

文章编号: 2095-2163(2021)11-0112-05

中图分类号: U11

文献标志码: A

# 基于迭代自组织数据分析算法的城市地下物流节点选择

朱怡欣

(上海工程技术大学 管理学院, 上海 201600)

**摘要:** 地下物流系统(Underground Logistics System——ULS)是指通过建立类似地铁的地下管道,来满足城市内部或城市之间运输货物需求的网络系统。随着快递行业的飞速发展,地下物流供应网络系统可以很好的解决现阶段物流的两个难题,第一频繁的交通拥堵,第二高昂的物流成本。本文通过迭代自组织数据分析算法(ISODATA)对A市某区域的数据收集情况,构建该区域物流节点选择模型,借助matlab编程进行求解物流网络节点群。通过定量的方法,从交通拥挤指数、一级节点转运率、节点服务覆盖率、货运量指数、建设及运营成本等进行深入研究与探讨节点选择的合理性和有效性。为后续地下通道网络设计和进一步的优化模型提供一定的指导作用。

**关键词:** 地下物流; 节点选择; ISODATA 算法

## Selection of urban underground logistics nodes based on iterative self organizing data analysis

ZHU Yixin

(School of Management, Shanghai University of Engineering Science, Shanghai 201600, China)

**[Abstract]** Underground logistics system refers to a network system that meets the needs of transporting goods within or between cities by establishing underground pipelines similar to subways. With the rapid development of the express industry, the underground logistics supply network system can well solve the two problems of logistics at this stage, the first frequent traffic congestion, and the second-high logistics cost. In this paper, the iterative self-organizing data analysis algorithm (ISODATA) is used to construct a logistics node selection model in a certain area of City A based on the collection of receipts in a certain area of the city, and to solve the logistics network node group with the help of matlab programming. Through quantitative methods, conduct in-depth research and explore the rationality and effectiveness of node selection from traffic congestion index, first-level node transfer rate, node service coverage rate, freight volume index, construction and operating costs. Provide certain guidance for the subsequent underground passage network design and further optimization model.

**[Key words]** ULS; node selection; ISODATA

## 0 引言

当前,世界各大城市交通都遇到的一个“困局”就是交通拥堵。交通拥堵不仅占用道路资源还导致了严重的石油资源浪费。目前国内主流的策略是通过不断建设交通设施,满足日益增长的交通需求,但在多数交通设施饱和的地区仍然存在较大的交通压力。因此“统筹规划地上地下空间开发”势在必行,地下物流系统逐渐被国内外学者重视起来。

地下物流系统拥有不可多得的优势,其中包括迅速高效、减轻环境污染、促进城市可持续发展等等。当然也存在建造成本高、风险大等问题。目前,中国地下工程技术水平已基本满足要求。正确选择地下物流网络的节点群、地下通道的网络设计和地下通道筹建,是当前多国政府正考虑和着手落实的重点领域,而地下物流节点的选择是其中的关键性

步骤。因此,本文通过迭代自组织数据分析算法(ISODATA),对A市某区域的数据收集情况构建该区域物流节点选择模型,借助matlab编程进行求解物流网络节点群,深入研究与探讨节点选择的合理性和有效性。

## 1 研究方法与基本假设

### 1.1 研究方法

本文利用收集的数据关于南京市仙林区的交通货运区域划分图和相应的货运OD(Origin Destination)流量矩阵(只考虑始发地和目的地的货运流量,不考虑在此之间的途经地)、各区域中心点及区域面积、各区域交通拥挤系数(为简化计算,部分数据经过处理)在保证一级节点转运率尽量低的情况下,为实现交通拥挤指数的改善和节点服务覆盖率尽量大的聚类目标,采用改进的ISODATA算

**作者简介:** 朱怡欣(1996-),女,硕士研究生,主要研究方向:物流管理。

**收稿日期:** 2021-05-30

法进行聚类。要保持地下物流的畅通该节点群需要满足以下要求:交通拥堵可以得到缓解、节点的转运率相对较小、节点覆盖面积尽量大、一二级节点的收发货量上限等。笔者须在满足上述要求的基础之上建立合理的节点物流网络节点群,因此可以考虑成是带约束的聚类问题,选择迭代自组织数据分析算法(ISODATA)。相关约束量为:一、二级节点的货运量上限、通车的地下节点之间的距离受航班和车辆行驶速度的约束<sup>[1]</sup>、进出 4 个物流园区的货物尽可能放入地下运输。在满足以上约束的条件下完成对节点的选取。

