Abstract— the traditional Air Transportation metrics, such as flight delays and flight cancellations, do not accurately and completely reflect the actual flight experience of passengers, and underestimate time penalties of cancellations on passenger trip time.

This paper describes and demonstrates algorithms calculating statistical measures of total passenger flight times between an origin/destination pair. The algorithms convert flight data to passenger trip data, based on a large quantity of public assessable historical data. A passenger-based metric, estimated passenger delay, is proposed to measure on-time performance from a passenger’s perspective. Ranking at origin-destination pairs, airport, and other results of analysis of 2004 data are provided.

Index Terms—Passenger Flow, On-Time Performance Measurement, Metrics, Network Properties

I. INTRODUCTION

The flight-based on-time performance metrics do not accurately reflect delays on passengers [1]. Bratu and Barnhart collected passenger complaints and negative reports for period 1995 to 2000, and compared them with flight-based performance metrics. They discovered the flight-based metrics could not explain the sharp increase of complaints and negative reports in 2000.

In Figure 1, we compare the flight-based metrics, such as number of delayed flights and number of flight cancellations, with passenger complaints collected by DOT on flight delays and cancellations in period 2002-2004. In year 2002, the air transportation system has 868,225 delayed arrivals, 65,143 cancellations, 829 complaints on flight delays, and 720 complaints on cancellations. We use those metric values in 2002 as baseline and plot the ratios of year 2003 and 2004. As the number of delayed and cancelled flights (the first two bars) increased in 2003 and 2004, the corresponding complaints (the third and fourth bars) decreased. The discrepancy between flight-based metrics and passenger feedback is clearly observed.

Delays, missed connections and cancellations are the reasons caused difference between flight experience and passenger experience. Firstly, flight-based metrics are constrained by the unit they use (per flight). They do not consider passenger related factors, like load factor, aircraft size, etc. So it is hard for them to explain the complicated scenarios of passenger flow and travel disruptions. Secondly, flight-based metrics underestimated the serious impacts of cancellations on passenger trip time. In August 2004, passengers on cancelled flight had been delayed by 13 hours averagely.

Consider the following two cases. Day 1 and day 2 both have 1000 scheduled flights and 10 of them are cancelled. The on-time performance of these two days will have no difference according to flight-based metrics. However, if the ten cancellations in day 1 happened in the late afternoon, within a short time period, and from one single airline, they will generate enormous amount of passenger delays due to shortage of resources. The ten cancellations in day 2 have a much scattered distribution through the day, and they are from different airlines that have lower load factors. Then passenger delays in day 2 could be much less compared with day 1.

“Flight delay has been one of the key indicators of system performance, and will continue to be an important indicator, more sophisticated delay metrics are needed to provide a more complete picture of performance (Bolczak and Hoffman 1997)”. To passengers, delay and cancellation are essentially the same. They both impose time penalties on passenger’s travel time. In this paper, we proposed a passenger-based on-time performance metric, passenger delay, to evaluate the on-time performance of domestic non-stop flights.

This paper is organized as follows. The background information and previous research are included in Section 2. Detailed algorithms involved in converting flight data to passenger trip data are discussed in Section 3. Major results are

Fig. 1 Comparison between flight-based metrics and passenger’s complaints on disrupted activities

* Y axis, “ratio” means the metric value in year 2003 or 2004 divided by the value of the same metric in baseline 2002. For example: # of delayed flights in year 2003 / # of delayed flight in baseline 2002 = 1.25
presented in Section 4, and finally conclusion and references are presented in Section 5 and 6.

II. BACKGROUND

Transportation systems are commonly represented using networks as analogy for their structure and flows. Physically, the air transportation network consists of routes and airports. Functionally, it has route characteristics like flight frequency, distance, etc associated with each route, and airport characteristics like airport capacity, runway layouts, etc. associated with each airport. The network as a whole can be affected by systematic characteristics like delay propagation, system capacity limitation, etc. The routes in the network have specific behaviors and these behavior patterns form the network properties.

The operational view of the air transportation network performance has been well developed. There are:
-- Flight metrics such as flight delays, flight cancellations, etc. used to measure the performance.
-- Simulation tools, such as TAAM, Vensim, etc., built to simulate network structure and to predict impacts on network performance given potential changes.
-- Research paper concerned in network structure and performance [3]-[12], etc.

The passenger or user’s view of the network performance is deficient. Major reason could be the lack of passenger trip data, which can only provided by airlines.

