Passenger Flow Simulation In A Complex Networked Transportation System

Danyi Wang  
Metron Aviation Inc.  
Herndon, VA, U.S.A.  
wang@metronaviation.com

Lance Sherry  
CATSR  
George Mason Univ.  
lsherry@gmu.edu

George Donohue  
CATSR  
George Mason Univ.  
gdonohue@gmu.edu

Bert Hackney  
Metron Aviation Inc.  
Herndon, VA, U.S.A.  
hackney@metronaviation.com

Abstract—Passenger trip time performance is positively correlated with passenger satisfaction, airfare elasticity, and airline profits. Researchers have demonstrated that flight metrics are a poor proxy for passenger trip experience. Trip delays experienced by passengers due to missed connections and cancelled flights are not negligible.

This paper describes a passenger flow simulation which captures the asymmetric and unique passenger trip on-time performance and reflects the complexity and significance of the impact of a small set of cancelled flights and missed connections on passenger trip delays. It measures system performance from the flying public’s view. Furthermore, it enables researchers to conduct experiments outside the range of historical data.

The results of this research provide decision makers with improved metrics for future investment decisions and better tools to manage the system. The passenger flow simulation model also provides the means to perform analysis for proposed changes to the system.

Keywords—on-time performance; passenger flow; performance metrics; passenger trip time

I. INTRODUCTION

The purpose of the Air Transportation System (ATS) is to provide safe and efficient transportation service of passengers and cargo. The on-time performance of a passenger’s trip is a critical performance measurement of the Quality of Service (QoS) provided by any Air Transportation System. QoS has been correlated with airline profitability, productivity, customer loyalty, and customer satisfaction [1].

Bratu et al. have shown that official government and airline on-time performance metrics (i.e. flight-centric measures of air transportation) fail to accurately reflect the passenger experience and underestimate the disruption on passenger trip time caused by cancelled flights and missed connections [2] [3] [4]. Flight-based metrics do not include the trip delays accrued by passengers who were re-booked due to cancelled flights or missed connections. Also, flight-based metrics do not quantify the magnitude of the delay (only the likelihood) and thus fail to provide the consumer with a useful assessment of the impact of a delay [5].

Research on passenger trip delay is limited because of the unavailability of proprietary airline data, which is also protected by anti-trust collusion concerns and civil liberty privacy restrictions. Wang et al. developed a set of algorithms designed to compute estimated passenger trip delay (EPTD) based on publicly available databases [6] [7] [8]. Results show disproportionately high passenger trip delays generated by cancelled flights. Cancelled flights accounted for only 1.4% of total scheduled flights in 2006, but they generated 39% of total EPTD. On average, passengers scheduled on cancelled flights in 2006 experienced 607 minutes of delay, while passengers scheduled on delayed flights experienced a much lower delay of 56 minutes. Except for the disproportionately high EPTD due to cancelled flights, Wang et al. proved passenger trip delay is a stochastic phenomenon that has asymmetric performance in terms of routes, airports, and time of year. Half of the total EPTD is generated by a smaller portion of routes (17%), airports (26%), and months (42%). Altogether, passenger behavior in the passenger tier of the air transportation system differs from flight behavior in the vehicle tier of the system.

Wang et al. designed the algorithm based on segment data, which doesn’t contain flight connection information. As a consequence, the analysis does not include the passenger trip delay caused by missed connections. Moreover, expansion of the air transportation system is trending out of the historical operation range with record high load factors, operations, and enplanements. This trend prohibits using historical data for analysis, since historical data cannot predict the impact of future policy changes on passenger trip time. In this paper, a passenger flow simulation (PFS) is developed to perform “future option design evaluation.” The PFS enables researchers to conduct experiments outside the range of historical data and estimates passenger trip delay not only due to delayed and cancelled flights, but also due to missed connections.

Section II of the paper describes the underlying concepts of the passenger flow simulation, PFS hierarchy, structure, algorithm, and results. Section III describes the experimental design for the PFS to identify significant factors for passenger trip performance and to perform sensitivity analysis.

