Flight Delay Dynamics: Unraveling the Impact of Airport-Network-Spilled Propagation on Airline On-Time Performance

Yajun Lu

Assistant Professor of Analytics & Operations Management
Jacksonville State University
ylu@jsu.edu

Joint work with:

Yi Tan[a] and Lu Wang[b]

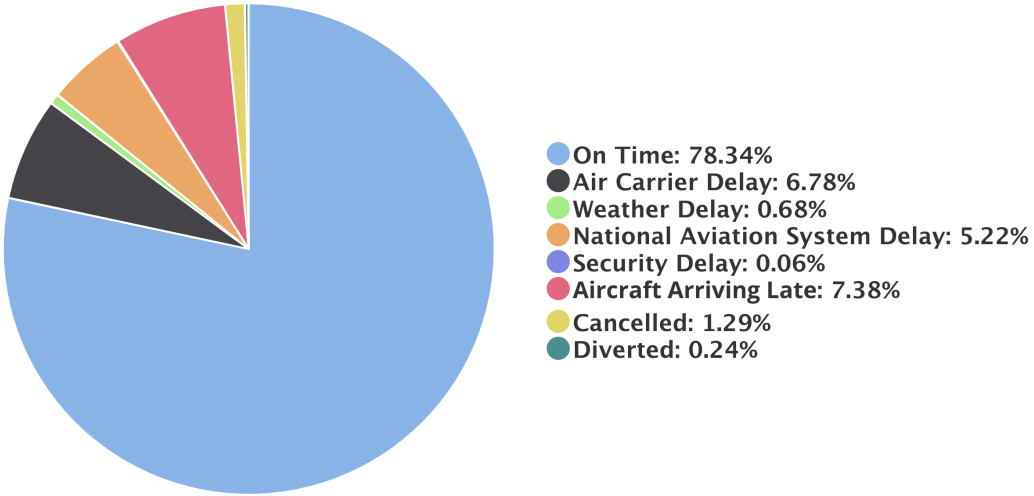
[a]University of Alabama in Huntsville, [b]Ball State University

May 10, 2025

Outline

Outline

On-Time Arrival Performance in 2023



Source: Bureau of Transportation Statistics (BTS)

Impact of flight delay

Domestic flight delays cost the U.S. economy around $32.9 billion annually (Ball et al., 2010).

- Passengers
  - Lost time
  - Additional expenses
  - Overall dissatisfaction

Impact of flight delay

- Airlines: potential penalties and operational costs
  - New DOT rule *49 USC 42305 mandates refunds for cancelled or significantly delayed/changed flights*
  - *Effective October 28, 2024*

