

文章编号: 1005-0523(2020)01-0061-09

基于 PSO-BP 算法的联盟航线市场份额预测

陈 娴, 朱金福, 刘 月

(南京航空航天大学民航学院, 江苏 南京 211106)

摘要: 联盟背景下航线产品与单个航班相比, 属性复杂、合作风险高。产品属性差异导致市场分担率不同, 航线市场份额的精准预测是联盟合作伙伴选择与协同价值计算的基础。因此, 充分利用联盟运输数据, 在达美航空 QSI 指标基础上, 加入了航线竞争程度、联盟合作关系等指标, 建立了适用于联盟背景下国际航线产品市场份额预测的指标体系。考虑到影响因素与 QSI 值间的非线性关系, 引入 BP 神经网络进行市场份额与影响因素的机器学习, 建立了两者间的非线性映射。为防止 BP 神经网络陷入局部极小值, 再引入粒子群算法确定神经网络初始权值, 建立 PSO-BP 算法的 QSI 模型。结果表明, 基于 PSO-BP 算法的 QSI 模型能更好预测国际航线市场份额, 绝对误差控制在 1%。

关键词: 市场份额预测; 神经网络; 粒子群; 联盟; QSI

中图分类号: [U8] **文献标志码:** A

DOI: 10.16749/j.cnki.jecjtu.2020.01.009

国际航线网络竞争日益加剧, 为适应全球合作趋势, 一些航空公司选择加入航空联盟, 航空联盟是两家及两家以上航空公司通过航线网络协同、联盟销售、资源共享、联合采购等合作方式, 在不增加资源投入的前提下, 实现自身业务范围拓展及经济价值提升的重要手段。联盟合作形式主要有联运、联营、投资参股, 其中联营的合作方式有代码共享、包租舱位、常旅客计划、市场营销合作、特许经营等^[1]。本文所涉及的联盟合作产品为代码共享产品, 代码共享是指航班使用两家或多家航班代号, 而实际仅由一家航空公司承运的合作方式, 实际执行航班的一方通常称为运营方, 另一方称为伙伴方或市场方。

一家航空公司主要承运门户机场之间的国际航线运输任务, 无法运输其他国家/地区内部航线。受到双边管制的约束, 任何一家航空公司都无法单独的将旅客送往其他国家非门户机场。国际航线网络资源分布不均, 各国对国际航线市场的准入、运营、退出都有着严格的控制, 以保护本国航空公司的整体利益, 且由于航空公司运力资源有限, 一家航空公司无法将旅客运送到世界各地。航空联盟的出现解决了航空管制的壁垒, 联盟合作成为了各航司拓展航线网络、提高航线市场份额的重要手段, 也为民航旅客全球出行提供安全、便捷、快速的“无缝隙”服务。2018 年航线数据显示, 三大联盟——星空联盟、天合联盟、寰宇一家总共占有全球客运市场超过 60% 的市场份额。

航空领域常用的市场份额预测方法: 服务质量指数 (quality of service index, QSI) 模型、航班频率“S 曲线”市场份额模型、离散选择模型。Gregory (2003)^[2] 采用集计多项 Logit 模型, 以产品服务评分、路径衔接质量、航空公司市场占有率、机票价格、机型等为自变量预测航线产品市场份额。Coldren (2004)^[3] 又考虑到不同航线路径产品的潜在竞争程度不一, 采用广义极值模型 (GEV 模型) 从出发时间、承运人、服务产品质量三方面描述产品竞争程度。Judith (2018)^[4] 在预测航空旅客路径市场份额时进一步考虑多机场交通的便利性、起飞到达机场的延误情况、枢纽机场偏好等机场因素。Marc (2016)^[5] 利用 QSI 模型以巴塞罗那——东亚航线为例说明新飞机引进能改变非枢纽二级机场在长航线市场中的地位。国内胡皓月^[6] 通过双层 QSI 模型建立航线市场份额预测模型,

收稿日期: 2019-04-03

基金项目: 南京航空航天大学创新基金资助项目 (kfj20180723)

作者简介: 陈娴 (1995—), 女, 硕士研究生, 研究方向为交通运输经济。

通讯作者: 朱金福 (1955—), 男, 教授, 博士, 主要研究方向为机场运输规划。

赵晓松^[7]考虑联盟变量的影响,但他们在研究航线市场份额时采用线性回归模型来拟合指标系数,误差较大。

1 问题描述

联盟协同效应是指联盟后两个或两个以上的航空公司的总体效益大于各个航空公司独立运营效益之和。联盟航线合作的协同价值表现为合作干线上旅客流量的增加,通过联盟合作产生多个合作产品(代码共享的航班及航班组合),这些合作产品所带来的航空公司市场份额增加的汇总就是联盟国际航线协同价值。