II. PASSENGER FLOW SIMULATION

The operational evolution plan (OEP) 35 airports are the nation’s busiest airports defined by the FAA [9]. They have the greatest number of operations and account for 73% of total enplanements and 79% of total operations in the air transportation system [10]. The passenger flow simulation is a closed network formed by 34 of the OEP-35 airports. Honolulu International Airport (HNL) is excluded due to its geographic location and negligible impact on the network.
A. The Underlying Concept

Air transportation simulations of flight movement do not capture the passenger flow and connecting process. In the air transportation network, passengers cluster together into groups to fly from one airport to another. After arrival at the destination airport, this group of passengers breaks up: nonstop passengers make connections to ground transportation, and connecting passengers continue their trips by re-clustering with other passengers. Compared with flight movement, passenger movement

- Simulates passenger behavior instead of flight behavior;
- Converts flight information, such as arrival time and origin and destination airports, into attributes of passengers or groups of passengers;
- Converts flight schedules into clustering and scattering rules followed by passengers.

The passenger flow simulation is built to simulate this dynamic clustering and scattering process of passenger flow in the system. Air transportation simulations of flight movement do not capture the passenger flow and connecting process.

B. Colored Petri Net Modeling Tool

A Petri Net is a graphical and mathematical modeling tool. It is well-suited to modeling public transportation networks [11] and has been used to model the passenger connecting process in a public bus transportation system [12].

For accurate modeling of a complex transportation system like the air transportation system, a more complicated extension of Petri Net is required. In this paper, a hierarchical, timed, Colored Petri Net (CPN) is built using CPN Tools to simulate passenger flow and connecting processes in the system. CPN Tools is a graphical user interface for editing, simulating, and analyzing Colored Petri Nets [13]. CPN Tools can model the complex level of interactions in the air transportation system visually by creating nodes, transitions, and arcs in the model environment. This visual modeling environment allows users to track and understand the behavior of each passenger easily.

The concept of “color” distinguishes tokens (or resources) in the net. “PaxGroup” is defined as a color in PFS:

\[
\text{Color: } \text{PaxGroup} = (\text{Origin}) \times (\text{Dest}) \times (\# \text{ of Pax Loaded}) \times (\text{Aircraft Size}) \times (\text{SchDepTime}) \times (\text{SchArrTime}) \times (\text{Carrier}) \times (\text{FlightIndex}) \times (\# \text{ of Local Pax}) \text{ timed;}
\]

For example, the PaxGroup (DCA, ORD, 165, 200, 730, 850, 13, 45, 165)@+750 in Figure 1 represents a group of 165 passengers, loaded on United Airlines flight 45 with 200 seats, scheduled to depart from DCA at system time 730, and arrived at ORD at system time 850. However, this flight actually departed at system time 750, which is 20 minutes later than scheduled.

Tokens in places (circles) represent available resources to enable a transition (rectangle). The left part of Figure 1 shows one group of passengers in place “local pax” and two groups of connecting passengers in place “conn pax”. The first group of 20 connecting passengers arrived at the gate at time 715, and the second group of 15 connecting passengers arrived at the gate at 770. When the flight departed at time 750, the first group of 20 connecting passengers were loaded on time, whereas the second group of 15 connecting passengers missed their connections, since they arrived at the gate after the flight departed. The departing flight was scheduled to load 35 connecting passengers and depart with 200 total passengers at system time 730, but it actually loaded 20 connecting

---

1 Local passengers are passengers who have just appeared in the air transportation system. They could either be nonstop passengers from DCA to ORD or connecting passengers whose first leg flight is from DCA to ORD.
passengers and departed at system time 750 with 185 total passengers. The right part of Figure 10 shows the CPN after the transition: one group of 15 connecting passengers missed their connecting flight, and a group of 185 passengers (165 local + 20 connecting) were ready to gate out.

C. PFS Overview

Figure 2 depicts the correlations between algorithms and PFS. The algorithm section above the dotted line targets the “historical analysis.” In this section, different algorithms are designed to manipulate the data in different data processing phases. The “historical analysis” section sets the stage for “future option design evaluation.” As shown in Figure 2, processed data, algorithms, and the analysis report are embedded into the passenger flow simulation model as parameters, logical structure, and initial tokens. In other words, the parameter setting and passenger flow control in the PFS are based on historical statistics calculated by algorithms.

