

文章编号: 1004-4353(2020)03-0226-07

基于复杂网络的危险品运输领域 研究热点的实证分析

蔡婷婷, 刘祥伟

(安徽理工大学 经济与管理学院, 安徽 淮南 232001)

摘要: 为获取我国危险品运输领域的研究热点,对我国知网数据库中收录的606篇样本文献的关键词信息进行实证分析.首先,使用Bicomb软件统计关键词的数量并导出关键词共现矩阵数据.然后,基于复杂网络理论构建了一个包含50个节点、267条边的无向加权网络模型.最后,使用Pajek软件和中心性分析法分别对热点网络的拓扑结构和研究热点进行分析显示,热点网络结构符合无标度网络和小世界网络特征,前3位的研究热点分别是公路运输、道路运输和层次分析法.

关键词: 危险品运输; Bicomb软件; 复杂网络; Pajek软件; 研究热点

中图分类号: X951

文献标识码: A

An empirical analysis of the hot spots in the field of dangerous goods transportation based on complex network

CAI Tingting, LIU Xiangwei

(School of Economics and Management, Anhui University of Science and
Technology, Huainan 232001, China)

Abstract: In order to obtain the research hotspot in the field of dangerous goods transportation in China, the keyword information of 606 sample documents collected in CNKI full-text database of Chinese journals is analyzed empirically. First of all, Bicomb software is used to count the number of keywords and derive the co-occurrence matrix data of keywords. Then, an undirected weighted network model with 50 nodes and 267 edges is constructed based on the complex network theory. Finally, Pajek software and central analysis method are used to analyze the topology and research hotspot of hot spot network. The results show that the hot spot network structure conforms to the scale-free network and small world network characteristics. The top three research hotspots are highway transportation, road transportation and hierarchical analysis.

Keywords: dangerous goods transportation; Bicomb software; complex network; Pajek software; research hotspot

0 引言

目前,各类危险品生产和使用中所存在的安全问题已经成为社会关注的重大问题.由于危险品的生产地和需求地通常是位于不同的区域,因此危险品的运输安全问题也成为学者关注的重要

对象^[1].为此,许多学者对危险品的运输进行了研究.例如:在危险品运输路径优化方面,张萌等^[2]构建了以最大事故后果最小、运输成本最少为双目标的危险品运输车辆路径优化模型,并给出了求解该模型的精确算法;李清等^[3-6]分别以运输路

收稿日期: 2020-03-21

作者简介: 蔡婷婷(1995—),女,硕士研究生,研究方向为物流系统规划设计.

基金项目: 国家自然科学基金资助项目(61402011);安徽省哲学社会科学规划项目(AHSKQ2014D50)

径最短、运输风险最小、运输成本最少、运输时间最短为优化目标建立了多目标优化模型,并分别采用改良的遗传算法、蚂蚁算法、粒子群算法对模型进行了求解。在危险品运输网络设计方面,Xin等^[7]采用极大极小后悔准则将危险品运输网络设计问题模型转化为边缘风险不确定性下的双层整数规划公式,并提出了一种鲁棒启发式算法来对其求解;项寅等^[8]构建了一个以政府的网络设计问题为上层规划、承运商的车辆路径选择问题为下层规划的双层规划模型,并给出了对其求解的算法。在危险品运输风险分析方面,Leung等^[9]基于扩展 Dempster-Shafer 证据理论,提出了一种危险品运输事故概率估计方法,并通过实例验证了该方法的有效性;Ke等^[10]提出了一种考虑多级模糊事件率的双收费策略,该策略可降低危险品在运输过程中的风险。以上学者们虽然在危险品运输路径优化、危险品运输网络设计及危险品运输风险分析方面进行了研究,但这些研究都仅是在微观层面上的,而在宏观层面对危险品运输领域的研究热点的分布情况研究得较少;因此,本文基于复杂网络理论对危险品运输领域研究热点进行实证分析。