Bratu’s research on flight schedule reliability [2] has made a breakthrough in passenger on-time experience. He was provided with a few months’ passenger booking data by a single legacy airline. Combined these passenger booking data with ASPM data, he analyzed the passenger delays generated by missed connections and flight cancellations, and did sensitivity analysis of the relationship between factors like load factor, flight frequency, and passenger delays. Unfortunately, Bratu’s research is constrained by a few months’ data from a single airline, hence is not sufficient to be expanded to a system level analysis.

Table 1 shows the difference between Bratu’s research and the algorithms described in this paper. They are differentiated by four major aspects.

III. METHODOLOGY

Passenger delays are caused by three kinds of disrupted activities: delays, missed connections and cancellations. The goal of methodologies is to estimate passenger delay caused by these disrupted activities, using public accessible flight-based databases.

Figure 2 illustrates the converting process from flight data to passenger trip data. Raw data is collected from BTS public accessible databases:
- Airline on-time performance database – It contains “departure delays and arrival delays for non-stop domestic flights by major air carriers, and provides such additional items as origin and destination airports, flight numbers, cancelled or diverted flights and etc.” (Bureau of Transportation Statistics)
- Air carrier statistics database – It contains domestic non-stop segment data by aircraft type and service class for passengers, freight and mail transported, available capacity, scheduled departures, departures performed and aircraft hours, etc.” (Bureau of Transportation Statistics)

The ATOP data is flight-based, which means one record in the database represents a single flight in the system. It contributes operational flight data to the final passenger trip database. The T-100 data is aggregated data, which contributes

<table>
<thead>
<tr>
<th>TABLE I DIFFERENCE BETWEEN BRATU AND WANG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspects</td>
</tr>
<tr>
<td>Data</td>
</tr>
<tr>
<td>Goal</td>
</tr>
<tr>
<td>Perspective</td>
</tr>
<tr>
<td>Model</td>
</tr>
</tbody>
</table>

* AOTP is the Airline On-Time Performance data (from BTS)
* T-100 = Air Carrier Statistics (Form 41 Traffic) domestic data (from BTS)

Fig. 2 Converting Process from flight data into passenger trip data
passenger and seat data to the final passenger trip database.

Data processing and data joining algorithms have the following functions, as showed in Table 2.

After the data has been filtered, reorganized and rejoined, the most important algorithm, delay conversion algorithm is applied on it to convert flight data into passenger trip data. Passenger delays are estimated in this process.

### A. Delay Conversion Algorithm

Total passenger trip delay is a function of aircraft flight delays and cancellations. Tracing each passenger, we are able to find whether he/she arrived on time. If his/her flight was delayed, how long he/she has been delayed. If the flight was cancelled, which flight he/she was relocated to, and compared to the original schedule, how late he/she arrived at destination. Unfortunately, passenger travel data is not available to public. Therefore, we developed PATSP algorithm, which is designed to estimates passenger delays caused by flight delays and flight cancellations.

A flight here defined as a departure-arrival process. If a flight $i$ is delayed, the associated passenger delays are calculated as:

$$\text{PassengerDelay}(i) = (\# \text{ of passenger loaded on flight } i) \cdot (\text{ActualArrTime}(i) - \text{SchArrTime}(i))$$

We assume passengers in cancelled flights will be relocated to the nearest available flights belong to the same carrier and have the same origin-destination pair, if the available flights still have empty seats to fit in more passengers. Generally passengers from a cancelled flight will be relocated to 2 or 3 different flights due to limited empty seats on each available flight. For a specific passenger from the cancelled flight, the delay he or she experienced is calculated as the time difference between actual arrival time of the flight he or she relocated to, and scheduled arrival time of the original scheduled flight, that is the cancelled flight. Finally, the total passenger delays caused by this cancellation equal to the summation of passenger delays each passenger in the cancelled flight experiences separately.

$$\text{TotalPassengerDelay}(i) = \sum_{j=1}^{n} p(i) \cdot (\text{ActualArrTime}(j) - \text{SchArrTime}(i))$$

Here, $i = \text{cancelled flight}$  
$j = \text{available flight}$  
$n = \text{total number of available flights needed to finish relocating passengers on cancelled flight } i$  
$p(j) = \text{number of passengers relocated from flight } i \ \text{to flight } j$  
$\text{ActualArrTime}(j) = \text{actual arrival time of flight } j$  
$\text{SchArrTime}(i) = \text{scheduled arrival time of cancelled flight } i$

The major process of the algorithm is shown in Figure 3.

