Research Article

A Method of Reducing Flight Delay by Exploring Internal Mechanism of Flight Delays

Yakun Cao1, Chenping Zhu1, Yanjun Wang2, and Qingyun Li2

1College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China
2Civil Aviation Institute, Nanjing University of Aeronautics and Astronautics, Nanjing 211100, China

Correspondence should be addressed to Yakun Cao; cyk0418@sina.com

Received 19 March 2019; Revised 30 July 2019; Accepted 30 August 2019; Published 30 December 2019

Academic Editor: Luigi Dell’Olio

Copyright © 2019 Yakun Cao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper explores the internal mechanism of flight departure delay for the Delta Air Lines (IATA-Code: DL) from the viewpoint of statistical law. We roughly divide all of delay factors into two sorts: propagation factor (PF), and nonpropagation factors (NPF). From the statistical results, we find that the distribution of the flight departure delay caused by only NPF exhibits obvious power law (PL) feature, which can be explained by queuing model, while the original distribution of flight departure delay follows the shift power law (SPL). The mechanism of SPL distribution of flight departure delay is considered as the results of the aircraft queue for take-off due to the airports congestion and the propagation delay caused by late-arriving aircraft. Based on the above mechanism, we develop a specific measure for formulating flight planning from the perspective of mathematical statistics, which is easy to implement and reduces flight delays without increasing operational costs. We analyze the punctuality performance for 10 of the busiest and the highest delay ratio airports from 155 airports where DL took off and landed in the second half of 2017. Then, the scheduled turnaround time for all flights and the average scheduled turnaround time for all aircraft operated by DL has been counted. At last, the effectiveness and practicability of our method is verified by the flights operation data of the first half of 2018.

1. Introduction

Flight delay is one of the major issues in aviation systems all over the world. Such delay events downgrade the functioning of airlines and cause tremendous loss in human life, economy and traffics [1, 2]. To alleviate the harm of flight delay, considerable work has been done [3–9]. Actually, the air transportation system is a rather complex system, which have been traditionally described as graphs with vertices representing airports and edges direct flights during a fixed time period [10]. These graphs are called aviation network. Recently, many research has been carried out from the viewpoint of complex network [11–13], which propose almost all kinds of aviation network features.

Many networks in nature display rather complex structures, that often seem random and unpredictable. Barabási and Albert discovered that many realistic networks [14, 15] exhibit the scale-free feature, which the vertex connectivity follows a PL distribution. The fundamental mechanisms leading to the PL distribution are considered to be growth and preferential attachment [16, 17]. On the other hand, in Ref. [18], the author proposed a SPL model with a parameter which controls the relative weights between the power-law and exponential behaviors. Empirical investigation for many real world networks [19–21] also shows SPL distribution. These work provide an effective theoretical support for us to explore the internal mechanism of flight delay and propose effective measures to alleviate the harm of flight delay.

There are many factors that cause flight delay, the Bureau of Transportation Statistics (BTS) classifies them into five categories [22]: (1) aircraft arriving late, (2) national aviation system (NAS) delay, (3) air carrier delay as a result of crew, baggage loading or maintenance problems, (4) extreme weather conditions such as hurricanes or blizzards and (5) security-related delays. If one flight is delayed, then a subsequent flight might also be delayed because it is awaiting that inbound aircraft. This kind of delay is called propagation delay [3, 5, 22–25], which is also quite substantial (more than one-third of the delays) [3]. On the other hand, since the schedule of one aircraft is quite tight, the on-route absorption of
departure delay of the last flight is very limited and the delay in subsequent flights is relatively predictable, while the delay caused by NPFs is hard to predict. Thus, quantitative research of propagation delay is great significance, which helps to come up with solutions.

