

# 基于大数据的快递物流选址研究及实证分析

赵越, 张爱华

(北京邮电大学经济管理学院, 北京 100876)

**摘要:** 近年来, 我国电商交易规模迅速扩大, 与此同时, 消费者们在生产和生活服务上不断提高需求, 共同推动了快递物流业与智慧物流模式的迅速发展, 快递物流业也面临着新的难题, 如管理理念落后, 信息化水平与服务水平低, 仓储用地稀少等, 亟待解决。为解决以上难题, 本研究在大数据+智慧物流背景下, 建立选址逻辑距离模型, 将路况大数据、地理大数据代入模型, 使其更符合现代城市复杂多变的交通与路网规划, 更具实用性; 改良 k-means 聚类算法为 LK-means 算法, 将传统欧氏距离修改为逻辑距离模型, 在原有点中扩建选址而非增添新的中心选址点, 贴合现代城市仓储用地稀少的情况, 降低固定建设成本。代入具体算例, 基于 K-means 算法与 LK-means 算法使用 python 与 matlab 编程得出选址方案并对比分析验证其优越性。

**关键词:** 管理科学与工程; 选址; K-means 算法; 大数据

**中图分类号:** C93

## Research and empirical analysis of express logistics site selection based on big data

ZHAO Yue, ZHANG Aihua

(Beijing University of Posts and Telecommunications, School of Economics and Management)

**Abstract:** With the rise of social software and self-media, the profession of "Internet celebrity" came into being, which has produced great social influence and economic benefits. Therefore, this paper decides to discuss the Internet celebrity economic model based on the Internet platform, which can be understood to a certain extent as using the traffic effect brought by Internet celebrity characters or behaviors to achieve the purpose of word-of-mouth marketing, and the analysis of relevant factors affecting word-of-mouth marketing from this perspective. Through the analysis of the research status of the characteristics and strategies of domestic Internet celebrity economy and online word-of-mouth, it was decided to study word-of-mouth marketing under the background of online celebrity economy based on the 5T model, put forward 8 hypotheses, innovatively put forward the amplification and neglect hypotheses in word-of-mouth marketing, issued questionnaires, analyzed and studied through factor analysis methods and correlation analysis methods, and concluded the amplification and neglect of following Internet celebrities in word-of-mouth marketing, and provided certain management suggestions.

**Key words:** Management Science and Engineering; Internet celebrity economy; Internet Word of Mouth Marketing; SPSS

**作者简介:** 赵越 (1998-), 女, 无, 服务运营管理

**通信联系人:** 张爱华 (1964-), 女, 教授、硕导, 服务科学与运营管理/用户体验与用户行为分析. E-mail: aihuazhang@bupt.edu.cn

## 0 引言

### 0.1 研究背景

45 随着互联网技术的高速发展与居民物质水平的提高,目前,我国电商企业呈现出整体蓬勃发展的态势,交易规模逐年增长未现颓势。2022年天猫双11狂欢日全天成交额共达5571亿元,再创新记录;物流订单量突破单日5亿大关。庞大的数据量标志着我国电商市场的巨大能量及潜力,网上购物以其方便迅捷、物美价廉逐步占据实体市场额度。

50 随着技术水平的不断提高,用户对于物流的需求越来越多,越来越多的快递物流企业通过数据中心的构建更方便的管理物流过程中产生的海量数据,挖掘数据隐藏的价值,从而提升物流智能化与信息化水平,提高企业利润。在过去,由于处理工具与方法上的局限性,对于数据的采集运用仍处于原始阶段,大数据的价值并没有体现出来。2009年,IBM第一次提出具有互联、先进和智能三大特点的“智慧供应链”概念,之后,中国物流技术协会信息中心、《物流技术与应用》等组织由此概念定义了“智慧物流”,《2017中国智慧物流大数据发展报告》提出“智慧物流大数据发展指数”量化评价快递物流业的智能化和数据化水平,60 基于大数据的快递物流活动,物流效率得到明显提高,物流资源得到有效利用,同时也降低了供应链整体成本,是快递物流业发展的新方向<sup>[1]</sup>。

### 0.2 研究目的

相比传统研究方法,大数据处理技术可以更好地分析和整理数据,使其具有一定实用价值,从而促使物流企业通过大量的数据作出更好的决策。结合应用实例,根据数据分析60 结果结合实际情况进行该地区基于大数据的快递物流资源优化配置分析管理决策。

## 1 文献综述

### 1.1 物流+大数据行业研究现状

由于人们对快递物流业长期认识有一定的偏差,对快递物流业如何与新时代结合,从而转变发展方式,进一步提高发展质量的思考不多,中国快递物流整体行业水平仍然滞后。孙65 正萍(2018)在文献中指出,大数据时代的到来和电商产业的迅速发展为现代物流业提供了新的发展方向,即以智能化、数据化、信息化与共享化为特点的智慧物流新模式<sup>[1]</sup>。符瑞光(2018)基于大数据与现代物流的背景对新零售模式进行分析,探索零售业在此背景下具有的主要特点,表明基于快递物流信息平台建设的新零售生态发展的时代将逐渐来临<sup>[2]</sup>。

70 大数据、云计算等技术在近几年得到迅猛的发展,但是在实际操作层面,国内与国外高端企业之间还存在很大差距。国内的物流企业在进行大数据分析时,仍以分析结构化数据为主,而在美国,很多物流企业已经开始基于特定环境分析非结构化数据,且从供应链的整体角度去研究大数据对物流行业的优化。Berit Dangaard Brouer等人(2016)讨论如何利用大数据的大规模优化方法对海运物流进行优化,通过分解子问题,提出数学规划来举例说明适用于大数据的先进启发式方法<sup>[3]</sup>。

75 邱晗光等人(2018)进行基于大数据对共配流程进行改进,从多方面介绍大数据的应用场景,如配送时间窗、配送路径、车辆调度,降本提效等方面<sup>[4]</sup>。陶巍等人基于大数据构建

了五个层次的智慧物流系统架构,从研发设计角度探讨物流的优化方向<sup>[5]</sup>。

## 1.2 大数据+物流选址研究现状

在快递物流资源的宏观配置方面,大数据可以为物流配送中心的选址提供一定的决策支撑。在这一领域,主要有定性研究与定量研究这两种研究方法。吴家豪(2018)等人基于配送中心的概念及选址原则,分析了现存选址问题的各种方法,并主要对 Baumol- Wolfe 法做了基本介绍,基于物流配送成本、运营成本、固定建设成本等因素建立模型<sup>[6]</sup>。骆军伟(2013)提出基于模糊理论的 AHP 研究方法,首先根据模糊层次分析法计算出配送中心选址的相关因素优先级,根据优先级结果建立配送中心选址理论体系,从而进行决策<sup>[7]</sup>。胡利利(2016)以顺丰在长沙的配送中心选址为例,构建了经济型、社会性与综合性评价体系,运用层次分析法和模糊综合评价法对选址结果与备选方案进行评价<sup>[8]</sup>。以上研究主要是通过定性分析方法。关于定量分析方法,欧阳浩等人(2014)基于仓库运营成本问题建立选址模型,在一定的迭代次数内,通过遗传算法计算得到最优或局部最优解决方案,从而保证仓库运营成本最低<sup>[9]</sup>。李妮娜(2015)以二级中转场为研究目标,以北京市为例从交通与环境条件以及经济需求这三个角度构建快递物流网点中心性评价体系,利用信息熵模型对备选点进行中心性指数测算得到最终选址结果<sup>[10]</sup>。

