

◎工程与应用◎

外贸集装箱运输 CFS 货运站内陆延伸的聚类分析

肖红¹,杨东援²,陈幼林³XIAO Hong¹,YANG Dong-yuan²,CHEN You-lin³

1.重庆交通大学,重庆 400074

2.同济大学,上海 200092

3.交通部科学研究院,北京 100029

1.Chongqing Jiaotong University,Chongqing 400074,China

2.Tongji University,Shanghai 200092,China

3.China Academy of Transportation Sciences Ministry of Communications,Beijing 100029,China

E-mail:xiaohonggood@sina.com

XIAO Hong,YANG Dong-yuan,CHEN You-lin.Research on Container Freight Station position to extend inside with clustering analysis.Computer Engineering and Applications,2007,43(18):184-187.

Abstract: The effect of foreign trade container goods are analyzed,the scale of the less-than container goods will be increase quickly,along with the imports and exports batches of foreign corporation decreased and modern corporation supply chain management.It's bring huge traffic pressure on the hub port highway.Using cluster analysis,corresponding implementation strategy is also put forward to ensure traffic node move freely.

Key words: foreign trade container;less-than container goods;Container Freight Station(CSF) position;cluster analysis

摘要: 对外贸集装箱的拆拼箱行为的影响因素进行分析,通过外贸生产企业的批量变化和现代企业供应链管理,外贸集装箱的拆拼箱规模迅速增长,由此对港口城市带来很大的交通压力。为了确保上海市运输节点的畅通,论文运用聚类分析提出解决方案。

关键词: 外贸集装箱;拼箱货物;CFS 货运站;聚类分析

文章编号:1002-8331(2007)18-0184-04 文献标识码:A 中图分类号:U492.3+34

随着上海洋山港的开港,上海港口的航线航班将逐渐外移。上海的两个主要集装箱港区—外高桥港区和洋山港区分布在上海市的东部,其腹地上海市、浙江、江苏都在城市西部,这些地方的货物经沪宁、沪杭高速公路后需要穿过市区,才能到达上海港。由此造成货物东西向的大流动,增加了城市道路的负荷。对于整箱运输而言,港区和货源地一旦确定,货主或货代公司选择货物运输路线应该是确定的,而对于拼箱运输来说,虽然港区和货源地已确定,但是由于集装箱的 CFS 货运站的不同,其运输线路也会发生改变,对城市道路的影响显然不同。

从图 1 可以看出,目前上海市流量密集的公路主要集中在外环线附近,外环线全线的交通量都在 30 000 pcu/日以上。外环线不仅是城市客运的主要干道,也是目前城市货运的主要干道。洋山港的通航会逐渐分流外高桥港区的流量,使外环线西段以外的交通压力得到缓解。

公路货运通道的规划与城市交通是一对辩证的关系,既互相促进,又相互制约。上海经济腹地主要集中于西向,进出港区交通与城市交通相互干扰比较严重。另外,上海市区客运车辆通行压力的不断增大,客货车辆的交叉通行,道路拥堵的现状,都无法应付未来集装箱道路集疏运量的增加。这使得笔者思考采取各种可能的手段来缓解交通压力。



图 1 上海市公路网交通量分布(2005年3月)

1 上海港拆拼箱量的快速增长

1.1 拆拼箱量的因素分析

影响上海港腹地的拼箱量的主要因素是出口货物体积、运

输费用和运输时间。

(1)和体积的关系

本文以 20ft 的标箱为集装箱的研究对象。20ft 箱 IC 型干货集装箱箱内容量为 33.2 m³,自重一般为 2 317kg,载重吨为 17.9t,总载重量 24t。货代公司对进出口货物收费的标准是体积和重量,在重量不超重的情况下以体积计费。通过笔者对上海港 34 家货代公司的调查了解到,当进出口货物小于 15 m³时,选择拼箱的概率较大。当进出口货物大于 15 m³时,选择整箱出口的概率较大。

(2)和费用的关系

进出口货物选择拼箱的运输费用比选择的整箱运输的费用要少很多,虽然它要多一部分的拼箱费用和拆箱费用,但是它由几个货主共同承担长距离的海上运输费用,总的费用就要少很多。这也是货主选择拼箱运输的主要原因。

