

# 基于数据挖掘的航班延误预警管理分析

罗凤娥 张成伟 刘 安

(中国民用航空飞行学院空中交通管理学院 广汉 618307)

**摘要** 目前国内对航空公司航班延误状态的描述缺乏统一的计算模型,对航班延误结果的评价也不是很明确。基于航空公司大量历史运行数据,结合数据挖掘中的预测模型建立方法,采用层次分析法(AHP)提出了以延误率、平均延误时间、延误旅客人数为评价指标的航班延误状态描述方法,并运用马尔可夫(Markov)链对评价指标进行预测。利用模糊层次分析法(Fuzzy Analytic Hierarchy Process, FAHP)得到各指标权重,结合模糊综合评判法对航班延误状况进行综合评价,建立航班延误预警指标体系。仿真实验与结果分析表明:预警指标能较准确地反映航班延误状况,评价结果客观,可为航空公司航班延误预警管理提供理论与方法支持。

**关键词** 数据挖掘,航班延误,预警指标体系,马尔可夫模型,模糊综合评判

中图分类号 F560 文献标识码 A

## Flight Delays Early Warning Management and Analysis Based on Data Mining

LUO Feng-e ZHANG Cheng-wei LIU An

(College of Air Traffic Management, Civil Aviation Flight University of China, Guanghan 618307, China)

**Abstract** It is still a hard task to model domestic flight delay and to evaluate flight consequence in an authorized way at present. In this paper, according to a large number of airlines' historical operation data, a mathematical flight delay forecast model was derived to estimate the delay of each flight by analyzing the main factors influencing the flight delay, the consideration of the scheduled flight delay rate, the average time of delay and the delay passengers, and the Markov theory which is based on data mining prediction model establishment method and is used to forecast the evaluation index. The weights are gained by using fuzzy analytical hierarchy process (FAHP), and the condition of flight delay is evaluated based on fuzzy comprehensive evaluation. Finally, the flight delay alarming index system was constructed. The experiment of the developed delay forecast model shows that these alarming indexes can reflect the condition of flight delay accurately, and the outcome is objective, so it can provide support for flight delay early warning management theory and method.

**Keywords** Data mining, Flight delay, Alarming index system, Markov model, Fuzzy comprehensive evaluation

航班延误是航空运输业普遍存在的现象,在所有导致航班运行不正常的原因中,属于航空公司的原因占据相当大的一部分。以中国民航总局《2014年民航行业发展统计公报》中航班正常率统计数据为例(见表1),全国航空公司共执行航班312.6万班次,其中正常航班213.7万班次,不正常航班93.7万班次,平均航班正常率为68.37%。2014年,全国客运航班平均延误时间为19分钟,同比减少2分钟<sup>[1]</sup>。因此,对延误的预警及管理成为了航空公司亟待解决的问题。

表1 2014年航班不正常原因分类统计

指标	占全部比例(%)	比上年相比增加百分点
航空公司原因	25.9	-11.2
空管原因	24.0	-
天气原因	24.8	2.7
其他	25.3	-

在不正常航班预警管理方面,吕宗平、丁建立建立了基于大型枢纽机场的航班延误预警指标体系与预警量级<sup>[2]</sup>;王红、曹卫东提出了以航班延误率、平均延误时间及延误损失为评价指标的航班延误状态描述方法<sup>[3]</sup>;石丽娜以预警管理理论为基础,在对多家航空公司航班延误情况进行调查的基础上,给出了航空公司航班延误预警评价指标<sup>[4]</sup>;徐肖豪、李雄研究了航班地面等待模型中的延误成本计算问题,建立了单元受限地面等待问题的数学模型<sup>[5]</sup>。

目前国内关于航空公司航班延误预警管理方面的研究较少,缺乏统一的标准来评价航班延误整体的状态。以航空公司的航班延误管理为研究目标,基于数据挖掘中的预测方法,运用层次分析法和马尔可夫(Markov)链的预测原理提出了航班延误预测的计算模型,并给出了模型的仿真实验与分析;从航空公司的角度出发分析航班延误的影响规模、影响程度、

本文受中国民航飞行学院科研基金学生科技活动基金项目(X2014-33),民航局引导资金项目(MHRD200926),民航局安全能力建设项目(FSD-SA0033)资助。

罗凤娥(1972—),女,教授,硕士生导师,主要研究方向为航空公司运行管理;张成伟(1990—),男,硕士生,主要研究方向为航空公司运行管理;刘安(1990—),男,硕士生,主要研究方向为航空公司运行管理。

影响范围,运用模糊层次分析法(FAHP)和模糊综合评判法建立航班延误预警指标体系,以期改善航空公司的航班延误预警处理工作提供支持。

## 1 航班延误概述

通过查阅分析中国民航局发布的《民航航班正常统计办法》,归纳总结 2012、2013、2015 年颁发的航班延误相关标准,按照不同年份规定将其列于表 2 中。

表 2 《民航航班正常统计办法》公布的相关概念区别

定义	航班正常	航班延误时间	延误时间计算公式
2012 年	1) 在计划关舱门时间后规定的机场地面滑行时间之内起飞,且不发生返航、备降等不正常情况 2) 不晚于计划开舱门时间后 10 分钟落地	反映航班延误程度的指标,实际起飞时间晚于计划关舱门时间与机场地面滑行时间之和的时间	航班延误时间 = 实际起飞时间 - (计划关舱门时间 + 机场地面滑行时间)
2013 年	1) 在计划离港时间前后 5 分钟之内撤轮档,且按航班运行正向进程起飞,不发生滑回、中断起飞、返航、备降等特殊特殊情况 2) 不晚于计划到港时间挡轮档	反映航班延误程度的指标,实际离港时间晚于计划离港时间 5 分钟之后的时间长度航班延误时间以分钟为单位	航班延误时间 = 实际离港时间 - (计划离港时间 + 5 分钟)
2015 年 征求意见稿	1) 在计划离港时间后 10 分钟之内离港,不发生滑回及起飞后不发生返航、备降等特殊特殊情况 2) 不晚于计划到港时间后 10 分钟到港	航班延误时间是反映航班延误程度的指标,即实际离港时间晚于计划离港时间 10 分钟之后的时间长度,以分钟为单位	航班延误时间 = 实际离港时间 - (计划离港时间 + 10 分钟)

