Industrial Engineering and Innovation Management (2022)
Clausius Scientific Press, Canada
DOI: 10.23977/ieim.2022.050702
ISSN 2522-6924 Vol. 5 Num. 7

Research on Flight Adjustment Optimization Based on Flight Delay Forecast

Luo Feng'e[1,a], Kan Xiangyu[1,b], Qi Fang[1,c], Wang Bo[1,d]

[1]*Civil Aviation Flight University of China, Air Traffic Management College, Guanghan, China*
[a]*lucylilyluo@vip.sina.com*, [b]*2983350084@qq.com*, [c]*517195416@qq.com*, [d]*17862071235@163.com*

Keywords: Flight delay forecast, Logistic model, Time series analysis model, Flight adjustment and optimization

Abstract: Flight delays are not only a single point effect, but also affect related flights and airports through the air transportation network, causing large-scale flight delays. This article is aimed at Advance control of flight delays, that is, the control of the spread delay of the flight plan, through the delay prediction of the flight, provide advance reference for the subsequent flight plan optimization, on the basis of the flight delay prediction, adjust and optimize the spread delay flight, use The aircraft path spread and delay optimization model based on the column generation algorithm adjusts the delayed flights.

1. Introduction

China's civil aviation industry is getting stronger and stronger, but the problem of flight delays has become more serious, affecting the development of civil aviation. Flight delays are not only a single point effect, but also affect related flights and airports through the air transportation network, causing large-scale flight delays. This article focuses on flight delays. Pre-control, that is, the research on the control of the spread and delay of the flight plan, through the delay prediction of the flight, it provides a pre-reference for the subsequent optimization of the flight plan, and adjusts and optimizes the spread and delay flights on the basis of the flight delay prediction, and uses the column-based generation The algorithm's aircraft path spread and delay optimization model adjusts the delayed flights[1].

In 2004, Khaled F. Abdelgani and Sharmila S. Shah and others [2] developed a model to predict flight delays and airline early warnings based on the spread of flight delays by analyzing the impact of flight resources such as crew and aircraft on flight operations. Michael O. BALL and others[3] analyzed the main factors affecting the distribution of departure delays and combined non-parametric methods with genetic algorithms to create a departure delay distribution estimation model to predict traffic congestion. Lu Zonglei and Wang Jiandong[4] established a large-scale flight delay early warning model based on machine learning technology by classifying airport delay data. Wang Hong, Liu Jinlan and others [5] proposed a flight delay condition description method based on flight delay loss, delay frequency and average delay time as evaluation indicators, and proposed a flight delay early warning management calculation model based on Markov chain. Xu Tao, Honor and others [6] developed and implemented a SOA (Service Oriented Architecture)-based flight delay decentralized analysis and early warning system to provide software support for early warning of flight delays. Sharpe analyzed the reasons for flight delays, combined flight delay data with airline operation management models, and studied information flow and design methods to warn flight delays.

2. Logistic model

The next subsections provide instructions on how to insert figures, tables, and equations in your document.

2.1 Logistic model introduction

The generalized linear model extends the linear model[7]. If the dependent variable does not meet the normality assumption, the generalized linear model can be used for analysis. It analyzes the relationship between the "average value of the dependent variable" and the "dependent variable" to solve non-linear problems.

Suppose X is a continuous random variable, and X obeys Logistic distribution, that is, the distribution function of X is:

$$F(X) = P(X \leq x) = \frac{1}{1+e^{-\frac{x-\mu}{\gamma}}} \tag{1}$$

Among them, μ is the position parameter, and γ is the shape parameter. The distribution function F(X) is the logistics function. It can be seen from this figure that the image of the function F(X) is an S-shaped curve, a graph with central symmetry, and the symmetry point is (μ, 0.5). It is steep near the center of symmetry and has a faster growth rate, while at both ends it is relatively stable and has a slower growth rate. The magnitude of γ determines the speed at which the curve grows near the center. The smaller the γ, the faster the speed.