(3)和时间的关系

相对于费用和体积来说,时间因素并不是很敏感,主要原因是海上运输的时间较长。在现代物流业的竞争条件下,需要拼箱的货物都没有等待的时间。在笔者的调查中,货代公司一般是提前在航班离港前 48 个小时或 36 个小时通知客户将货物运到货代公司的仓库,这个时间包括了拼箱时间和报关时间等,拼箱需要的作业是工人的包装和刷唛等工序,仅仅是用于货物的拼箱时间并不多。

通过以上的分析,进出口货物是否选择拼箱运输取决于进出口货物的体积和费用,主要体现在进出口货物的批量、选择拼箱和整箱的费用差两个指标上。

1.2 拆拼箱量的变化趋势分析

笔者随机调查了 401 家外贸生产企业,主要是对货主的出口批量进行了调查。调查结果显示,目前 70%的出口加工业的出口批量明显比原来要少,20%的出口加工业的出口批量和原来相比差别不明显,10%的出口加工业对这个问题不是很清楚。而对于今后出口批量变化的预测,75%的企业估计会呈下降的趋势,25%的企业不是很清楚。

现在的国际市场竞争很激烈,从上海港腹地的拼箱和整箱货种来看,纺织品和机电产品几乎占了上海港集装箱运输的一半以上。纺织品的样式、颜色都随着市场变化很快,因此决定了纺织品和服装行业批量呈下降的趋势。机电产品的变化也是很快的,不管是机器设备还是电子产品,在市场竞争中,只有满足了客户的要求才能有销路。现在商品都推崇个性化需求,信息化的快速发展为这种需求变化提供了可能,一旦客户提出订单的需求,企业都能以最快的速度研制和生产。因此,企业的生产模式应是少批量多批次的生产,企业出口批量的整体变化趋势应该是减少的。

企业供应链管理不仅仅是把供应商、制造商和顾客简单的连接起来,而是通过信息技术和现代管理技术,根据市场上的变化,在公司策略目标下对客户的需求进行快速的自我调整,而生产企业要达到这种快速的调整,供应商、制造商和顾客就需整合在一起。通过现代企业的供应链管理,可以保证生产企业的 JIT 采购、JIT 配送和 JIT 分拨物流来实现同步过程,保证了生产和销售的最小库存,由于库存量的减少势必会影响到出口加工业的采购和生产销售的批量减少。

通过以上两点的分析,笔者可以看出外贸企业的进出口货物批量呈下降的趋势,个体的变化将对整体的变化产生影响,因

此,从长江三角洲腹地的集装箱生成量来看,进出口的拼箱量的比例呈上升的趋势。据上海港口管理局的资料得知,2004 年上海港完成的集装箱吞吐量 1 455 万标箱,进出口重箱为 1 062.1 万标箱,拆拼箱量为 222 万标箱,占进出口重箱的 21%。近 5 年来,上海港的集装箱吞吐量均以 30%的速度递增,随着上海港腹地集装箱生成量的增加和拆拼箱比例的增加,今后上海港腹地的拼箱量将增长很快,在上海港港区附近的拆拼箱作业将迅速增加。

2 CFS 货运站内陆延伸的聚类分析

2.1 上海港 CFS 货运站的布局

集装箱运输的拆拼箱作业主要是在货代企业的 CFS (Container Freight Station 简写 CFS)货运站和物流中心进行。CFS 货运站是货代公司集结货物的主要场所。物流中心不是政府机构,只是由政府提供这样一个平台,由货代公司进驻。以集装箱为货物运输载体而建立的物流中心主要是进行集装箱的拆拼箱作业、货物的仓储包装、集装箱的堆放,因此,可以这样认为,CFS 货运站和堆场只是物流中心一部分。

目前上海市货代公司 70%—80%的 CFS 仓库分布在宝山的 SCT 港区附近,有一部分是在上海港浦东集装箱物流园区内和靠近外高桥港区的高东镇,还有一部分企业在外高桥保税物流园区内有自己的仓库。

