

# 公路货运物流的个性化推荐研究

陈 靛, 赵芬霞, 李欣曼, 潘小卓

天津商业大学理学院, 天津

收稿日期: 2023年5月14日; 录用日期: 2023年6月4日; 发布日期: 2023年6月16日

## 摘 要

近年来为改善公路物流行业环境, 涌现了一批物流公共信息平台, 这些物流公共信息平台虽然在一定程度上改善了货源信息匹配度不高的问题, 但仍需要货物运输需求方通过关键词检索主动进行操作, 信息匹配效率低下。针对上述情况, 本文提出了一种基于内容的货运信息个性化推荐方法。先运用余弦相似度计算车货信息的匹配程度, 再运用AHP方法对匹配程度高的车源物流公司进行评判, 最终为相关货源推荐出最佳车源信息, 为车主和货主寻找交易方提供了便利。

## 关键词

公路货运物流, 余弦相似度, 基于内容的推荐, AHP方法

# Personalized Recommendation Research on Road Freight Logistics

Jing Chen, Fenxia Zhao, Xinman Li, Xiaozhuo Pan

School of Science, Tianjin University of Commerce, Tianjin

Received: May 14<sup>th</sup>, 2023; accepted: Jun. 4<sup>th</sup>, 2023; published: Jun. 16<sup>th</sup>, 2023

## Abstract

In recent years, in order to improve the environment of the road logistics industry, a number of logistics public information platforms have emerged, although these logistics public information platforms have improved the problem of low matching of source information to a certain extent, they still require the cargo transportation demander to actively operate through keyword retrieval, and the information matching efficiency is inefficient. In view of the above situation, this paper proposes a content-based personalized recommendation method for freight information. Firstly, the cosine similarity is used to calculate the matching degree of vehicle and cargo information, and then the AHP method is used to evaluate the vehicle source logistics company with high matching

degree, and finally the best vehicle source information is recommended for the relevant cargo source, which provides great convenience for car owners and cargo owners in finding trade partners.

## Keywords

Road Freight Logistics, Cosine Similarity, Content-Based Recommendations, AHP

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

公路货运行业是国家基础性、战略性的产业，是社会经济发展的重要组成部分，也是保障产业供应链稳定的关键环节[1]。随着我国互联网、物联网的快速发展，公路货运行业与互联网技术的融合加快了公路货运行业的数字化、信息化建设进程[2]。近年来涌现出大量基于互联网和信息技术的车货匹配平台型企业[3]。

个性化推荐是根据用户的兴趣特点和购买行为，向用户推荐用户感兴趣的信息和商品[4] [5] [6]。车货供需信息的有效匹配推荐不仅可以降低匹配成本，提高双方交易效率，还能够提高用户粘性，是物流平台能够长远发展的核心[7] [8] [9] [10]。许多学者针对车货匹配的个性化推荐进行了研究。

傅怡[11]提出了一种基于参数自适应调整的改进蚁群算法，提升了物流的运作效率，使物流运力资源得以合理配置。叶子[12]建立了一对多成本最小化的车货匹配模型，采用禁忌搜索算法得到了模型的最优解。李蒙[13]应用深度学习模型构建了物流服务交易群智推荐算法。张会凤[14]以托运意愿和承运意愿最大化作为目标函数构建车货匹配模型，验证了匹配模型的有效性和实用性。高艺[15]给出了基于 AHP 的物流供应商物流能力的评价方法，据此完成了货运信息的个性化推荐。张青杰[16]综合运用图论中顶点着色法、量子进化算法等对物流供需信息的组合匹配模型进行了研究。穆文忠[17]采用 AHP 和系统动力学的因果分析法得到了每个评价指标的权重并确定了相应指标的数据来源，实现了对物流配送方案服务质量的数值评估。方超[18]通过 AHP 法优化构建了用户偏好模型，将用户偏好模型与协同过滤算法绑定产生了推荐结果。钱勇[19]利用 AHP 方法建立了层次结构的物流服务提供商指标体系，确定出了最佳的物流服务提供商。