By comparing the conditional probability judgment represented by the conditional probability P(Y|X), the observation results are divided into various categories.

$$P(Y=1|x) = \frac{\exp(\omega \cdot x + b)}{1+\exp(\omega \cdot x + b)} \tag{2}$$

$$P(Y=0|x) = \frac{1}{1+\exp(\omega x \cdot + \mathrm{b})} \tag{3}$$

Among them, ω∈Rn and b∈R are parameters, ω is the weight vector, b is the deviation, and ω*x is the inner product of ω and x. For any given instance x, calculate the corresponding conditional probabilities P (Y = 1 | x) and P (Y = 0 | x) according to formulas (2) and (3). Compare the values of P (Y = 1 | x) and P (Y = 0 | x), and classify x as a category with a larger probability value. For simplicity, let ω=(ω(1),ω(2),...,ω(m),b)T,x =(x(1),x(2),...,x(n), 1) T becomes a logistic regression model:

$$P(Y=1|\mathrm{x}) = \frac{\exp(\omega \cdot \mathrm{x})}{1+\exp(\omega \cdot \mathrm{x})} \tag{4}$$

$$P(Y=0|\mathrm{x}) = \frac{1}{1+\exp(\omega \cdot \mathrm{x})} \tag{5}$$

The probability of an event is the logarithm:

$$\text{logit}（p）=\log\frac{p}{1-p} \tag{6}$$

In logistic regression, the probability of an event occurring at Y=1 under the condition of instance x is:

$$\log\frac{P(Y=1|x)}{1-P(Y=1|x)}=\omega\cdot x \tag{7}$$

That is, in the logistic regression model, the log probability of the event "divide Y by τ=1" has a linear relationship with the independent variable x, and the log probability of the event "multiply X by Y=1" has a linear relationship. It can be determined by the linearity represented by the τ function. It can be seen from formula (4) that when the value of ω*x is close to positive infinity, P(Y = 1 | x) is close to 1, and when the value of ω*x is close to negative infinity, P(Y = 1 | x) tends to be close to 0.

Suppose P(Y=1|x)=π(x), P(Y=0|x)=1-π(x), the likelihood function is:

$$\prod_{i=1}^{N}[\pi(x_i)]^{y_i}[1-\pi(x_i)]^{1-y_i} \tag{8}$$

The log likelihood function is

$$L（\omega）=\sum_{i=1}^{N}[y_i(\omega\cdot x_i)-\log(1+\exp(\omega\cdot x_i)] \tag{9}$$

Find the maximum value of L(u) to get the estimated value of ω. Assuming that the maximum likelihood estimate of ω is, then

The logistic regression model is:

$$P（Y=1|x）=\frac{\exp（\hat{\omega}\cdot x）}{1+\exp（\hat{\omega}\cdot x）} \tag{10}$$

$$P（Y=0|x）=\frac{1}{1+\exp（\hat{\omega}\cdot x）} \tag{11}$$

2.2 Flight delay classification model

Flight delays are divided into two types: delayed and non-delayed, which are two-category variables, so the value of the dependent variable is as follows:

$$Y=\begin{cases}1（Delay）\\0（nodelay）\end{cases} \tag{12}$$

3. Flight delay forecast model based on time series analysis

This part establishes the time series of flight delays[8], analyzes the characteristics of chaos, and realizes the prediction of flight delays based on the chaotic prediction method. For a specific airport, first use small data methods to identify the chaos of the delay time series. Then the sequence is analyzed according to the recursive graph, and finally a chaotic prediction model based on extreme learning machine is established to realize the prediction of various flight delays.

3.1 Construction of Time Series of Airport Flight Delay

Airport flight operation information usually includes flight plan data and OOOI data (out, off, on, in time date) of each flight taking off and landing at the airport, that is, each data record contains the flight number and model and the departure aircraft. Landing at the airport, planned departure and departure time, actual departure time, estimated arrival time. Information such as actual arrival time and walking time. If the actual departure time of the flight exceeds 15 minutes (including 15 minutes) after the scheduled departure time, or the actual departure time of the flight exceeds 15 minutes (including 15 minutes) after the scheduled arrival time, the flight is deemed to be delayed.