In order to alleviate the delay of propagation, researchers have proposed to modify schedule departure time so as to re-allocate the existing slack in the flight schedule [3, 6, 26-28]. These studies share a similar research methods: they allow schedule departure time to vary within a time window, then establish an objective function with several constraints, and finally obtain the optimal solution. They focus on the impact of schedule modification on system performance to maximize the utilization of aviation resources. But we are more concerned about how to reduce flight delay ratio and hope to propose the concrete practicing method. In the follow, we propose a specific implementation method, not an objective function, although we used the same idea as the previous studies, that is, modify schedule departure time. We take advantage of the predictability of propagation delay and assume that there is no newly formed delay (delay caused by NPFs) after changing the plan, the effectiveness and practicability of our method is verified by the flights operation data of the first half of 2018.

The structure of this paper is organized as follows: Section 2 presents a statistical law for airline of DL and explores internal mechanism of flight delay. Section 3 contains analysis for operation performance evaluation of different airports and statistical results of the scheduled turnaround time for all flights and the average scheduled turnaround time for every aircraft. And the specific method is put forward. Section 4 presents and discusses the empirical results. In Section 5, conclusions and some hints for future research are given.

### 2. Statistical Law and Internal Mechanism

We collect primary records of flights operation from July 1, 2017 to December 31, 2017 for the Delta Air Lines. The data of flights operation were downloaded from the website of the Bureau of Transportation Statistics (BTS) [29]. Our analysis focuses on the departure delay rather than the arrival delay, because the arrival delay is approximately linearly related to the departure delay [30]. In general, the departure delay is commonly measured as the difference between the scheduled and the actual flight departure time. The Federal Aviation Administration (FAA) defines the flight departure delay as the flights departure at least 15 minutes behind schedule. The detail information for primary data is listed in Table 1.

![Figure 1: Log–log plots of PDF of the flight departure delay.](image)

#### Table 1: The information for DL airline in the second half of 2017.

<table>
<thead>
<tr>
<th>Airline</th>
<th>Number of flights</th>
<th>Number of airports</th>
<th>Number of aircraft</th>
<th>Flight departure delay ratio (delay more than 15 minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DL</td>
<td>470267</td>
<td>155</td>
<td>841</td>
<td>11.97%</td>
</tr>
</tbody>
</table>

For each flight, we can estimate the PDF of the flight’s departure delay distribution. The delay distribution shows attenuation trend, which is faster than the linear attenuation in double logarithmic chart. Therefore, we consider the departure delay distribution is well approximated by SPL:

\[
p(l) = c\cdot(l+\hat{\beta})^{-\gamma}.
\]

Shown in Figure 1, the fitting function \(p(l)\) of SPL can describe the empirical data very well. Statistical data shown as black filled circles, while red fitting line in panel describes the fitting result of Formula (1), in which the corresponding parameters \(c = 132.43, \hat{\beta} = 25.83,\) and \(\gamma = 2.74.\) The constants of \(c, \hat{\beta},\) and \(\gamma\) are estimated in the way of the least square fitting, and the goodness of fit is about \(R^2 \approx 0.999.\)

To explore the internal mechanism of flight departure delay, we first investigate the factors causing flight delays. As shown before, delay factors include five categories, we consider these five kinds of factors can be roughly divided into two sorts: the propagation factor (PF), i.e., category (1) aircraft arriving late, and the nonpropagation factor (NPF) which include all other four. Flight delays caused by NPFs are more accidental, while delay propagation has more direct relevance. Delay propagation occurs when late arrivals at an airport cause late departures, which in turn cause late arrivals at the destination airports. In general, the air traffic controller will set appropriate turnaround buffer time to prevent propagation delay when formulating flight planning [7], although this method reduces revenue-marking flight time and incurs schedule time costs. From the follow statistical results, we find that current measure of setting buffer time does not play a prominent role.

Actually, a key challenge to explore the internal mechanism of flight delay is extracting effective information from
the raw data. Because the existing data do not provide direct information to distinguish between the different types of delay factors [23]. The other reason is that flight delay may be not merely attributed by a late arrival of the flight immediately preceding it, but also be attributed by one or more other factors (NPFs). In order to quantitatively study the propagation delay and simplify the cause-explanation of late-arrival in the present work, we consider that: a delayed flight with the time between the last actual arrival and the current schedule departure less than $T_{tur}$ is attributed by PF. We know that the schedule turnaround time is consisting of two portions, namely the schedule buffer time and the standard ground service time [31]. For different types of aircraft, the required standard ground service time is about 30–50 minutes (generally speaking, the larger the passenger capacity of the aircraft, the longer the necessary ground service time). That means if the time between the last actual arrival and the current schedule departure is less than 30–50 minutes, it can be attributed to propagation delay.