### 1.2 模型构建基本原则

本文研究了与地下物流模型构建相关文献,结合本地区实际情况,总结出以下基本原则:

(1) 网络系统化原则:通过物流系统各组织要素之间的协调、配合实现整体目标,遵循网络化原则。

(2) 社会化原则:符合社会经济发展总体战略目标,推动区域经济可持续发展。

(3) 协调统一性原则:充分利用原有物流系统相关资源,实现地上地下物流系统整合兼容。

### 1.3 基本假设

假设两节点之间的距离取直线距离,认为两节点间的地质等因素不影响建设施工。

假设南京市的地铁线路对于地下物流系统网络规划不产生影响。

假设班车返回途中也会运载同样多的货物量。

假设班车在两节点之间一直不间断发车,每天 18 h,每小时 5 班次。

## 2 地下物流节点选择模型的建立

本节主要根据该区域的实际情况,建立节点选择模型。在基本假设基础之上,构建节点选取指标体系,分析模型聚类目标及约束条件,最终实现节点群能够满足交通拥堵得到缓解、节点转运率相对较小、节点覆盖面积尽量大、一二级节点的收发货量上限等条件。

### 2.1 地下物流节点的选择指标

(1) 交通拥堵指数(Traffic Performance Index(TPI))

交通拥堵指数顾名思义表示道路网络拥堵程度。该指数取值范围为 0~10,数值越高表示拥堵程度越高。由于该地区人口稠密,因此可以推测该地区的交通拥堵与该地区的进出口量直接相关。通过对已知数据进行拟合,得到拟合正比例函数为:

$$TPI(S_i(t)) = 0.00267 S_i(t) \quad (1)$$

式中,  $S_i(t)$  表示第  $i$  区域节点第  $t$  天的货运量。

(2) 一级节点转运率( $\varphi$ )

一级节点的转运率,是指从物流园区经由最近的一级节点,转运至其它所有一级节点的货物量,所占到物流园区总出货量的百分比<sup>[1]</sup>,其表达式为式(2):

$$\varphi = \frac{\sum_{\min dy_i=1}^4 S_{\min dy_i}}{\sum_{q_i=1}^4 S_{q_i}} \quad (2)$$

式中,  $S_{\min dy_i}$  表示最近的一级节点  $\min q_{yi}$  转运至其它所有的节点的货物量,  $S_{q_i}$  为第  $q_i$  个工业园区的货运量。

(3) 节点服务覆盖率

节点服务覆盖率,是指节点服务覆盖面积占区域总面积的百分比表达为式(3):

$$INSS_i = \frac{\sum_{i=1}^n Ss_i}{S_{qure}} \quad (3)$$

式中,  $Ss_i$  表示第  $i$  个节点服务覆盖面积,  $S_{qure}$  为该区域的总面积。

(4) 节点强度

考虑加权网络边带的权重,将点的强度  $S_i$  定义为:

$$S_i = \sum_{i \in N_i} \omega_{ij} \quad (4)$$

式中,  $\omega_{ij}$  表示点  $i$  和点  $j$  的权重。两点之间的权重由货运量和之间距离共同决定如式(5):

$$\omega_{ij} = \mu_1 S(i, j) + \mu_2 S_{dis}(i, j) \quad (5)$$

式中,  $S(i, j)$  表示点  $i$  和点  $j$  之间的货运量,  $S_{dis}(i, j)$  表示点  $i$  到点  $j$  之间的距离。

南京市作为高等教育重点地区,人口密度虽不大,但货物吞吐量较大。根据货运 OD 数据,典型日货物运输总量达 32.7 万吨。综合考量地下物流系统的建设成本以及相应货物运输量,对应的权重取为  $\mu_1 = \mu_2 = 0.5$ 。