3.2 Analysis of Chaotic Characteristics of Delayed Time Series

Usually, the delayed coordinate method is used to reconstruct the time series [9]. Assuming the delay time series {xi,i = 1,2,...,N}, the reconstructed phase space is:

$$X_j = [x_j, x_{j+\tau}, \ldots, x_{j+(m-1)\tau}] \qquad (13)$$

In the formula, j = 1, 2, ..., M, M = N-(m-1)τ. In this work, the average mutual information method is used to calculate the appropriate time delay τ, and the pseudo-neighbor method is used to calculate the optimal embedding dimension m. According to the principle of the average mutual information method, take two time series S = {si, i = 1, 2,..., n} and Q = {qi, i = 1, 2,..., n}, then The average mutual information of S and Q is:

$$I(Q,S) = \sum_i \sum_j P_{sq}(s_i, q_i) \log_2 [\frac{P_{sq}(s_i, q_i)}{P_s(s_i) P_q(q_j)}] \qquad (14)$$

Taking si as xi and q as x+π, then I(Q,S) = I(τ). According to I(τ), the first minimum value is usually selected as the best delay time when reconstructing the phase space.

3.3 Chaos Recognition of Delayed Time Series

An important criterion for whether the delayed time series[10] is chaotic is whether its largest Lyapunov exponent λ is greater than 0. If> 0, the long-term change of the delay time series is sensitive to the initial conditions, and the series is in a state of chaos.

1) According to the τ and m of the delay time series, the phase space { Xj, j=1, 2, ..., M} is reconstructed.
2) In phase space, find the nearest neighbor of point and limit the short-term component
3) In the formula, P is the average period of the sequence.
4) Calculate the distance between Xj and Xj after i discrete time steps,
5) For each i, find the average value of lnd; (i) of all j
6) Here, q is the number of dj(i) not equal to 0; Δt is the discrete time interval. Use the method of least squares to create a regression line whose slope is the sequence λ. The phase space of the delay time series is reconstructed, and the maximum Lyapunov exponent λ is calculated.
7) Select that table's row, then right-click the row and select "Table Properties";
8) In the Table Properties window, click the Row tab and select the box that says "Repeat as header row at the top of each page."

3.4 Analysis of Chaotic Characteristics and Predictability of Delay Time Series Based on Recursive Graph

Analyze the predictability of the sequence by drawing a recursive graph and the randomness and determinism mechanism contained in the sequence. The recursive graph is drawn according to the following formula:

$$R_{ij} = \theta(\varepsilon - \|X_i - X_j\|) \quad (15)$$

Where: i, j=1, 2,..., N; ε is the given threshold; ||·|| is the distance between two points in the phase space; θ(·) is the Heaviside function, that is, θ(a <0)=0, θ(a≥0)=1. When Rij=1, mark the black dot at the graph coordinates (i, j), representing recursion.

4. Research on Optimization of Delayed Flight Adjustment

4.1 The Conditions for Establishing the Airport Transit Time Model Based on Bayesian Network

According to relevant civil aviation documents, the models are divided into four categories according to the passenger capacity. The first category is spacecraft with less than 60 seats. The second category is the 61-150 spacecraft; the third category is the 151-250 spacecraft. The four types are spacecraft with more than 250 seats.

Passing time: There are 12 kinds of conditions when passing the station. "<X" means that the passing time is below; "X-Y" means that the passing time is between and ">X" means that the passing time exceeds.

Forward flight delay time[11] The forward flight delay time is divided into 9 situations, and the division method is similar to the transit time.

First departure delay [12]: There are multiple flights every day at the same airport, and the airport of the departure station will delay the first flight. The first delay within 15 minutes is not regarded as a delay, 15-30 minutes is regarded as a slight delay, 30-an hour is regarded as a slight to mild-to-moderate delay, and 60-120 minutes is regarded as a slight-to-moderate delay to a more serious delay, more than 120 minutes is regarded as a serious delay. Delay of the last flight: There are multiple flights every day at the same airport, and the delay level of the last flight arriving at the destination airport has been implemented. A delay within 10 minutes after the arrival of the last flight is not considered a delay, 10-30 minutes is considered a slight delay, 30-60 minutes is considered a slight to moderate delay, and 60-120 minutes is considered a moderate to severe delay. And 120 minutes or more is considered a severe delay.