我国目前实施的统计办法与国际上通用的办法存在一定差异,表 3 比较了我国现行的办法以及国外的相关统计办法,希望提供一些改进建议供参考<sup>[6]</sup>。

表 3 国外航班正常性统计办法的比较

国家	统计标准	统计范围	不正常原因
美国	航班如果在计算机订座系统显示的计划时间后 15 分钟内离(到)港,则该航班统计为离(到)港正常。这里的离(到)港时间是指航班的撤(挡)轮档时间,不是离地(落地)时间	航班量占国内定期航班总量 1% 以上的 14 家主要航空公司,在航班量占国内定期航班总量 1% 以上的 29 个大型机场之间的定期国内航班正常情况,不包括国际航班以及其他性质飞行	不正常原因分为 5 大类: 1) 航空公司 2) 恶劣天气 3) 国家民航系统 4) 前班飞机晚到 5) 公共安全
澳大利亚	同美国	国内主要航空公司在国内主要航线上的航班准点表现	同美国
日本	只统计航班的离港正点率,航班如果在计划时间后 15 分钟内离港,则该航班统计为离港正常。这里的离(到)港时间是指航班的撤(挡)轮档时间,不是离地(落地)时间	国内主要航空公司	不正常原因分为 4 大类: 1) 天气 2) 设备故障 3) 设备重复 4) 其他原因

## 2 航班延误预警评价指标体系的建立

为提高航班正点率,航空公司研究分析可能会造成航班延误的原因,尽早发现并解决问题,必要时发布预警信息,将延误造成的损失降到最小。航班延误的预警管理是通过确定延误状态的评价指标,实施航班延误状态的评价,对延误发展趋势进行预测,从而确定延误的警限,并发出预警信息的过程。

### 2.1 航班延误状态评价指标的确定

2015 年民航局最新发布的民航航班正常统计办法征求意见稿中将航班延误原因分为天气、航空公司、航班时刻安排、军事活动、空管、机场、联检、油料、离港系统、旅客、公共安全等 11 个大类。通过对某航空公司 2013 年航班运行数据进行预处理,分析归纳总结出 8 大类某航空公司航班延误原因(见图 1),各类延误原因在 2013 年航班总数 137063 中所占比重分别为:流量控制原因(31817 个,占 23.21%)、天气原因(9427 个,占 6.8%)、航空公司原因(7956 个,占 5.8%)、机场原因(969 个,占 0.7%)、旅客原因(839 个,占 0.61%)、军事活动原因(622 个,占 0.45%)、公共安全原因(136 个,占 0.1%)、联检原因(39 个,占 0.03%)。

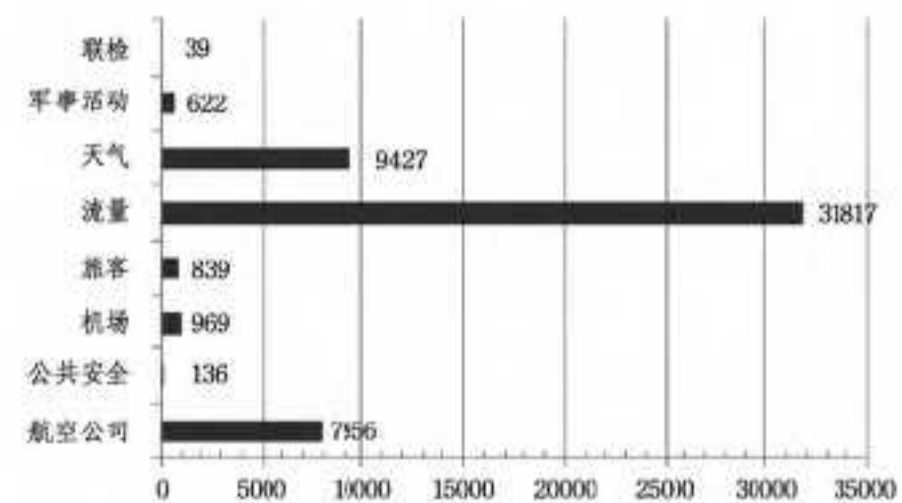


图 1 某航空公司 2013 年航班延误数量分布

通过以上分析得出,航班延误的主要影响因素之间相互独立,但又相互影响。当天气原因导致大面积的航班延误发生时,也可能进一步导致流量控制原因导致的延误问题。本文的研究数据完全来自航空公司运行控制中心真实的历史运行数据,分析航班延误原因时主要从航空公司运营角度考虑排名前 5 位的流量控制原因、天气原因、航空公司原因、机场原因和旅客自身原因。

航班延误预警指标体系,是一系列相互联系并能反映航班延误状态的指标构成的有机整体。航班延误状态的确定需要充分考虑延误评价指标的灵敏性、广泛性、可测度与相对独立性。为此,通过分析研究国内航班延误数量、程度和损失,综合设计了适合航空公司延误管理的 3 个评价指标。