从航空公司和合作伙伴航线关系角度划分,航空联盟可以分为三种类型:互补型航空联盟、平行型航空联盟以及混合型航空联盟。随着航空公司航线网络的不断扩张和航空联盟合作关系的进一步深入,联盟关系最终都会朝着混合型航空联盟发展。图1为典型混合型联盟航线示意图,本航空公司与合作伙伴航空公司在相互重合的部分干线上进行平行联盟合作,且通过互补联盟合作各自增加了3个以远点,6条新的销售路径,在已有的航线市场上如A-Hub1通过代码共享也可以拓宽销售渠道增加市场份额。

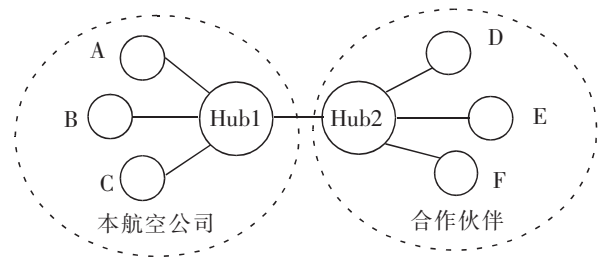


图1 联盟合作网络图

Fig.1 Airline alliance cooperation network

联盟国际航线运输产品,由联盟内合作伙伴一体化运作,但各航节的运输任务由合作伙伴分别承运。联盟环境下航线产品是某一OD对市场中的航班或航班组合形式,与传统的航线市场份额相比,联盟

背景下航线产品数量繁多,旅客在选择航班时,会充分比较OD对市场中的可行的产品,他们更愿意直飞产品。联盟合作产品由于行李直挂、贵宾厅共享、常旅客计划里程积分、统一的服务等便利可提升旅客的乘机体验。为更好的计算联盟合作产品在OD对市场中的市场份额,通过对OD产品市场份额影响因素如航班频率、起飞时刻、中转时间、距离绕航率、联盟合作关系等分析,建立了联盟背景下航线市场份额预测的QSI模型。

2 模型建立

2.1 指标选取及分析

以中国东方航空公司中美国际航线市场为例,选取CKG—JFK、TAO—SFO、XIY—LAX、SZX—SFO、CKG—LAX、CKG—SEA等OD对市场,共810个航线产品为研究对象,数据来源于2018年中美航线市场OD对数据(包括OAG航班计划数据、航空联盟表、代码共享航班表、MIDT数据等)。通过OD产品生成模块(Connection Builder)确定符合中转衔接、绕航率等要求的合作产品。借鉴达美航空QSI模型行业指标,结合联盟合作产品的实际情况,初步选择的影响因素如下:

- 1) 航班频率:路径产品中各航段频率的最小值,商务旅客更注重航班频率。
- 2) 机型座位数:大机型舒适性高,OD路径各航段机型座位数最小值。
- 3) 航空公司服务:根据skytrax星级用分值(0~10分)表示,分值越大,服务越好。
- 4) 中转次数:记录OD路径中航班衔接的中转次数。
- 5) 中转时间:记录OD路径中航班衔接的总中转时间。
- 6) 总行程时间:记录OD路径中旅客行程花费总时间。
- 7) 航线竞争程度:用赫斯曼——赫芬达尔指数(HHI)计算各OD对市场产品竞争程度: $HHI=S_1^2+S_2^2+S_3^2$,其中 S_1^2 、 S_2^2 、 S_3^2 分别为同一航空公司中转、不同航空公司中转、其他方式三种航线产品市场份额。

8) 联盟合作关系:联盟合作关系评分,分值越大,联盟合作越紧密。

9) 距离绕航率:用OD对航线产品总飞行距离与同一OD对最小飞行距离比值表示

10) 出发时间偏好:表明每一个航线产品对潜在旅客的吸引程度,根据调查得到不同时间段旅客的选择偏好,用分值(0-10分)表示。

为描述联盟合作产品与一般联运产品在常旅客计划服务、航班延误后的服务补救、行李直挂等中转服务、销售网络拓展方面对于服务质量指数(QSI值)的影响,引入联盟合作关系指标,联盟合作关系指标由常旅客计划、销售网络、服务补救、中转服务4个细分指标按照权重综合计算得到。依据路径种类和合作方式为每一航线产品合作关系评分。根据航线产品的实际情况将合作方式分为仅国内段代码共享、有航班衔接的代码共享、无双边合作,按路径种类分为直航、同公司中转、同一联盟不同公司中转、不同联盟不同公司中转、合资企业。设计调查问卷并回收有效问卷198份,其中8份问卷由中国东方航空公司负责国际事务的专家进行评分,在实际计算指标及指标权重结果时,8份专家评分问卷结果比重为1:5。反馈后的结果显示,各项评分结果较为集中,因而选择旅客选择每项指标结果的众数作为该项指标的评分结果,如表1所示。联盟合作关系的总得分取4项指标评分结果平均值。