### Table 1: Overview of Passenger Flow Simulation (PFS) Structure

<table>
<thead>
<tr>
<th>Aspects</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airports</td>
<td>OEP-34 airports (excluding HNL)</td>
</tr>
<tr>
<td>Routes</td>
<td>1,030 routes formed by OEP-34 airports</td>
</tr>
<tr>
<td>Carriers</td>
<td>17 major carriers</td>
</tr>
<tr>
<td>Daily Flights</td>
<td>8,500</td>
</tr>
<tr>
<td>Daily Enplanements</td>
<td>900,000</td>
</tr>
<tr>
<td>PFS Modes</td>
<td>Deterministic and Stochastic</td>
</tr>
<tr>
<td>Hierarchy</td>
<td>3-level</td>
</tr>
<tr>
<td>Places</td>
<td>580</td>
</tr>
<tr>
<td>Transitions</td>
<td>343</td>
</tr>
<tr>
<td>Initial Tokens</td>
<td>20,000</td>
</tr>
<tr>
<td>Functions</td>
<td>42</td>
</tr>
</tbody>
</table>

The network structure of PFS is formed by 34 of the OEP-35 airports and the 1,030 routes between pairs of these airports. Passengers flow from one airport to another through the existing routes. Table 1 gives an overview of the PFS structure.

The PFS has two modes: deterministic and stochastic. They share the same PFS structure but use different functions and parameter values. In the deterministic PFS, passengers scheduled on a specific flight (e.g., UAL123 on route ORD and PFS DCA) will arrive at the actual arrival time (e.g., 1145). But in the stochastic PFS, the flight time is determined by a set of -

2 Not all of the 1122 possible city pairs are served by direct flights.
random number generators with specific means and standard deviations (e.g. flight time for UAL flights on route ORD-DCA is a normal random variable with $\mu=100$ minutes and $\sigma=15$ minutes). Thus the arrival time for flight UAL123 is a stochastic value generated as “DepTime + NormRNG ($\mu=100$, $\sigma=15$)”. In summary, the deterministic PFS is a pure conversion between flight performance and passenger performance, whereas the stochastic PFS allows flexibility and is more suitable for future option design evaluation.

D. PFS Hierarchy

The PFS has three levels, as shown in Figure 3. Airport and En Route subnets are represented as substitute transitions in the top-level net of PFS. These figures will be decomposed and described in the following paragraphs. Locations of the zoomed-in subfigures are labeled in Figure 3 (Figure 3.1 ~ 3.6). In addition, large figures showing the top-level, second-level and the third-level nets are available in Appendix A.

As shown in Figure 3.1, the top-level page depicts 34 airport substitute transitions, a single en route substitute transition, and the 68 ports connecting them. The zoomed-in inset shows an aggregated departure gate and an aggregated arrival gate for Washington-National Airport (DCA); there is one of these for each airport. An airport substitution transition is directionally connected to and from en route substitution through two ports, one representing an arrival gate and the other representing a departure gate (Zoomed-in figure is available in Appendix A).

In the second-level airport subnet, the flow process of passengers inside the airport boundary is divided into three steps: (1) splitting PaxGroup, (2) re-clustering of PaxGroup and (3) loading PaxGroup. In step 1, a PaxGroup arriving at the airport is split into two subgroups. The group of connecting passengers use the airport as a connecting hub, whereas the group of non-connecting passengers terminate their itineraries and leave the system at the airport. Different routes have different splitting rates for connecting and non-connecting passengers. In the example shown in Figure 3.2, 22% of passengers coming from ATL to ORD connect to another flight at ORD, while 78% terminate their trips at ORD.

In step 2 (Figure 3.3), connecting passengers are re-clustered to form a new group for the second-leg flights. There are two functions involved in the simulation code. One of the functions returns the minimal connecting time (MCT) required for passengers between gates. The other function divides a group of connecting passengers into several subgroups and sends them to different gates [6].

Finally, the newly formed PaxGroup is loaded onto their flights and ready to gate-out (Figure 3.4). The detailed loading process, which is the 3rd level subnet, will be explained in Figure 3.6.
All the functions, statistics, and ratios written in the transition code segment and on the arcs are responsible for guiding the passenger flow according to historical statistics. In this specific PFS, all the statistics are obtained from 2006 historical data provided by BTS. There are three BTS databases involved: AOTP, T-100 and DB1B. A detailed explanation of how to calculate the statistics is available in reference papers [6] and [8].

In the second-level en route subnet, each PaxGroup goes through taxi-out, air time, taxi-in, and finally reaches the arriving gate at destination airports (Figure 3.5). Functions in this subnet are responsible for reading attributes of PaxGroup, transporting passengers on the correct route (gate-to-gate), and assigning correct taxi-out, en route, and taxi-in times to the PaxGroup. The taxi-out time, air time, and taxi-in time are generated by random number generators, following some distributions with specific means and standard deviations calculated using 2006 data.