1) 航班延误率:在某一时间段内,实际航班延误的数量与执行航班数量之比,用于衡量航班延误的规模。延误率越高,反映航班延误波及范围越大。

2) 航班平均延误时间:反映航班总体延误程度的指标,即计划航段班次总延误时间与计划航段班次之比。

3) 延误旅客人数:延误一旦发生则不可避免地会影响到旅客,波及对象的数量反映了延误的影响范围大小。

### 2.2 航班延误层次分析模型的建立

层次分析法是将与决策总是有关的元素分解成目标、准则、方案等层次,在此基础上进行定性和定量分析的决策方法。该方法是美国运筹学家匹茨堡大学教授萨蒂于 20 世纪

70年代初应用网络系统理论和多目标综合评价方法提出的一种层次权重决策分析方法。

航班延误状态的确定需要充分考虑延误评价指标的灵敏性、广泛性、可测度与相对独立性。为此,在分析研究国内航班延误数量、程度和损失的基础上,综合设计了适合航空公司延误管理的3个评价指标。

以航空公司的航班延误管理为研究目标,采用层次分析法提出了以航班延误率、平均延误时间及延误旅客人数为评价指标的航班延误状态描述方法,并都通过所选取的准则层进行一致性检验。建立针对航空公司的航班延误层次模型,如图2所示。

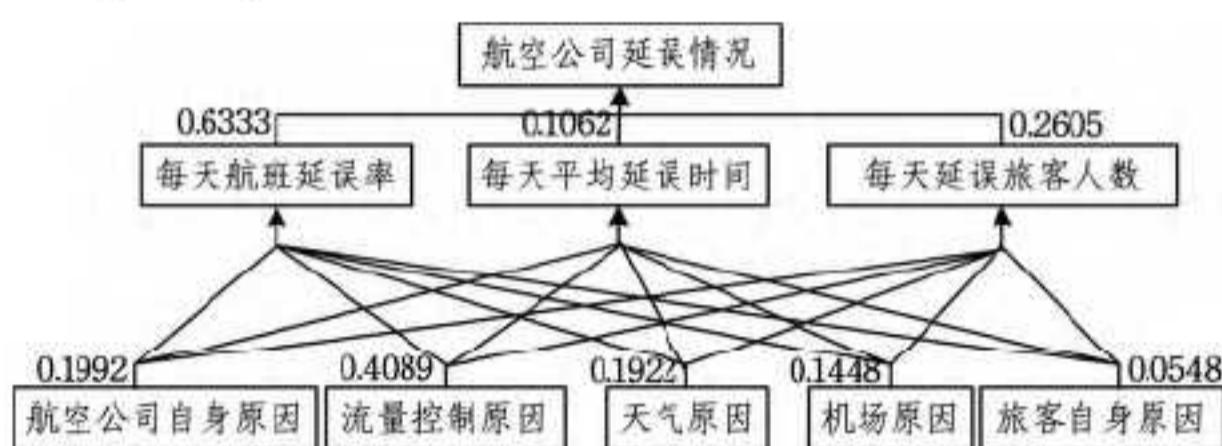


图2 基于AHP的航班延误评价体系结构

根据建立的该层次分析模型,可以从方案层(延误原因产生的5个方面)分别考察各个评价指标(每天的航班延误率、平均延误时间和延误旅客人数)的发生与发展情况,从而为预警提供分析基础。

### 2.3 评价方法的确定与预警量级的构建

航班延误的评价是一个复杂的多目标决策问题,采用模糊层次分析法得到各指标权重,利用模糊综合评判法对航班延误状况进行综合评价。FAHP在进行判断目标的总体评价时,缺乏一个统一的、具体的指标量化方法。因此,需要将模糊层次分析法与模糊综合评判法相结合,即首先用模糊层次分析法计算各指标权重,然后用模糊综合评判法进行综合评价。

具体步骤如下:

- 1) 建立具有层次结构的航班延误指标体系,如图3所示。



图3 航班延误预警指标体系

- 2) 得出两两因素比较的隶属度,构造模糊一致矩阵的取值,如表4所列。

表4 0.1~0.9数量标度

标度	定义	说明
0.5	同等重要	两元素相比较,同等重要
0.6	稍微重要	两元素相比较,一元素比另一元素稍微重要
0.7	明显重要	两元素相比较,一元素比另一元素明显重要
0.8	重要得多	两元素相比较,一元素比另一元素重要得多
0.9	极端重要	两元素相比较,一元素比另一元素极端重要
0.1,0.2, 0.3,0.4	反比较	若元素 $c_i$ 与元素 $c_j$ 相比较得到判断 $r_{ij}$ , 则元素 $c_j$ 和元素 $c_i$ 相比较得到的判断矩阵为 $r_{ji}=1-r_{ij}$

- 3) 计算各层次指标的权重  $w$ , 如表5所列。

$$w_i = \frac{1}{n} - \frac{1}{2a} + \frac{1}{na} \times \sum_{k=1}^n r_{ik}, i \in N \quad (1)$$

其中,  $n$  为模糊一致矩阵的阶数, 取  $a=(n-1)/2$ 。

表5 各指标之间的模糊一致矩阵及权重的确定

	延误率	平均延误时间	延误旅客人数	权重
延误率	0.50	0.55	0.25	0.27
平均延误时间	0.45	0.50	0.15	0.2
延误旅客人数	0.75	0.85	0.50	0.2

- 4) 计算综合评判结果,  $B=W \cdot R$ , 其中  $B=[b_1, b_2, \dots, b_n]$ , 是系统总体性能对应评价集  $V$  的隶属向量,  $W$  是权重向量,  $R$  是指标的评判矩阵。

航班延误预警量级的构建是根据各指标包含的具体预警要求, 给定评判集  $V=\{I \text{ 级、II 级、III 级、IV 级、V 级}\}$ 。其中 I 级表示正常, II 级表示蓝色预警, III 级表示黄色预警, IV 级表示橙色预警, V 级表示红色预警。各指标相对不同预警等级的具体取值可参考表6。