Table 3 gives an example of estimating passenger delays caused by cancellation. Assume a flight with 100 passengers is cancelled. Its scheduled arrival time is 12:00 pm. The first available flight has 30 empty seats and it arrives at 2:00 pm. The second available flight has 45 empty seats and it arrives at 3:00 pm. The third available flight has 40 empty seats and it arrives at 4:00 pm. So the passengers relocated to the first available flight will be delayed by 2 hours each. The passengers relocated to the second available flight will be delayed by 3 hours each, and 4 hours each for those relocated to the third available flight. Therefore the passenger delay caused by the cancelled flight is $2(\text{hr}) \times 30 + 3(\text{hr}) \times 45 + 4(\text{hr}) \times 25 = 295$ hours.
Assumptions for delay conversion algorithms are:

- To be conservative, we set 15 hours as the cap of passenger delays [1].
- Carrier and its subsidiaries will help each other relocating passengers from cancelled flights. For example, if a flight of American Airline (AA) is cancelled, passengers in this flight will be relocated to the nearest available flights, no matter it is from American Airline (AA) or its subsidiary American Eagle (MQ).

IV. RESULTS

We focus our research on domestic flights through OEP35 airports. OEP35 airports have the greatest number of operations and are heavily traveled. They account for 73% of total enplanements and 69% of total operations in the air transportation system [12]. The closed network formed by OEP35 airports generates 1044 directed routes in 2004.

Note: results shown in this section are all from the closed network formed by OEP35 airports in 2004.

A. Cancellations Disproportionally Generate High Passenger Delays

The disruption to passenger trip time caused by cancellations is underestimated. Simple metric, like number of cancellations, does not tell the complexity of passenger relocation process and the huge delays on passengers. According to our estimation results, passengers on cancelled flights in August 2004 were averagely delayed by 13 hours.

We depict the passenger delays at OEP 35 airports in 2004 in Figure 5. Total passenger delays consist of two parts: passenger delays caused by cancellations and passenger delays caused by flight delays. Averagely, 40% of total passenger delays were caused by flight cancellations, while cancellations only accounted for 1.7% of total flights. Among OEP35 airports, ORD generated the largest number of highest total passenger delays. At some of the OEP35 airports, such as MCO, TPA, CVG, IAD, and etc., cancelled flight which accounted for less than 3% of total flights, produced more than half of the total passenger delays.

As noted in Figure 5, passenger delays at different airports have specific patterns. For a single airport, passenger delays are seasonal dependent. We normalized the passenger delays by enplanements in order to prevent impacts from different level of enplanements in each season. Figure 6 shows the monthly normalized passenger delays at ORD in 2004.

Fig. 6 Example: seasonal change of passenger delays at ORD in 2004

As shown in Figure 6, normalized passenger fell below 400 thousand hours in low season such as April and October, but began to grow through the next 3 months until they peaked in midseason like January and May. The seasonal changes in passenger delays caused by cancellations are clearly evident. In general, passenger delays caused by cancellations are more sensitive to seasonal changes. Cancellations happen in midseason might have stronger impacts on passenger delays than cancellations in low season.

Flight experience describes the operational view of the air transportation system performance, while passenger experience describes the passenger view of the air transportation system performance. They review the system performance from different perspectives. Combining them together could provide us more complete and accurate views of the air transportation system performance.

B. Validate Consistency between Passenger’s Complaints and Estimated Passenger Delays

In Section 1, we depicted discrepancy between flight-based
metrics and passenger’s flight experience. We proposed a passenger-based metric, “passenger delay”, to measure on-time performance from a passenger’s perspective. Can this metric accurately reflect passenger’s experience? A consistency check is shown in Table 4.

Averagely there was one complaint on flight cancellations reported to DOT for every 94 cancellations, and one complaint on flight delays reported for every 1896 delayed flights. Though the number of delayed flights was much higher than the number of cancellations, average complaints on cancellations were 20 times more than complaints on flight delays. This can be well explained using passenger delays. As long as passengers experience an average delay of 98 minutes, they reported complaints on the disrupted activities, not matter the delays were caused by flight delays or by cancellations. Figure 6 confirms that passenger delays consist well with passenger’s on-time experience.