表 1 联盟合作关系指标评分结果
Tab.1 Score results of alliance cooperation relationship

| 路径种类 | 合作方式 | 销售网络 | 常旅客计划 | 中转服务 | 服务补救 | 总得分 |
|-------------|------------|------|-------|------|------|------|
| 直航 | 代码共享 | 6 | 10 | 10 | 10 | 8.28 |
| | 无双边合作 | 0 | 10 | 10 | 10 | 5.7 |
| 同公司中转 | 仅国内段代码共享 | 4 | 10 | 9 | 10 | 7.25 |
| | 有航班衔接的代码共享 | 6 | 10 | 9 | 10 | 8.11 |
| | 无双边合作 | 0 | 10 | 9 | 10 | 5.53 |
| 同一联盟、不同公司中转 | 仅国内段代码共享 | 8 | 8 | 6 | 7 | 7.56 |
| | 有航班衔接的代码共享 | 10 | 8 | 6 | 7 | 8.42 |
| | 无双边合作 | 5 | 8 | 6 | 6 | 6.17 |
| 不同联盟、不同公司中转 | 仅国内段代码共享 | 8 | 0 | 3 | 3 | 4.25 |
| | 有航班衔接的代码共享 | 10 | 0 | 3 | 3 | 5.11 |
| | 无双边合作 | 5 | 0 | 3 | 3 | 2.96 |
| 合资企业 | 双边合作 | 6 | 10 | 10 | 10 | 8.28 |
| 权重 | | 43% | 30% | 17% | 10% | |

考虑到不同的起飞时刻对旅客的吸引力不同,将一天24小时每隔一小时划分一个时间区间,共划分为24个区间,让旅客对各航班时刻区间进行评分,了解旅客对出发时刻的偏好,调查得到的数据显示,评分最高的时间区间为10:00至10:59,18:00至19:59。

为进一步了解各影响因素与市场份额间的相关程度,表2给出了各指标与市场份额的相关性分析结果。根据分析结果剔除在0.01水平显著不相关的影响因素航班频率、起飞时间偏好指标,考虑到国际航线市场航班频率较低且飞行时间较长,起飞时刻对于旅客的吸引力不强,因此剔除相关性较弱的这两个指标具有合理性。最终选择建立QSI模型的8项指标分别为(1)机型座位数 X_1 , (2)航空公司服务 X_2 , (3)中转次数 X_3 , (4)中转时间 X_4 , (5)总行程时间 X_5 , (6)航线竞争程度 X_6 , (7)联盟合作关系 X_7 , (8)距离绕航率 X_8 。

表2 各指标与市场份额相关性分析
Tab.2 Correlation analysis between indicators and market share

| 指标 | Pearson 相关性 | 显著性水平(双侧) |
|--------|-------------|-----------|
| 航班频率 | 0.027 | 0.568 |
| 机型座位数 | 0.235** | 0.000 |
| 航空公司服务 | 0.299** | 0.000 |
| 中转次数 | -0.272** | 0.000 |
| 中转时间 | -0.065 | 0.174 |
| 总行程时间 | -0.228** | 0.000 |
| 航线竞争程度 | 0.128** | 0.007 |
| 联盟合作关系 | 0.262** | 0.000 |
| 距离绕航率 | -0.164** | 0.001 |
| 出发时间偏好 | 0.030 | 0.528 |

注:** 为在 0.01 水平(双侧)上显著相关。

2.2 QSI 模型

采用 QSI 模型通过对影响旅客选择航线产品的诸多因素分析,应用系数量化每个 OD 产品的“质量”。对于任意 OD 对,将每一个产品的 QSI 值表达为机型座位数 X_1 、航空公司服务 X_2 、中转次数 X_3 、中转时间 X_4 、总行程时间 X_5 、航线竞争程度 X_6 、联盟合作关系 X_7 、距离绕航率 X_8 等 8 个指标的函数。

$$Q_i = F(X_1, X_2, \dots, X_8) \quad (1)$$

式中: Q_i 为第 i 个合作产品的 QSI 值; X_1, X_2, \dots, X_8 为上述影响产品服务质量的因素; F 为影响因素与因变量 QSI 值间的函数关系。

$$M_i = \frac{Q_i}{\sum_{j \in J} Q_j} \quad (2)$$

式中: M_i 为第 i 个航线产品的市场份额; Q_i 为第 i 个合作产品的 QSI 值, j 为第 j 个航线产品, J 为该 OD 对市场所有的航线产品集合。