表6 航班延误预警量级

预警等级	分数	延误率(%)	平均延误时间(min)	延误旅客人数
正常	0	<5	<20	<4000
蓝色预警	1	5~10	20~40	4000~8000
黄色预警	2	10~20	40~60	8000~12000
橙色预警	3	20~30	60~120	12000~16000
红色预警	4	>30	>120	>16000

模糊综合评判隶属函数的构建, 在区间上为一线性函数, 则:

延误率的隶属函数  $U(x_1)$  为:

$$U(x_1) = \begin{cases} 0, & x_1 \leq 5 \\ \frac{x_1 - 5}{5}, & 5 < x_1 \leq 10 \\ \frac{x_1 - 10}{20}, & 10 < x_1 \leq 30 \\ 1, & x_1 > 30 \end{cases} \quad (2)$$

平均延误时间的隶属函数  $U(x_2)$  为:

$$U(x_2) = \begin{cases} 0, & x_2 \leq 20 \\ \frac{x_2 - 20}{40}, & 20 < x_2 \leq 60 \\ \frac{x_2 - 60}{60}, & 60 < x_2 \leq 120 \\ 1, & x_2 > 120 \end{cases} \quad (3)$$

延误旅客人数的隶属函数  $U(x_3)$  为:

$$U(x_3) = \begin{cases} 0, & x_3 \leq 4000 \\ \frac{x_3 - 4000}{12000}, & 4000 < x_3 \leq 16000 \\ 1, & x_3 > 16000 \end{cases} \quad (4)$$

## 3 基于 Markov 链的航班延误状态预测

### 3.1 数据挖掘

数据挖掘是一个迭代的过程, 在这个过程中, 通过自动或手工方法取得的进步用“发现”来定义。在探测性分析方案中, 无法预测出“有趣的”结果包含什么东西, 此时数据挖掘就非常重要, 它从大量的数据中搜寻有价值的、非同寻常的新信息, 是人和计算机合作的结果。在实践中, 数据挖掘的两个基本目标往往是预测和描述。预测是使用数据集中的一些变量或域来预测其他相关变量的未知值或未来的值, 而描述是找出描述可由人类解释的数据模式[8]。因此, 可以把数据挖掘活动分成两类:

- 1) 预测性数据挖掘,生成给定数据集所描述的系统模型。
- 2) 描述性数据挖掘,在可用数据集的基础上生成新的、非同寻常的信息。

预测性数据挖掘的目标是得出一种以可执行码来表示的模型,该模型可以用于执行分类、预测、评估或者其他相似的任务。而描述性数据挖掘的目标是发现大型数据集中的模式和关系,从而理解所分析的系统。对不同的数据挖掘应用,预测和描述的相对重要性可能不大相同。预测和描述的目标都是通过数据挖掘技术来实现的。本文将采用预测性数据挖掘的思想和方法,选取马尔可夫(Markov)预测方法建立航班延误的预测模型。

### 3.2 Markov 链的基本理论

马尔可夫链是离散状态的马尔可夫过程,是指系统的未来状态仅与现在的状态有关而与以前状态无关的随机过程,它具有无后效性的特点。马尔可夫链的预测原理就是根据某些变量的现在状态及其变化趋势,预测在未来某一特定期间内可能出现的状态,从而为决策提供依据。

在建立航班延误预测分析模型的过程中,结合海量航班数据离散的特点,在  $m$  个时刻中以实际状态的概率向量与理论计算的状态的概率向量的误差平方和达到最小为准则,建立航班延误预测最优化模型<sup>[7]</sup>。

设  $\partial(t) = (p_i(1), p_i(2), \dots, p_i(n))$  是时刻  $t(t=0, 1, 2, 3, \dots, m)$  航班延误系统在  $n$  个状态下的概率向量,设一步状态转移概率矩阵为  $P = (p_{ij})_{n \times n}$ ,实际上由于客观环境的变化,相邻时刻的一步状态转移概率矩阵并不完全相同,因此  $\partial(t+1)$  与  $\partial(t)P(t=0, 1, 2, 3, \dots, m-1)$  之间总存在误差。由误差平方和达到最小的准则,建立如下航班延误预测最优化模型,记为模型(5)。

$$\begin{aligned} \min f(P) &= \sum_{t=0}^{m-1} \|\partial(t+1) - \partial(t)P\|^2 \\ &= \sum_{t=0}^{m-1} (\partial(t+1) - \partial(t)P)(\partial(t+1) - \partial(t)P)^T \quad (5) \end{aligned}$$

$$\text{s. t. } \begin{cases} \sum_{j=1}^n p_{ij} = 1, i=1, 2, \dots, n \\ p_{ij} \geq 0, i, j=1, 2, \dots, n \end{cases}$$

### 3.3 预测过程

按照模型(5)建立的优化模型,根据第一周航班延误的实际数据,对第二周航班延误率、平均延误时间、延误旅客人数进行预测,并将预测结果与第二周实际数据进行比较。以某航空公司 2013 年 3 月 4 日—2013 年 3 月 17 日两周航班延误真实数据为例给出了预测的实现过程,仿真计算过程利用 Matlab 软件对模型进行求解。

#### (1) 延误率的预测

某航空公司 2013 年 3 月 4 日—2013 年 3 月 10 日一周 7 天的航班延误率情况如表 7 所列。

表 7 一周航班延误率的实际值(%)

星期	航空公司	流量控制	天气	机场	旅客自身	合计
一	3.14	25.71	0.00	2.86	2.00	33.71
二	7.80	21.73	0.00	0.84	0.28	30.65
三	9.72	20.83	3.61	1.11	1.11	36.38
四	6.25	16.76	3.98	0.00	1.99	29.83
五	17.31	17.86	4.12	1.10	1.92	42.31
六	7.26	20.95	8.94	0.00	0.84	37.99
日	10.41	16.99	2.47	0.27	0.00	30.14
平均	8.84	20.12	3.30	1.00	1.16	34.43