然而，公路货运的车货配载率低、成本一直居高不下等问题，依然是制约公路货运行业发展的重要因素。因此，针对当前物流公共信息平台服务不足的问题，本文以中国物通网数据为研究对象，结合物流服务商的综合评价，构造了车货匹配的个性化推荐模型，以期提升车货配载效率，提高物流公共信息平台的服务质量。

## 2. 基于内容的推荐系统

基于内容推荐方法的理论依据主要来自于信息检索和信息过滤，所谓基于内容的推荐方法就是根据用户过去的浏览记录来向用户推荐用户没有接触过的推荐项目。过程一般包括以下三步：

- 1) 内容表征：为每个项目抽取一些特征来表示此项目；
- 2) 特征学习：利用一个用户过去喜欢的项目的特征数据，来学习出此用户的喜好特征；

3) 生成推荐列表：通过比较用户的喜好特征与候选项目的相似度，为此用户推荐一组相关性最大的项目。

在推荐系统中，关于相似度有多种度量可以应用，而余弦相似度的计算相对简便且准确性较高，因此被广泛使用。

设  $\alpha, \beta$  是两个  $n$  维向量， $\alpha = (a_1, a_2, \dots, a_n)^T, \beta = (b_1, b_2, \dots, b_n)^T$ ，则  $\alpha$  与  $\beta$  夹角  $\theta$  的余弦可表示为

$$\cos \theta = \frac{\alpha \cdot \beta}{\|\alpha\| \|\beta\|} = \frac{\sum_{i=1}^n a_i b_i}{\sqrt{\sum_{i=1}^n a_i^2} \sqrt{\sum_{i=1}^n b_i^2}}$$

0 度角的余弦值是 1，而其他任何角度的余弦值都不大于 1；因此余弦值越接近 1，两者之间的夹角越接近  $0^\circ$ ，也就是两个向量相似度越高[17] [18]。

余弦相似度计算源于向量余弦计算的思想，通过测量两个向量内积空间夹角的余弦值来度量它们之间的相似性。本文将应用余弦相似度计算车源信息与货源信息的匹配程度。

### 3. 公路货运物流的车货匹配推荐

#### 3.1. 获取数据

本文的研究数据选自中国物通网。中国物通网是专业的一站式物流信息平台，由多位物流及电子商务专家根据国内外企业高速发展的现状，商品流通进入速度快、方位全、地域性广等特征，专业设计、精心制作的物流行业信息门户网站。

以出发地为北京、目的地为广州的货物 A 为例，为其进行车货匹配，货物 A 的货源信息如表 1 所示。

**Table 1.** Source information of goods A

**表 1.** 货物 A 的货源信息

	需求车型	货物重量
货物 A	6.8 米 高栏车	10 吨

首先在物通网车源服务页面收集以北京为出发地、广州为目的地的车源信息。为方便后期综合评判，仅选择车主为物流公司的车源信息；同时为保证数据的有效性，所有的车源信息的车辆参数选择为“6.8 米以上”，限载重量选择为“10 吨以上”，车源信息如表 2 所示。

**Table 2.** Vehicle source information

**表 2.** 车源信息

出发地	目的地	车辆参数	限载重量 (吨)	出发地	目的地	车辆参数	限载重量 (吨)
北京	广州	高栏车 12.5 米	28	北京	广州	高栏车 13 米	28
北京	广州	平板车 17.5 米	30	北京	广州	高低板车 17.5 米	30
北京	广州	高栏车 9.6 米	18	北京	广州	高栏车 12.5 米	28
北京	广州	高栏车 9.6 米	18	北京	广州	轿车运输车 13.5 米	10
北京	广州	半挂车 13 米	30	北京	广州	高栏车 13 米	30
北京	广州	轿车运输车 13.5 米	10	北京	广州	高低板车 17.5 米	20