1 数据来源及研究方法

1.1 数据来源

在中国知网 CNKI 数据库网站(<https://www.cnki.net/>)中按高级检索条件(篇名:危险品、运输)、发表时间不限、文献类型(中文文献)检索,共检索出 908 篇文献。通过手动筛选、去除重复文献和没有关键词的文献(如报纸、会议论文),最终得到 606 篇有效文献。

将筛选所得的 606 篇文献按照年度发表的数量绘制折线图,如图 1 所示。由图 1 可知,1990—2019 年我国危险品运输领域的论文数量呈先增后降的趋势,其中研究高峰出现在 2014 年,达到 62 篇。依据各年发表的文献数量可将其划分为 3 个阶段:1990—2005 年为第 1 阶段,此阶段发表的文献数量较少,共发表 23 篇;2006—2014 年为第 2 阶段,此阶段发表的文献数量呈逐年增长趋势,共发表 364 篇,年均约 40 篇;2015—2019 年为第 3 阶段,此阶段发表的文献数量呈下降趋势,

共发表 219 篇,年均约 44 篇。综上,1990—2019 年我国危险品运输领域的研究有了较大发展,但文献的发表数量仍相对较少。

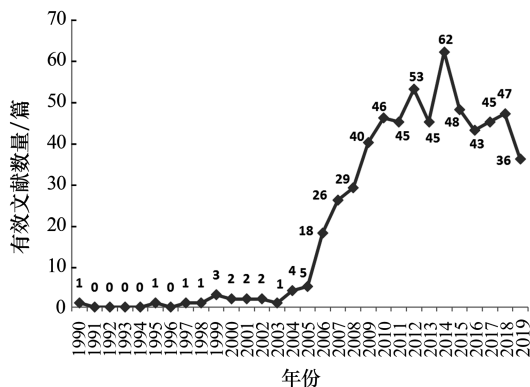


图1 1990—2019年危险品运输领域发表的有效文献数量

1.2 研究方法

复杂网络理论将复杂系统中的个体视为网络中的“节点”,将个体之间的联系或相互作用的关系视为网络中连接节点的“边”,并由此建立一个可抽象表征复杂系统的网络模型^[11]。分析复杂网络特征的常用统计指标有:度与度分布、平均路径长度、聚集系数、介数。

1) 度与度分布。度是描述单个节点属性的重要概念。在网络中,与节点 i 相接的边的总数称为节点 i 的度(k_i)。各节点度的平均值为网络的平均度,其表达式为

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i.$$

度分布是指图中各个节点度的分布情况,常用分布函数 $P(k)$ 来表示一个任意选择的节点其度为 k 的概率^[12]。

2) 平均路径长度。任意两个节点之间的平均距离值即为网络的平均路径长度 L ,其表达式为

$$L = \frac{1}{\frac{1}{2}N(N-1)} \sum_{i \neq j} d_{ij}.$$

其中 d_{ij} 为网络中任意两个节点 i 和 j 间的距离。

3) 聚类系数。聚类系数是用于衡量网络节点集聚情况的参数。网络中一个度为 k_i 的节点 v_i 的聚类系数 $C(i)$ 可表示为

$$C(i) = \frac{2E_i}{k_i(k_i - 1)},$$

其中 E_i 是节点 v_i 的 k_i 个邻节点之间实际存在的边数^[13].

4) 介数. 介数通常分为边介数和节点介数两种, 由于边在本文中并没有实际意义, 而节点代表的是关键词, 具有研究意义, 所以本文中只讨论节点之间的关系. 节点 i 的介数是指网络中经过 i 的所有最短路径的数量, 它反映的是节点 i 对其他节点之间联络的控制作用^[14]. 节点 i 的介数计算公式为

$$B_i = \sum_{i,j} \sum_{l \in S_{ij}} \delta_l^i,$$

其中 S_{ij} 为 (i, j) 之间最短路径的集合.