不论是定性方法或者定量方法,都是较为传统的选址方法。为适应信息化的物流时代,国内有部分学者将大数据与物流结合,进行宏观选址分析。合肥工业大学的武方方(2015)用 Dijkstra 距离更换原有的欧氏距离或曼哈顿距离,优化 K-Means 算法,并结合 MapReduce 模型进行并行化处理。将优化后的 k-means 算法的距离权重设定为包含需求量、交通条件、市场效益在内的多因素分析后的“虚拟距离”<sup>[11]</sup>。李捷承等人(2018)为解决配送中心选址的 NP-hard 问题,设计基于 BIRCH 聚类的物流选址算法,融合了 BIRCH 聚类算法和重心法,可避免同时决策选址与确定路径两个难题<sup>[12]</sup>。刘威(2017)结合 K-means 聚类算法与证据推理法,解决已知备选点数量的多配送中心选址优化问题,把单区域内多选址问题转换为多区域单选址问题<sup>[13]</sup>。Ting L. Lei 等人(2018)基于 VAOMP 模型开发了一种有效的统一求解器,用于位置分配分析,目的是开发一种基于禁忌搜索(TS)元的快速启发式算法,可适用于在实时环境中获得 VAOMP 的最优或近似最优解的消息传递接口(MPI)<sup>[14]</sup>。干彬等研究者提出基于改进粒子群优化算法的优化选址方法,建立了优化选址经济模型和辅助模型;在对模型进行求解时,先对物流配送选址进行编码,反映粒子之间的合作与竞争的隐性知识,有效改善了粒子的搜索性能;同时利用边界缓冲墙方法对越界的粒子进行有效处理,加快了收敛速度,且显著提高了寻优精度<sup>[15]</sup>。

## 1.3 配送中心选址研究综述

物流配送中心的选址由于面向的层次、类型、对象不同,加上其本身的变化性、复杂性的特征,因此国内外针对配送中心选址的方法研究种类多种多样。

### 一、构建评价指标体系

#### (一) 层次分析法及其与模糊综合评价相结合

此方法首先指将与目标相关的各种重要因素划分为不同的层次,即目标、准则、方案等,在划分后的层次中开展定性和定量分析。梁非坤(2011)建立了层次分析法的理论模型并代入计算矩阵,并指出选址评价因素及其权重赋值应当结合具体案例进行分析,以保证选址结果的科学合理和一定的实用性<sup>[16]</sup>。李胜馨(2012)提出了配送中心选址原则,列举了现有

的常用选址方法,在此背景下提出结合综合建设成本最低的模型与层次分析法进行选址决策<sup>[17]</sup>。

## (二) 指标评价体系法

120 评价指标体系是根据评价对象各方面的特征及需要研究的方面构建有一定联系的多个指标,通过指标对某对象进行评价即指标评价体系法。指标评价体系法可用专家打分法赋予权重,也可用客观方法赋予权重。易玉枚等人(2011)为实现危险化学品配送中心选址,利用专家打分法对各子目标,如配送中心建设成本、配送成本等设置一定的权重进行打分,考虑危险化学品的风险特性、时间属性等因素,从而形成评价指标体系<sup>[18]</sup>。

125 除此之外,也可以使用其他方法对权重赋值,如李婷婷等人(2011)建立使仓储成本、运输成本与建设成本的总和最低的配送中心选址模型,使用实际算例,分别使用精确算法与近似算法进行计算分析<sup>[19]</sup>。肖岳峰等人(2011)运用了两阶段选址法,根据专家评判法的结果先选择合理的备选点,以车辆满载率、地价成本、配送成本、运输时间、固定建设成本、物流服务质量作为评价指标进行分析,构建合理选择模型并以实例验证方法优越性<sup>[20]</sup>。

## 二、考虑空间地理因素的选址方法

130 重心法将物流网络中的需求点和配送点看做分布某一范围内的散点,各点的需求量或吞吐量看作散点的重量,利用物理学中求物体重心的方法求出所有点的重心作为配送中心选址点。作为配送中心选址的基础方法,部分学者直接使用重心法进行选址研究,也有一部分学者基于重心法进行一定的优化进行选址,更贴合实际情况。任永昌等人(2010)结合重心法与经济分析法进行单配送中心选址研究,首先构造重心模型,通过不断的迭代求解模型,再对选址结果进行分析得到最终选址点<sup>[21]</sup>。谭阳波等人(2013)在某地区实际状况的背景下,利用迭代重心法进行配送中心选址,并从环境、经济条件、交通状况等多方面对选址结果进行评价<sup>[22]</sup>。

## 三、多目标配送中心选址

### (一) 多目标规划法

140 当需要对两个或更多非正相关的相互影响的目标进行决策时,需要寻找在一定约束条件下使目标和最优的解,或在不同目标间进行权衡取舍找到符合决策者期待的唯一解决方法。王瑞等人(2013)建立选址决策模型,在多个配送中心已建立的情况下进行单配送中心选址,在该模型的基础上使用 Lingo 求解无约束非线性规划,并使用实际算例进行验证<sup>[23]</sup>。

### (二) 双层规划法

145 同时考虑两个决策问题并设置一定的约束条件,建立相关模型并求解即双层规划法。Eiichi Taniguchi 等人(1999)为确定公共物流终端的最佳规模和位置,考虑了网络中的交通状况建立了双层规划模型,其上层目标为总成本最小,下层规划是一定需求下的物流选址模型,然后通过遗传算法求解,并成功应用于实际道路网络中<sup>[24]</sup>。

## 四、其他研究方法

150 配送中心选址研究方法还有许多。孟祥海等人(2006)采用遗传算法进行求解,以快递物流需求量产生的净利润最大化为目标函数确定选址地点<sup>[25]</sup>。刘璇等人(2012)结合 GIS 与空间分析技术对配送中心进行选址,将多因素代入决策,建立基于 GIS 技术的配送中心选址模型,经过计算确定最佳选址地点<sup>[26]</sup>。张雷等人(2012)针对特定一类的配送中心选址问题建立数学模型,通过优化后的和谐搜索算法进行选址,并与最优保存算法和遗传算法对比,验证该优化算法的优越性和实用性<sup>[27]</sup>。

各个研究者对配送中心的选址研究从不同的维度出发,每一种方法都应当结合实际情况

与实际问题分析, 本文也决定从特定的维度出发, 即二级配送中心的城市内选址, 建立用于此情况的通用解法。

## 2 模型构建

### 160 2.1 模型参数研究

配送中心选址的相关因素主要有物流配送成本、配送中心需求量、配送中心建设成本以及时间窗这四个方面。

165 (1) 关于需求量, 需要获取的数据为各末端快递物流网点的需求量与实际地址。快递物流网点的服务范围一般是以配送量为标准进行划分的, 因此, 假定某市整体物流业务量已知, 某市各快递网点服务范围大致相等, 可根据该市各区域人口密度与网点数目求解, 得到各区域快递网点需求量与该区域总需求量;

(2) 关于时间窗, 由于车载记录感知大数据 (RFID) 记录的行驶数据难以获得, 故决定基于本级配送中心到下一级网点的实际距离与路况信息进行拟合, 得到某型号货运车辆行驶的时间窗大数据;

170 (3) 关于物流配送成本, 拟合为运输距离\*常数 (油耗);

(4) 关于配送中心建设成本, 假定同一层次的配送中心面积类似, 仅需考虑不同区域的地价数据。

### 2.2 配送中心选址考量因素

175 (1) 物流配送中心选址是一个多层级的选址问题, 随着社会的发展, 现代物流业大多采用分级物流体系, 且在快递物流运输过程中不会出现跨层配送。

(2) 同级物流设施所服务的快递数量的均衡性。若在同一级中, 大量的快件集中于其中一个物流设施, 容易导致该设施快件积压, 效率极低, 也影响了客户体验。同时, 其他物流设施由于快递数量过少造成人力物力浪费。因此, 在考虑其他成本时也需要考虑各配送中心的位置保证其所负责快递量的大致均衡。

180 对特点 (2) 的一个可行的解决办法: 先进行配送范围的划分, 根据配送范围确定配送设施的数量, 由此确定各配送设施的位置, 同时, 通过约束条件保证配送中心的快递量在某一范围内波动。决策出配送设施的位置后, 可根据当天具体路况信息与订单数据决策配送路径, 避免同时决策位置和路径两个 NP-hard 问题。

185 基于这两个考虑的因素, 本研究决定以改进后的 K-means 算法对选址成本问题进行建模求解, 使之更符合城市配送中心选址的实际应用场景。

### 2.3 模型假设

给定某一区域所有快递物流网点的经纬度坐标, 从中选出一定数量的点建立配送中心。主要考虑配送中心需求量、物流运输成本、固定建设成本以及时间窗成本这四个因素, 得出各项分成本的表达式的家和为逻辑距离成本, 并使总成本达到最小或接近最小。

190 为了使模型能够更容易求解, 又能有一定的实用价值, 对模型作如下的限制与假设:

- 1) 同一级别的快递网点业务量与服务半径大致均衡;
- 2) 为便于计算, 只考虑单程货运车辆从本级配送中心到下一级网点的时间窗;
- 3) 在一定的备选范围内选出一定数量的配送中心;