C. Differentiate Service on Routes

Routes in the system are not homogeneous. Each route has specific behavior. Routes have different levels of enplanements and flights due to limited airport and airspace capacities. We normalized the passenger delays by enplanements and flight delays by flights in order to achieve a fair comparison among routes. Figure 7 shows rankings of routes in terms of normalized passenger delays. Route JFK-ORD generated the biggest normalized passenger delays, 98 minutes per enplanement, in 2004. Large normalized passenger delays don’t necessarily indicate large normalized flight delays, since factors affecting passenger delays, like cancellations, aircraft size and load factor, do not affect flight delays. As observed in Figure 7, routes with the largest normalized passenger delays, such as JFK-ORD and PHL-EWR, are not necessary the routes with the largest normalized flight delays.

Scheduled trips of passengers might be seriously disrupted when the passenger delays are large, especially for connecting passengers. Risk analysis of passenger being delayed more than an hour is provided in the last column of Figure 7. Risk of long delays is a different concept to normalized passenger delays. It emphasizes more on number of passengers being severely delayed, instead of mean of passenger delays. Figure 7 confirms that high risk of long delays does not always followed by large normalized passenger delays.

Figure 8 gives another example of the divergence of passenger experience and flight experience on different routes. The top 20 most crowded routes are ranked by total enplanements in 2004, and the top 20 busiest routes are ranked by total number of departures in 2004. Large aircrafts are flown on routes with most enplanements. Conversely, routes with high flight frequency prefer smaller aircrafts than large aircrafts, especially on those short haul routes with commuter flights, such as SAN-LAX, LGA-BOS and DCA-BOS.

After normalizing the passenger delays with total enplanements on routes, we observed that flights on some of the short haul routes in Eastern area produced the largest normalized passenger delays, minutes per enplanement, in 2004. Large normalized passenger delays don’t necessarily indicate high large normalized flight delays, since high flight frequency prefer smaller aircrafts than large aircrafts, especially on those short haul routes with commuter flights, such as SAN-LAX, LGA-BOS and DCA-BOS.

In addition, those routes not only produced the highest normalized passenger delays, but also have highest risk of passenger being delayed more than an hour. The top 3 routes with highest probability of passenger being delayed more than an hour are EWR-BOS with probability 0.31, BWI-ROS with probability 0.23, and BWA-MEM with probability 0.22. Passenger delays propagated disproportionally from those

---

**TABLE 4**

**EXAMPLE: CONSISTENCY CHECK BETWEEN P ESTIMATED PASSENGER DELAYS AND PASSENGER’S COMPLAINTS (AUGUST 2004)**

<table>
<thead>
<tr>
<th>Flight-Based Metrics v.s. Passenger’s Complaints</th>
<th>Estimated Passenger Delays v.s. Passenger’s Complaints</th>
</tr>
</thead>
<tbody>
<tr>
<td># of cancellations per complaint on cancellations</td>
<td>Average PaxDelays caused by cancellations per complaint on cancellations</td>
</tr>
<tr>
<td>94</td>
<td>98</td>
</tr>
<tr>
<td># of delayed flights per complaint on flight delays</td>
<td>Average PaxDelays caused by delayed flights per complaint on flight delays</td>
</tr>
<tr>
<td>1896</td>
<td>98</td>
</tr>
</tbody>
</table>
eastern airports to other connected airports. Based on our estimation, 26% of the total passengers arrived at OEP35 airports in 2004 were transported from airports in eastern area\(^1\). But these 20% passengers propagated 35% of total passenger delays to their destinations.

V. CONCLUSIONS

This paper shows the divergence of flight on-time experience and passenger on-time experience. In order to achieve a more accurate and complete measurement of the air transportation system performance, a passenger-based metric, “estimated passenger delay”, is proposed in the paper.

Passenger delay measures the time penalties on passenger trip time due to both flight delays and flight cancellations. It captures the large delays caused by small amount of cancellations, and well reflects the actual delays to passengers. Compared with passenger delay analysis conducted by Bratu and Barnhart, algorithms in this paper founded on public accessible databases, and extend the analysis from restricted level (single airline, short time period) to system-level.

Future work is to apply the passenger relevant properties on different routes to build a simulation model that tracks and predicts passenger flow though the air transportation network.

Advent of internet and web-based ticketing system has enabled passengers to place a greater emphasis on ticket value. We hope to adjust passengers’ behavior by providing them with disruption-risk analysis of the routes and airlines they chose. Helping passengers, the direct buyers and users of the service, to differentiate themselves by providing both ticket value and disruption-risk prediction, could prompt the airlines to differentiate their services.

\(^1\) Airports in eastern area include: PHL, LGA, EWR, JFK, IAD, DCA, BWI, CVG, PIT and ORD

VI. REFERENCES