洋山港的开港,上海港的航线航班将外移,SCT 港区今后不再作为集装箱码头,由于拼箱运输是货主将货物运到货代公司的 CFS 货运站,从 CFS 到港区堆场这笔费用应由货代承担,上海港区的外移将增加货代公司的费用,虽然货代公司可以通过涨价来解决这个问题,从经济学的角度,货代公司仍然要支付一部分费用,因此从长远来看,如果政府不提供相应的政策,货代公司必将仓库外移。近期,由于考虑到仓库的搬迁费用,可能有一部分企业暂不考虑将仓库外移。

根据笔者对上海市 34 家货代企业的调查了解,一个标箱的货物平均由 10 个唛头组成,通常情况下 1 个唛头是一个客户的货物,由一辆 5 吨散货车运输。这就意味着拆拼箱量的规模越大,需要的散货车数量越多,对货运站周围的城市道路的交通影响越大。今后如果政府不提供相应的措施,近洋的 CFS 货运站将外移到外高桥港区附近,这样对外环隧道带来很大的交通压力。如何缓减由于集装箱运输的拆拼箱行为对上海市道路特别是运输节点带来的压力?

笔者针对这个问题和一些专家讨论过,有的专家提出将 CFS 货运站迁移到上海市的其它地方,笔者并不赞同这个观点。原因如下:上海市的西北物流园区建在上海市的西北角,临近沪宁和沪杭高速公路,最初的初衷就是将沪宁和沪杭高速公路的货车货物以最近的距离运到西北物流园区进行拼箱作业,经过这么多年的发展,实际效果并不理想。笔者对西北物流园区调查的实际结果是,目前西北物流园区做的并不是集装箱拼箱作业,而是在做城市物流的配送。进出口货代公司并没有把他们的 CFS 仓库进驻在西北物流园区,主要的因素是运输费用和运输时间。从 CFS 到港区的堆场的运输费用由货代承担,对于货代来说是不愿意承担这笔额外的费用的;运输时间也比在港区附近的 CFS 货运站运输货物的时间长。

于是笔者提出能否将长江三角洲地区的 CFS 货运站的空间分布往长三角的内陆城市延伸?应该如何延伸?延伸到哪里

比较合适?

2.2 CFS 货运站内陆延伸的聚类分析

2.2.1 聚类分析的定义及方法

聚类分析(Cluster Analysis)是统计学所研究的“物以类聚”问题的一种方法,能够将一批样本数据按照它们在性质上的亲疏程度在没有先验知识的情况下自动进行分类。是人们认识和探索事物之间内在联系的有效手段,它既可以用作独立的数据挖掘工具,来发现数据库中数据分布的一些深入信息,也可以作为其他数据挖掘算法的预处理步骤。

常用的聚类方法有层次法、图论法和目标函数法。

(1)层次法

层次法是根据待聚类事物相似性的程度,对其进行逐层的分割和归并,从而形成一个多层的类别树状图。

(2)图论法

图论法是根据待聚类对象的模糊图的连通性以及按最小跨度树的割边进行子图的划分。

(3)目标函数法

对于每一个中心点,可以用一个目标函数来描述待选聚类的好坏程度,其中最常用的方法就是 c -均值聚类算法。这种算法定义了一个聚类的中心,通过反复迭代,极小化围绕中心的点的总的分散程度,最终确定元素所属于的类别。

目标函数法 c -均值聚类包括清晰 c -均值聚类和模糊 c -均值聚类两种。

在集装箱货运站布局的模糊聚类分析中,可供选择的城市有很多,需要通过反复迭代,确定城市所属于的类别。因此,目标函数法适合所研究问题的特征。模糊 c -均值聚类法适用于影响聚类分析的因素及其影响程度不完全确定,具有一定程度的模糊性的条件下的聚类分析。根据前文所分析的主要影响集装箱货运站的规划与建设的因素,在信息收集、描述、处理的过程中,存在一定程度的不确定性和模糊性的特征,因而选择模糊 c -均值聚类法研究集装箱货运站布局的聚类分析。