To explore the impact of PF on the statistical law of the departure delay, we remove the departure delay causing by PF from the raw data. Since the data we collected without the information about the passenger capacity for different aircraft, we plot the departure delay (remove the delayed flights causing by PF) distribution $p_z(l)$ by setting $T_{tur} = 30, 40,$ and 50 for all aircraft in Figure 2.

It exhibits a PL distribution instead of a SPL distribution, given by
number of delayed flights caused by PF, and the fewer the number of delayed flights caused by NPFs. On the other hand, the smaller the $u_1D447/u1D461/u1D462/u1D45F$ value, the better the fit of the curve using the PL function.

In the statistical process, we use different thresholds $u1D447/u1D461/u1D462/u1D45F$ to obtain PL distributions, which shows that the distribution of flight delay caused by NPFs does exhibit the characteristics of PL distribution. To understand the origin of this observed PL distribution, we have to realize that the airport runway restrictions and the take-off queue size as the significant causal factors that affect the actual departure time [32]. One delayed flight caused by NPFs, such as extreme weather, the flights behind this at the same airport usually delay too. When emergencies return to normal, the waiting aircraft’s takeoff is a queuing process. Therefore, the distribution characteristic shown in Figure 2 can be regarded as the consequence of a decision-based queuing process [17, 33, 34]: when some perceived priority has been executed, the time of the planes waiting for take-off will show the characteristic of PL, with most flights rapidly take off, whereas a few experience very long waiting times.

Therefore, the mechanism of SPL distribution of flight departure delay is considered as the results of aircraft queue for take-off due to the airports congestion and the propagation delay caused by late-arriving aircraft.

3. Method

According to the previous mechanism of the flight delay, we can deal with the flight delay from two aspects, namely the airports congestion and the propagation delay. The most effective way to reduce queuing time is building multiple airport runways. However, it is a huge investment. From the perspective of statistics, a new method is developed to improve the flight on-time performance. This method consists of two stages: (1) data statistics and summarization; (2) implementation steps.

3.1. Data Statistics and Summarization. Due to the airports congestion, delay originating from these airports spreads to downstream flights. So the operation performance of airports plays a vital role in the punctuality ratio of airlines. The data that we collected not only contains the message of time for

![Figure 3: Log–log plots of departure delay distribution $p_3(l)$ for the 10 busy airports with the highest delay ratio.](image)

$$p_3(l) = c_2 \cdot \Gamma^{\gamma_2},$$

where $c_2$ is a constant and $\gamma_2$ is a constant parameter of the distribution known as the exponent or scaling parameter. We obtain the value of $c_2$ and $\gamma_2$ by the way of the least square fitting (after taking the log of the two sides and $\gamma_2$ becomes the slope of the line). As shown in Figure 2, the main part of the distributions fit well with the fit function of Formula (2), while the tail of distributions (larger delay) do not appear to be captured by it. However, the goodness of fit of $R^2$ for all distributions with different $T_{tar}$ is bigger than 0.99. From the data, we find that the number of delayed flights with delay $l$ bigger than 500 minutes is about 400–500, accounting for only 0.085–0.106% of the total number of flights. The fact that the scaling spans close to two orders of magnitude, from minutes to hours, indicates that most flight delays (70.51% for DL) are within less than one hour. With the increasing of the $T_{tar}$ value, the value of the distribution function is smaller. Obviously, the longer the necessary ground service time, the more the number of delayed flights caused by PF, and the fewer the number of delayed flights caused by NPFs. On the other hand, the smaller the $T_{tar}$ value, the better the fit of the curve using the PL function.