使用 QSI 模型进行航线产品市场份额预测的关键是建立各影响因素与 QSI 值之间的函数关系,为反映各影响因素与 QSI 间是否为线性关系,进一步采用 810 个合作产品样本绘制各影响因素与 QSI 值间的矩阵散点图,由图 2 可知,距离绕航率与行程总时间呈正相关,航线中转产品在中转次数相同时中转时间仍有较大差异,QSI 值与距离绕航率负相关,与机型座位数呈正相关,与行程总时间、中转时间呈负相关。除行程总时间外,其余影响因素与 QSI 值间均为非线性关系,因此初步认定 QSI 值与影响因素间为非线性关系。

以机型座位数、中转时间、联盟合作程度为自变量,QSI 值为因变量,采用多元线性回归方法拟合参数,结果显示 R 方和调整后的 R 方分别为 0.399,0.387,模型拟合效果不佳。因此,确定各影响因素与 QSI 值间为非线性函数关系,下文将进一步采用粒子群算法与神经网络拟合 QSI 值与 8 个指标间的非线性函数关系。

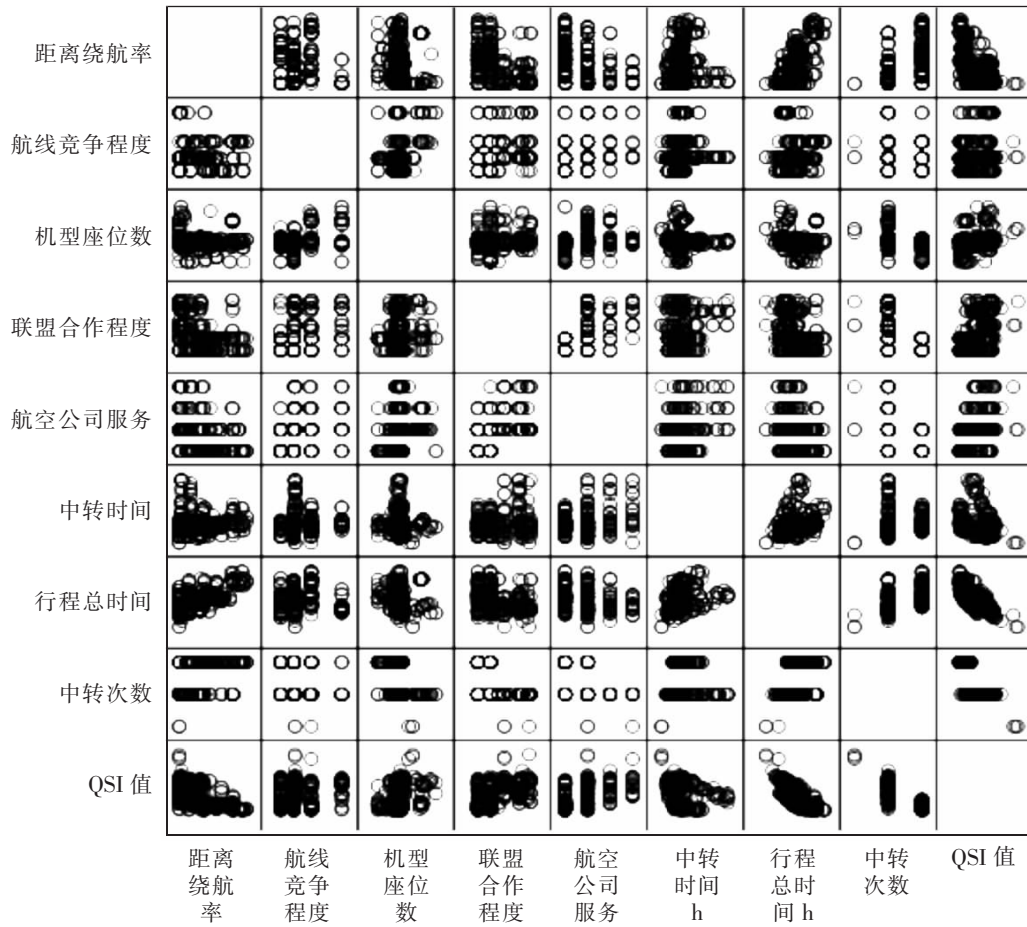


图 2 各影响因素与 QSI 值矩阵散点图
 Fig.2 Matrix scatter plot of various factors and QSI

2.3 基于粒子群算法与 BP 神经网络的市场份额预测

常用的可表达因变量与自变量间非线性关系的方法主要分为两类,一类是参数法,包括非线性回归、指数平滑法、趋势外推法等,另一类是非参数法,包括神经网络法、支持向量机法等,参数法需要首先根据实际数据做出散点图,再由散点图确定变量间函数的类型,最后通过数据拟合函数关系式的未知参数。无论多元线性回归或者是指数平滑法都需要预先确定函数类型,而联盟航线市场份额预测问题由于影响因素多,无法直接确定 QSI 值与指标间的函数类型,因此采用参数法不可行。非参数法中的支持向量机方法虽然在解决小样本、高维度以及非线性方面有一定的优势,但其对于大规模样本难以实施,而国际航线市场中 OD 产品路径众多,因此为很好的解决大样本数据间的非线性拟合问题,考虑使用人工神经网络方法。

人工神经网络(ANN)在理论上可实现任何非线性映射,BP 神经网络是一种按误差逆传播算法训练的典型多层前馈型神经网络^[8],基于粒子群算法与 BP 神经网络的市场份额预测基本思想:以机型座位数、联盟合作程度、航空公司服务等 8 个影响因素作为输入数据,产品的服务质量指数值为输出数据,通过输入与输出数据的机器学习,自适应调节网络权值与阈值,使得网络输出的 QSI 值与期望输出的 QSI 值均方误差最小。