2.2.2 模糊 c -均值聚类

(1)关于模糊的 c -划分

将 μ_{ik} 扩展为元素 x_k 属于类别 \tilde{S}_i 的隶属度,可以将清晰的 c -划分推广为模糊的 c -划分。

定义 1 对于一个给定的整数 $c(2 \leq c < n)$, 设 $V_{cn} = \{ \text{所有 } c \times n \text{ 的实矩阵} \}$, 如果:

$$\mu_{ik} \in [0, 1] \quad \forall i, k \quad (1)$$

$$\sum_{i=1}^c \mu_{ik} = 1 \quad \forall k \quad (2)$$

$$0 < \sum_{k=1}^n \mu_{ik} < n \quad (3)$$

那么,模糊矩阵 $\tilde{U} = [\mu_{ik}] \in V_{cn}$ 是一个模糊的 c -划分,并记为

$$M_{fc} = \{ \text{所有 } X \text{ 上的模糊 } c\text{-划分} \}$$

(2)基于模糊的 c -划分的聚类中心及模糊 c -均值聚类法

定义 2 已知类别数, $2 \leq c < n, \tilde{U} \in M_{fc}$, 设

$$v_i = \frac{\sum_{k=1}^n \mu_{ik} x_k}{\sum_{k=1}^n \mu_{ik}} \quad \forall i \quad (4)$$

则称 v_i 是类别 i 的模糊的 c -划分聚类中心。

基于模糊的 c -划分,可以获得模糊 c -均值聚类法。模糊目标函数 c -均值聚类法亦是以极小化类别方差为目标函数,进行聚类分析的方法。模糊目标函数 c -均值聚类法的模型为

$$\min z(\tilde{U}) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|x_k - v_i\|^2 \quad (5)$$

$$\text{s.t. } v_i = \frac{\sum_{k=1}^n (\mu_{ik})^m x_k}{\sum_{k=1}^n (\mu_{ik})^m} \quad \forall i \quad (6)$$

$$\tilde{U} \in M_{fc} \quad (7)$$

其中, $m \geq 1$ 是一个给定的数。

注 1 式(5)-(7)称为 m -加权模型, m 是指数权重。 m 越大使得聚类中心更靠近隶属度较高的元素。式(5)为模糊目标函数, $\tilde{U} = [\mu_{ik}]$ 为模糊分类矩阵(且 $\mu_{ik} \in [0, 1] \quad \forall i, k$), v_i 是类别 i 的聚类中心 ($i=1, 2, \dots, C$)。

注 2 设 G 为一个对称正定的 $p \times p$ 的矩阵, 定义 R^p 空间的 G -范数为

$$\|x_k - v_i\|_G^2 = (x_k - v_i)^T G (x_k - v_i) \quad (8)$$

这里, $x_k \in R^p, v_i \in R^p$ 。

注 3 当 $G=I$ 时, 范数蜕化为一般范数, 即

$$\|x_k - v_i\|_G^2 = \|x_k - v_i\|^2 \quad (9)$$

因此, 更一般的模糊 c -均值聚类法的模型可表达为以下数学规划

$$\min z(\tilde{U}; V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|x_k - v_i\|_G^2 \quad (10)$$

$$\text{s.t. } \tilde{U} \in M_{fc} \quad (11)$$

$$v_i \in R^p, \quad i=1, 2, \dots, c \quad (12)$$

其中 $m \geq 1$ 和 G 为已知的权重和权矩阵。

目标函数式(10)的约束条件为 $\sum_{i=1}^c \mu_{ik} = 1$, 引入拉格朗日乘数 λ

$$z(\tilde{U}; V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik})^m \|x_k - v_i\|_G^2 + \lambda \left(\sum_{i=1}^c \mu_{ik} - 1 \right) \quad (13)$$

对式(13)求对于 λ, μ_{ik} 和 v_i 的偏导, 可以得到局部最优点的必要条件为

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik})^m x_k}{\sum_{k=1}^n (\mu_{ik})^m}, \quad i=1, 2, \dots, c \quad (14)$$

$$\mu_{ik} = \frac{\left(\frac{1}{\|x_k - v_i\|_G^2} \right)^{\frac{1}{m-1}}}{\sum_{j=1}^c \left(\frac{1}{\|x_k - v_j\|_G^2} \right)^{\frac{1}{m-1}}}, \quad i=1, 2, \dots, c; k=1, 2, \dots, n \quad (15)$$