2 基于复杂网络理论构建危险品运输领域的研究热点网络

2.1 选取节点

首先使用 Bicom (书目共现分析系统) 软件对 606 篇样本文献进行数据挖掘, 并合并“危险品道路运输”“道路运输”“道路危险品运输”等 synonym 后, 共得到 2 555 个关键词. 在统计过程中, 为了得到更准确的实证分析结果, 本文对关键词进行校对 (删除一些错误的关键词), 并将筛选出的前 50 个 (按频数排名) 关键词定义为节点, 如表 1 所示.

表 1 高频关键词统计表

序号	关键词	出现频次	百分比/%	累计百分比/%	序号	关键词	出现频次	百分比/%	累计百分比/%
1	危险品	246	9.628 2	9.628 2	26	管理	11	0.430 5	36.829 7
2	危险品运输	184	7.201 6	16.829 7	27	级联失效	11	0.430 5	37.260 3
3	道路运输	49	1.917 8	18.747 6	28	安全工程	10	0.391 4	37.651 7
4	运输	36	1.409 0	20.156 6	29	无线传感器网络	9	0.352 3	38.003 9
5	航空运输	33	1.291 6	21.448 1	30	高速公路	9	0.352 3	38.356 2
6	安全管理	31	1.213 3	22.661 4	31	复杂网络	9	0.352 3	38.708 4
7	公路运输	29	1.135 0	23.796 5	32	多目标优化	9	0.352 3	39.060 7
8	风险评价	27	1.056 8	24.853 2	33	网络设计	8	0.313 1	39.373 8
9	路径优化	26	1.017 6	25.870 8	34	应急救援	7	0.274 0	39.647 7
10	风险分析	25	0.978 5	26.849 3	35	风险控制	7	0.274 0	39.921 7
11	运输管理	21	0.821 9	27.671 2	36	车辆调度	7	0.274 0	40.195 7
12	双层规划模型	21	0.821 9	28.493 2	37	运输风险	7	0.274 0	40.469 7
13	监控系统	20	0.782 8	29.275 9	38	优化选线	7	0.274 0	40.743 6
14	RFID	19	0.743 6	30.019 6	39	运输路径	7	0.274 0	41.017 6
15	层次分析法	19	0.743 6	30.763 2	40	GPRS	7	0.274 0	41.291 6
16	危险品运输网络	18	0.704 5	31.467 7	41	时变条件	7	0.274 0	41.565 6
17	安全	17	0.665 4	32.133 1	42	安全评价	7	0.274 0	41.839 5
18	GPS	16	0.626 2	32.759 3	43	安全监管	7	0.274 0	42.113 5
19	铁路运输	15	0.587 1	33.346 4	44	遗传算法	7	0.274 0	42.387 5
20	物联网技术	15	0.587 1	33.933 5	45	模糊综合评价	6	0.234 8	42.622 3
21	路径选择	14	0.547 9	34.481 4	46	权重	6	0.234 8	42.857 1
22	GIS	13	0.508 8	34.990 2	47	风险	6	0.234 8	43.092 0
23	水路运输	13	0.508 8	35.499 0	48	军用危险品	6	0.234 8	43.326 8
24	对策	12	0.469 7	35.968 7	49	危险品运输车辆	6	0.234 8	43.561 6
25	恐怖袭击	11	0.430 5	36.399 2	50	网络优化	6	0.234 8	43.796 5

由表 1 可以看出, 高频关键词的分布主要分为以下几类:

1) 危险品运输方面. 相关研究的关键词主要

有道路运输、航空运输、公路运输、铁路运输、水路运输等, 但由于危险品货物的特殊性, 因此道路运输仍然是危险货物的主要运输方式.

2)危险品运输管理方面.相关研究的关键词主要有安全管理、运输管理、管理、安全监管.

3)危险品运输风险方面.相关研究的关键词主要有风险评价、风险分析、风险控制、运输风险.

4)危险品运输车辆路径选择方面.相关研究的关键词主要有路径优化、路径选择、优化选线、运输路径、车辆调度.

5)危险品运输安全方面.相关研究的关键词主要有安全管理、安全工程、安全评价、安全监管、恐怖袭击.