BP 神经网络训练的关键在于初始权值的选取以及隐含层节点个数的确定。为解决 BP 神经网络收敛速度的问题,通过粒子群算法追随当前搜索到的最优值来找到全局最优值优化神经网络初始权值,减小神经网络陷入局部极小值的可能性,具体算法步骤如下:

- 1) 样本数据预处理:为减少由于数据量纲差异而造成的误差,通过 MatlabR2016a 工具箱中 mapminmax

函数将不同量纲的8个指标数据全部转化到[0,1]。

2) BP神经网络结构的确定:8个指标对应输入层节点数为8,网络输出为OD对路径产品的QSI值,输出层对应1个节点。隐含层分别设置节点数4,5,6,7,8,9,10,15。激活函数分别选择双曲正切S型函数tansig,S型对数函数logsig对比训练误差。图3为选择不同的隐含层节点数与激活函数的网络训练误差,通过不断的试验最终确定最优的隐含层节点数为7、隐含层激活函数采用S型对数函数logsig。训练算法采用Levenberg-Marquardt算法(trainlm)。

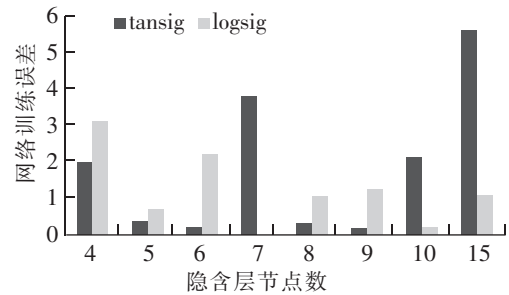


图3 隐含层节点数与激活函数误差
Fig.3 Error of different implicit layer node number and activation function

3) 粒子群算法的参数设置及粒子群位置初始化:在 n 维搜索空间内随机初始化 m 个运动粒子的位置及位置,每一个粒子代表一个可行解,设第 m 个粒子的位置、速度分别为 $X_i=(X_i^1, X_i^2, \dots, X_i^n)^T, V_i=(V_i^1, V_i^2, \dots, V_i^n)^T$,个体极值为 $P_i=(P_i^1, P_i^2, \dots, P_i^n)$,全体极值为 $P_g=(P_g^1, P_g^2, \dots, P_g^n)$ 。在本文的算例中初始化粒子数 $M=40$,本文中BP神经网络结构为8—7—1,其中8为输入的指标个数,7为隐含层最优神经元个数,1为输出数据维度,则设置空间维度为网络权值数与阈值数的和,具体计算搜索空间纬度 $n=(8+1) \times 7+(7+1)=71$,由于本例中BP神经网络的初始权值与阈值在[0,1]之间,所以PSO算法中的初始集也设置在[0,1]之间,设置最大循环迭代次数为400。