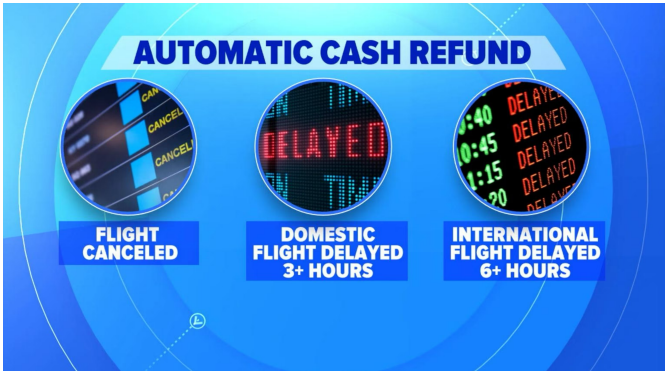


Image Source: ABC news

Impact of flight delay

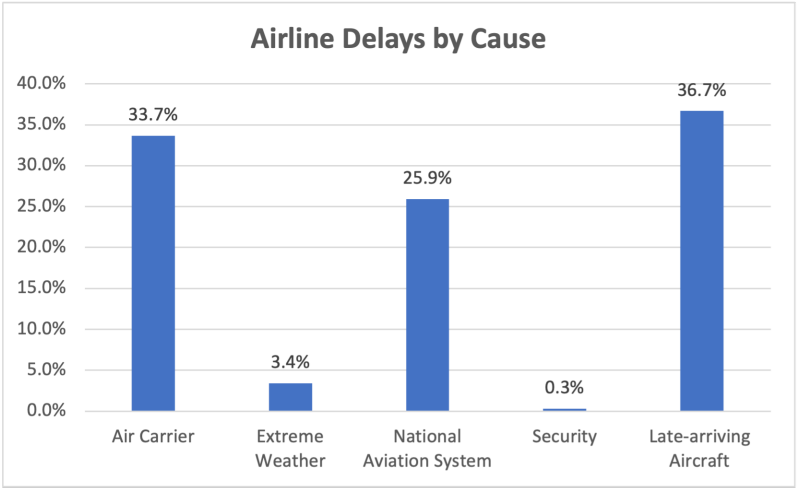
- Airports: diminished terminal and runway efficiency

Image Source: travelfreak.com

Benefits of Early Flight Delay Detection

- Improve airport resource allocation and operation efficiency (Rebollo and Balakrishnan, 2014).
- Enhance airline and air traffic controller decision-making (Wang et al., 2003).
- Save cost (Ball et al., 2010).

Causes of flight delay



BTS categorize flight delays into five categories

Late-arriving aircraft

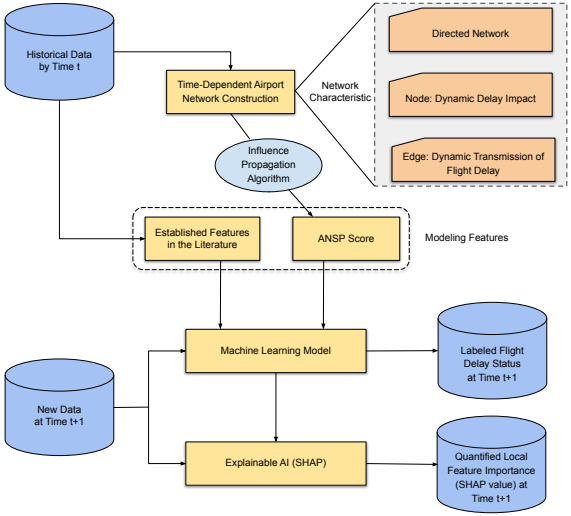
Existing work focus on the propagation of delays between subsequent flights

We investigate the propagation of flight delays across airports.

- *Airport-Network-Spilled Propagation* (ANSP) Mechnism.

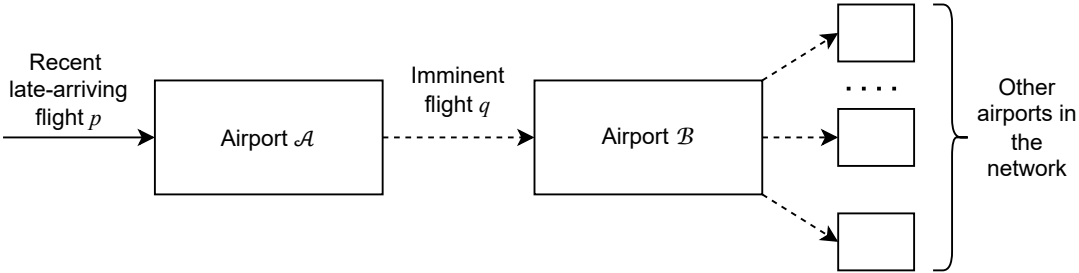
  - Flight delays at one airport can be transmitted to connected airports through flights.
- We develop a novel time-dependent, network-based approach to model the ANSP mechanism.

Decision Support System for Flight Delay Detection

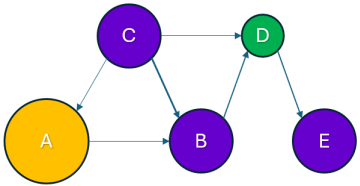


Outline

A simplified scenario of ANSP Mechanism



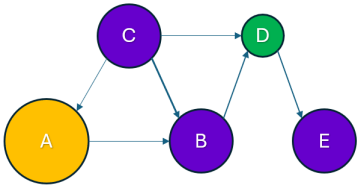
Time-dependent Airport Network Model



Weighted directed network

- Airports are represented as nodes.
  - The node size indicates the delay influence.
- Scheduled imminent flights are represented as edges.