### 2.2 地下物流节点选择模型聚类目标

为选择出符合条件的物流节点群,本文采取自下而上的节点规划思路。由于二级节点直接服务于终端用户,二级节点选择决定该地区货运需求能否满足,因此本文以二级节点运行最优为目标,筛选出数量最少、服务范围最广的二级节点群。

(1) 一级节点转运率( $\varphi$ ) 尽量小

$$F_{\varphi Y} = -\varphi = -\frac{\sum_{\min dy_i=1}^4 S_{\min dy_i}}{\sum_{q_i=1}^4 S_{q_i}} \quad (6)$$

式中,  $S_{\min dyi}$  表示最近的一级节点  $\min dyi$  转运至其它所有的节点的货物量,  $S_{qi}$  表示第  $qi$  个工业园区的货运量。

(2) 节点服务覆盖率尽量大

节点服务覆盖面积占区域总面积的百分比:

$$F_{JY} = \frac{\sum_{i=1}^n Ss_i}{Squire} \quad (7)$$

式中,  $Ss_i$  表示第  $i$  个节点服务覆盖面积,  $Squire$  为该区域的总面积。

(3) 交通拥堵指数改善情况

$$F_1 = \sum_{i=1}^n [TPI_i(0) - TPI_i(t)] \quad (8)$$

式中,  $TPI_i(t)$  表示第  $i$  个节点在第  $t$  天的交通拥堵指数,  $TPI_i(0)$  表示初始当天的交通拥堵指数。

综上, 目标函数可表示为:

$$\max F_{xBest} = (F_{\varphi Y} + F_{JY} + F_1) = \frac{\sum_{i=1}^n Ss_i}{Squire} + \sum_{i=1}^n [TPI_i(0) - TPI_i(t)] - \frac{\sum_{\min dyi=1}^4 S_{\min dyi}}{\sum_{qi=1}^4 S_{qi}} \quad (9)$$

### 2.3 地下物流节点选择模型约束条件

(1) 一级节点的货物量上限约束

一级节点的地面收发货量上限为 4 000 t

$$SHY(y_i) \leq 4\,000 \quad (10)$$

(2) 二级节点的地面收发货量上限为 3 000 t

$$SHT(t_i) \leq 3\,000 \quad (11)$$

(3) 通车的地下节点之间的距离受航班和车辆行驶速度的约束(单位:m)

$$S(i, j) \leq 9\,800 \quad (12)$$

(4) 进出 4 个物流园区的货物尽可能放入地下运输

$$y_n \leq \frac{\sum_{qi=1}^4 S_{qi}}{S_{yi\max}} \quad (13)$$

节点选择模型聚类目标及约束为:

$$\max F_{Xbest} = (F_{\varphi Y} + F_{JY} + F_1)_{\max} = \frac{\sum_{i=1}^n Ss_i}{Squire} + \sum_{i=1}^n [TPI_i(0) - TPI_i(t)] - \frac{\sum_{\min dyi=1}^4 S_{\min dyi}}{\sum_{qi=1}^4 S_{qi}} \quad (14)$$

$$\text{s.t. } SHY(y_i) \leq 4\,000$$

$$SHT(t_i) \leq 3\,000$$

$$S(i, j) \leq 9\,800$$

$$y_n \leq \frac{\sum_{qi=1}^4 S_{qi}}{S_{yi\max}}$$

### 3 节点选择聚类模型的求解

由于聚类分析算法能够挖掘数据之间的内在联系, 广泛应用在数据归类处理等问题中, 相关的衍生算法应运而生。本文基于 ISODATA 算法思想, 加入一级节点聚类点的寻优约束, 尽可能保证一级节点的个数最少的同时能够有较低的一级节点转运率, 对节点选择模型进行求解。

#### 3.1 ISODATA 算法

ISODATA 算法基本思路: 首先选择某些初始值, 将样本按指标分配到各个聚类中心中, 计算各类中诸样本的距离指标函数; 按给定的要求, 将前一次获得的聚类集进行分裂与合并处理, 从而获得新的聚类中心; 重新进行迭代运算, 计算各项指标, 判断聚类结果是否符合要求; 经过多次迭代后, 若结果收敛则运算结束。具体流程如图 1 所示, ISODATA 算法参数设置详见表 1。