4) 粒子的位置与速度更新:粒子在空间内运动时通过计算个体极值与全体极值更新粒子位置,粒子速度及位置更新公式如下

$$V_{in}^{k+1}=wV_{in}^k+c_1r_1(P_{in}^k-X_{in}^k)+c_2r_2(P_{gn}^k-X_{in}^k) \quad (3)$$

$$X_{in}^{k+1}=X_{in}^k+V_{in}^k \quad (4)$$

式中: V_{in}^{k+1}, V_{in}^k 分别为 $k+1, k$ 次迭代时粒子的速度; P_{in}^k 为 k 次迭代时粒子自己搜索的历史最优值; P_{gn}^k 为 k 次迭代时全部粒子搜索到的最优值; X_{in}^{k+1}, X_{in}^k 分别为 k 次迭代时粒子的位置; in 为 n 维中的第 i 个粒子; gn 为 n 维中的全局最优粒子; w 为惯性权重系数,在很大程度上影响粒子全局和局部搜索能力; c_1, c_2 为学习因子; r_1, r_2 是[0,1]之间的随机数。由程序反复实验确定在本例中PSO算法采用0.9~0.3线性递减的惯性权重策略,具体的公式如下

$$w=w_{start}-\frac{w_{start}-w_{end}}{k_{max}}k \quad (5)$$

式中: k 为当前迭代次数; w_{start}, w_{end} 分别为初始和终止权重值; k_{max} 为最大迭代次数; k 为当前迭代次数。在空间内,需要将粒子的速度限定在 $[-v_{max}, v_{max}]$ 范围内。在本文中采用自适应时变策略的学习因子,分别设置 c_1 为2.5~0.5线性减小, c_2 为0.5~2.5线性增加。

5) 计算粒子的适应度值:选用BP神经网络训练的均方误差作为适应度函数,计算每一粒子的适应度值,寻找空间内个体最优及全局最优。

$$f=\frac{1}{n} \sum_{i=1}^n (y_{facti}-y_i)^2 \quad (6)$$

式中: f 为适应度值, y_{facti} 为第 i 个航线产品的QSI实际值, y_i 为BP神经网络预测得到第 i 个航线产品的QSI值, n 为航线产品的总样本数。最终算法将停止在适应度值最低的粒子位置上。

6) 迭代终止条件判断:全局最优值达到了预设的精度或者迭代次数达到设定的最大迭代次数则迭代终止,否则转到第四步更新粒子的位置及速度。

7) 将粒子群算法得到的最优值代入BP神经网络的初始权值进行神经网络的二次训练与学习,最终建立联盟背景下基于粒子群算法与BP神经网络的QSI模型。

3 预测结果及分析

对于联盟背景下国际航线市场份额问题,随机选取中美市场由中国的非枢纽门户机场至美国非枢纽门户机场航线、中国非枢纽门户机场至美国门户机场航线,选取 CKG—JFK、TAO—SFO、XIY—LAX、SZX—SFO、CKG—LAX、CKG—SEA 等 OD 对市场,共 810 个航线产品为研究对象,其中 80%的航线产品数据用于 QSI 模型的网络训练,10%的航线产品数据作为测试数据,10%的产品数据用于实际预测。表 3 给出其中的一部分网络训练数据便于参照,810 个航线产品数据中 576 个为联盟合作产品,联盟合作关系指标能较好区分联盟合作产品与一般联运产品在服务质量上的差异。

表 3 网络训练数据
Tab.3 Network training data

| OD 对 | 路径 | 距离 绕航率 | 航线竞 争程度 | 中转 时间/h | 中转 次数/次 | 行程 总时间/h | 航空公司 服务 | 联盟合作 关系 | 机型 座位数/个 |
|------|-----------------|-----------|------------|------------|------------|-------------|------------|------------|-------------|
| 1 | CKG—JFK | 1.00 | 1.04 | 0.00 | 0 | 14.92 | 10 | 5.7 | 212 |
| 1 | CKG—SFO—JFK | 1.24 | 1.04 | 1.42 | 1 | 23.40 | 6 | 2.96 | 149 |
| 1 | CKG—SFO—JFK | 1.24 | 1.04 | 1.42 | 1 | 23.40 | 6 | 7.56 | 146 |
| 1 | CKG—SFO—JFK | 1.24 | 1.04 | 0.90 | 1 | 23.20 | 6 | 2.96 | 159 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 1.67 | 1 | 17.58 | 6 | 4.25 | 291 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 1.42 | 1 | 17.42 | 6 | 8.11 | 167 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 1.42 | 1 | 17.42 | 6 | 7.56 | 167 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 1.42 | 1 | 17.42 | 6 | 5.53 | 167 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 1.42 | 1 | 17.42 | 6 | 8.42 | 167 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 1.42 | 1 | 17.42 | 6 | 7.56 | 167 |
| 1 | CKG—PEK—JFK | 1.02 | 1.04 | 0.75 | 1 | 16.58 | 6 | 5.53 | 167 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 6 | CKG—PEK—SEA | 1.00 | 0.36 | 1.83 | 1 | 15.93 | 6 | 4.25 | 193 |
| 6 | CKG—PEK—SEA | 1.00 | 0.36 | 1.83 | 1 | 15.93 | 6 | 8.42 | 193 |
| 6 | CKG—PEK—SEA | 1.00 | 0.36 | 1.75 | 1 | 15.93 | 4 | 4.25 | 167 |
| 6 | CKG—PEK—SEA | 1.00 | 0.36 | 0.75 | 1 | 15.02 | 6 | 5.11 | 176 |
| 6 | CKG—PVG—SEA | 1.05 | 0.36 | 0.75 | 1 | 15.33 | 4 | 4.25 | 179 |
| 6 | CKG—KMG—SZX—SEA | 1.20 | 0.36 | 2.58 | 2 | 19.25 | 4 | 4.25 | 167 |
| 6 | CKG—SFO—LAS—SEA | 1.28 | 0.36 | 2.40 | 2 | 24.07 | 4 | 2.96 | 143 |