- 195
- 4) 配送中心可供货到多个下一级网点;
  - 5) 不考虑配送中心之间相互供货, 不考虑跨级别供货;
  - 6) 运输成本与交通道路拥堵指数正相关;
  - 7) 配送中心建设成本与该区域地价正相关;
  - 8) 配送中心的容量、服务范围以及个数受限制;
  - 9) 基于配送时间窗建立时间惩罚机制。

200 **2.4 模型构建**

**2.4.1 模型参数**

- N——第二级配送中心的个数;
- K——第三级快递网点个数;
- $X_{ij}$ ——第 i 个节点到第 j 个节点的两点距离;

205 
$$X_{ij} = \sqrt{(K_{ix} - K_{jx})^2 + (K_{iy} - K_{jy})^2} \quad (2-1)$$

- $N_{ij}$ ——第 i 个节点到第 j 个节点的单位运输成本;
- $d_{ij}$ ——第 i 个节点到第 j 个节点的实际距离 (地图道路距离);
- V——某型号货运车辆的平均运输速度;
- $\theta$ ——某型号货运车辆的单位路程油耗成本;
- 210  $Y_n$ ——第 n 个区域的交通拥堵程度, 区域点根据坐标值确定;
- $Q_{1i}$ ——第 i 个配送中心的平均容量;
- $Q_{2j}$ ——第 j 个三级快递网点的平均需求量;
- $F_n$ ——第 n 个区域的土地价格, 区域点根据坐标值确定;
- $A_n = \{(K_{1x}, K_{1y}), (K_{2x}, K_{2y}) \dots (K_{ax}, K_{ay})\}$ ——坐标点在第 n 个区域的集合, 第 n
- 215 个区域内共有 a 个节点。
- S——配送中心的建设面积;
- R——配送中心的服务半径;
- $T_{max}$ ——根据不同城市的实际情况, 第 i 个节点到第 j 个节点的配送时间窗最大值;
- $T_w$ ——第 i 个节点到第 j 个节点的总时间窗成本;
- 220  $T_{ij}$ ——第 i 个节点到第 j 个节点的运输时间;

**2.4.2 模型变量**

- $K_{ix}$ 、 $K_{iy}$ ——第 i 个节点的 x 坐标与 y 坐标。

**2.4.3 目标函数**

规划期内总费用由三部分组成:

- 225 (1) 运输费用
- 货物从第 i 个节点到第 j 个节点所需的总运输里程矩阵  $d_{ij}$  为

$$\begin{bmatrix} d_{11} & \cdots & d_{1j} \\ \vdots & \ddots & \vdots \\ d_{i1} & \cdots & d_{ij} \end{bmatrix}$$

235 考虑配送中心的配送半径问题。由于城市配送区域是城市商业和人口聚集的区域，需求点密集，可看作是在该区域内均匀分布的，考虑利用中心地理论和六边形法则确定配送中心的数量，提出理想选址范围是以配送中心为图形中心的大小相同的正六边形。

假定配送中心的服务半径为 R，规划区面积为 B，R 和 B 的取值结合不同城市的实际情况综合给出。那么每个配送中心的配送范围即以该配送中心为圆心，R 为半径的圆的内接正六边形，其面积 S 为

$$S = \frac{3\sqrt{3}}{2} \cdot R^2 \quad (2-2)$$

235 因此，在配送中心规划中，确定了规划区域的面积 B 和该地区配送中心的服务半径 R，就可以确定该区域内配送中心数目 N，具体为：

$$N = \frac{B}{S} = \frac{2\sqrt{3}s}{9R^2} \quad (2-3)$$

故，货物从第 i 个节点到第 j 个节点所需的总运输成本为  $\sum_i \sum_j d_{ij} \cdot \theta$ ；

$$d_w = \begin{cases} \sum_i \sum_j d_{ij} \cdot \theta, & X_{ij} \leq R; \\ \sum_i \sum_j d_{ij} \cdot \theta \cdot 200, & X_{ij} > R; \end{cases} \quad (2-4)$$

240 根据坐标点  $(K_{ix}, K_{iy})$  判断第 i 个点到第 j 个点的地图实际距离途径哪些区域，根据该区域的交通拥堵指数与距离进行计算，得到第 i 个节点到第 j 个节点的总时间窗（ $\sum$  交通拥堵指数 \* 实际距离 / 速度）矩阵  $T_{ij}$  为

$$\begin{bmatrix} T_{11} & \cdots & T_{1j} \\ \vdots & \ddots & \vdots \\ T_{i1} & \cdots & T_{ij} \end{bmatrix}$$

245 故，货物从第 i 个节点到第 j 个节点所需的总时间窗为  $\sum_i \sum_j T_{ij}$ 。根据白玉凤（2016）对北京市海淀区的配送车辆与油耗调研，估计配送时间  $T_{max}$  为 120min，配送员早到一分钟的机会成本约为 2 元<sup>[28]</sup>，晚到的机会成本设置为较大值 200，故总时间窗成本为

$$T_w = \begin{cases} (120 - \sum_i \sum_j T_{ij}) * 2, & T_{ij} \leq 120; \\ (\sum_i \sum_j T_{ij} - 120) * 200, & T_{ij} > 120; \end{cases} \quad (2-5)$$

则总费用为上述两项的和，根据经济性原则，需要使总费用最小。

#### 2.4.4 约束条件

250

##### 1) 供需约束

所有第三级快递网点需求的货物总量不能超过所有配送中心总容量能力：

$$\sum_{j=1}^K Q_{2j} \leq \sum_{i=1}^N Q_{1i}, i = 1, 2, \dots, N, j = 1, 2, \dots, K; \quad (2-6)$$

2) 非负约束

255 模型中的变量必须大于或等于零, 即:

$$N_{ij}, d_{ij} \dots \geq 0; \quad (2-7)$$

3) 需求约束

第  $i$  个二级配送中心配送的  $m$  ( $m < K$ ) 个第三级快递网点需求的货物总量不超过其容量。

260 
$$\sum_{j=1}^m Q_{2j} \leq Q_{1i}, i = 1, 2 \dots N, j = 1, 2 \dots K; \quad (2-8)$$

### 2.4.5 模型形式

综合上述分析, 建立配送中心的选址模型为:

265 
$$\min E = \min(d_w + T_w) \quad (2-9)$$

S. T:

$$\sum_{j=1}^K Q_{2j} \leq \sum_{i=1}^N Q_{1i}, i = 1, 2 \dots N, j = 1, 2 \dots K; \quad (2-6)$$

$$N_{ij}, d_{ij} \dots \geq 0; \quad (2-7)$$

$$\sum_{j=1}^m Q_{2j} \leq Q_{1i}, i = 1, 2 \dots N, j = 1, 2 \dots K; \quad (2-8)$$

## 270 2.5 算法优化

### 2.5.1 K-means 算法

K-means 聚类算法起源于 1967 年, 由 J. B. Mac Queen 提出。因 K-means 算法的方便快捷, 因此它也被称作是经典的基于划分规则的聚类算法。传统 K-means 算法的距离 cost 主要是基于欧氏距离或曼哈顿距离, 因此在面对在实际物流规划问题时, 无法根据城市道路的规划而灵活应变, 缺乏实用性。因此, 本文引用了基于成本的城市内逻辑距离, 相对传统 K-means 算法有更强的针对性。

275

### 2.5.2 LK-means 算法

为贴合实际需求, 用地图实际距离替换一般选址方法中的理论距离。故考虑到当今时代, 大型城市的仓储用地越来越少的实际情况, 为减少计算量与复杂程度, 对算法优化方向提出以下要求:

280

新的配送中心的选址点为已存在的第三级快递网点, 故不需要考虑重复建设后的地价信息;

交通拥堵指数只考虑行政区整体拥堵程度, 而不深入到街道拥堵指数;

两点之间实际距离数据与欧氏距离数据已知;

285 基于上述要求与传统 k-means 算法的逻辑, 改进后的 k-means 算法称为 LK-means (Logical K-means) 算法。

根据论文实际需求与算法改进思路, 将优化后的算法称为 LK-means 算法 (逻辑 k-均值算

法), LK-means 算法的具体步骤为:

---

输入: 以经纬度坐标形式表示的所有快递网点数据集  $\text{dataSet} = x_1, x_2, x_3, \dots, x_m$ ,

聚类簇数  $K$ ;

过程: 函数  $\text{KMeans}(\text{dataSet}, K)$

1: 从  $\text{dataSet}$  中随机选择  $K$  个样本作为初始“簇中心”向量:  $\mu_1, \mu_2, \mu_3, \dots, \mu_K$ ;

2: repeat

3: 令  $C_j = \emptyset (1 \leq j \leq K)$ ;

4: for  $j=1, 2, \dots, k$  do;

5: 计算样本  $x_m$  与各“簇中心”向量  $\mu_j (1 \leq j \leq K)$  的逻辑距离;

6: 根据距离  $m$  最近的“簇中心”向量确定  $x_m$  的簇编号  $C_{[m,j]}$ ;

7: 将样本  $x_m$  划入相应的簇并存入该簇的集合中:  $PC_j = C_{[m,j]} \cup x_m$ ;

8: end for

9: for  $i = 1, 2, \dots, k$  do;

10: 计算“假簇中心”向量:  $(\mu_j)' = \frac{1}{|C_j|} \sum_{x \in C_j} x$ ;

11: 从  $\text{dataSet}$  中选择距离当前假“簇中心”欧氏距离最近的点  $\mu_m$ ;

12: 将  $\mu_m$  设置为新的“真簇中心”  $(\mu_j)' = \mu_m$ ;

11: if  $(\mu_j)' \neq \mu_j$  then

13: 将当前“簇中心”向量  $(\mu_j)$  更新为“簇中心”的  $\mu_j$

14: else

15: 保持当前均值向量不变

16: end if

17: end for

18: else

19: until 当前“簇中心”向量均未更新

输出: 簇划分  $C = C_1, C_2, \dots, C_K$

---

为避免处理时间过长, 设置最大迭代次数为 100 次, 当迭代达到 100 次则停止运行。

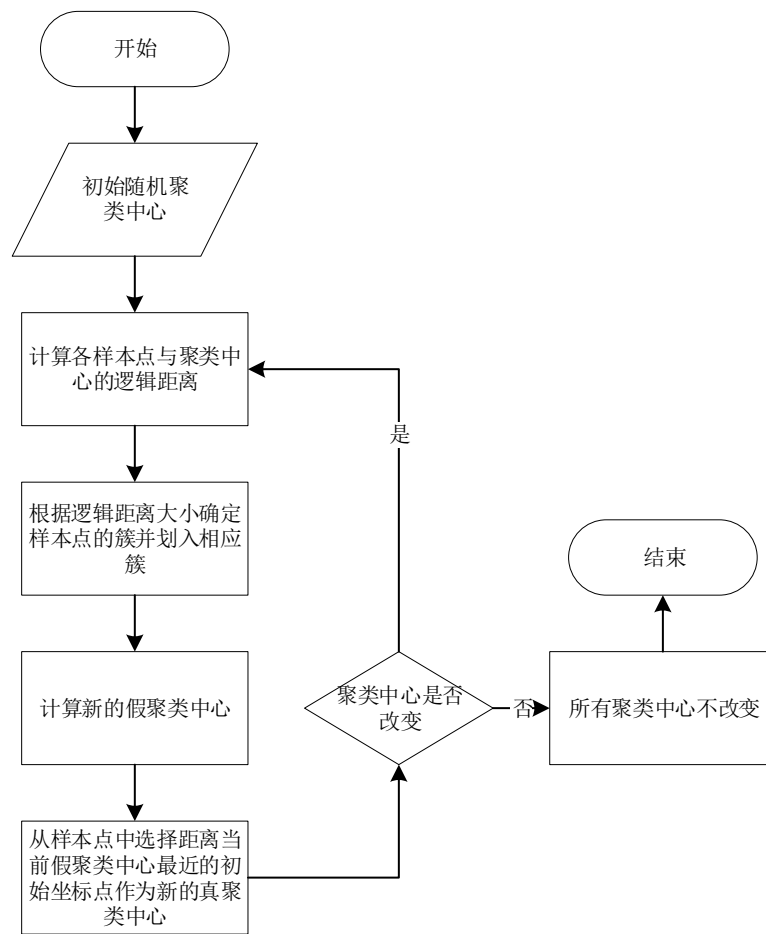


图 1 LK-means 算法流程图

### 2.5.3 LK-means 算法难点解决方案

295 根据算法逻辑及算法框架，决定采用 python 语言编程。在算法优化的过程中出现了一系列问题并寻求相应解决方法。

300 (1) LK-means 算法中实时调用百度 API 并返回数据导致算法的时间复杂度提升，降低数据了读取与处理速度。随着结点数量的持续增大，相关方面的开销成平方次急剧增加，因此针对这个问题考虑提前调用百度 API 进行所有原始节点的地图实际距离的读取并存入 excel 表格中，供 LK-Means 算法在运行时读取，提高 LK-means 聚类算法的计算效率与处理速度，从而降低 LK-means 算法对内存和内核处理能力的要求。

(2) 由于采用严格上三角矩阵，故算法在进行数据读取的过程中如果遍历至下三角的数据将返回“none”，导致程序报错。故修改遍历部分的代码逻辑使读取至下三角部分的数据返回其关于对角线对称的数据，即交换下标使  $i=j, j=i$  从而能够正确读取有效数据。

305 (3) 从假质心到真质心的转换逻辑为遍历全部样本寻找距离假质心欧氏距离最近的结点作为每一次迭代新的质心。但是在转换过程中可能出现距离假质心最近的点不在该簇内的情况，从而导致实现转换后该簇不存在任何质心，程序报错。出现该问题的主要原因是构成 cost 的两个成本无法确定权重。影响快递物流配送中心选址的因素多种多样，根据当地的

310 实际交通、运输、地理位置等情况分配权重,为实际数据集分配合理的权重将决定质心的位置是否合理。因此,决定将代入成本计算的运输成本与时间窗成本做归一化处理,拟定系数均为1,即对某城市来说,交通因素和距离因素在配送中心选址的考量中同等重要。

### 3 实例分析

#### 3.1 案例背景

315 根据北京市邮政管理局发布的《2021年北京市邮政行业发展统计公报》显示,2021年,北京市邮政企业和快递服务企业业务收入共达384.29亿元。同时,2021年北京市快递业务量总计达22.10亿件,快递业务收入总计313.43亿元。庞大的数字表示,快递物流市场的增长未见颓势,且在人民的生活中占据越来越重要的地位。但是在迅速发展的同时,当前北京快递业仍存在诸多问题亟待解决。

##### (1) 北京市高消费成本严重制约快递企业发展

320 随着北京市房价与物价的上升,直接导致了位于北京的快递物流网点的经营成本的提高,如房屋租金、固定水电费,燃油费用等等,从而导致北京市快递企业的生存发展越来越困难,利润率越来越低。除高额的固定成本外,配送范围过大也成为北京市快递物流企业成本上涨的另一大原因。在党的十九大报告中,将疏解北京非首都功能作为新的发展战略,故北京市物流设施不断向城郊疏解,导致配送半径的增大,提高了人力与物流运输成本。

##### (2) 城市仓储用地较少

325 近几年来,由于北京市规划布局受国家发展战略的影响发生变更,原有物流用地迅速缩减,新增申请也更加困难,根据头等仓发布的仓储物流地产大数据报告显示,北京市A类仓库和B类仓库的库存量空置率极低,证明了物流用地资源的紧缺。由于北京市政府缺乏对物流产业设施的统一规划布局,导致其后期发展性较弱,严重制约了物流企业的后期发展,更导致了物流资源的浪费。

#### 3.2 数据与研究范围确定

335 近几年来,随着北京市经济发展水平的不断提高,常住人口数量持续增加,种种原因均导致了北京市快递服务需求的区域性差异分布。第一,根据人口数量确定快递需求分布。快递行业服务于客户,人口数量及密度越大的地区,经济发展速度与区域现代化进程均会有一定程度的提高,从而导致快递行业更加发达。根据北京市统计局发布的《2020年北京市统计年鉴》显示,2020年朝阳区,海淀区的常住人口分别为345.1万人和313.2万人,是北京市常住人口最多的两个区域。其次,地区生产总值同样对快递需求分布一定的影响。地区生产总值越高表明了经济更加繁荣,消费需求更发达,对居民快递物流需求水平也有一定的提升。