Continued

北京	广州	高低板车 17.5 米	28	北京	广州	轿车运输车 13.5 米	10
北京	广州	高栏车 9.6 米	10	北京	广州	高低板车 17.5 米	30
北京	广州	厢式货车 6.8 米	10	北京	广州	高低板车 17.5 米	30
北京	广州	高栏车 12.5 米	28	北京	广州	厢式货车 6.8 米	10
北京	广州	高栏车 13 米	30	北京	广州	高栏车 13 米	10
北京	广州	高栏车 13 米	30	北京	广州	厢式货车 6.8 米	10
北京	广州	高低板车 17.5 米	30	北京	广州	半挂车 13 米	30
北京	广州	高栏车 13 米	30	北京	广州	轿车运输车 13.5 米	10
北京	广州	高栏车 13 米	30	北京	广州	高栏车 13 米	28
北京	广州	高栏车 6.8 米	13	北京	广州	厢式货车 6.8 米	10
北京	广州	高栏车 13 米	40	北京	广州	半挂车 13 米	30
北京	广州	半挂车 13 米	30	北京	广州	高栏车 13 米	18
北京	广州	厢式货车 6.8 米	10	北京	广州	高栏车 6.8 米	13
北京	广州	高低板车 17.5 米	28	北京	广州	厢式货车 6.8 米	15
北京	广州	高栏车 12.5 米	28	北京	广州	轿车运输车 13.5 米	10
北京	广州	高低板车 17.5 米	30	北京	广州	高栏车 12.5 米	28
北京	广州	高栏车 12.5 米	10	北京	广州	高栏车 9.6 米	18
北京	广州	厢式货车 6.8 米	10	北京	广州	高低板车 17.5 米	30
北京	广州	平板车 17.5 米	30	北京	广州	高低板车 17.5 米	30

### 3.2. 信息量化

由于收集到的数据属性的种类繁多复杂，因此首先对数据进行预处理。

数据清理：为了使数据整洁有效，所有的货源与车源信息发布时间取在一周之内，过早或者过晚的信息将不再数据集中。

数据量化：将需求车型、车辆参数按照一定的范围进行分类，用数字分出等级，进行量化。车辆类型主要以车长分类，分别为 2~5 米、5~8 米、8~12 米、12~15 米、15~17.5 米、17.5 米以上，依次用数字 1~6 表示。

依据以上量化原则对货源信息及收集到的 50 条车源信息进行量化，结果如表 3、表 4 所示。

### 3.3. 计算项目之间的相似度

下面以量化后的车源信息(车辆参数、限载重量)与货源信息(需求车型、货物重量)为研究对象，运用余弦相似度计算货源信息与车源信息的相似度，继而筛选出分值最高的 N 条车源信息作为候选车源，应用到后续发货用户选择货车的过程中。

Table 3. Information data of goods A after quantization

表 3. 货物 A 量化后信息数据

	需求车型	货物重量
货物 A	2	10

**Table 4.** Vehicle source information data after quantization**表 4.** 车源量化后信息数据表

车辆编号	车辆参数	限载重量 (吨)	车辆编号	车辆参数	限载重量 (吨)	车辆编号	车辆参数	限载重量 (吨)
车源 1	4	28	车源 18	4	30	车源 35	2	10
车源 2	5	30	车源 19	2	10	车源 36	4	10
车源 3	3	18	车源 20	5	28	车源 37	2	10
车源 4	3	18	车源 21	4	28	车源 38	4	30
车源 5	4	30	车源 22	6	30	车源 39	4	10
车源 6	4	10	车源 23	4	10	车源 40	4	28
车源 7	5	28	车源 24	2	10	车源 41	2	10
车源 8	3	10	车源 25	5	30	车源 42	4	30
车源 9	2	10	车源 26	4	28	车源 43	4	18
车源 10	4	28	车源 27	5	30	车源 44	2	13
车源 11	4	30	车源 28	4	28	车源 45	2	15
车源 12	4	30	车源 29	4	10	车源 46	4	10
车源 13	5	30	车源 30	4	30	车源 47	4	28
车源 14	4	30	车源 31	5	20	车源 48	3	18
车源 15	4	30	车源 32	4	10	车源 49	5	30
车源 16	2	13	车源 33	5	30	车源 50	5	30
车源 17	4	40	车源 34	5	30			