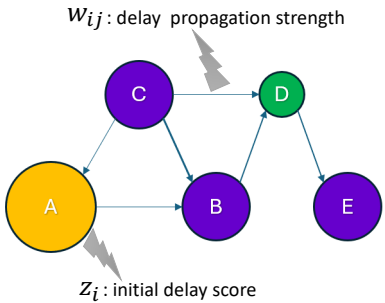
Time-dependent Airport Network Model



Weighted directed network

- Airports are represented as nodes.
  - The node size indicates the delay influence.
- Scheduled imminent flights are represented as edges.
  - The edge weight indicates the propagation strength.

Time-dependent Airport Network Model



- Initial delay score for each airport i: $z_i = \sum_{k=1}^{m} e^{-\gamma t_k}$
  - γ: decay constant
  - t_k: time elapsed since the delay occurrence
- Delay influence transmission strength: $w_{ij} = e^{-\beta s}$
  - β: intensify rate
  - s: time interval to next schedule arrival
- Influence propagation algorithm (Page *et al.* 1999) to derive ANSP scores:

$$\vec{\xi} = \alpha \mathbf{W}_{norm} \vec{\xi} + (1-\alpha)\vec{z} \qquad (1)$$

  - ξ: vector of ANSP score for each airport
  - $\mathbf{W}_{norm}$: normalized delay transmission matrix
  - Solved with power-iteration method (Tong et al., 2006)

Outline

Data

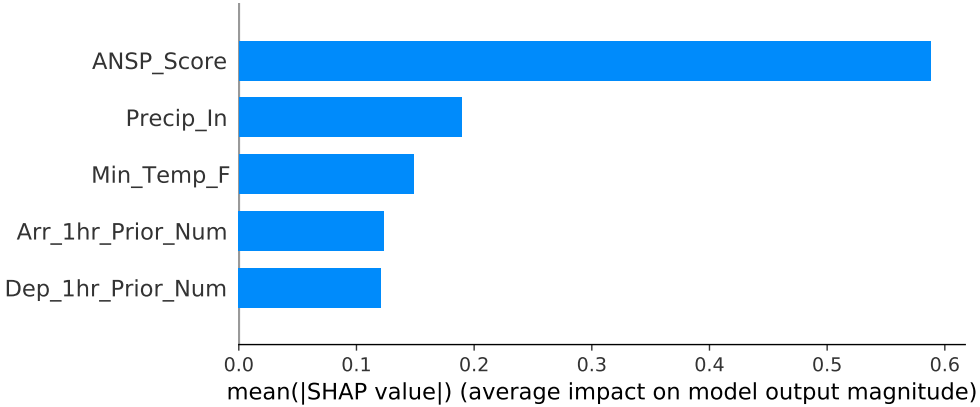
- Data sources: FAA, Bureau of Transportation Statistics (BTS), the Iowa Environmental Mesonet (IEM)
- Domestic flights among Large Hub airports (30).
- Data period from June 1 to July 31, 2023.

	Delayed		Not Delayed		All	
Variable	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Arr_1hr_Prior_Num	26.6351	17.8025	23.4015	18.6038	24.4297	18.4145
Arr_1hr_Post_Num	27.5002	17.9473	27.9598	18.7795	27.8136	18.5202
Dep_1hr_Prior_Num	27.8005	17.3124	24.6531	16.9450	25.6539	17.1255
Dep_1hr_Post_Num	26.7346	18.0315	27.3462	17.5380	27.1517	17.6987
Max_Temp_F	87.0774	10.5877	86.2735	11.1112	86.5291	10.9539
Min_Temp_F	68.5045	8.9009	67.4464	9.2473	67.7828	9.1518
Avg_Wind_Speed_Kts	6.7452	2.2925	6.7508	2.3261	6.7490	2.3155
Precip_In	0.1442	0.4125	0.0846	0.2874	0.1036	0.3335
Scheduled_Turnaround_Lessthan60	0.7945	0.4041	0.7129	0.4524	0.7388	0.4393
Betweenness_Centrality	0.0023	0.0010	0.0022	0.0010	0.0022	0.0010
Closeness_Centrality	0.9558	0.0590	0.9554	0.0593	0.9555	0.0592
ANSP_Score	1.9211	1.1385	1.2893	1.0620	1.4902	1.1260
N	134,508		288,521		423,029	