式(14)和式(15)组成一个由 $(n+1) \times c$ 个方程构成的非线性方程组。对这 $(n+1) \times c$ 个方程组的求解是一个复杂的问题。

模糊目标函数 c -均值聚类法的算法步骤。

模糊目标函数 c -均值聚类算法要求优化如下输入数据:

- (1) 类别数 $c, 2 \leq c < n$;
- (2) 指数权重 $m, 1 < m < \infty$;
- (3) 确定 G -范数的 $p \times p$ 的对称正定矩阵 G ;
- (4) 确定一种产生初始隶属度矩阵, 即 c -划分 $\tilde{U}^{(0)}$ 的方法;
- (5) 给定停止准则 $\Delta = \|\tilde{U}^{(l+1)} - \tilde{U}^{(l)}\|_c \leq \varepsilon$ 中的 ε 。

具体计算步骤如下:

步骤 1 选定 $c (2 \leq c < n), m (1 < m < \infty)$ 和矩阵 G , 初始化 $\tilde{U}^{(0)} \in M_c$, 设迭代指标 $l=0$;

步骤 2 按公式(14)由 $\tilde{U}^{(l)}$ 计算 c 个模糊聚类中心 $\{v_i^{(l)}\}$;

步骤 3 按公式(15), 当 $x_k \neq v_i^{(l)}$ 时, 由 $\{v_i^{(l)}\}$ 计算新的隶属度矩阵 $\tilde{U}^{(l+1)}$, 否则令

$$\mu_{ik} = \begin{cases} 1, & j=i \\ 0, & j \neq i \end{cases}$$

步骤 4 计算停止准则 $\Delta = \|\tilde{U}^{(l+1)} - \tilde{U}^{(l)}\|_c$, 若 $\Delta > \varepsilon$, 令 $l=l+1$ 转步骤 2; 若 $\Delta \leq \varepsilon$, 结束计算;

步骤 5 采用最大隶属原则确定事物确切的归属类别, 即如果某 i_0 满足条件 $\mu_{i_0 k} = \max_{V_i} \{\mu_{ik}\}$, 则 i_0 为 x_k 的归属类别。

对于模糊目标函数 c -均值聚类算法, 从输入数据看, 需要研究的问题是:

- (1) 如何选择最优的类别数 c ;
- (2) 如何选择最优的指数权重 m , 当 $m \rightarrow \infty$ 时, 将导致模糊

c -划分蜕化为 $\tilde{U}=[1/2]$, 对 m 的经验取值一般为 $1.1 \leq m \leq 5$;

(3) 应该如何选取权矩阵 G , 权矩阵 G 确定了类别的形态, 例如选择各个特征的方差的倒数作为对应分量的权, 即取 $G = [\text{diag}(\sigma_j^{-2})]^{-1}$, 这样重新定尺度后, 可以更均衡地考虑每个特征值的差别;

- (4) 应该如何选择初始的模糊划分 $\tilde{U}^{(0)}$ 。

2.2.3 CFS 货运站内陆延伸的聚类分析

本论文选择长江三角洲地区的 16 个城市进行聚类划分。这 16 个城市有上海、南京、杭州、苏州、扬州、无锡、镇江、舟山、宁波、南通、绍兴、湖州、台州、泰州、嘉兴和常州。选择的评价指标有国内生产总值、社会消费品零售总额、固定资产投资、货运总量和政策导向系数。其中前 4 个指标的原始数据来源于《中国城市统计年鉴 2002》。政策导向系数的确定根据国家对于城市的投资情况和城市本身的发展前景。

在本论文的研究中, 沿海城市的政策导向系数为 3, 内河城市为 2, 内陆城市为 1。另一个需要说明的问题是把这 16 个城市聚为三类。因为在 16 个城市中既有沿海港口城市, 又有内河港口城市和内陆城市, 如果聚为两类, 聚类之后的效果可能不会很明显。事实上, 用 MATLAB 编写的程序将城市聚为两类, 除了上海是一类城市之外, 其他城市均为二类, 因此, 做这样的聚类就没有什么意义了。再者, 如果把这些城市聚为四类, 效果也不会很明显。所以基于以上考虑, 将选取的长三角地区的 16 个城市聚为三类。