6)危险品运输涉及的技术与算法方面.相关研究的关键词主要有双层规划模型、RFID、层次分析法、GPS、物联网技术、GIS、无线传感器网络、复杂网络、多目标优化、网络设计、GPRS、遗传算法、模糊综合评价、网络优化.

2.2 边的确定

本文将关键词之间的共现关系定义为边,将出现共现关系的频数定义为权值,以此构建一个50×50的邻接矩阵N^[15]:

$$N = \begin{bmatrix} 0 & a_{1,2} & \cdots & a_{1,50} \\ a_{1,2} & 0 & \cdots & a_{2,50} \\ \cdots & \cdots & \ddots & \cdots \\ a_{50,1} & a_{50,2} & \cdots & 0 \end{bmatrix}.$$

若 2 个关键词同时出现在同一篇文献中,记为 1,表示它们之间有联系;否则记为 0.

2.3 网络拓扑模型的构建

首先对从 Bicom 软件中导出的关键词共现邻接矩阵数据进行预处理,然后将所得数据改写为 Pajek 软件可直接读取的 .net 文件格式.由于 Pajek 软件将所有矩阵格式的网络文件都视为有向图,所以需要创建一个新的网络将有向图转换成无向图.其步骤是:首先通过依次点击 Net→Transform→Arcs→Edges→Bidirected only→Min Value 菜单命令将有向图转换成无向边;然后依次点击 Info→Network→General 菜单命令就可得到网络的全部信息,如图 2 所示;最后利用 Pajek 的可视化功能的绘图命令(Draw)即可得到无向加权网络图(包含 50 个节点、267 条边),如图 3 所示.为了提高图片的清晰度,图 3 中未显示边的权重.

Number of vertices (n): 50		
	Arcs	Edges
Number of lines with value=1	0	124
Number of lines with value#1	0	143
Total number of lines	0	267
Number of loops	0	0
Number of multiple lines	0	0

图 2 研究热点的网络信息

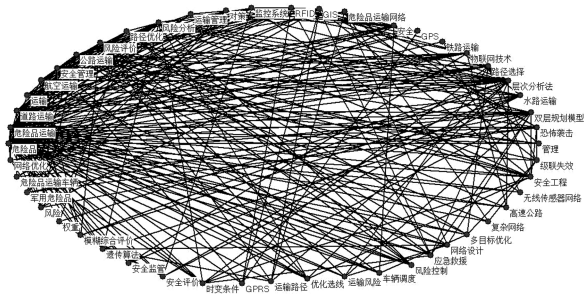


图 3 研究热点的网络结构图

3 网络结构特征分析

3.1 度与度分布

依次点击 Net→Partitions→Degree→All 菜单命令,将节点按度值进行分类,统计结果如表 2 所示.

表 2 节点度值的分布优化

节点度值	节点数目	所占比例/%	节点度值	节点数目	所占比例/%
42	1	2	12	6	12
40	1	2	11	3	6
21	2	4	9	7	14
20	1	2	8	1	2
19	1	2	7	8	16
18	1	2	6	6	12
17	1	2	5	4	8
15	1	2	4	5	10
13	1	2			

由表 2 可以看出,研究热点网络中最大的节点度值为 42,最小为 4. 点击编辑网络图标显示,度值排名前 3 的关键词分别是“危险品”“危险品运输”和“公路运输”. 这表明上述 3 个节点与其他节点的联系最紧密,在研究热点网络中处在重要的位置.

经计算,研究热点网络的平均度值是 10.68,表明网络中每个关键词均与其他 10 个关键词保

持紧密联系.对比表 2 中的各节点度值可知,度值小于平均度值的节点数(31 个)明显多于度值大于平均度值的节点数(19 个),表明网络的节点度分布得很不平衡.

图 4 为研究热点网络在普通坐标下的度分布情况,其中横坐标 k 为度值,纵坐标 $P(k)$ 表示度值为 k 的概率.由图 4 可以看出,节点度值服从幂律分布,其拟合函数为 $P(k)=0.4122k^{-0.902}$, $R^2=0.4922$.对图 4 中各节点的横、纵坐标取对数,得到该网络在对数坐标系下的节点度分布,如图 5 所示^[11].由图 5 可以看出,研究热点的网络节点在对数坐标系下的度分布呈直线递减形式,表明该热点网络具有显著的无标度网络特征^[11].