在联盟背景下中美航线市场份额预测的算例中,经过 220 次迭代,PSO 算法得到最优适应度值 10.5953,与不使用粒子群算法相比,神经网络的训练次数由 589 次减少为 42 次,有效地避免了神经网络陷入局部极小值。图 4 为采用 PSO—BP 神经网络预测值与实际值的对比图,图中为圆原点折线为 PSO—BP 神经网络预测输出值,十字折线为实际 QSI 值,从预测结果中来看,预测值与实际值的基本波动趋势一致。但有些合作产品计算出来的 QSI 值还存在一定的误差,分析产生误差的原因影响服务质量指数的因素中未考虑机票价格因素。同一航线产品机票价格随时间波动大,而价格数据属于商业机密难以获得,因此在 QSI

值计算时存在误差。

实验分别采用 PSO—BP 神经网络、BP 神经网络与线性回归对比实验,结果表明 PSO—BP 神经网络预测精度较高,神经网络训练误差 2.6×10^{-4} ,达到训练的误差要求。经计算,BP 神经网络绝对误差为 3.21,线性回归的绝对误差为 4.98,粒子群算法与神经网络结合的绝对误差为 1.01。为进一步论证联盟背景下航线产品市场份额的预测,加入联盟合作关系指标的合理性,分别采用机型座位数 X_1 ,航空公司服务 X_2 ,中转次数 X_3 ,中转时间 X_4 ,总行程时间 X_5 ,航线竞争程度 X_6 ,联盟合作关系 X_7 ,距离绕航率 X_8 等 8 个指标为输入数据与机型座位数 X_1 ,航空公司服务 X_2 ,中转次数 X_3 ,中转时间 X_4 ,总行程时间 X_5 ,航线竞争程度 X_6 ,距离绕航率 X_7 等 7 个指标为输入数据。使用 PSO—BP 神经网络方法进行对比实验,实验结果表明加入联盟合作关系绝对误差由 4.69 减少为 1.01。进一步分析原因若不引入联盟合作关系,对于联盟合作产品的预测 QSI 值与 QSI 实际值相比偏小,因此使用 PSO—BP 神经网络预测航线市场份额且考虑引入联盟合作关系指标具有较高的精度。

$$MAE = \frac{\sum |e_i|}{n} \quad (7)$$

式中:MAE 为绝对误差; e_i 为预测 QSI 值与实际 QSI 值间的误差; n 为样本数目。

采用训练完成的 PSO—BP 算法预测中美市场深圳至旧金山航线路径产品的市场份额,表 4 是各航线产品实际市场份额与预测市场份额的对比结果,预测的结果误差均控制在 1%,PSO—BP 算法预测联盟背景下航线市场份额可用于航空公司航线市场份额的预测,为进一步提高预测精度,可以在模型预测结果的结果上,由航空公司工作人员根据实际的运行数据及经验进一步的修正航线市场份额。

表 4 SZX—SFO 航线市场份额预测结果对比
Tab.4 Market forecasting of SZX—SFO

| | | | | | | | | | | |
|--------|-------|-------|-------|-------|-----|-------|-------|-------|-------|-------|
| 实际市场份额 | 2.56% | 2.56% | 1.71% | 2.56% | ... | 1.29% | 1.86% | 1.29% | 2.22% | 8.51% |
| 预测市场份额 | 2.59% | 2.59% | 1.45% | 2.57% | ... | 1.30% | 1.94% | 1.30% | 2.23% | 8.78% |