根据以上一周的航班延误率实际数据,按照模型(5)计算

出转移概率矩阵  $P_1$ :

$$P_1 = \begin{bmatrix} 0.0000 & 0.3617 & 0.4250 & 0.0000 & 0.2132 \\ 0.1676 & 0.8324 & 0.0000 & 0.0000 & 0.0000 \\ 0.4096 & 0.0000 & 0.0000 & 0.0000 & 0.5904 \\ 0.0000 & 0.0000 & 0.0000 & 0.0243 & 0.4694 \\ 0.0000 & 0.0000 & 0.5306 & 0.3697 & 0.6061 \end{bmatrix}$$

以该周最后一天(周日)的数据作为初始状态,即  $\partial(0) = (10.41\%, 16.99\%, 2.47\%, 0.27\%, 1.37\%, 31.51\%)$ ,由  $\partial(m) = \partial(0)P^m$  可以计算未来一周的延误率预测值,如表 8 所列。

表 8 未来一周航班延误率预测值(%)

星期	航空公司	流量控制	天气	机场	旅客自身	合计
一	3.86	17.91	4.49	0.62	0.63	27.51
二	5.84	16.30	1.87	1.98	1.52	27.51
三	5.50	17.32	2.37	3.22	0.10	28.51
四	6.54	14.02	1.83	3.92	1.20	27.51
五	5.10	15.95	1.90	4.61	0.95	28.51
六	2.95	11.90	1.75	5.16	1.75	23.51
日	3.71	10.97	1.73	5.68	1.42	23.51
平均	4.79	14.91	2.28	3.60	1.08	26.65

该航空公司下一周 7 天即 2013 年 3 月 11 日—2013 年 3 月 17 日的航班延误率实际情况与预测值的对比分析如图 4 所示。

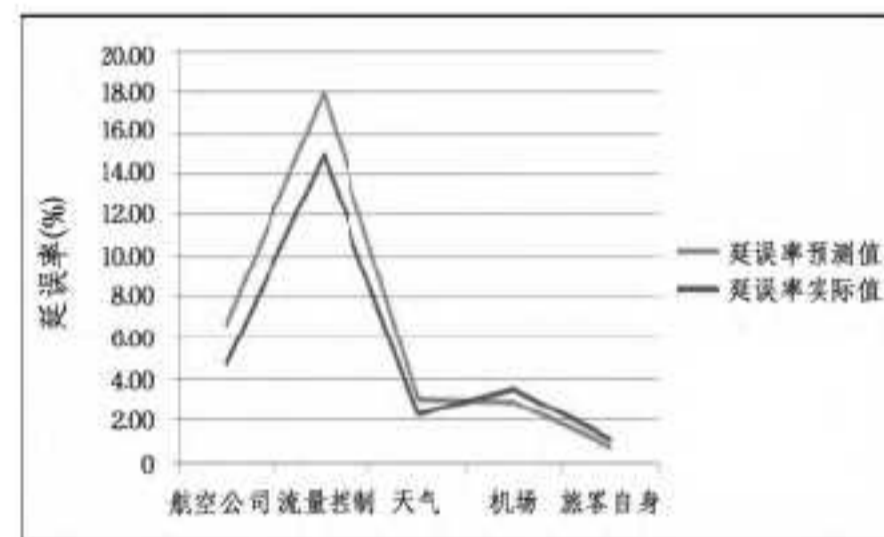


图 4 未来一周航班延误率预测值与实际值的对比

#### (2) 平均延误时间的预测

某航空公司 2013 年 3 月 4 日—2013 年 3 月 10 日一周 7 天航班平均延误时间实际情况如表 9 所列。

表 9 一周航班平均延误时间实际值(min)

星期	航空公司	流量控制	天气	机场	旅客自身	合计
一	70	52	0	85	51	258
二	75	41	0	30	39	185
三	38	42	114	40	63	297
四	55	37	180	0	116	388
五	114	50	161	29	67	421
六	60	101	75	0	39	275
日	93	39	61	36	0	229
平均	72	52	84	31	54	293

根据以上一周的航班平均延误时间实际数据,采用相同的方法进行平均延误时间的预测,按照模型(5)计算出转移概率矩阵  $P_2$ :

$$P_2 = \begin{bmatrix} 0.0000 & 0.3972 & 0.0858 & 0.0000 & 0.5170 \\ 0.4920 & 0.0125 & 0.1716 & 0.3240 & 0.0000 \\ 0.0000 & 0.0000 & 0.4805 & 0.0000 & 0.5195 \\ 0.0000 & 0.4436 & 0.0000 & 0.5488 & 0.0076 \\ 0.0000 & 0.0000 & 0.0568 & 0.4000 & 0.5759 \end{bmatrix}$$

以该周最后一天(周日)的数据作为初始状态,即  $\partial(0) = (93, 39, 61, 36, 0)$ ,由  $\partial(m) = \partial(0)P^m$  可以计算未来一周的延误率预测值,如表 10 所列。

表 10 未来一周航班平均延误时间预测值(min)

星期	航空公司	流量控制	天气	机场	旅客自身	合计
一	32	66	60	14	58	230
二	57	66	74	35	56	288
三	38	65	86	45	57	291
四	53	61	52	49	75	290
五	55	57	52	53	76	293
六	24	53	57	55	56	245
日	33	49	89	58	48	277
平均	42	60	67	44	61	273