4.2 Airport Transit Time Model Based on Bayesian Network

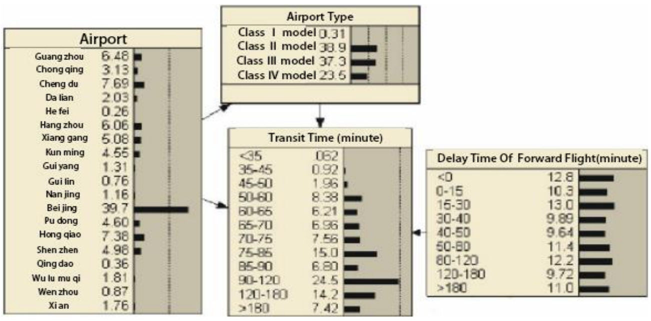

Figure 1: Bayesian network parameter learning results of single airport transit time

Obtain aircraft flight data information from a database query of a large and medium-sized international airline in China to create a Bayesian network structure model [13][14], and learn, train and train according to basic parameters to obtain the Bayesian network probability distribution function of the transit time Han number. As shown in Figure 1.

It can be seen from the figure that the airport, aircraft type and forward flight delay time all have an impact on flight time, so these three connection points are the parent nodes of time. Relatively speaking, large and medium-sized core regional airports have more types of three and four types of airports for medium-sized airports, so the airport is the parent node of the model category. The number of aircraft flying at Beijing Capital Airport accounts for 39.7% of the total number of aircraft. The analysis was carried out at a specific Capital International Airport, and the minimum transit time required in the relevant civil aviation documents was used as a reference to indicate the probability distribution functions of other airports. This situation may be similar to the Capital International Airport.

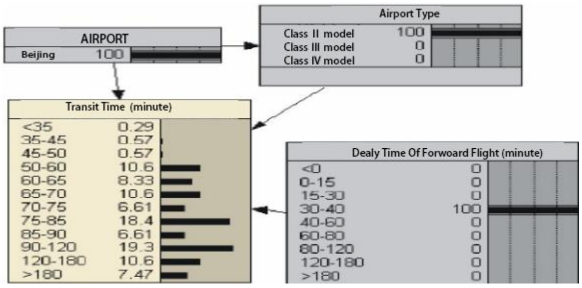

Figure 2: The distribution of transit time of 30-40min before flight delay

Flights that take off and land within 15 minutes of the departure time published on the scheduled flight schedule are normal flights. Therefore, the case where the delay time of forward flight is greater than 15 minutes is considered. Taking Capital Airport Type II aircraft as the research object, Figure2 and Figure3 show the distribution of transit time when the forward flight delay time is 30-40 min and 60-80 min, respectively.

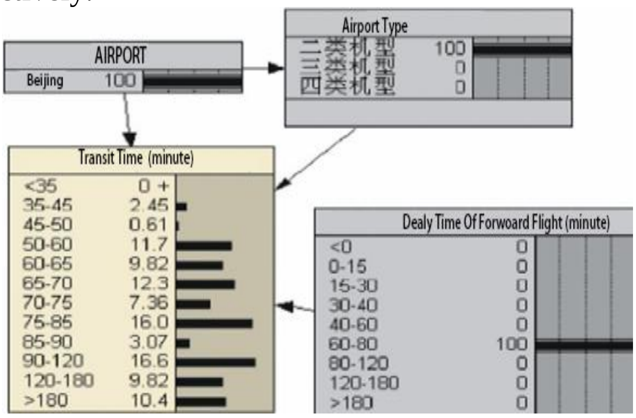

Figure 3: The distribution of transit time of 60-80min before flight delay

According to the relevant requirements of civil aviation, the shortest transit time of the second category airport of the Capital International Airport is 65 minutes. Figure2 shows that when the forward flight delay time is 30 to 40 minutes, 20.36% of the airports will pass the stop time. Adjusting the transit time to less than 65 minutes, 12.03% of the airports adjusted the transit time to less than 60 minutes; Figure3 shows that when the forward flight delay time is 60 to 80 minutes, 24.58% of the airports will the transit time is adjusted to 65 minutes. Next, 14.76% of airports adjusted the transit time to less than 60 minutes. Therefore, the longer the forward flight delay time, the greater the possibility and intensity of adjusting the airport transit time.

Table 1 lists the adjustment status of the completion time of the second type of airport under different forward flight delay conditions. In the table, the forward flight delay time, the special crossing time and the remaining time are all in minutes. The percentage of the total number of aircraft flights in Table I refers to the delay time of the airport and the forward flight, and the transit time obtained in the table is the percentage of the number of aircraft flights performed during the special transit time. The excess time refers to the deviation between the frequently required transit time (here 65min) in civil aviation documents and the special transit time.