In the statistical process, we use different thresholds $T_{tar}$ to obtain PL distributions, which shows that the distribution of flight delay caused by NPFs does exhibit the characteristics of PL distribution. To understand the origin of this observed PL distribution, we have to realize that the airport runway restrictions and the take-off queue size as the significant causal factors that affect the actual departure time [32]. One delayed flight caused by NPFs, such as extreme weather, the flights behind this at the same airport usually delay too. When emergencies return to normal, the waiting aircraft’s takeoff is a queuing process. Therefore, the distribution characteristic shown in Figure 2 can be regarded as the consequence of a decision-based queuing process [17, 33, 34]: when some perceived priority has been executed, the time of the planes waiting for take-off will show the characteristic of PL, with most flights rapidly take off, whereas a few experience very long waiting times.

Therefore, the mechanism of SPL distribution of flight departure delay is considered as the results of aircraft queue for take-off due to the airports congestion and the propagation delay caused by late-arriving aircraft.

Table 2: Statistical results of 10 airports with the highest delay ratio.

<table>
<thead>
<tr>
<th>Airports (IATA-code)</th>
<th>Number of flights</th>
<th>Number of delayed flights (delay more than 15 minutes)</th>
<th>Total delay (in minutes)</th>
<th>Delay ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFO</td>
<td>6753</td>
<td>1298</td>
<td>82510</td>
<td>19.22%</td>
</tr>
<tr>
<td>EWR</td>
<td>2779</td>
<td>514</td>
<td>39920</td>
<td>18.49%</td>
</tr>
<tr>
<td>JFK</td>
<td>14472</td>
<td>2536</td>
<td>197146</td>
<td>17.52%</td>
</tr>
<tr>
<td>LGA</td>
<td>11289</td>
<td>1795</td>
<td>144010</td>
<td>15.90%</td>
</tr>
<tr>
<td>MIA</td>
<td>4668</td>
<td>727</td>
<td>54702</td>
<td>15.57%</td>
</tr>
<tr>
<td>PBI</td>
<td>2936</td>
<td>455</td>
<td>29964</td>
<td>15.49%</td>
</tr>
<tr>
<td>ORD</td>
<td>4200</td>
<td>644</td>
<td>44935</td>
<td>15.33%</td>
</tr>
<tr>
<td>BOS</td>
<td>7827</td>
<td>1170</td>
<td>97251</td>
<td>14.94%</td>
</tr>
<tr>
<td>LAX</td>
<td>16661</td>
<td>2464</td>
<td>139844</td>
<td>14.7%</td>
</tr>
<tr>
<td>SEA</td>
<td>12102</td>
<td>1679</td>
<td>94217</td>
<td>13.87%</td>
</tr>
</tbody>
</table>
departure and arrival, but also the carrier, tail number of aircraft and the airports for departure and arrival. Next, we assess the operation performance of each airport and compute the scheduled turnaround time for all flights and the average schedule turnaround time for every aircraft.

While recent studies on air traffic delays focus primarily on operation performance for the different airlines [22, 35], we are interested in operation performance for the different airports. As we know, airports are distributed in different locations, the punctuality ratio for different airports are very different due to the weather conditions and other regional factors. From our statistical results, we find there are 44 airports which have more than 2,000 taking-off flights in the second half of 2017 and 10 of 44 airports with the highest delay ratios are reported in Table 2. We can see that, airport of SEA has more delayed flights than BOS, but the total delay is smaller. That means the flight delay of airport BOS is mostly larger than SEA, so delay at airport BOS will have a greater impact on subsequent flights.

Initial delays affect the downstream flights, but small delays do not have much impact due to the scheduled turnaround buffer time. The study of delay distribution for various airports is necessary, not only delay ratio. In Figure 3, we compare the flight departure delay distributions of 10 airports. From Table 2, we know that airports of JFK, LGA, LAX, and SEA concentrate a large part of Delta Airline’s flights, but the characteristics of their delay distributions are not very different from each others. The shape of the delay distribution of different airports is similar, but small difference can only be observed when one focuses on EWR airport. The EWR airport shows a bias toward larger delays and may have a greater impact on subsequent flights than other airports.