该航空公司下一周 7 天即 2013 年 3 月 11 日—2013 年 3 月 17 日的航班平均延误时间实际情况与预测值的对比分析如图 5 所示。

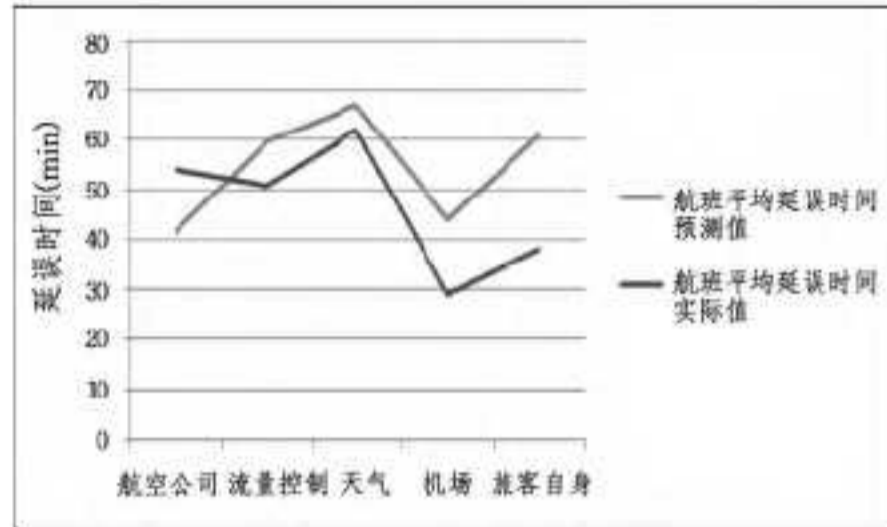


图 5 未来一周航班平均延误时间预测值与实际值的对比

(3) 延误旅客人数的预测

某航空公司 2013 年 3 月 4 日—2013 年 3 月 10 日一周 7 天航班延误旅客人数实际情况如表 11 所列。

表 11 一周航班延误旅客人数实际值(人)

星期	航空公司	流量控制	天气	机场	旅客自身	合计
一	1614	13227	0	1603	987	17431
二	3846	11829	0	237	132	16044
三	5369	11262	1889	668	642	19830
四	3458	8515	1618	0	1129	14720
五	8127	9138	1479	645	808	20197
六	3467	11021	4068	0	379	18935
日	5012	9232	1500	146	0	15890
平均	4413	10603	1508	471	582	17578

根据以上一周的航班平均延误时间实际数据,采用相同的方法进行航班延误旅客人数的预测,采用同样的预测方法,按照模型(5)计算出转移概率矩阵  $P_3$ :

$$P_3 = \begin{bmatrix} 0.0000 & 0.3718 & 0.2625 & 0.0000 & 0.3657 \\ 0.1433 & 0.8567 & 0.0000 & 0.0000 & 0.0000 \\ 0.3632 & 0.0000 & 0.0226 & 0.0000 & 0.6142 \\ 0.0000 & 0.0000 & 0.0000 & 0.1874 & 0.8126 \\ 0.0000 & 0.0000 & 0.0000 & 0.0629 & 0.9371 \end{bmatrix}$$

以该周最后一天(周日)的数据作为初始状态,即  $\partial(0) = (5012, 9232, 1500, 146, 0)$ , 由  $\partial(m) = \partial(0)P^m$  可以计算未来一周的延误率,如表 12 所列。

表 12 未来一周航班延误旅客人数预测值(人)

星期	航空公司	流量控制	天气	机场	旅客自身	合计
一	3826	9772	1350	274	92	15314
二	3697	7939	1437	126	103	13302
三	3366	7331	1438	207	105	12447
四	3142	6763	1320	297	105	11627
五	2899	6244	1235	305	105	10788
六	3684	7765	1145	305	210	13109
日	3264	7180	1492	227	216	12379
平均	3411	7571	1345	249	134	12709

该航空公司未来一周 7 天即 2013 年 3 月 11 日—2013 年

3 月 17 日的航班平均延误时间实际情况与预测值的对比分析如图 6 所示。

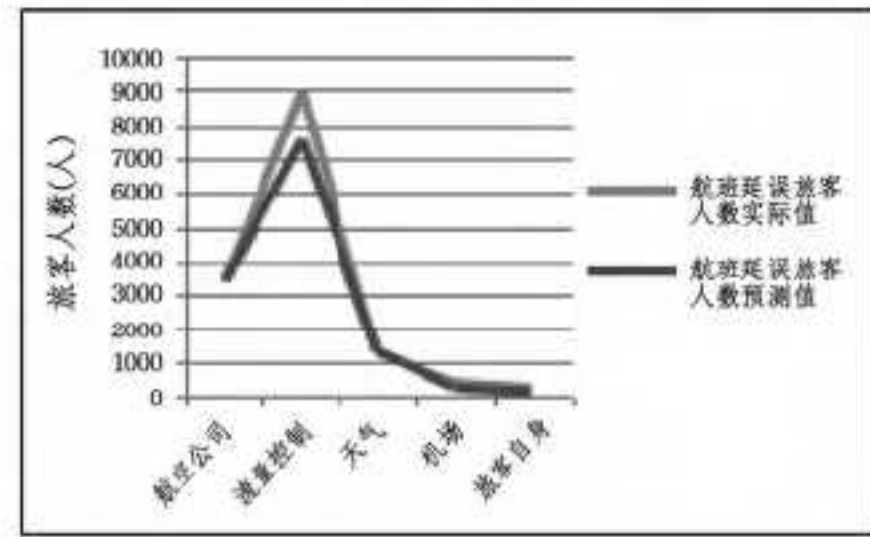


图 6 未来一周航班延误旅客人数预测值与实际值的对比

(4) 预测结果分析

仿真实验结果可以从航空公司影响航班延误的 5 类因素分别进行比较,表 13 给出了一周平均情况的预测结果与实际值的误差分析。

表 13 预测结果的误差分析

误差	航空公司	流量控制	天气	机场	旅客自身	平均值
延误率(%)	1.89	3.05	0.81	0.65	0.38	1.36
平均延误时间(min)	12	9	5	15	18	11.8
延误旅客人数(人)	20	1384	59	177	147	357