In the third-level passenger loading subnet (Figure 3.6), general connecting passengers flow to the upper branch and then are loaded onto the scheduled flights, while disrupted passengers (due to missed connections and cancels), flow to the lower branch and wait to be re-booked. Flights finished loading general connecting passengers will check for disrupted passengers before they depart. If disrupted passengers are detected, flights with available empty seats will load them until either no more seats are available or there are no more disrupted passengers. The general connecting passengers (upper branch) have higher priority than cancelled or missed.
connection passengers (lower branch). Passengers on both branches are sorted into first-come-first-serve airline queues to ensure passengers will be loaded on the correct flights (purchased flights) or will be re-booked by the same airline if disrupted (Appendix A Figure A.6).

**E. Passenger Missed Connection Algorithm in PFS**

Each experiment of PFS has two scenarios, the base scenario and the experimental scenario. The base scenario simulates passenger flow in an ideal environment without disruptions such as flight delays or cancellations. The goal of running the base scenario is to obtain passenger connecting information given a flight schedule. Passenger connecting information is then fed to the experimental scenario, which simulates delays, cancellations, and missed connections. As shown in Figure 4, passenger connecting information provided by the base scenario enables us to conduct research on missed connections in the experimental scenario. The simulation results of the experimental scenario estimate EPTD not only due to delayed and cancelled flights but also due to missed connections.

**F. PFS Sample Results**

July 6, 2005 is a randomly chosen weekday in summer 2005. Flight performance on July 6, 2005 was as follows:

- Scheduled Flights = 8,540;
- Delayed Flights = 1,764 = 21% of Scheduled Flights;
- Cancelled Flights = 176 = 2% of Scheduled Flights.

We used PFS to simulate passenger flow and calculate passenger trip delay on July 6, 2005. As shown in Figure 5, 2% of cancelled passengers generated 30% of total EPTD, 1% of missed connection passengers generated 15% of total EPTD, and 21% of delayed passengers generated 55% of total EPTD. On average, passengers scheduled on cancelled flights experienced 403 minutes of delay, missed connection passengers experienced 341 minutes of delay, and passengers scheduled on delayed flights experienced 64 minutes of delay.

**III. EXPERIMENTAL DESIGN FOR PFS**

The purpose of the experiment design is to identify and rank the significant factors that have strong impacts on passenger trip time and to analyze the sensitivity of EPTD given changes in these factors.

Based on experience and literature, six items are chosen as initial significant factors. These factors are shown in Table 2. Results of the experiments will prove how good the initial “guess” of significant factors is, and at what level they affect the passenger trip delay. A full factorial design for six factors, assuming a linear response function, needs $2^6 = 64$ total runs, and each run requires two PFS scenarios (base and experimental scenarios). In total, 128 PFS models need to be built and executed for a full factorial design. Concerned about time, we performed a fractional factorial design with six factors, two levels (high and low) and $\frac{1}{8}$ fraction. Table 2 lists the six factors and their high and low levels.

**TABLE 2 HIGH AND LOW LEVEL SETTINGS FOR FACTORS**

<table>
<thead>
<tr>
<th>Factors</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td># Passengers Loaded</td>
<td>Increased by 5%</td>
<td>Decreased by 15%</td>
</tr>
<tr>
<td>Aircraft Size (# of seats)</td>
<td>Increased by 15%</td>
<td>Decreased by 5%</td>
</tr>
<tr>
<td>Airline Cooperation Policy</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Flight Delay</td>
<td>+ 15 minutes</td>
<td>- 15 minutes</td>
</tr>
<tr>
<td>Cancellation Time</td>
<td>Cancelled four hours earlier</td>
<td>Remain the same cancellation time</td>
</tr>
<tr>
<td>Minimal Connecting Time</td>
<td>+ 15 minutes</td>
<td>- 15 minutes</td>
</tr>
</tbody>
</table>

The high and low levels of “# of passengers” and “aircraft size” are designed to keep load factor in the range of [61%, 92%]. The highest value of load factor (92%) occurs in experiments with “# pax” = H and “aircraft size” = L, while the lowest value of load factor (61%) occurs in experiments with “# pax” = L and “aircraft size” = H. Airline cooperation policy indicates whether airlines on the same route cooperate with each other on re-booking disrupted passengers. If not, disrupted passengers must stick with the same airline for re-booking.