The insufficient schedule turnaround time is another important factor for causing the propagation delay. The schedule turnaround time stands for the time spent by an aircraft on ground from scheduled arrival to scheduled departure from the gate, which is used for an aircraft to absorb last flight delay, complete full off-loading and loading maintenance of aircraft and where required, catering and cabin cleaning procedures. This measure is associated with airport operational efficiency and is used to improve the planning of flight connectivity and the robustness of flight plan. In our method, we will modify the existing flight schedule and redistribute part of the scheduled buffer time in the flight schedule without changing total slack time of the day and total daily number of flights.

In order to properly reset the slack, we count the scheduled turnaround time for all flight and the average scheduled turnaround time for all aircraft operated by DL in the second half of 2017. Since there are typically no flights between 0 and 6 o’clock, we do not take into account this longer time when calculating the scheduled turnaround time. On the other hand, records available in BTS are not always complete for all aircraft. To promote the quality of statistics, we take 100 flights within 6 months as the filtering threshold, which means that aircraft with their taking-off records smaller than 100 will not be counted into our statistics in the present work. After filtering, a total of 728 aircrafts are counted, and the total number of turned around for these aircraft is 347073.

The scheduling of aircraft turnarounds is a consequence of both the operational policies and the scheduling strategies of an airline. For different airlines, the average scheduled turnaround time is quite different, Southwest Airlines in the USA shows a low average aircraft turn time of 17 minutes and United Airlines an average turn time of 50 minutes [36]. In Ref. [36], we know that Delta Airlines shows an average turnaround time of 46.7 minutes, in which the database includes information from September 1987 to May 1994. According to our statistics, the average scheduled turnaround time $\bar{t}_{\text{dur}}$ of all flights is about 75.3 minutes and standard deviation $\sigma$ is about 92.9. This shows that the scheduled turnaround time of flights has increased greatly nowadays, it is particularly advantageous to our method of redistributing part of the schedule buffer time. Number distribution of the schedule turnaround time is shown in Figure 4(a), almost all flights’ scheduled turnaround time is longer than 30 minutes. So we set the minimum necessary turnaround time to be 30 minutes in our method.

![Figure 4: (a) Number distribution of the schedule turnaround time for 347073 flights. (b) Number distribution of the average schedule turnaround time for a total of 728 aircraft.](image-url)
Speciﬁc measures are as Figure 5, where \( t \) means schedule departure time, \( t' \) means schedule arrival time, \( t'_{\text{act1}} \) means actual departure time, \( t_{\text{act2}} \) means actual arrival time, \( t_{\text{buf}} \), \( t_{\text{ser}} \), and \( t_{\text{tur}} \) means scheduled buffer time, standard ground service time and scheduled turnaround time, respectively.

One aircraft flies from airport 1 to airport 4, if airport 1 belongs to one of the 10 busiest airports in the previous statistics, then we delay the scheduled departure time of flight 2 from \( t_2 \) to \( t'_{\text{act2}} \), and the amount of delay is equal to the scheduled turnaround time \( t_{\text{tur}} \) between the flight 2 and the flight 3 minus the necessary turnaround time \( T_{\text{tur}} \). All in all, if the time interval \( \Delta t \) between the actual arrival time of flight 1 and the schedule departure time of flight 2 is larger than required ground service time, flight 2 will take off on time.

3.2. Implementation Steps. The overall approach is based on the flight delay mechanism where newly formed delays usually occur at busy airports due to airport/airspace capacity constraints and they spread to downstream flights by the same aircraft. From our data, it is possible to trace the propagation of delay from airport to airport: if a particular aircraft is scheduled to fly from airport A to airport B and then to airport C and departs from A with a long delay, part or all of that delay will be propagated downstream and result in departure delay at B and, possibly, subsequently at C. In this section, we will develop a new method for formulating flight planning by using the previous statistical results.