从预测误差值可以看出,用航空公司过去一周航班延误率、平均延误时间及航班延误旅客人数预测未来一周的情况的平均延误误差较小,分别为 1.36%、11.8min 与 357 人。

4 实例分析

以某航空公司 2013 年 3 月 7 日(周四)全天航班实际运行数据为算例,来验证航班延误预警综合评判结果。将数据分别代入式(2)~式(4)得到所对应的  $U$  值,如表 14 所列。

表 14 航班延误预警指标实际数据和对应隶属度

指标	延误率	平均延误时间	延误旅客人数
实际数据	29.83%	52min	14720 人
$U$ 值	0.99	0.8	0.89
权重	0.27	0.2	0.2

评判集的模糊等级采用三角形隶属函数,其隶属度如图 7 所示[9]。

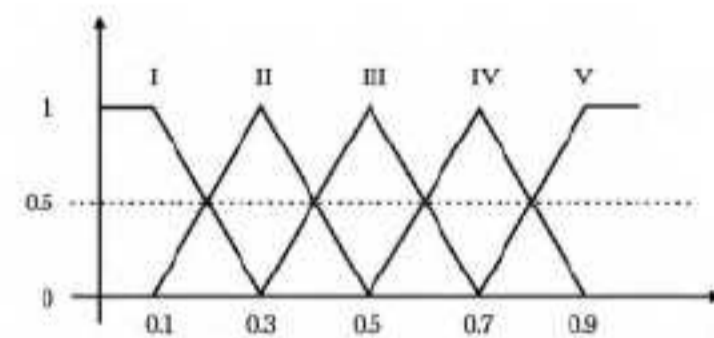


图 7 三角形隶属函数与危险等级的隶属度

对航班延误状况进行综合评判,用模型  $M(\cdot, +)$  计算得:

$$B = W \cdot R = [0.27, 0.2, 0.2] \cdot \begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 1.0 \\ 0.0 & 0.0 & 0.0 & 0.5 & 0.5 \\ 0.0 & 0.0 & 0.0 & 0.1 & 0.9 \end{bmatrix} = [0, 0, 0, 0.12, 0.55]$$

评判集  $V = \{I \text{ 级、II 级、III 级、IV 级、V 级}\}$ , 数量化表示为  $V = \{0, 1, 2, 3, 4\}^T$ 。

(下转第 557 页)

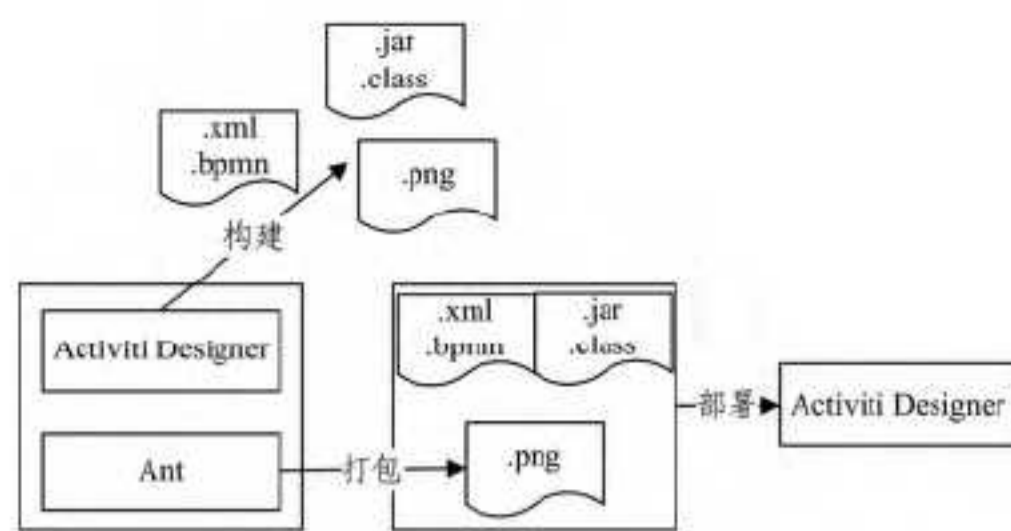
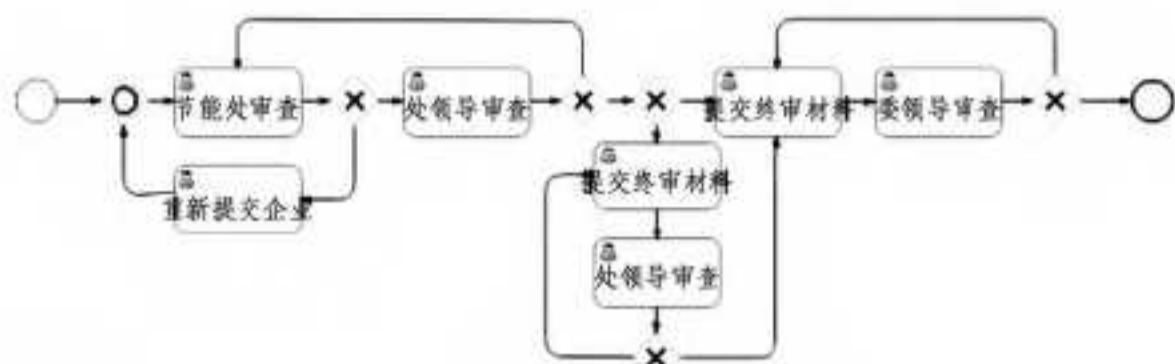