340 综合北京市区域经济发展水平、人口分布情况的数据可以看出,北京市朝阳区、海淀区是市内快递需求较为旺盛的两个地区。由于研究时间有限,且聚类算法在连续范围内最为有效,故本文选取北京市东城区、西城区、朝阳区、海淀区这四个区域三通一达(中通、圆通、申通、韵达)共419个快递网点进行选址研究。

### 3.3 数据调研与分析

#### 3.3.1 物流运输成本

为降低算法复杂度,根据数学建模公式将物流运输成本拟合为物流运输里程\*常数(油耗),为更贴近北京市具体情况将物流运输里程定为地图间两点实际距离。

利用八爪鱼软件爬取百度地图北京市三通一达所有第三级快递网点地址,经过数据筛选后可知北京市三通一达所有第三级快递网点的数量为 937 个,东城区、西城区、朝阳区、海淀区共有 419 个网点。将所有网点地址信息代入百度坐标拾取系统后,按照行政区域划分并得到所有快递网点的经纬度,见图 2。

序号	名称	地址	坐标	经度	纬度	行政区域
208	韵达快递(国奥公司)	北京市朝阳区区域清街2号院2号楼1层世茂	116.373305,40.023431	116.373305	40.023431	朝阳区
209	韵达	北京市朝阳区驹南街	116.555957,39.964389	116.555957	39.964389	朝阳区
210	韵达	姚家园将台洼东风公园鹿队大院甲一号	116.520148,39.966732	116.520148	39.966732	朝阳区
211	韵达	北京市朝阳区左安路	116.454727,39.874508	116.454727	39.874508	朝阳区
212	韵达	北京市朝阳区东四环北路10号院4附近	116.493651,39.966304	116.493651	39.966304	朝阳区
213	韵达(远洋服务站)	北京市朝阳区甘露园南里路西附近	116.517267,39.921338	116.517267	39.921338	朝阳区
214	韵达网点	北京市朝阳区三里屯街道朝阳区中纺里	116.460422,39.934329	116.460422	39.934329	朝阳区
215	韵达	北京市朝阳区金台西路甲2号	116.475622,39.925702	116.475622	39.925702	朝阳区
216	韵达	王四营镇新东郊南城南门	116.540542,39.872438	116.540542	39.872438	朝阳区
217	北京韵达石佛营公司	北京市朝阳区八里庄南里8号楼	116.50496,39.928577	116.50496	39.928577	朝阳区
218	申通快递	北京市东城区三环以内北三环东路36号	116.416834,39.973621	116.416834	39.973621	东城区
219	申通快递	北京市东城区北岗子街8号	116.432393,39.896497	116.432393	39.896497	东城区
220	北京和平里申通快递	和平里北街和平里七区10号楼底商	116.426704,39.966663	116.426704	39.966663	东城区
221	申通快递东直门分公	北京市东城区和平里七区第九小学	116.426712,39.966673	116.426712	39.966673	东城区
222	申通快递(北京新世界	国瑞东路40号	116.431783,39.903911	116.431783	39.903911	东城区
223	圆通快递	北京市东城区东皇城根南街34-1	116.41339,39.927791	116.41339	39.927791	东城区
224	圆通快递	北京市东城区安定门街道北三环东路36	116.416766,39.973456	116.416766	39.973456	东城区
225	圆通快递	北京市东城区东直门外小街甲2号正东	116.452835,39.949983	116.452835	39.949983	东城区
226	圆通快递	青年湖南街与青年湖西里交叉口北100米	116.404602,39.961182	116.404602	39.961182	东城区
227	圆通快递(双花园商业	北京市东城区忠实里5号院	116.452316,39.908495	116.452316	39.908495	东城区
228	圆通快递(新世界分部	东直门12号	116.434635,39.896644	116.434635	39.896644	东城区
229	圆通速递	北京市东城区东直门外小街34号附近	116.447253,39.948198	116.447253	39.948198	东城区

图 2 北京市部分区域三通一达网点地址及经纬度

利用 matlab 软件生成北京市所有快递网点坐标散点图(按行政区域划分)如图 3,不同的颜色及类型的点代表北京市不同的行政区域。

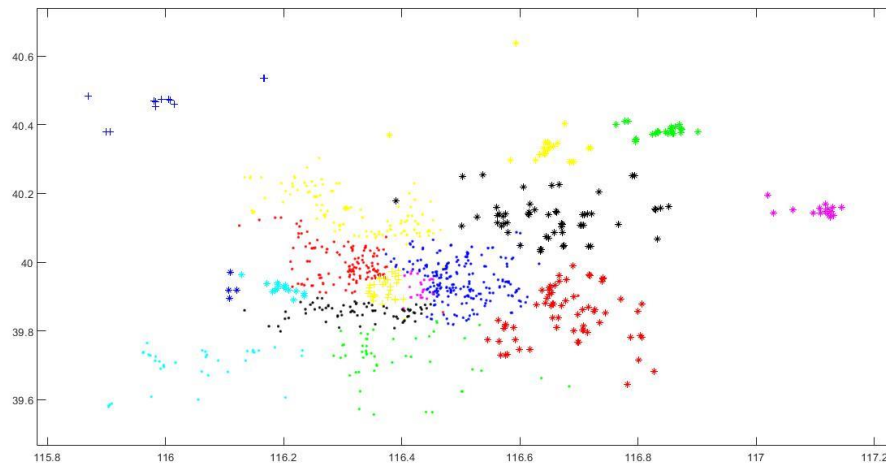


图 3 北京市全区域三通一达网点坐标散点图

利用百度地图开放平台的驾车路线规划,批量算路等 API 进行数据处理,单次取一对地点的经纬度坐标数据向百度 API 发送处理请求,返回结果后生成距离矩阵并存储在文件中。由于研究经费的限制,取东城区、西城区、朝阳区、海淀区在内的 419 个快递网点进行地址计算,共需计算 9 万次左右,生成的数据矩阵部分如图 4 所示。

156.112389.40	156.481198.39	156.429943.39	156.442453.39	156.601373.39	156.56791.40	156.492796.40	156.475602.39	156.482081.39	156.483423.39	156.463813.39	156.483541.39	156.483948.39	156.423238.39	156.588491.40	156.533239.40	156.479261.39	156.466418.39	156.482901.39	156.430613.39	156.514411.39	156.389644.39	156.432208.39	156.432791.39	156.588478.39
0	22991	20720	24255	22465	8720	9679	6181	6102	12664	11224	12137	10018	6188	7817	3747	4874	6186	12038	23074	11748	14296	12630	11377	16964
156.481188.39	0	9245	3901	15501	17931	13056	6641	7279	956	3978	1505	7346	6985	10300	15261	10866	9005	1203	12644	1975	14011	18997	10593	10296
156.429943.39	0	7464	24417	15931	9011	5409	7381	9532	7767	10415	2640	2789	17498	12969	6793	4679	10011	18906	12491	4677	21295	1765	19177	
156.442453.39	0	18250	18087	18244	7689	9662	2667	1642	1214	4761	6680	20234	13726	10910	6187	4239	15443	6913	12123	12412	10412	8690	18394	
156.601373.39	0	20187	23384	21779	20397	15541	16269	14877	13844	22941	22941	18454	21871	23521	20969	15113	15161	14846	29744	18166	26286	7310	15823	
156.56791.40	0	9149	11818	11539	17478	16661	18039	15455	13625	2011	2807	10111	13573	17835	2811	15790	19736	8238	17414	12823	16929	16929	16929	
156.492796.40	0	8799	8720	13312	11842	13875	9420	7362	10661	8509	2909	8754	12871	19862	12770	12317	18673	9999	18673	9999	18673	9999	18673	
156.475602.39	0	2077	7182	6568	7716	4742	2912	13163	8726	4929	3860	7612	16918	9207	8978	13991	6647	15656	6647	15656	6647	15656	6647	
156.482081.39	0	7478	7151	9039	6825	4995	13091	6635	4446	4063	7835	19001	8200	11118	15042	8630	14657	8630	14657	8630	14657	8630	14657	
156.483423.39	0	4029	795	7759	7759	7411	17907	15715	11219	8458	591	12389	3548	14474	18996	11046	10026	11046	10026	11046	10026	11046	10026	
156.463813.39	0	4005	8189	8528	16263	11827	9411	3370	3900	13007	6081	12739	18916	9048	12762	9048	12762	9048	12762	9048	12762	9048	12762	
156.483948.39	0	8876	8094	18444	16998	12002	1541	738	11897	4005	15197	19493	11729	10686	11729	10686	11729	10686	11729	10686	11729	10686	11729	
156.423238.39	0	2191	16678	10349	6496	3608	8488	4488	17027	10763	7817	20675	3889	17444	3889	17444	3889	17444	3889	17444	3889	17444	3889	
156.588491.40	0	19096	10520	5106	2324	7098	17061	10281	7064	18701	3661	18911	3661	18911	3661	18911	3661	18911	3661	18911	3661	18911	3661	
156.533239.40	0	5971	12175	15807	17984	28238	17475	15900	6589	18978	11114	15978	11114	15978	11114	15978	11114	15978	11114	15978	11114	15978	11114	
156.479261.39	0	6970	11266	16183	5106	2324	7098	17061	10281	7064	18701	3661	18911	3661	18911	3661	18911	3661	18911	3661	18911	3661	18911	
156.466418.39	0	5974	15880	8234	9778	17213	6350	14936	14996	15914	7445	17874	7445	17874	7445	17874	7445	17874	7445	17874	7445	17874	7445	
156.430613.39	0	10024	3895	14781	19303	11393	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	10936	
156.389644.39	0	14303	13037	20278	20784	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	20402	
156.432208.39	0	17180	18764	13731	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	7662	
156.432791.39	0	24581	3828	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	23504	
156.432791.39	0	22276	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	12524	
156.432791.39	0	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	21005	