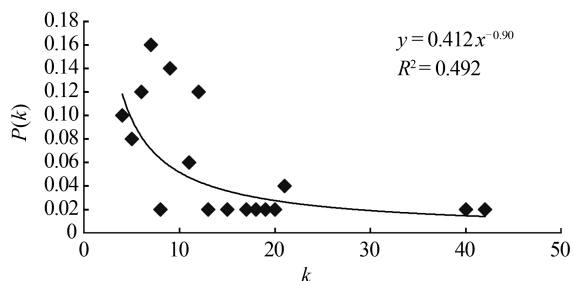


图 4 研究热点的网络度分布图

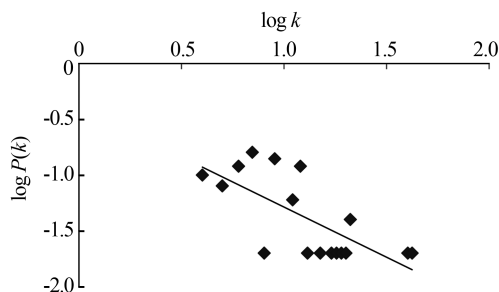


图 5 对数坐标系下的研究热点的网络度分布图

3.2 平均路径长度

通过依次点击 Net→Paths between 2 vertices→Distribution of Distances→From All Vertices 菜单命令,在报告窗口中可以看到网络的平均路径长度为 1.824 49,如图 6 所示.这说明,任意两个关键词之间平均有两条边相连.

网络中任意两个节点之间的距离的最大值称为网络的直径,记为 D .通过依次点击 Net→Paths between 2 vertices→Diameter 菜单命令,可以求得网络中距离最大的两个节点的距离.在报告窗口中可以看到,最长的一条线是从道路运

输(3)到复杂网络(31),直径为 3,如图 7 所示.这说明,在研究热点的网络中任意两个关键词之间的距离不会超过 3 条边.

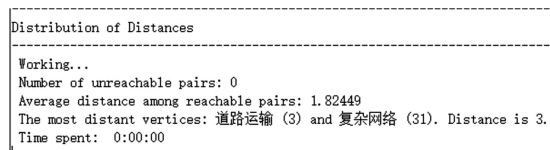


图 6 研究热点的网络平均路径长度

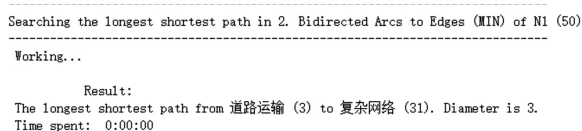


图 7 研究热点的网络直径图

3.3 聚类系数

在危险品运输研究热点网络中,网络的聚类系数(CC_1)能够反映关键词之间的联系紧密程度.通过依次点击 Net→Vector→Clustering Coefficients→ CC_1 菜单命令,可得到研究热点网络中各个节点的聚类系数 $CC_1(i)$.聚类系数 CC_1 的计算公式^[16]为

$$CC_1 = \sum_{i=1}^N CC_1(i) / N,$$

其中 N 为整个复杂网络的节点数, i 为节点编号.为方便统计,聚类系数统一只保留 4 位小数,结果如表 3 所示.

由表 3 可知,各关键词的聚类系数 $CC_1(i)$ 在 0~1 范围之内. $CC_1(i)$ 越接近 1,该顶点的集聚性越强,相反则越弱.经计算,研究热点网络的聚类系数 CC_1 为 0.556 7,表明该研究热点网络的集聚性较强.