Table 1: Capitial airport transit time adjustment table

Flight delay time	**Stop time**	**Percentage of delayed flights**	**Redundant time**
15—30	45——65	15.93%	0—20
30—40	50——65	18.93%	0—15
40—60	45——65	20.44%	0—20
60—80	50——65	21.52%	0—15
80—120	45——65	28.43%	0—20
120—180	50——65	23.12%	0—15
＞180	50——65	34.88%	0—15

The probability distribution functions of other models are basically the same, and the number of articles is relatively limited. This formula calculates that the airport transit time includes two parts: the road service project time and the sinking time. The data information in the table shows that the minimum crossing time required in the document involves a certain sinking time [15]. Therefore, in order to better reduce the spread of mid and downstream aircraft delays caused by forward flight delays, in order to better reduce The spread of delays to midstream and downstream aircraft flights has been adjusted to varying degrees during the remaining time of the sinking.

Acknowledgment

Airline delays are currently the primary problem to be solved urgently by major airlines, which have a huge negative impact on airline costs, passenger travel convenience, and the development of the aviation industry. Aiming at the control problem of spread delay, this paper studies the control method of spread delay in flight planning through data mining and model building. Flight delay control is a complex issue. This article has limitations in some aspects and some issues that need to be resolved later.

References

[1] Civil Aviation Administration of China. Statistical Bulletin on the Development of Civil Aviation Industry in 2019[R].2020.

[2] Arıkan U, Gürel S, Selim Aktürk M. Integrated aircraft and passenger recovery with cruise time controllability [J].Annals of Operations Research, 2016.

[3] Tu Y, Ball M O, Jank W S. Estimating flight departure delay distributions—a statistical approach with long-term trend and short-term pattern[J].Journal of the American Statistical Association, 2018, 103(481):112-125.

[4] Zonglei L, Jiandong W, Guansheng Z.A new method to alarm large scale of flights delay based on machine learning[C]. In: Knowledge Acquisition and Modeling, 2008. KAM'08.International Symposium on. IEEE, 2018:589-592.

[5] Wang Hong, Liu Jinlan, Cao Weidong, Xun Xiuxia. Airline flight delay early warning management model and analysis[J].Computer Simulation,2009,26(04):292-296.

[6] Xu Tao, Glory, Wang Jiandong. Analysis and early warning system of civil aviation flight delay spread based on SOA[J].Computer Science, 2009,36(07):157-160.

[7] Fang Kuangnan, Zhang Guijun, Zhang Huiying. Personal credit risk early warning method based on Lasso-logistic

model[J]. Quantitative Economics and Technical Economics, 2014, 31(02): 125-136.
[8] Wen Shao-Chun,Yang Cheng-Hsiung. Time series analysis and prediction of nonlinear systems with ensemble learning framework applied to deep learning neural networks [J]. Information Sciences, 2021,572.
[9] Ma Lan, Ou Shangheng. Chaotic characteristics analysis and short-term prediction of flight delay time series[J]. Chinese Science and Technology Paper, 2020, 15(06): 671-677.
[10] Miroslava Igondová,Dušan Cibulka. Precipitable Water Vapour and Zenith Total Delay time series and models over Slovakia and vicinity [J]. Contributions to Geophysics and Geodesy, 2010, 40(4).
[11] Derek J McKillop, Peter Auld. National turnaround time survey: professional consensus standards for optimal performance and thresholds considered to compromise efficient and effective clinical management [J]. Annals of Clinical Biochemistry, 2017, 54(1).
[12] Wu Weiwei, Meng Tingting, Zhang Haoyu. Research on optimization of flight planning based on airport delay prediction[J]. Transportation System Engineering and Information, 2016, 16(06): 189-195.
[13] Lida Huang, Guoray Cai, Hongyong Yuan, Jianguo Chen. A hybrid approach for identifying the structure of a Bayesian network model[J]. Expert Systems With Applications, 2019, 131.
[14] Bushra K. Ayshathil, Lasrado Sunitha, Prasad K.S. Detection of Lung Cancer by Modified Irregular Tree Structure Bayesian Network Model Based Image Segmentation[J]. Materials Today: Proceedings, 2019, 11(P3).
[15] Yong Tian, Bojia Ye, Lili Wan, Minhao Yang, Dawei Xing. Restricted Airspace Unit Identification Using Density-Based Spatial Clustering of Applications with Noise[J]. Sustainability, 2019, 11(21).