The rank order of significant factors in terms of the absolute value of coefficients for NAS-wide total EPTD is depicted in Figure 6. The most significant factor to total EPTD is flight delay, which is obvious, since more than half of the total EPTD is due to delayed flights. Along with flight delay, number of passengers, flight cancellation time, and airline cooperation policy also have significant impacts on total EPTD.
The rank order changes from case to case. For example, the rank order of factors in terms of EPTD due to cancelled flights is: airline cooperation policy, aircraft size and number of passengers (or load factor), flight cancellation time, as shown in Figure 7. These three factors have stronger impacts on EPTD due to cancelled flights than any other factors, since they are directly related to re-booking flexibility and resource availability.

As shown in Figure 8, the rank order changes for total EPTD due to missed connections. The most significant factors affecting total EPTD due to missed connections are: flight delay, cancellation time, and minimal connecting time. The risk of missing a connecting flight increases if a previous flight leg is delayed. Affected passengers, whether due to flight cancellation or missed connections, compete for limited resources. As a consequence, cancellation time has a strong impact on EPTD due to missed connections. If the airport is poorly designed, connecting passengers may need longer minimal connecting time to travel from one gate to another, and this may result in missing connecting flights.

In summary, the significant factors for different cases are as follows:

- To reduce total EPTD: decrease flight delay, encourage airline cooperation, earlier cancellation time, and lower load factor
- To reduce EPTD due to cancelled flights: encourage airline cooperation, lower load factor, and earlier cancellation time
- To reduce EPTD due to missed connections: decrease flight delay, decrease minimal connecting time required and encourage earlier cancellation time
- To reduce EPTD due to delayed flights: less flight delay and fewer passengers loaded.

A simple sensitivity analysis is done for a better understanding of the impact of factors on EPTD. As shown in Table 3, change in a single factor can result in 8% to 24% less total EPTD, thereby saving millions of dollars per day.

<table>
<thead>
<tr>
<th>Changes in a single factor</th>
<th>Compared with Total EPTD on July 6, 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Decrease in total EPTD (hours per day)</td>
</tr>
<tr>
<td>Reduce flight delay</td>
<td>Decreased by 15 minutes</td>
</tr>
<tr>
<td>by 15 minutes</td>
<td></td>
</tr>
<tr>
<td>Encourage airline cooperation</td>
<td>Decreased by 12%</td>
</tr>
<tr>
<td>Cancel flights 4 hrs earlier</td>
<td>Decreased by 10%</td>
</tr>
<tr>
<td>Reduce load factor</td>
<td>Decreased by 8%</td>
</tr>
<tr>
<td>from 83% to 70%</td>
<td></td>
</tr>
</tbody>
</table>

Officials, operators, and service providers should consider the combined effect of factors on EPTD, which helps to achieve the strategic goals with minimal changes or costs.
IV. CONCLUSIONS

The goal of air transportation service is to provide safe, affordable, and convenient transport for passengers and cargo. As a consequence, the top level performance measures of the ATS should include the trip delays experienced by airline passengers. Passenger-based metrics, together with flight-based metrics, can give a more accurate and complete description of the ATS performance.

The passenger flow simulation captures the asymmetric and unique passenger trip on-time performance and reflects the complexity and significance of the impact of a small set of cancelled flights and missed connections on passenger trip delays. Major findings of this research are listed as follows:

1) High passenger trip delays are disproportionately generated by cancelled flights and missed connections.

2) Passenger-based metrics are needed to capture the passenger travel experience, since flight-based metrics can unintentionally distort the actual performance of the system and effectively “hide” explanatory and diagnostic system behavior.

3) Congestion flight delay, load factor, flight cancellation time, and airline cooperation policy are the most significant factors affecting total EPTD in the system. The combined effect of multiple factors should be investigated and used to support the decisions made by officials, policymakers, and researchers.

4) Passengers should treat trip time as a stochastic phenomenon that can be assigned a probability of occurrence but cannot be avoided entirely in any systematic manner. Simple strategies can be used by passengers to reduce the probability of occurrence, such as choice of departure airport and route. For example, for a trip from Washington, D.C. to Chicago, flights from DCA to MDW had a 5% probability of more than one hour delay, whereas flights from DCA to ORD had 12% probability of more than one hour delay.

REFERENCES

Appendix A

Figure A.1

Figure 3.1

Top Level

Airport Substitute Transition

Enroute Substitute Transition

2nd Level

Airport Subnet

En-Route Subnet

3rd Level

Passenger Loading Subnet

Figure A.2
Figure A.3
Figure A.6