根据本文对模糊聚类方法的叙述, 运用 MATLAB 编写程序, 得到聚类中心的各特征值如表 1, 表 2 所示。

根据表 2 的计算结果, 各城市的聚类情况为:

表 1 $m=1.5$ 模糊聚类中心

$m=1.5$	国内生产总值/亿元	社会消费品零售总额/亿元	固定资产投资/万元	货运总量/万吨	政策导向系数
第一类	4 893.00	3 921.00	19 947 000.00	49 499.00	3.00
第二类	842.00	962.00	3 711 500.00	12 216.00	1.73
第三类	201.00	146.00	1 225 700.00	5 504.00	1.47

表 2 各城市对应类别的隶属度

	第一类	第二类	第三类
上海	1.000 00	0.000 00	0.000 00
杭州	0.000 00	0.994 92	0.005 07
南京	0.000 00	0.999 97	0.000 05
无锡	0.000 00	0.198 67	0.801 32
苏州	0.000 00	0.997 18	0.002 82
宁波	0.000 00	0.995 74	0.004 26
扬州	0.000 00	0.000 01	0.999 99
常州	0.000 00	0.000 00	1.000 00
南通	0.000 00	0.000 00	1.000 00
湖州	0.000 00	0.000 11	0.999 89
镇江	0.000 00	0.000 20	0.999 80
绍兴	0.000 00	0.000 10	0.999 90
泰州	0.000 00	0.000 12	0.999 88
台州	0.000 00	0.000 01	0.999 99
舟山	0.000 00	0.004 14	0.995 86
嘉兴	0.000 00	0.022 35	0.977 65

一类城市: 上海;

二类城市: 杭州、南京、苏州、宁波共 4 个;

三类城市: 无锡、扬州、常州、南通、湖州、绍兴、镇江、台州、泰州、舟山、嘉兴共 11 个。

因此, 可以考虑选择第一、第二类城市作为在长三角地区建设公路 CFS 集装箱货运站节点的首选城市。

苏南地区外向型经济非常活跃, 外贸出口的机电产品和高新产品较多, 适箱货很大。近洋运输需经外高桥港区出港, 苏南地区和上海宝山、嘉定等集装箱生成量中的集装箱货物必须经过上海的外环隧道, 为了缓减其对上海市带来的交通压力特别是对外环隧道的压力, 笔者建议在靠近上海的苏州建立公路集装箱出口 CFS 货运站, 这样同时可以缓减沪宁高速公路的交通压力。集装箱的拼箱运输存在着规模经济效益, 如果拼箱量小, 货代企业的成本加大, 拼箱的时间和费用增加, 是不利于拼箱作业的, 由于苏南地区外向型经济非常活跃, 有足够的拼箱量, 完全可以保证集装箱 CFS 货运站在当地及周围地区拼箱货的拼箱作业。

对于进口分拨的 CFS 货运站的空间分布问题, 由于进口分拨是由不同目的地的货物拼在一起进口, 如一个 20 ft 的标箱进口上海港, 其货物的目的地有苏州、宝山和宁波, 所以, 进口分拨的作业点不应设在长三角的其它地区, 应该设在港口所在城市。另外, 进口分拨的操作流程需在海关监管的仓库进行, 也就是说进口分拨的货物不仅从港区到 CFS 货运站需要在海关监管, 进口分拨的操作也需要海关监管。因此, 为了保证外环隧道的畅通, 在考虑到上述因素的情况下, 笔者提出, 近洋进口分拨的 CFS 仓库应尽量设在上海的宝山区, 如果将其设在外高桥港区附近, 将对外环隧道带来很大的压力。

3 结论

本文通过分析今后上海港外贸拆拼箱的规模, 由此带来的城市道路的压力, 提出通过改变 CFS 的空间分布来缓减

(下转 201 页)