Empirical Results

Classifier	Feature Set	AUC
XGBoost	Baseline	0.7282
	Baseline + ANSP Score	0.7456
Random forest	Baseline	0.7071
	Baseline + ANSP Score	0.7243
Bagged logistic	Baseline	0.6487
	Baseline + ANSP Score	0.7036
ANN	Baseline	0.6913
	Baseline + ANSP Score	0.7287

SHAP Analysis



Aggregated SHAP scores

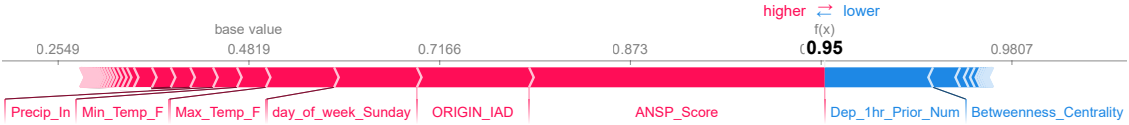
A Snapshot of Two Example Flights

Route	Flight Date	Day of week	CRS (Scheduled) Departure Time	Actual Departure Time	CRS (Scheduled) Arrival Time	Actual Arrival Time
IAD to BNA	7/16/2023	Sunday	22:20	0:10 (+1 day)	23:05	0:53 (+1 day)
PHX to AUS	7/6/2023	Thursday	6:15	6:13	10:25	10:16

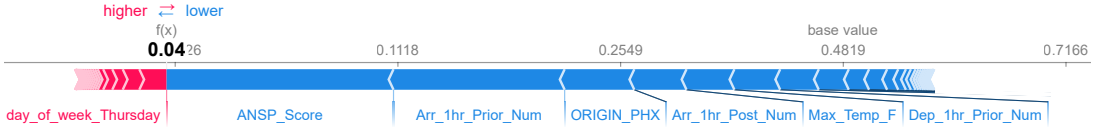
Operating Carrier	ANSP score	Actual Delay Status	Arr_1hr_Prior_Num	Arr_1hr_Post_Num	Dep_1hr_Prior_Num	Dep_1hr_Post_Num
G4 (Allegiant)	3.6261	1	23	17	0	0
WN (Southwest)	0.1940	0	0	4	18	19

Min_Temp_F	Max_Temp_F	Precip_In	Avg_Wind_Speed_Kts	Betweenness_Centrality	Closeness_Centrality	Scheduled_Turnaround_Lessthan60
71	92	0.14	3.2152	0.0019	0.9063	1
85	115	0	6.2566	0.0015	0.9667	0

Local Explanations



IAD – BNA



PHX – AUS

Outline

Conclusion

- The ANSP mechanism that describes how flight delays influence propagate throughout the entire airport network.
- A novel time-dependent, network-based approach to model the ANSP mechanism.
- A decision support system for early detection of departure delays and identification of the determining factors.
- Identifying important features in flight departure delay detection.

Q & A

ylu@jsu.edu
http://yajunlu.com

Reference I

M. Ball, C. Barnhart, M. Dresner, M. Hansen, K. Neels, A. Odoni, E. Peterson, L. Sherry, A. Trani, and B. Zou. Total delay impact study: a comprehensive assessment of the costs and impacts of flight delay in the united states. `https://rosap.ntl.bts.gov/view/dot/6234`, 2010. Accessed: September 23, 2024.

J. J. Rebollo and H. Balakrishnan. Characterization and prediction of air traffic delays. *Transportation research part C: Emerging technologies*, 44:231–241, 2014.

H. Tong, C. Faloutsos, and J.-Y. Pan. Fast random walk with restart and its applications. In *Sixth international conference on data mining (ICDM'06)*, pages 613–622. IEEE, 2006.

P. T. Wang, L. A. Schaefer, and L. A. Wojcik. Flight connections and their impacts on delay propagation. In *Digital Avionics Systems Conference, 2003. DASC'03. The 22nd*, volume 1, pages 5–B. IEEE, 2003.