利用表 3 和表 4 中的信息计算货物 A 与系统车源的相似度，结果如表 5 所示。

**Table 5.** Similarity between goods A and vehicle source of the system**表 5.** 货物 A 与系统车源相似度

	车源 1	车源 2	车源 3	车源 4	车源 5	车源 6	车源 7	车源 8	车源 9	车源 10
货物 A	0.998460	0.999480	0.999480	0.999480	0.997898	0.983282	0.999786	0.995580	1.000000	0.998460
	车源 11	车源 12	车源 13	车源 14	车源 15	车源 16	车源 17	车源 18	车源 19	车源 20
货物 A	0.997898	0.997898	0.999480	0.997898	0.997898	0.998999	0.995229	0.997898	1.000000	0.999786
	车源 21	车源 22	车源 23	车源 24	车源 25	车源 26	车源 27	车源 28	车源 29	车源 30
货物 A	0.998460	1.000000	0.983282	1.000000	0.999480	0.998460	0.999480	0.998460	0.983282	0.997898
	车源 31	车源 32	车源 33	车源 34	车源 35	车源 36	车源 37	车源 38	车源 39	车源 40
货物 A	0.998868	0.983282	0.999480	0.999480	1.000000	0.983282	1.000000	0.997898	0.983282	0.998460
	车源 41	车源 42	车源 43	车源 44	车源 45	车源 46	车源 47	车源 48	车源 49	车源 50
货物 A	1.000000	0.997898	0.999774	0.998999	0.997898	0.983282	0.998460	0.999480	0.999480	0.999480

对计算得出的所有余弦相似度值取 TOP-10 作为候选推荐车源，结果为[0.999786, 1.000000, 1.000000, 0.999786, 1.000000, 1.000000, 1.000000, 1.000000, 1.000000, 0.999774]，其对应车源分别为车源 7, 车源 9, 车源 19, 车源 20, 车源 22, 车源 24, 车源 35, 车源 37, 车源 41, 车源 43。

### 3.4. 基于 AHP 方法的物流公司评判

针对初步推荐的 10 条车源信息, 对其所属物流公司其他信息进行收集, 包括运输的轻货价格及重货价格、运输时效、企业诚信指数、企业年限[20], 具体信息如表 6 所示。

Table 6. Logistics company information

表 6. 物流公司信息

车辆编号	车辆参数	轻货价格 (元/立方)	重货价格 (元/吨)	运输时效 (天)	诚信指数	企业年限
7	高低板车 17.5 米	100	300	3	96	3
9	厢式货车 6.8 米	150	430	3	190	11
19	厢式货车 6.8 米	160	600	3	140	6
20	高低板车 17.5 米	180	200	3	120	4
22	高低板车 17.5 米	90	240	3	120	4
24	厢式货车 6.8 米	150	500	2	100	5
35	厢式货车 6.8 米	170	470	4	170	9
37	厢式货车 6.8 米	120	500	2	100	2
41	厢式货车 6.8 米	170	470	4	118	5
43	高栏车 13 米	220	750	3	106	3

#### 3.4.1. AHP 方法确认指标权重

应用 AHP 方法求解指标权重。对 5 个指标进行两两比较, 得到指标因素判断矩阵如表 7 所示。

Table 7. Index factor judgment matrix

表 7. 指标因素判断矩阵

C	轻货价格 (元/立方) C1	重货价格 (元/吨) C2	运输时效(天) C3	诚信指数 C4	企业年限 C5
轻货价格(元/立方) C1	1	3/5	1/2	2/3	4/5
重货价格(元/吨) C2	5/3	1	4/5	1	5/4
运输时效(天) C3	2	5/4	1	5/4	3/2
诚信指数 C4	3/2	1	4/5	1	6/5
企业年限 C5	5/4	4/5	2/3	5/6	1

设

$$A = \begin{pmatrix} 1 & \frac{3}{5} & \frac{1}{2} & \frac{2}{3} & \frac{4}{5} \\ \frac{5}{3} & 1 & \frac{4}{5} & 1 & \frac{5}{4} \\ 2 & \frac{5}{4} & 1 & \frac{5}{4} & \frac{3}{2} \\ \frac{3}{2} & 1 & \frac{4}{5} & 1 & \frac{6}{5} \\ \frac{5}{4} & \frac{4}{5} & \frac{2}{3} & \frac{5}{6} & 1 \end{pmatrix}$$