图 4 两点距离计算矩阵

### 3.3.2 快递需求量数据

365

根据前文配送中心选址模型,可知需要获取的数据为北京市第三级快递物流网点的需求量。根据北京市邮政管理局全市邮政业运行情况数据,得出北京市近五年来各月份整体快递物流业务量以及年度峰值月份数据。

北京市邮政业务量 (万件)					
	2022	2021	2020	2019	2018
1 月	19028.82	18947.16	17730.39	18237.5	20025.73
2 月	15513.06	12117.56	13967.67	11712.57	12214.35
3 月	15513.06	19646.81	19533.22	17900	17933.34
4 月	14195.84	18474.44	19620.82	17323.9	16462.36
5 月	14218.41	18144.64	21443.86	17411.35	17820.78
6 月	14218.41	20701.11	23000	19400	19158.06
7 月	17700	17600	19400	18600	17782.6
8 月	16500	18500	18900	19400	18085.47
9 月	17700	19600	20200	19600	19184.21
10 月	16400	15900	18700	20000	17700
11 月	17600	21800	25100	25200	23300
12 月	——	19600	20600	23900	21200
快递业务量峰值	19028.82	21800	25100	25200	23300

图 5 北京市邮政业务量

370

综合实际情况与流转型配送中心功能,以近五年来每年峰值数据建立灰色预测模型,预测 2023 年北京全市整体快递物流业务量月度峰值数据为 29464.6349 万件。根据各区域的网点数量分布,确定北京市各区域的快递网点总需求量。

### 3.3.3 时间窗数据

375

根据白玉凤(2016)在北京市海淀区进行的实际调研结果显示,北京市中心城区配送主要使用如图 5-10 所示的小型货车。该型号货车净重 2 吨,装载量 3.5 吨,装载体积为 19 立方米,油箱尺寸为 60L,使用均价 5.54 元/升的 92 号汽油。每升汽油可供该车辆在公路上大概运输 8.5 公里,加满油后可运输共 480.5 公里,又可知该型号小货车一天的租赁成本约为 500 元。故可计算出,该型号车辆每公里的运输成本为 0.65 元,在北京市市内的行驶速度平

均 30 公里/小时，即 0.5 公里/分钟<sup>[28]</sup>。

380 交通拥堵指数即实际情况下平均一次出行耗费的时间与畅通情况下耗费的时间的比值。利用八爪鱼软件可爬取高德地图发布的实时路况拥堵数据，该数据是基于高德历史积累的北京市交通出行数据的挖掘计算得到的。收集近一个月以来北京市各区域早晨 7:00-9:00 的交通拥堵指数大数据，通过求均值，可得到基于两点之间实际距离数据与交通拥堵指数，根据模型假设对原始数据进行处理，可得到该型号货运车辆行驶的时间窗大数据。

### 385 3.4 传统 K-means 算法选址结果

#### 3.4.1 配送中心数量

通过前期资料搜集，可知东城区、西城区、朝阳区与海淀区的辖区面积见表 1。

表 1 北京市四区域辖区面积

区域	东城区	西城区	朝阳区	海淀区
辖区面积/km <sup>2</sup>	41.84	50.70	470.8	430.8

390 故四区总面积为 994.14km<sup>2</sup>。李妮娜（2015）基于中心地理论进行快递服务网点的布局研究，在论文中提出三级快递物流网络的配送中心服务半径为 R = 15km<sup>[28]</sup>。将 R = 15km 代入计算中，通过公式

$$N = \frac{B}{s} = \frac{2\sqrt{3}s}{9R^2} \quad (3-1)$$

求得 N = 4，故 K-means 算法集簇数为 4。

#### 3.4.2 matlab 输出结果

395 利用 matlab 软件进行算法分析，输出图 6 所示的结果。

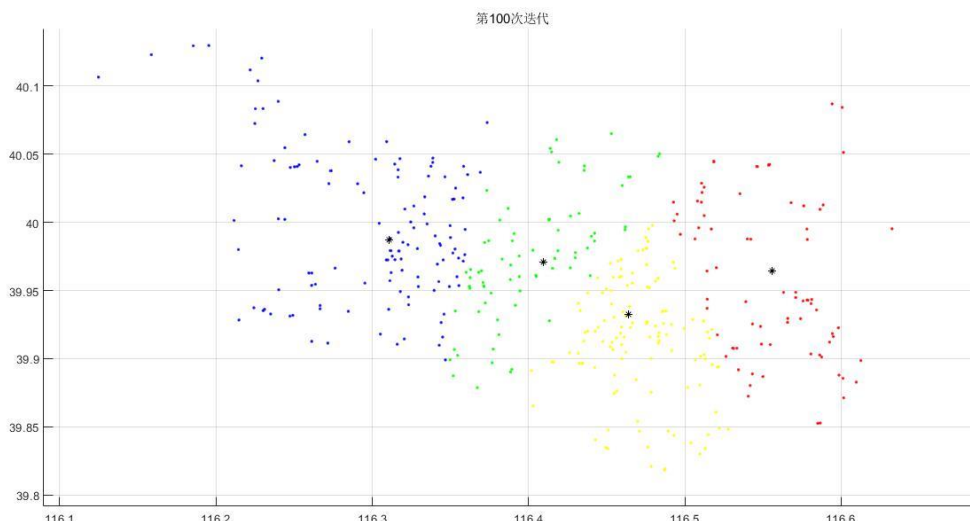


图 6 传统 K-means 算法选址散点图

根据第 100 次迭代的结果，拟定 4 个点为 k-means 聚类的新质心，代入百度拾取坐标系后得到最终的选址结果见表 2。

400 表 2 基于 K-means 算法的配送中心选址结果

经纬度坐标	实际地址
(116.310859,39.987404)	北京市海淀区苏州街 29 号-1 栋

(116.409605,39.97108)	北京市朝阳区安华路
(116.463813,39.932435)	北京市朝阳区白家庄路3号A106
(116.555957,39.964389)	北京市朝阳区东坝中路28号院