为了判断研究热点网络是否满足小世界特性,本文通过建立一个无向的 ER 随机网络并与研究热点网络进行比较来判断网络是否满足小世界网络特性.判断的依据为二者(随机网络与研究热点网络)的比值 σ . σ 的计算式为 $\sigma = \frac{C/C_r}{L/L_r}$,其中 C 和 C_r 分别表示具有相同节点数目和相同路径数目的研究热点网络的聚类系数和随机网络的聚类系数, L 和 L_r 分别表示具有相同节点数目和相同路径数目的研究热点网络和随机网络的平均路

径长度.若 $\sigma > 1$,则可认为本文提出的研究热点网络满足小世界特征.

无向的 ER 随机网络的创建方法为:点击 Pajek 软件中的 Net→Random Network→Total No. of Arcs 菜单命令,依次输入顶点数和边数(输

入的顶点数和边数与研究热点网络的顶点数、同边数相同).经计算,随机网络的聚类系数和平均路径长度分别为 0.108 2 和 2.421 19, $\sigma=6.828 1$. σ 大于 1 表明,本文提出的研究热点网络符合小世界网络特征.

表 3 研究热点网络中各关键词的聚类系数

关键词	聚类系数	关键词	聚类系数	关键词	聚类系数
危险品	0.220 7	GPS	0.809 5	风险控制	0.714 3
危险品运输	0.194 9	铁路运输	0.500 0	车辆调度	0.666 7
道路运输	0.278 9	物联网技术	0.469 7	运输风险	0.733 3
运输	0.393 9	路径选择	0.500 0	优化选线	0.821 4
航空运输	0.418 2	GIS	0.438 1	运输路径	0.466 7
安全管理	0.392 2	水路运输	0.416 7	GPRS	0.900 0
公路运输	0.333 3	对策	0.509 1	时变条件	0.694 4
风险评价	0.367 6	恐怖袭击	0.444 4	安全评价	0.619 0
路径优化	0.490 9	管理	0.600 0	安全监督	0.833 3
风险分析	0.415 2	级联失效	0.619 0	遗传算法	0.619 0
运输管理	0.800 0	安全工程	0.606 1	模糊综合评价	0.583 3
双层规划模型	0.500 0	无线传感器网络	0.666 7	权重	0.800 0
监控系统	0.424 2	高速公路	0.900 0	风险	0.638 9
RFID	0.484 8	复杂网络	0.833 3	军用危险品	0.476 2
层次分析法	0.342 9	多目标优化	0.666 7	危险品运输车辆	0.333 3
危险品运输网络	0.333 3	网络设计	0.571 4	网络优化	0.800 0
安全	0.583 3	应急救援	0.611 1		

3.4 中心性分析

中心性是判定网络中的节点或边是否重要的指标,它常用的分析指标有点度中心性和中介中心性^[11].

3.4.1 点度中心性 点度中心性是刻画节点中心性的最直接的度量指标,它通过求出每个节点的度数来判断节点在整个网络中所处的位置.一般情况下,某节点的度值越大表明该节点在网络中越重要.由 Pajek 软件求得的前 10 个(按度值排名)网络节点的关键词如表 4 所示.

由表 4 可以看出,度值排在前 3 的关键词分别是“危险品”“危险品运输”“公路运输、层次分析法”,其中前两名的度值明显高于第 3 名的度值.这表明,前两个节点所代表的关键词在网络中的地位最为重要.另外,“道路运输”“风险分析”和“安全管理”这 3 个关键词的度值与第 3 名的度值较为接近,这表明这 3 个关键词与其他关键词的联系也较为紧密.

表 4 研究热点网络中度值排名前 10 的关键词

序号	关键词	度值
1	危险品	42
2	危险品运输	40
3	公路运输、层次分析法	21
4	道路运输	20
5	风险分析	19
6	安全管理	18
7	风险评价	17
8	GIS	15
9	双层规划模型	13
10	运输、监控系统、RFID、物联网技术、路径选择、安全工程	12

3.4.2 中介中心性 中介中心性反映的是不同节点或边在整个网络中作为媒介的能力.由 Pajek 软件求得的前 10 个(按介数值排名)网络节点的关键词如表 5 所示.