首先将判断矩阵  $A$  每行元素相乘，得

$$M = (0.16, 1.66667, 4.6875, 1.44, 0.55556)^T,$$

对向量  $M$  的元素开 5 次方根得

$$\tilde{w} = (0.69314, 1.10757, 1.36203, 1.07565, 0.88909)^T,$$

将  $\tilde{w}$  标准化得

$$w = (0.13518211, 0.216005582, 0.265633889, 0.209781759, 0.173396659)^T,$$

进一步计算，得最大特征值为

$$\lambda_{\max} = \frac{1}{5} \sum_{i=1}^5 \frac{(Aw)_i}{w_i} = 5.001144$$

下面进行一致性检验。一致性指标为  $CI = \frac{\lambda_{\max} - 5}{5 - 1} \approx 0.00029$ ；一致性比率为  $CR = \frac{CI}{RI} = \frac{0.00029}{1.12} \approx 0.000255$ ，其中平均随机一致性指标  $RI$  如表 8 所示。

**Table 8.** Average random consistency index  $RI$   
**表 8.** 平均随机一致性指标  $RI$

维数 $n$	1	2	3	4	5	6	7	8	9	10
指标值 $RI$	0.00	0.00	0.52	0.89	1.12	1.26	1.36	1.41	1.46	1.49

一致性比率  $CR = 0.000255 < 0.1$ ，通过了一致性检验。因此，指标的权重向量为

$$w = (0.13518211, 0.216005582, 0.265633889, 0.209781759, 0.173396659)^T,$$

### 3.4.2. 计算物流公司综合得分

首先根据上述评判指标因素的评判标准建立评分等级，建立四级评判评语集  $V = \{v_1, v_2, v_3, v_4\}$ ，分别表示“不好”、“一般”、“较好”、“很好”，并设置评判等级分值分布如表 9 所示。

**Table 9.** Judge the grade score  
**表 9.** 评判等级分值

评价等级/得分	不好/30	一般/60	较好/75	很好/90
轻货价格(元/立方)	大于 200	150~200	100~150	80~100
重货价格(元/吨)	550~750	450~550	350~450	200~350
运输时效(天)	大于 4	4	3	2
诚信指数	90~110	110~140	140~170	170~200
企业年限	1~3	3~5	5~8	8~12

以 7 号物流公司为例，根据表 9 评判等级为物流公司的各个指标因素进行打分，得到 7 号物流公司的得分，如表 10 所示。



**Table 10.** Score of each index factor of No.7 Logistics Company**表 10.** 7 号物流公司各个指标因素得分

车辆编号	轻货价格(元/立方)	重货价格(元/吨)	运输时效(天)	诚信指数	企业年限
7	90	90	75	30	30

根据各指标权重与各指标具体得分结合计算加权平均，对 7 号物流公司的综合评分进行计算：

$$\begin{aligned}
 D_7 &= 0.13518211 \times 90 + 0.216005582 \times 90 + 0.265633889 \times 75 \\
 &\quad + 0.209781759 \times 30 + 0.173396659 \times 30 \\
 &= 63.02479
 \end{aligned}$$

7 号物流公司的综合得分为 63.02479。类似可得其余物流公司的综合分数，结果如表 11 所示。

**Table 11.** Each index factor score and comprehensive score of 10 logistics companies**表 11.** 10 家物流公司各个指标因素得分及综合分数

车辆编号	轻货价格 (元/立方)	重货价格 (元/吨)	运输时效(天)	诚信指数	企业年限	综合分数
7	90	90	75	30	30	63.02479
9	75	75	75	90	90	80.74768
19	60	30	75	60	75	60.10529
20	60	90	75	60	60	70.46468
22	90	90	75	60	60	74.52014
24	75	60	90	30	60	63.7033
35	60	60	60	75	90	68.34863
37	75	60	90	30	30	58.5014
41	60	60	60	60	60	60
43	30	30	75	30	30	41.95353