Since the newly formed delay was hard to predict when we formulated the flight planning, we simply assume that flights departure from these 10 of the highest delay ratio airports mentioned above will experience this kind of delay. Actually, we cannot reduce the newly formed delay by optimizing flight plans, but we can mitigate the propagation effects of last flight delay by postponing the scheduled departure time of subsequent flights. On the other hand, we have to keep the scheduled departure time of the next flight unchanged and reserve enough turnaround time (greater than the necessary turnaround time) for the next flight. This means that we can delay the scheduled departure time of the current flight, and the maximum amount of delay is equal to the schedule buffer time between the current and the next flight operated by the same aircraft. According to our statistical results, the scheduled turnaround time varies greatly between different flights or different aircraft, but the required standard ground service time is about 30–50 minutes, so we set the necessary turnaround time \( T_{\text{tur}} \) to 30, 40, and 50 minutes as mentioned earlier.

Specific measures are as Figure 5, where \( t \) means schedule departure time, \( t' \) means schedule arrival time, \( t_{\text{act1}} \) means actual departure time, \( t_{\text{act2}} \) means actual arrival time, \( t_{\text{buf}} \), \( t_{\text{ser}} \), and \( t_{\text{tur}} \) means scheduled buffer time, standard ground service time and scheduled turnaround time, respectively.

In Figure 4(b), we can see that almost all aircraft’s average scheduled turnaround time is about 50–140 minutes. If we set the necessary turnaround time too large, then the change to the flight plan is small, and the effect of restraining delay propagation will not be obvious.
4. Empirical Results

In order to verify the effectiveness and practicability of our method, we collect additional six-month data of flight operation in the first half of 2018. We will use the method of this article to adjust the flight planning and compare the number of delayed flights before and after adjustment for the first six months of 2018. From the previous statistical results, we know that the 10 airports with the busiest and the highest delay ratio are SFO, EWR, JFK, LGA, MIA, PBI, ORD, BOS, LAX and SEA. We assume that if one flight departs from one of these 10 airports, it will generate newly formed delay and cause another flight immediately after it with the same aircraft also to delay. However, strictly speaking, the latter flight delay may be not merely attributed by a late arrival of the flight immediately preceding it, but also be attributed by one or more of other factors. In other words, sometimes the actual departure delay is hard to predict when we change the flight plan by our method, while the delay only caused by PF is not. Therefore, to simplify the prediction of current flight delays in the present work, we do not take into account the newly formed delay when the last flight by the same aircraft departed from one of the 10 highest delay ratio airports.

The six-month data comprehends 463322 flight operation records, and a total of 84828 flights departing from the 10 highest delay ratio airports. Actually, since there are typically no flights between 0 and 6 o’clock, delay on the last flight of each day does not propagate to the first flight of the next day. Therefore, without considering the delayed propagation of the last flight per day, we only adjust the schedule departure time for 72902 flights instead of 84828 flights. Comparing the results before and after adjustment, we find that the departure delay ratio dropped from 13.91% to 12.06%, 12.25% and 12.39% with $T_{ur}$ equal to 30 minutes, 40 minutes and 50 minutes, respectively. The change in the number of delayed flights in each delay interval is presented in Figure 6.

Obviously, we can see that the number of delayed flights in almost all delay intervals has decreased. And the smaller the necessary turnaround time $T_{ur}$, the more the delay and delay ratio will be reduced. But we cannot set $T_{ur}$ too small in our method, because large aircraft require a relatively long turnaround time, small $T_{ur}$ does not correspond to actual. The other reason is that the operation of the flights is full of many uncertain factors, the slack time is reserved to help deal with some unexpected situations and improve the robustness of the flight plan. On the other hand, our method is pretty effective in reducing flight delay, although it is not significant for flights with larger delay.

In addition, our approach is based on the predictability of propagation delays and mathematical induction, which provides a new way to optimize flight schedules. Although this is by no means intended as an exhaustive study, it nonetheless provides a starting point to motivate future research, which is more accurate forecasting of the newly formed delays and finding the optimal amount of slack that we redistributed.

Data Availability

The data used to support the findings of this study can be found from the website of the Bureau of Transportation Statistics (BTS) at http://www.bts.gov.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work is partially supported by National Nature Science Foundation of China under Grant Nos. 11775111, 61773203, and U1833126. And we thank Dr. Bo Zhou for revising our manuscript.

References