对比表 1、表 4 和表 5 可以看出,大多数频次高的关键词,其对应的度数和介数也较高,但并不

完全一致.例如,“运输”排在频数表的第 4 位,但其度值与介数分别排在第 10 位和第 17 位;“GIS”排在频数表的第 22 位,但其度值与介数分别排在第 8 位和第 9 位.这表明,某些关键词虽然出现的次数多,但它与其他关键词的联系并不多.

表 5 研究热点网络中介数值排名前 10 的关键词

序号	关键词	介数值
1	危险品运输	0.241 6
2	危险品	0.234 1
3	公路运输	0.053 9
4	道路运输	0.040 2
5	层次分析法	0.039 9
6	安全管理	0.025 9
7	风险分析	0.025 1
8	风险评价	0.022 9
9	GIS	0.020 3
10	监控系统	0.014 1

综上所述,危险品运输领域的研究热点为公路运输、道路运输和层次分析法.

4 结论

本文基于复杂网络理论构建了危险品运输领域的研究热点网络模型,并通过 Pajek 软件的可视化功能得到了研究热点网络的全景图.通过对热点网络的节点度分布、平均路径长度、聚类系数、介数等统计指标进行分析得出,本文提出的网络符合无标度网络特征和小世界网络特征.综合节点的频次、度值、介数可知:公路运输、道路运输和层次分析法是该领域的研究热点,且研究热点的分布较为均匀,相互之间的联系较为紧密.本文研究结果可为学界对危险品运输领域的深入研究提供参考.

参考文献:

[1] 王博玉.我国危险化学品安全管理现状[J].化工设计通讯,2020,46(8):180.
[2] 张萌,王能民.重大事故规避的危险品运输车辆路

径优化研究[J].运筹与管理,2018,27(8):1-9.
[3] 李清,尹宇起,胡志华.基于多目标遗传算法的危险品运输路径优化[J].合肥工业大学学报(自然科学版),2016,39(9):1278-1282.
[4] 种鹏云,帅斌,尹惠,等.基于连通可靠性的危险品运输路径选择问题[J].中国安全科学学报,2014,24(5):92-97.
[5] 宋洋,徐桢,王燕青.基于蚂蚁算法的危险品运输路径优化研究[J].安全与环境工程,2014,21(1):148-152.
[6] 薛翔,朱小林.基于优先权重粒子群算法的危险品运输路径优化建模[J].大连海事大学学报,2016,42(4):112-118.
[7] XIN C L, LETU Q G, WANG J M, et al. Robust optimization for the hazardous materials transportation network design problem[J]. Journal of Combinatorial Optimization, 2015,30(2):320-334.
[8] 项寅,姜冠群.路径-脆弱点最短距离最大化的危险品运输网络设计[J].中国安全生产科学技术,2019,15(2):57-63.
[9] LEUNG Y, LI R R, JI N N. Application of extended Dempster-Shafer theory of evidence in accident probability estimation for dangerous goods transportation[J]. Journal of Geographical Systems, 2017,19(3):249-271.
[10] KE G Y, ZHANG H W, BOOKBINDER J H. A dual toll policy for maintaining risk equity in hazardous materials transportation with fuzzy incident rate[J]. International Journal of Production Economics, 2020,227:107650.
[11] 姚红光.物流网络的延误扩散与鲁棒性分析[M].北京:国防工业出版社,2015:1-68.
[12] 王欢.基于复杂网络理论的宝鸡公交网络构建与分析[J].电气自动化,2015,37(4):24-26.
[13] 闫玲玲,陈增强,张青.基于度和聚类系数的中国航空网络重要性节点分析[J].智能系统学报,2016,11(5):586-593.
[14] 马夏夏,蔡永明,刘贵香.基于 LDA 的股市复杂网络研究热点分析[J].情报探索,2016(11):15-19.
[15] 王绍卜.基于复杂网络的港口物流领域研究热点实证分析[J].计算机系统应用,2014,23(7):246-251.
[16] 李茂.北京产业关联网络的拓扑特征研究[J].北京社会科学,2016(5):57-67.