根据综合得分进行排序，得分最高的 TOP5 作为最终候选车源，依次分别为车源 9、车源 22、车源 20、车源 35、车源 24。

#### 4. 结论

本文以物通网数据为研究对象，首先应用余弦相似度计算货源信息与车源信息的相似度，筛选出了 TOP-10 候选车源；其次通过 AHP 方法确定物流公司的评判指标权重，给出了 TOP-10 候选车源的综合得分，基于此得到了最佳车源匹配的个性化推荐。此方法既满足了货主对个性化货运信息推荐的需求，同时也提高了车辆资源的利用率，为物流行业信息化水平的发展提供了一定的理论基础。与此同时，由于数据获取艰难以及模型相对简单，推荐的精准度还有待提升。利用协同过滤模型将车源信息和货源信息的匹配时多次模拟以提高准确性是未来需要进一步研究的方向。

#### 资助项目

大学生创新训练计划项目(202110069006)。



## 参考文献

- [1] 欧仁昆, 杨鹏. 绿色公路货运物流发展对策研究[J]. 中国市场, 2021(1): 161-163
- [2] 何鑫. 公路货运论坛及其帖子推荐系统的设计与实现[D]: [硕士学位论文]. 南京: 南京大学, 2019.
- [3] 贾兴洪, 海峰, 董瑞. 车货匹配双边平台单归属用户比率提升控制设计[J]. 计算机集成制造系统, 2017, 23(4): 903-912.
- [4] 张兀. 平台个性化推荐研究评述及展望[J]. 经济研究导刊, 2022(6): 118-120.
- [5] 李学超, 张文德, 曾金晶, 余芳. 推荐系统领域研究现状分析[J]. 情报探索, 2019(1): 112-119.
- [6] 刘春霞, 武玲梅, 谢小红. 推荐系统评估研究综述[J]. 现代计算机(专业版), 2018(24): 11-15+20.
- [7] 胡丽娜. 大数据时代下的电子商务物流个性化信息服务模式研究[J]. 中国储运, 2021(3): 157-158.
- [8] 刘馨蔚. 基于大数据时代下电子商务个性化推荐的研究与应用[J]. 商业流通, 2022(5): 23-25.
- [9] 孙秋天, 张凯, 王玄平, 崔宏兵. 基于余弦相似度的货品属性关联度研究[J]. 起重运输机械, 2021(13): 28-32.
- [10] 刘大帅, 张杰, 徐伟杰, 赵宇. 集成模糊 Kano 模型和余弦相似度的顾客需求满意度测评[J]. 包装工程, 2022, 43(14): 166-171.
- [11] 傅怡. 基于改进蚁群算法的货运物流平台车货匹配优化方法[D]: [硕士学位论文]. 合肥: 合肥工业大学, 2020.
- [12] 叶子. 基于信用评价体系的一对多车货匹配研究[D]: [硕士学位论文]. 成都: 西南交通大学, 2021.
- [13] 李蒙, 李文敬. 基于深度学习的物流服务交易群智推荐算法的研究[J]. 现代计算机, 2021, 27(34): 1-10.
- [14] 张会凤. 基于双方交易意愿的车货匹配模型研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2018.
- [15] 高艺. 物流信息平台中货运信息的个性化推荐方法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2017.
- [16] 张青杰. 基于物流供需信息的组合匹配模型研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2017.
- [17] 穆文忠. 面向物流配送领域的个性化推荐系统设计与实现[D]: [硕士学位论文]. 哈尔滨: 哈尔滨工业大学, 2014.
- [18] 方超. 基于物流数据的个性化推荐系统的研究与实现[D]: [硕士学位论文]. 南京: 南京邮电大学, 2017.
- [19] 钱勇, 李金铮, 费亚鹏, 井雪婷. 基于层次分析法下国际贸易公司物流服务提供商的选择研究[J]. 中国商论, 2021(19): 79-81.
- [20] 赵芬霞, 李坤燕, 宋爽, 王雨嫣. 基于 AHP 模糊综合评判的公路货运物流供应商能力研究[J]. 绿色科技, 2022, 24(17): 259-263.