4 结论

基于联盟背景下国际航线市场合作产品的特征,补充达美航空提出的用于市场份额预测的 QSI 指标,通过大量的联盟数据分析机型座位数、航班时刻、航班频率、航线竞争程度、联盟合作关系、中转时间、总行程时间等多个指标对服务质量指数的影响,建立了联盟背景下航线市场份额预测指标体系。为了改变 QSI 模型依靠人工经验以及采用简单的线性回归确定系数的现象,精确拟合指标与航线产品市场份额间的非线性关系,采用 BP 神经网络完成输入指标与 QSI 值间的非线性映射。为解决 BP 神经网络初始权值确定的随机性造成网络陷入局部极小值的问题,在 BP 神经网络的基础上加入粒子群算法,极大的提高了预测精度,结果表明基于 PSO—BP 算法的 QSI 模型为联盟航线市场份额预测提供有效的方法。

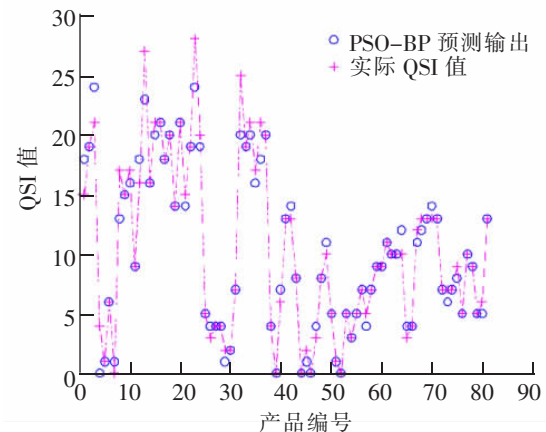


图 4 PSO—BP 测试结果对比图

Fig.4 Testing results comparison of PSO—BP

参考文献:

- [1] 胡皓月,蔡碧金,朱金福. 航空公司联盟合作伙伴评估与选择研究[J]. 航空计算技术,2016(2):42-46.
- [2] COLDREN G M,Koppelman F S,Kasturirangan K,et al. Air travel itinerary share prediction: logit model development at Major US Airline[J]. Journal of Air Transporty Management,2003.
- [3] COLDREN G M,KOPPELMAN F S. Modeling the competition among air-travel itinerary shares: GEV model development[J]. Transportation Research,Part A (Policy and Practice),2005,39(4):0-365.
- [4] BUSQUETS J G,ALONSO E,EVANS A D. Air itinerary shares estimation using multinomial logit models[J]. Transportation Planning and Technology,2018,41(1):3-16.
- [5] TEMBLEQUE V M,SUAU S P. A model to analyse the profitability of long-haul network development involving non-hub airports: The case of the Barcelona - Asian market[J]. Case Studies on Transport Policy,2016,4(2):188-197.
- [6] 胡皓月,朱金福,葛伟,et al. 基于多层 QSI 的航空联盟网络优化研究[J]. 武汉理工大学学报(交通科学与工程版),2016,40(5):880-884.
- [7] 赵晓松,朱金福,葛伟. 联盟环境下国际航线市场份额预测方法研究[J]. 交通运输系统工程与信息,2018,18(2):33-39.
- [8] 王超英,钟辉. 基于粒子群算法优化 BP 神经网络的产品质量预测分析[J]. 计算技术与自动化,2017,36(3):92-95.

Market Share Forecasting of Airline Alliance Route Based on PSO-BP Algorithm

Chen Xian, Zhu Jinfu, Liu Yue

(College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Under the context of alliances, airline products have complex attributes and high risk of cooperation compared to individual flights. Differences in product attributes lead to different market share. Airline alliance partner selection and collaborative value calculation are based on accurate forecasting of airline market share. Therefore, based on the QSI (Quality of Service Index) indicators of Delta Airline, this paper made full use of the alliance transportation data and added indicators such as route competition degree and alliance cooperation relationship to establish market share forecasting for international alliance route products. When fitting the nonlinear relationship between the eight factors and QSI value, considering that the parameter method needs to determine the form in advance, it introduced BP neural network to fit the nonlinear relationship between market share and factors. In order to solve the problem that the BP neural network falls into the local minimum value, the particle swarm optimization algorithm was introduced to determine the initial weight of the neural network and finally the QSI model of the PSO-BP algorithm was established. The results show that the QSI model based on PSO-BP algorithm can better predict the market share of international routes and the absolute error is below 1%.

Key words: market share forecasting; neural network; particle swarm optimization; alliance; QSI