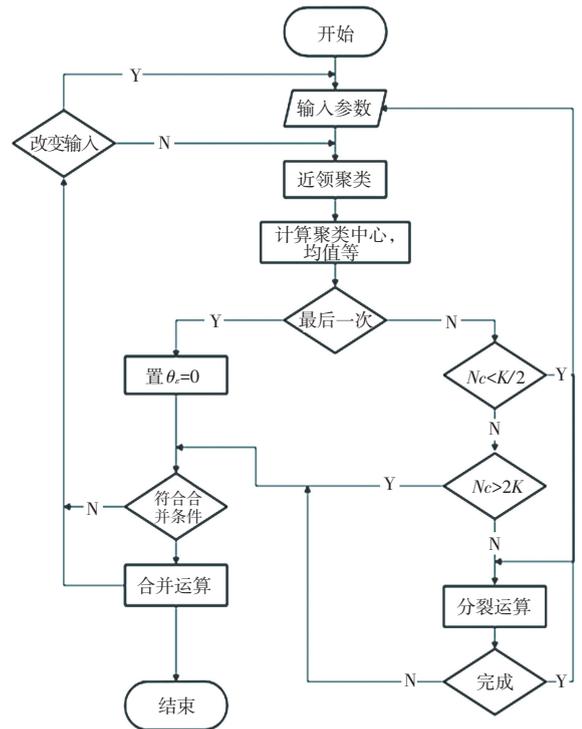


图 1 ISODATA 算法流程图

Fig. 1 ISODATA Algorithm flowchart

表 1 ISODATA 算法参数设置  
Tab. 1 ISODATA parameter settings

名称	值
预期的聚类中心数目	$K = 3$
每一聚类域中最少的样本数目	$\theta_N = 3$
一个聚类域中样本距离分布的标准差	$\theta_s = 1.8$
两个聚类中心间的最小距离	$\theta_c = 4$
一次迭代运算中可以合并的聚类中心的最多对数	$L = 1$
迭代运算的次数	$I = 100$

3.2 改进的 ISODATA 算法求解过程

依据 ISODATA 算法流程,对包含约束和目标的节点选择模型进行求解。本文基于 ISODATA 算法思想,加入一级节点聚类点的寻优约束,尽可能保证在一级节点个数最少的同时,能够有较低一级节点转运率,最终得到二级物流节点及其服务范围。求解步骤如下:

**Step 1** 样本共 110 个模式  $\{x_i = 1, 2, \dots, 110\}$ , 选取 8 个作为初始聚类中心  $\{Z_1, Z_2, \dots, Z_8\}$ 。

**Step 2** 110 个模式样本分给最近的聚类  $N_i$ , 假若  $D_j = \min \{ \|x - Z_i\|, i = 1, 2, \dots, N_c \}$ , 即  $\|x - Z_i\|$  距离最小, 则  $x \in N_i$ 。

**Step 3** 如果  $N_i < q_N$ , 则取消该样本子集, 此时  $N_c = N_c - 1$ 。

**Step 4** 修正各聚类中心:

$$Z_j = \frac{1}{N} \sum_{x \in N_i} x, j = 1, 2, \dots, N_c \quad (15)$$

**Step 5** 计算各聚类域  $S_i$  中模式样本与各聚类中心间的平均距离<sup>[2]</sup>:

$$\bar{D}_j = \frac{1}{N} \sum_{x \in N_i} \|x - Z_j\|, j = 1, 2, \dots, N_c \quad (16)$$

**Step 6** 计算全部模式样本和其对应聚类中心的总平均距离<sup>[2]</sup>:

$$\bar{D} = \frac{1}{N} \sum_{j=1}^N N_j \bar{D}_j \quad (17)$$

**Step 7** 判别分裂、合并及迭代计算一级节点转运率 ( $\varphi$ ), 直到找出各个节点中  $\varphi$  最小的那个节点。若迭代次数大于 1 次, 则置  $\theta_c = 0$ , 跳到 Step11。若  $N_c \leq K/2$ , 则对已有聚类进行分裂处理<sup>[3]</sup>。若迭代运算的次数是偶数次, 或  $N_c \leq 2K$ , 则不再分裂处理, 否则进行分裂处理。