图6 部署流程图

部署时导入的文件可以为 bar 文件、xml 文件或 zip 文件。文件打包可以使用 Ant 脚本自动完成。验证部署结果如图 7 所示。



流程实例

ID	业务ID	发起人	启动流程的活动ID	已启动
1515		kermit	startevent1	Sat Jun 06 17:55:58 CST 2015
1521		kermit	startevent1	Sat Jun 06 18:03:30 CST 2015

图7 部署结果图

图 7 展示了部署的流程模型和启动的两个流程实例，其中，流程实例信息中展示了 ID、任务 ID、发起人、启动流程的活动 ID 以及启动时间。

结束语 对 Activiti 核心技术进行了介绍，并对在线流程审批流程中的概念和 Activities 中的概念进行对比，从而提

出了一种基于 Activiti 框架的在线流程审批算法，最后讨论了该算法在实际应用中的部署和验证。

## 参考文献

- [1] 鲁彦平. 我国行政审批制度改革的基本路径与系统反思[J]. 领导科学, 2016(2)
- [2] 薛晓东, 梁丹妮, 叶萍. ESIA 理论视域下地方政府投资项目行政审批程序优化研究[J]. 电子科技大学学报(社科版), 2015(1): 7-11
- [3] 黄国玲. 我国地方政府行政审批流程再造的对策建议[J]. 城市建设理论研究(电子版), 2015(16)
- [4] Rademakers T. Activiti in Action[M]// Manning, 2012
- [5] Baina K, Baina S. User experience-based evaluation of open source workflow systems; The cases of Bonita, Activiti, jBPM, and Intalio[C]// 2013 3rd International Symposium ISKO-Maghreb. IEEE, 2013; 1-8
- [6] Yang S L, Hu J P. Design of Task Workflow Based on Activiti Technology[J]. Applied Mechanics & Materials, 2015, 740: 802-805
- [7] Siaw-Teng L, Tam C W M. Research ethics and approval process; A guide for new GP researchers[J]. Australian Family Physician, 2015, 44(6): 419-22
- [8] Siaw-Teng L, Tam C W M. Research ethics and approval process; A guide for new GP researchers[J]. Australian Family Physician, 2015, 44(6): 419-22
- [9] 徐亦楠, 葛志辉, 潘海源. Activiti5 工作流在 OA 系统中的应用[J]. 大众科技, 2014(1): 5-7
- [10] 于永会, 刘争鸣, 唐军. 基于开源工作流 Activiti5 的制造执行管理系统的设计实现[J]. 工业控制计算机, 2014(9): 121-122

(上接第 546 页)

则航班延误对应评判集的等级总分为:

$$[0, 0, 0, 0, 0.12, 0.55] \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{matrix} = 2.56$$

得出评价结果，评价结果和第 IV 级最接近，即认为对于当时的航班延误状况，应当发布橙色预警。

通过仿真实验与结果分析，给出了航班延误预警管理计算模型，构建了航班延误预警指标体系。评价结果表明，预警指标能较准确地反映航班延误状况，评价结果客观，可作为航空公司航班延误预警参考使用。

结束语 航班发生延误不可避免地会对航空公司、旅客和机场的正常运作产生严重的影响，如果能在航班发生大面积延误之前就能将延误控制在一定范围内，或者能够提前做好充分的准备工作，将会减少很多不必要的损失。从航空公司运行角度出发，运用层次分析法给出了航班延误状态描述，通过算例的仿真实验与分析，验证了所构建的航班延误预测分析模型和预警管理评价指标体系能够有效地反映航班的延误状态，为航空公司不正常航班的处置提供决策依据。

但是在实际运行过程中，应当针对不同规模的航空公司考虑不同情况下的预警评价指标体系的建立，使得航班延误预警更具针对性。同时，本文中的预测过程没有考虑突发情况的发生对延误预警结果的影响，在后续的研究过程中还需

进一步深入研究，以提高预警的准确性。

## 参考文献

- [1] 2014 年民航行业发展统计公报[R]. 中国民用航空局, 2015
- [2] 吕宗平, 胡欣, 丁建立. 航班延误预警指标体系与预警量级构建[J]. 航空计算技术, 2010, 40(1): 1-4
- [3] 王红, 刘金兰, 曹卫东, 等. 航空公司航班延误预警管理模型与分析[J]. 计算机仿真, 2009, 26(4): 292-296
- [4] 石丽娜. 多等级模糊评价方法在航班延误中的应用[J]. 上海工程技术大学学报, 2006, 20(3): 276-279
- [5] 徐肖豪, 李雄. 航班地面等待模型中的延误成本分析与仿真[J]. 南京航空航天大学学报, 2006, 38(1): 115-120
- [6] 魏景焕. 国内外航班正常性统计办法的比较与启示[N/OL]. 民航资源网. <http://news.carnol.com/list/294/294552.html>
- [7] 何江宏, 陈启明. 基于 Markov 链的最优化预测模型及其应用研究[J]. 合肥学院学报(自然科学版), 2006, 16(1): 11-13, 2
- [8] 程华, 李艳梅, 罗谦, 等. 基于 C4.5 决策树方法的到港航班延误预测问题研究[J]. 系统工程理论与实践, 2014, 34(S1): 239-247
- [9] 吴抗抗. 航班延误及其衍生事件预测预警方法研究[D]. 南京: 南京航空航天大学, 2012
- [10] 罗凤娥, 张成伟. 基于时序数据挖掘的航班延误预测分析[J]. 现代电子技术, 2014(24): 52-55
- [11] 李频. 基于灰色动态马尔科夫的航班延误预测[J]. 上海工程技术大学学报, 2014, 28(4): 333-336, 346
- [12] 李频, 刘君强. 基于加权马尔科夫的航班延误预测研究[J]. 滨州学院学报, 2014(6): 50-54