Matlab 输出结果显示,这四个选址点均存在最远配送点,利用百度地图的距离测算每个选址点距离最远配送点的距离见图7,将结果整理见表3。

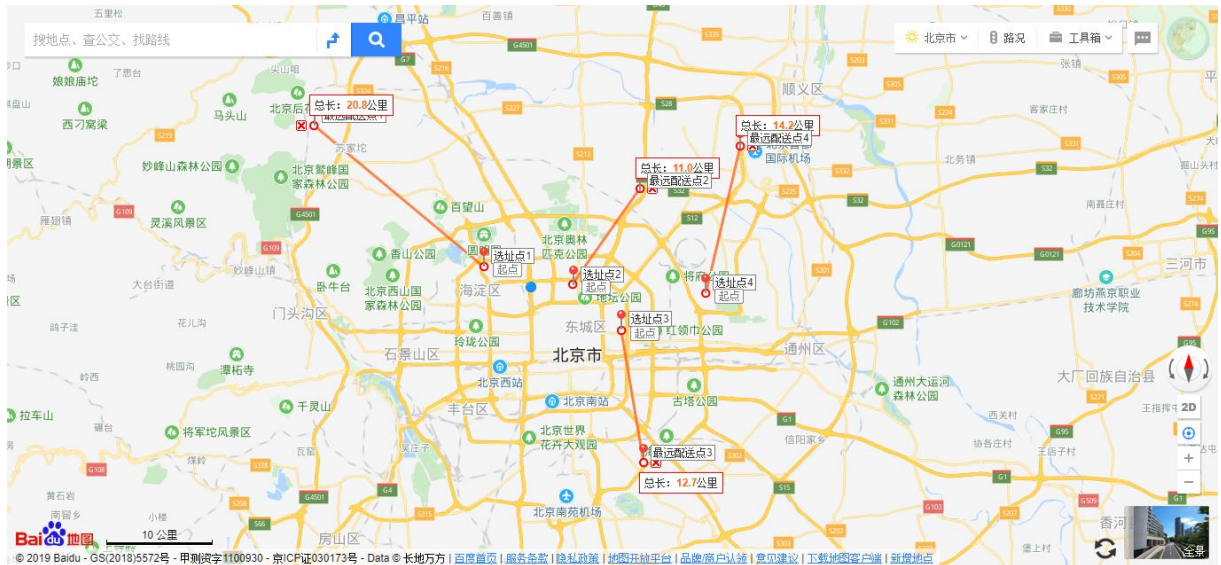


图7 配送中心离最远配送点的距离

405

表3 配送中心信息表

经纬度坐标	实际地址	最远配送点	最远配送距离/Km
(116.310859,39.987404)	北京市海淀区苏州街29号-1栋	(116.124869,40.106593)	20.8
(116.409605,39.97108)	北京市朝阳区安华路	(116.483832,40.050353)	11.0
(116.463813,39.932435)	北京市朝阳区白家庄路3号A106	(116.487369,39.819119)	12.7
(116.555957,39.964389)	北京市朝阳区东坝中路28号院	(116.594178,40.086961)	14.2

### 3.5 优化 LK-means 算法选址结果

#### 3.5.1 python 输出结果

将原始数据代入代码中,输出最终聚类结果见图8,图中四个菱形点为最终选取的四个配送中心选址点。

410

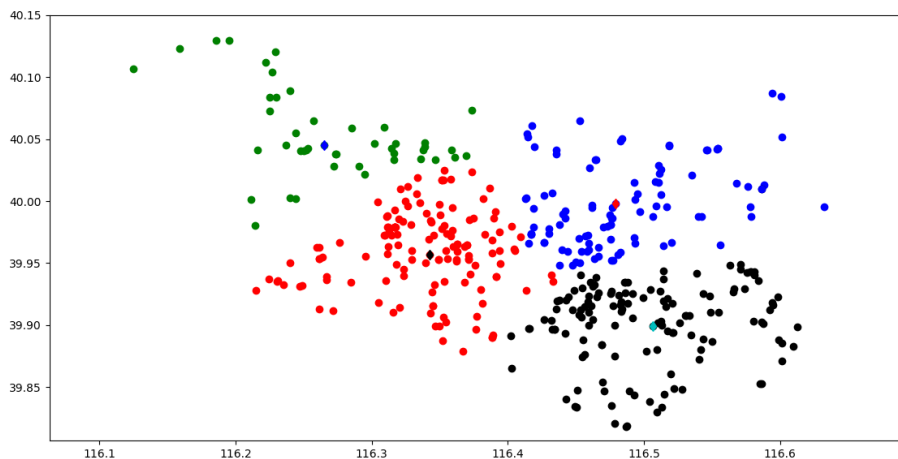


图 8 基于 LK-means 算法选址结果

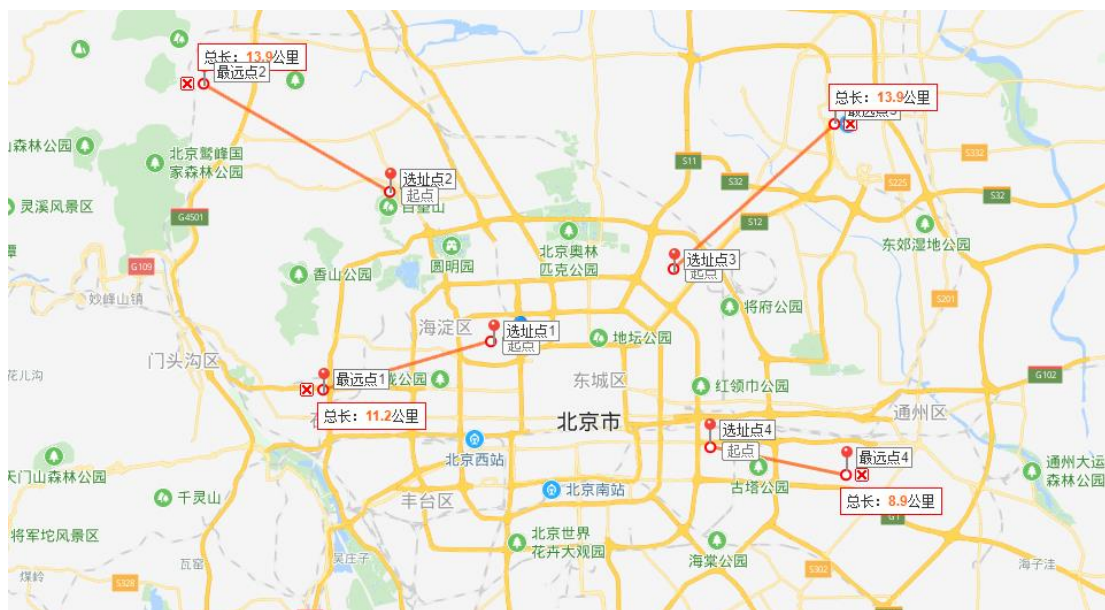
### 3.5.2 选址坐标

根据聚类最终结果，拟定 4 个点为 Lk-means 聚类的新质心，代入百度拾取坐标系统后得到最终的选址结果见表 4。

表 4 基于 LK-means 算法配送中心选址坐标

经纬度坐标	实际地址
(116.342564,39.956537)	北京市海淀区高粱桥斜街 11 号
(116.264791,40.044839)	北京海淀区西北旺黑山扈路 105 号
(116.479265,39.997868)	北京市朝阳区望京园 401 号
(116.506738,39.898951)	北京市朝阳区广渠东路 3-2-7 号

python 输出结果显示，这四个选址点均存在最远配送点，利用百度地图的距离测算每个选址点距离最远配送点的距离见图 9，将结果整理见表 5。



420

图 9 基于 LK-means 算法选址配送中心离最远配送点的距离

表 5 基于 LK-means 算法选址配送中心信息表

经纬度坐标	实际地址	最远配送点	最远配送距离/Km
(116.342564,39.956537)	北京市海淀区高梁桥斜街 11 号	(116.214745,39.928351)	11.2
(116.264791,40.044839)	北京海淀区西北旺黑山扈路 105 号	(116.124869,40.106593)	13.9
(116.479265,39.997868)	北京市朝阳区望京园 401 号	(116.600644,40.084412)	13.9
(116.506738,39.898951)	北京市朝阳区广渠东路 3-2-7 号	(116.60965,39.882754)	8.9

## 4 结论

### 4.1 结果分析

综合以上实例分析结果，可以发现优化后的 LK-means 算法相比 K-means 算法存在一定的优势。

#### (1) 经济性对比分析

根据上文中北京市快递物流服务业发展现状表明，近几年来，北京市物流仓储与中转用地以及物流设施的外扩现象明显，多集中于非中心城区。但本文提出的城市快递物流配送中心建设与物流运营外扩现象相反，在中心城区建立一定数量的流通型配送中心，为验证优化算法的可执行能力与合理性，本文对基于传统算法与优化算法的两种选址结果的以下因素进行对比。