**Step 8** 计算每个聚类中样本距离的标准差向量:

$$\sigma_j = (\sigma_{1j}, \sigma_{2j}, \dots, \sigma_{nj})^T \quad (18)$$

其中, 向量的各个分量为:

$$\sigma_{ij} = \sqrt{\frac{1}{N} \sum_{k=1}^{N_j} (x_{ik} - Z_{ij})^2} \quad (19)$$

式中,  $i = 1, 2, \dots, n$  为样本特征向量的维数;  $j = 1, 2, \dots, N_c$  为聚类数;  $N_j$  为样本个数。

**Step 9** 求每一标准差向量中  $\{\sigma_j, j = 1, 2, \dots, N_c\}$  的最大分量, 以  $[\sigma_{j \max}, j = 1, 2, \dots, N_c]$  表示。

**Step 10** 在任一最大分量集  $\{\sigma_{j \max}, j = 1, 2, \dots, N_c\}$  中, 若有  $\sigma_{j \max} > \theta_s$ , 同时又满足  $\bar{D}_j > D$  和  $N_j > 2(\theta_N + 1)$ , 即  $N_i$  中样本总数超过规定值一倍以上  $N_c \leq \frac{K}{2}$ , 则将  $Z_j$  分裂为两个新的聚类中心和, 且  $N_c$  加 1。

**Step 11** 计算全部聚类中心的距离:

$$D_{ij} = \|Z_i - Z_j\|, i = 1, 2, \dots, N_c - 1; j = i + 1, \dots, N_c \quad (20)$$

**Step 12** 将  $D_{ij} < \theta_c$  的值按最小距离次序递增排列, 即:

$$\{D_{i_1, j_1}, D_{i_2, j_2}, D_{i_3, j_3}, \dots, D_{i_L, j_L}\} \quad (21)$$

式中,  $D_{i_1, j_1} < D_{i_2, j_2} < \dots < D_{i_L, j_L}$ 。

**Step 13** 将距离  $D_{i_k, j_k}$  的两个聚类中心  $Z_{i_k}$  与  $Z_{j_k}$  合并, 得到新的中心为:

$$Z_K^* = \frac{1}{N_{i_k} + N_{j_k}} (N_{i_k} Z_{i_k} + N_{j_k} Z_{j_k}) \quad (22)$$

式中, 被合并的两个聚类中心向量分别以其聚类域内的样本数加权<sup>[4]</sup>, 使  $Z_K^*$  为真正的平均向量。

**Step 14** 若完成第 100 次迭代, 则算法结束; 否则, 根据参数调整迭代次数。如需更改输入参数, 转到第一步; 否则转到第二步。在此操作中, 每次重复操作的次数应增加 1。

4 算例及结果分析

4.1 地下物流节点的选择

根据改进的 ISODATA 算法进行聚类对南京市仙林区的交通货运区域进行分析, 最终得到一级节点 8 个, 二级节点 22 个, 具体节点详见表 2。根据货运量的情况对一级节点转运情况进行分析, 运用 MATLAB 编程得到一级节点的转运率, 相关一级节点转运率见表 3。

表 2 节点选择情况

Tab. 2 Node selection

一级节点	对应的二级节点	一级节点	对应的二级节点
894	898	870	864 855 836 876
888	891 886	832	826 839
803	792 799 796	897	896 899 892
817	814 793 827 828	853	844 857 859

表3 一级节点转运率

Tab. 3 The primary node transfer rate

距离物流园区最近的一级节点	一级节点转运率/%
817	34.12
803	30.42
888	29.14
853	77.91

#### 4.2 节点服务范围

通过聚类计算后,8个节点分属于不同的区域,如图2所示。不同区域之间的二级节点只能通过一级节点和其它区域的节点,通过地下物流系统网络进行货物的运输。由于所有的节点的服务半径在3公里范围内自由选择,通过图像覆盖的方法对服务范围进行求解,运行MATLAB编程得到各个节点实际服务范围如图3所示。

