图3 全部客户点分类

表2 各节点之间的最短距离OD矩阵(米)

0	0									
1	11495	0								
2	13474	3329	0							
3	13816	2373	1925	0						
4	16587	5199	3901	2900	0					
5	14241	2798	2346	542	2580	0				
6	14109	2647	2811	898	2798	587	0			
7	15289	3900	2603	1600	1323	1256	1475	0		
8	14891	3447	2951	1147	1867	795	1014	544	0	
	0	1	2	3	4	5	6	7	8	
X坐标		120.165	120.185	120.171	120.185	120.171	120.167	120.177	120.172	
Y坐标		30.229	30.243	30.25	30.255	30.254	30.255	30.26	30.261	
时间窗口	[480,1260]	[700,1140]	[680,1140]	[780,1140]	[880,1140]	[900,1140]	[860,1140]	[510,1140]	[780,1140]	
卸货时间		30min	31min	29min	26min	29min	39min	32min	32min	
需求量		(28条)	(29条)	(27条)	(24条)	(27条)	(37条)	(30条)	(30条)	

例子中,配送客户点都分布在市区,配送量不大,本文只选用了一种类型车辆,车辆的容量为100条卷烟,每公里的车辆损耗假设与距离呈线性关系;由于配送客户点都在市区,车辆驾驶员不需要加班工作,工作时间都在8小时内;由于客户点的时间窗的存在,驾驶员有时会遇到在12:00或者18:00还在工作的情况,这样驾驶员可以抽出30min的时间吃饭。本文对算法进行了50次运算,50次中得到卷烟仓库派遣4辆车的情况为8次,得到3辆车的情况为42次。8次中4辆车行程最短的为121971米,42次中3辆车行程最短的为92921.7米。很明显,即使不加上车辆的费用,3辆车的情况也比4辆车的情况好很多。50次中最优配送方案(见图4)为:

第1辆车的路线为:0->5->7->2->0

第2辆车的路线为:0->8->4->6->0

第3辆车的路线为:0->3->1->0

对第1辆车行驶时间进行分析,可知0号节点(长河仓库)到达5号客户节点(香溢卷烟商场)的时间为14241/500

=28min(汽车行驶30公里/小时);5号客户节点的时间窗为900,由此可知汽车可在872时从0号节点发车(即14:32发车);5号客户点的卸货时间为29min,车辆离开5号客户点到7号客户节点(银鸿卷烟经营部庆春店)的时间为929;从5号客户节点到7号客户节点的时间为1256.4/500=3min,到达7号客户节点的时间为931,此时在7号客户点的时间窗口之内,无需等待时间,在此的卸货时间为32min,离开7号客户点的时间为962;从7号客户点到2号客户点(明源卷烟店)的行驶时间为2602.6/500=5min,到达2号客户点的时间为967;在2号客户点的时间窗之内,无需等待,卸货时间为31min,此时离开2号客户点的时间为998;此时车辆返回仓库,返回行驶时间为13473.5/500=27min,到达仓库时间为1025;若出车过程中遇到780和1080,则还需加30min的就餐时间。对第1辆车线路进行需求量分析,总共需要X牌卷烟为27+30+29=86条,满足车载量要求(车载容量为100条)。其它两辆车作业时间分析也同理。经过多次运行计算,得到最优配送方案如图4所示。

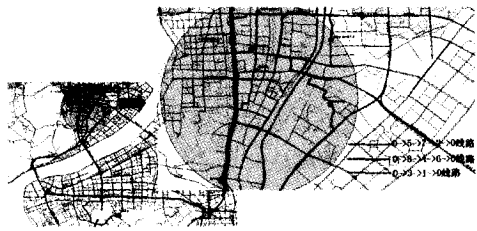


图4 50次运行中最优的3条线路配送方案

为了进一步验证所提方法的有效性,论文对 Solomon 的 VRPTW 标准问题集中的 RC1/RC108 数据进行分析求解。

如图5所示,已知客户点数  $n=100$  个,车辆总数  $K=25$  辆,车辆最大载重  $W=200$  吨。假设车辆均为匀速行驶,速度  $v_0=60\text{km/h}$ ,并令 RC1/RC108 数据中 coordinations 单位为公里,demand 单位为吨,ready time, due data 和 service time 单位为分钟。