在时间成本上，时间窗成本体现了快递配送的时效性，成本越低，配送速度越快。由于 LK 优化算法在计算时已经将时间窗成本列入逻辑距离中，且从选址结果可以看出，LK-means 优化算法配送中心的平均配送范围相比传统算法更小一些，更加均衡，人口密集区域的配送范围最小。可以预知，随着快递物流业务量的不断增加，这种差异将更加明显。

在建设成本上，由于 LK-means 算法不考虑新建配送中心，而是在原有网点的基础上适当扩容，故其建设成本可认为极低，不包含地价成本。而传统 k-means 算法的聚类中心是生成所有样本点以外的新的点，故需要在北京市中心城区新建配送中心，其建设成本远高于优化算法的建设成本。

在惩罚成本上，配送中心基本在服务范围内开展配送活动。从选址结果可以看出，传统 k-means 算法的选址结果明显超出配送中心的服务半径，所以所承担的惩罚成本增加。而 LK-means 算法的选址结果更加合理，均未超出配送中心的服务半径。

从运输成本看，LK-means 算法相比 k-means 算法，其优化点在于运输成本与时间窗成本代入考量中，故虽然两种算法的运行结果均位于中心城区，但 LK-means 算法的运输成本会低于传统 k-means 算法的运输成本。

450 综上所述,从总成本看,选址结果显示,LK-means 算法的总成本最低,相比 K-means 算法有效降低了快递物流行业和消费者所重视的建设成本、时间窗成本、配送运输成本和配送距离成本,同时,在中心城区开放配送中心,提高了货物流转效率,缩短了配送中心与需求点的距离,提升了快递物流的服务质量。

#### (2) 社会性分析

结合北京市地图可以看出,配送中心所在地属于北京市中心城区,路网密集,交通可达度较高。在中心城区设置配送中心能够提高快递配送效率,提高消费者满意度。

455 从北京市城区规划的角度看,北京市人口众多,地价高昂,仓储用地数量稀少。使用 LK-means 算法规划选址点贴合北京市中心城区仓储用地稀少的实际情况,不增添新的中心选址点,而选择在原有网点上改建,减少房屋建设成本投入;其次,配送点处于高新技术产业园区和大学地段,人口密度较大,各方面经济发达程度高,是快递物流需求增长的核心地区。将配送中心设置在这样的区域相对设置在城郊来说更有优越性。

460 综上所述,优化 LK-means 算法相比 K-means 算法有一定的优越性,更符合现代物流的需求。

## 4.2 总结

465 与传统的物流系统资源分配方法相比,本文研究的是基于大数据的快递物流资源优化配置问题,详细研究的是配送中心选址问题,信息化与智能化的特点符合现代物流的实际需求,更具有实操性、可行性和有效性。本研究的具体创新点如下:

使用 LK-means 算法规划选址点贴合北京市中心城区仓储用地稀少的实际情况,不增添新的中心选址点,而选择在原有网点上改建,减少房屋建设成本投入。

470 基于大数据进行资源配置研究,结合智慧物流模式的信息与业务数据化的特点,促进大数据技术与物流各业务模块的发展深度融合,从而提升物流系统整体的信息化水平与智能化水平,优化结果更加具有实际价值,更符合现代化需求,从而推动物流业的整体和高效发展,实现降本增效的目的。

475 最近几年,北京市物流用地与物流设施外扩现象明显,多集中于非中心城区。但本文提出的城市快递物流配送中心建设与物流运营外扩现象相反,在中心城区建立一定数量的流通型配送中心,为验证优化算法的可执行能力与合理性,本文对通过传统算法与优化算法的选址结果的各项成本进行对比。

采用地图实际距离代替两点距离,考虑到城市规划选址复杂多变的地形;

将时间窗成本加入选址考量因素中,更符合当代人对于快递时效的高标准,同时提升物流配送效率有利于物流企业降本提效。

## 480 [参考文献] (References)

- 485 [1] 孙正萍.基于大数据下的智慧物流[J].物流工程与管理,2018,40(09):71-72.  
[2] 符瑞光.大数据与现代物流深度融合下的新零售模式分析[J].商业经济研究,2018(23):73-75.  
[3] Brouer B.D., Karsten C.V., Pisinger D. (2016) Big Data Optimization in Maritime Logistics. In: Emrouznejad A. (eds) Big Data Optimization: Recent Developments and Challenges. Studies in Big Data, vol 18. Springer, Cham  
[4] 邱晗光,李海南,宋寒.需求依赖末端交付与时间窗的城市配送自提柜选址-路径问题[J].计算机集成制造系统,2018,24(10):2612-2621.  
[5] 陶巍,周长荣.基于大数据的智慧物流系统架构研究[J].中国物流与采购,2022(20):74-75.DOI:10.16079/j.cnki.issn1671-6663.2022.20.003.

- 490 [6] 吴家豪,余朋林.城市物流配送中心选址方法及应用研究[J].商场现代化,2018(13):47-48.  
[7] 骆军伟.基于模糊理论的物流中心选址研究[J].物流技术,2013,32(05):356-357+488  
[8] 胡利利. 顺丰速运长沙配送中心选址评价研究[D].中南林业科技大学,2016.  
[9] 欧阳浩,王萌,黄镇谨,陈波,戎陆庆.用遗传算法解决物流中的仓库选址问题[J].制造业自动化,2014,36(02):51-52+64.  
[10] 李妮娜. 基于中心地理论的城市快递服务网点选址研究[D].北京交通大学,2015.
- 495 [11] 武方方. 基于大数据的物流配送中心选址优化研究[D].合肥工业大学,2015.  
[12] 李捷承,陶耀东,孙咏,高岑.基于 BIRCH 聚类的物流配送设施选址算法[J].计算机系统应用,2018,27(09):215-219.  
[13] 刘威. 基于 K-means 及 ERA 的多配送中心选址优化[D].大连海事大学,2017.
- 500 [14] Ting L. Lei, Richard L. Church & Zhen Lei (2016) A unified approach for location-allocation analysis: integrating GIS, distributed computing and spatial optimization, International Journal of Geographical Information Science, 30:3, 515-534, DOI:  
[15] 干彬,李海东,龚皓,凌检.大数据下可视化物流配送优化选址仿真[J].计算机仿真,2018,35(08):293-296+419.10.1080/13658816.2015.1041959  
[16] 梁非坤.基于层次分析法的物流中心选址问题研究[J].财会通讯,2011(14):60-61.
- 505 [17] 李胜馨.基于配送中心优化选址问题的研究[J].物流工程与管理,2012,34(08):77-78+71.  
[18] 易玉枚,廖可兵,张佐钊,彭辉冰.危险化学品物流网络选址-运输优化研究[J].中国安全科学学报,2011,21(06):135-140.  
[19] 李婷婷,黄晓东,李珍萍.多配送中心选址问题的数学模型及算法[J].物流技术,2011,30(15):85-88.  
[20] 肖岳峰,张昭.基于 DEA 的区域连锁便利店配送中心选址的研究[J].商场现代化,2011(05):51-53.
- 510 [21] 任永昌,邢涛,赵国强.城市物流中心单点选址模型优化及方案评选[J].辽宁工程技术大学学报(自然科学版),2010,29(04):690-693.  
[22] 谭阳波,杨怀珍,卢广玲.基于迭代重心法的 H 公司河南区域配送中心选址研究[J].物流技术,2013,32(05):308-310+342.  
[23] 王瑞,胡洁琼.基于竞争的配送中心选址研究[J].物流科技,2013,36(07):48-49+54.
- 515 [24] Eiichi Taniguchi,Michihiko Noritake,Tadashi Yamada,Toru Izumitani. Optimal size and location planning of public logistics terminals[J]. Transportation Research Part E,1999,35(3).  
[25] 孟祥海,苗鑫.连续与离散相结合的单一物流中心选址方法[J].哈尔滨工业大学学报,2006(07):1055-1058.
- 520 [26] 刘璇,杨敏华,胡兵.GIS 环境下的物流配送中心选址模型研究[J].测绘与空间地理信息,2012,35(07):138-141.  
[27] 张雷,常敏慧.配送中心选址问题的和谐搜索算法[J].微型机与应用,2012,31(20):83-86.
- [28] 白玉凤. 共同配送下区域速递配送中心选址和配送路径优化[D].北京邮电大学,2016.