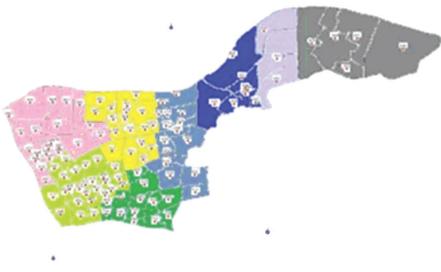


图2 ISODATA 聚类结果

Fig. 2 ISODATA Clustering results

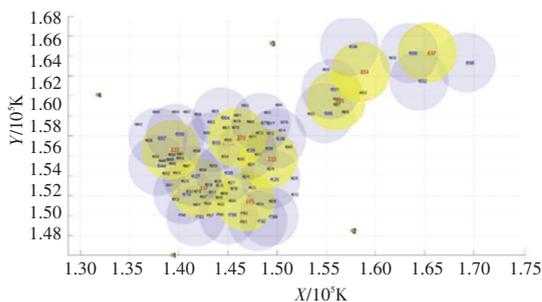


图3 各节点服务范围

Fig. 3 The scope of each node's service

#### 4.3 结果分析

南京市仙林区作为江苏省南京市高等教育产业的重点地区,具有人口密度中等、货运吞吐量较大的特点。为加快货运量的运输可在白象片区、总统府、安德门等交通枢纽及科技产业园区设置一级物流节点。优化后,整个物流网络除去重合面积有269.40 km<sup>2</sup>,各区域总面积293.76 km<sup>2</sup>,根据各节点的服务范围可求得覆盖率为91.71%,最高转运率可达到0.78,有效提高了物流网络服务的范围以及各节点转运率。

#### 5 结束语

本研究结合实际情况,运用合理的模型对地下物流节点选址问题进行分析处理,具有一定的实践应用价值。该网络采用的模型可推广到其它城市和地区的地下物流网络的建设,对于大多是地下交通类项目,如地铁规划、地下隧道规划等也具有一定的适用性,可以为政府有关部门提供合理的决策建议。

#### 参考文献

- [1] 童胜昌. 城市地下物流网络设计与优化研究[D]. 北京:华北电力大学(北京),2019.
- [2] 刘鹏. 基于高光谱技术的植物分类及状态监测方法研究[D]. 杭州:杭州电子科技大学,2019.
- [3] 郑文秀. 针对柔性电路板的AOI设备运动控制系统研究与设计[D]. 沈阳:东北大学,2017.
- [4] 朱卫星. 神经网络在电磁无损检测技术中应用的研究[D]. 哈尔滨:哈尔滨理工大学,2005.
- [5] 陈梓洋,王宇飞,钱侃,等. 自然场景下基于区域检测的文字识别算法[J]. 计算机技术与发展,2015,25(7):230-233.
- [6] 曾江源. ISODATA算法的原理与实现[J]. 科技广场,2009(7):126-127.
- [7] 李楠. 地下物流网络的节点选址和线路优化[D]. 武汉:江汉大学,2020.
- [8] IE ZEVGOLIS, A MAVRIKOS, D C KALIAMPAKOS. Construction, storage capacity and economics of an underground warehousing-logistics center[J]. Tunneling and Underground Space Technology, 2014,19: 165-173.
- [9] TANIGUCHI E, OOISHI R, KONO T. Development and future perspectives for underground freight transport systems in Japan [C]//2<sup>nd</sup> International symposium on underground freight transportation by capsule pipelines and other tube/tunnel systems, Delft. 2000.

(上接第111页)

- [10] XI J, MU K, DING T, et al. A macroscopic and hierarchical location model of regional road traffic disaster relief material repository[J]. Advances in Mechanical Engineering, 2019, 11(1).
- [11] 李栋学,刘茂. NSGA-II在应急物资储备库选址中的应用[J]. 工业安全与环保,2009,35(3):1-3.
- [12] 付德强,张伟,吴先锋,等. 基于NSGA-II的灾害随机环境下应急储备库多目标选址决策模型研究[J]. 软科学,2013,27

(11):48-52.

- [13] 周骞,刘军,邹桂兰,等. 基于多目标粒子群算法的应急物资储备库选址[J]. 长沙理工大学学报(自然科学版),2013,10(1):34-40.
- [14] 刘晋,邹瑞,韩琦,等. 基于自适应遗传算法的应急物资储备库选址及物资调配优化研究[J]. 安全与环境学报,2021,21(1):295-302.