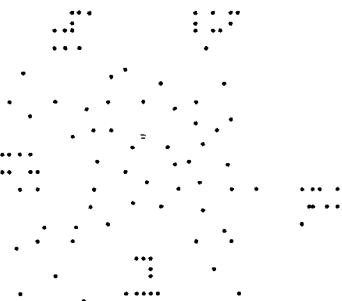


图5 客户点分布

利用  $N$  阶近邻算法确定聚类簇数  $K=13$ ,如图6所示,将100个客户点分为13类,最少的为4个客户点,最多的为11个客户点。

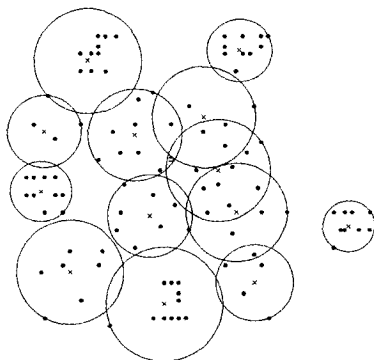


图6 全部客户点分类

应用上述算法进行计算后,得到配送路径如图7所示。

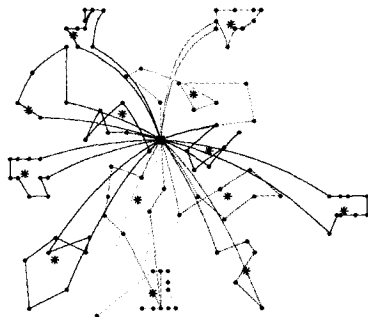


图7 配送路径图

**结束语** 本文针对解决带有时间窗口并具有较大规模配

送网点分布集的物流车辆配送线路问题进行研究。其中,建立融有道路等级权值、车辆运输损耗费用、驾驶员加班费用、就餐时间以及车辆早到配送网点的等待费用和晚到配送网点的延迟费用模型,使建立的模型尽可能贴近物流运输线路实际情况。在算法上,引入了  $N$  阶近邻自适应分类算法,该算法根据网点之间的距离分布可以很好地确定分类数  $K$  的值,并通过  $K$ -means 算法将配送网点集划分为若干个子网点集。最后,以杭州市实际路网矢量数据为依托,以配送 X 牌卷烟和 Solomon 的 VRPTW 问题集的数据为例子,通过对本文建立的模型和算法仿真,可知所建模型和设计的算法是有效的。

## 参考文献

- [1] 杨弋,顾幸生. 物流配送车辆优化调度的综述[J]. 东南大学学报:自然科学版,2003(z1):105-111
- [2] Lenstra J K, Kan A H G. Complexity of vehicle routing and scheduling problems[J]. Networks,1981,11(2):221-227
- [3] Solomon M M, Desrosiers J. Survey Paper-Time Window Constrained Routing and Scheduling Problems[J]. Transportation science,1988,22(1):1-13
- [4] Thangiah S R, Nygard K E, Juell P L. Gideon: A genetic algorithm system for vehicle routing with time windows[C]// Seventh IEEE Conference on Artificial Intelligence Applications, 1991. IEEE,1991,1:322-328
- [5] Blanton Jr J L, Wainwright R L. Multiple vehicle routing with time and capacity constraints using genetic algorithms[C]// Proceedings of the 5th International Conference on Genetic Algorithms. Morgan Kaufmann Publishers Inc.,1993:452-459
- [6] Hwang H S. An improved model for vehicle routing problem with time constraint based on genetic algorithm[J]. Computers & Industrial Engineering,2002,42(2):361-369
- [7] Timucin Ozdemir H, Mohan C K. Evolving schedule graphs for the vehicle routing problem with time windows[C]// Proceedings of the 2000 Congress on Evolutionary Computation,2000. IEEE,2000,2:888-895
- [8] Tan K C, Lee T H, Ou K, et al. A messy genetic algorithm for the vehicle routing problem with time window constraints[C]// Proceedings of the 2001 Congress on Evolutionary Computation,2001. IEEE,2001,1:679-686
- [9] Baker B M, Ayeche M A. A genetic algorithm for the vehicle routing problem[J]. Computers & Operations Research,2003,30(5):787-800
- [10] 戚铭尧,张金金,任丽. 基于时空聚类的带时间窗车辆路径规划算法[J]. 计算机科学,2014,41(3):218-222
- [11] Chen A Y, Peña-Mora F, Ouyang Y. A collaborative GIS framework to support equipment distribution for civil engineering disaster response operations [J]. Automation in Construction, 2011,20(5):637-648
- [12] Zsigraiova Z, Semiao V, Beijoco F. Operation costs and pollutant emissions reduction by definition of new collection scheduling and optimization of MSW collection routes using GIS. The case study of Barreiro, Portugal [J]. Waste management, 2013, 33(4):793-806
- [13] 张贵军,吴锡华. GIS 线形矢量图形最优路径算法研究及仿真实现[J]. 系统仿真学报,2003,15(4):551-553
- [14] 洪榛,张贵军,俞立. 基于  $N$  阶近邻分析的自适应差分进化算法[J]. 控制理论与应用,2012,28(11):1613-1620