Architecture Technology Corporation
Specialists in Computer Architecture

NOVEL, MULTIDISCIPLINARY GLOBAL OPTIMIZATION UNDER UNCERTAINTY PHASE II

NASA ROA-2015/D.2 Leading Edge Aeronautics Research for NASA Project (LEARN1)

Grant Number NNX1AR29A

Final Briefing for the 2015 LEARN Phase II

August 16, 2016

Version: 1

Architecture Technology Corporation
Dr. Ole Mengshoel, CMU Silicon Valley
Mr. Ralph Tamburro, PANYNJ
Captain William Cotton, Cotton Aviation Enterprises
Dr. Nikunj Oza, NASA Ames

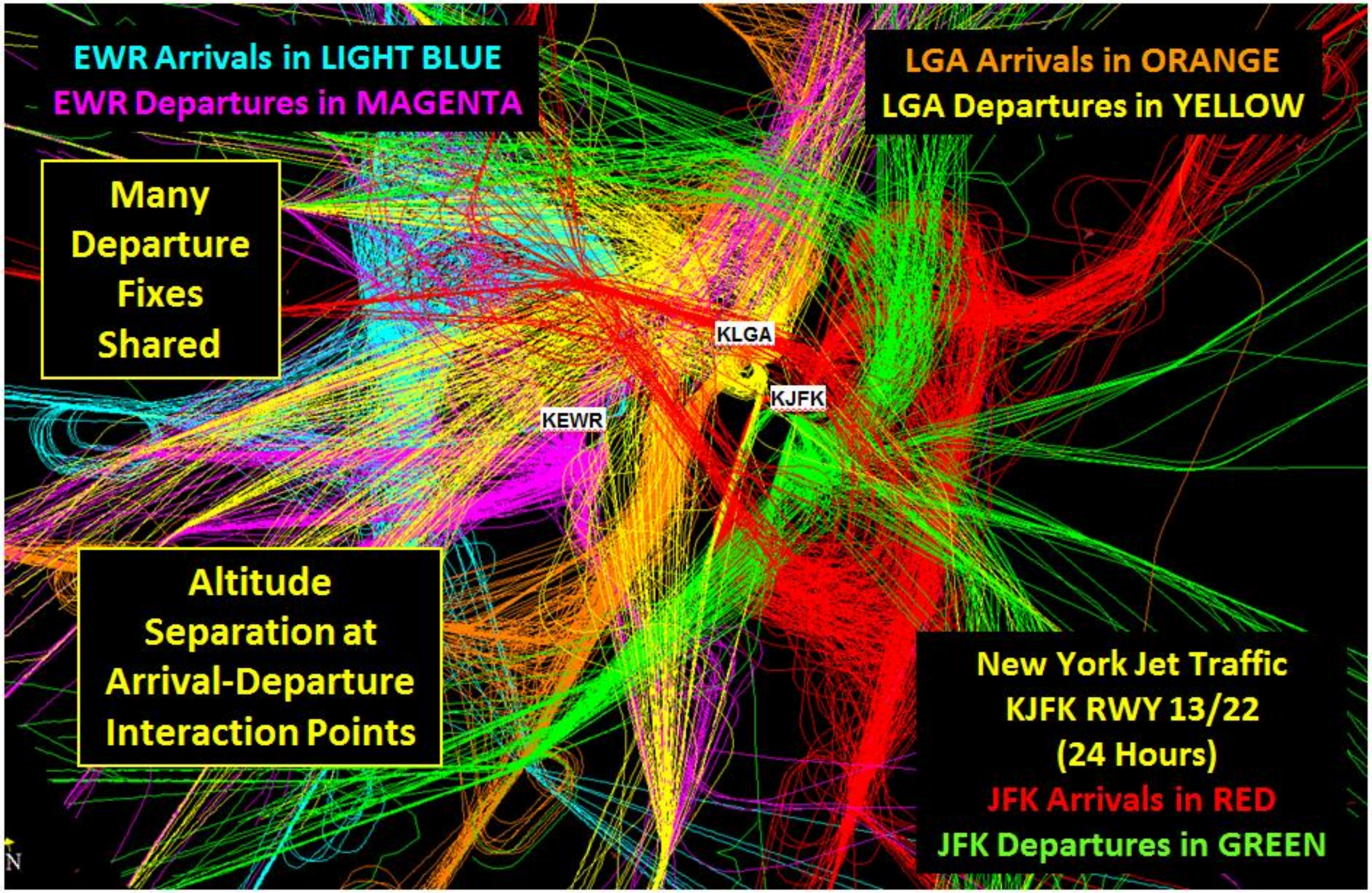
Outline

- Phase I Review and Summary
- Phase II Objectives
- Phase II Tasks
- SOSS Airport Models
- SOSS Traffic Scenarios
- Scheduling Algorithms
- Bayesian Network(BN) Evaluation and Development
- PROCAST Simulation and Evaluation
- Results, Discussion, and Future Work

The ATM Metroplex Problem

- Two or more busy airports in close proximity
- Shared entry/exit points to the terminal airspace
- Inter-dependent, crossing arrival and departure flows
- Several traffic control facilities involved

The New York Metroplex



Source: Georgia Tech, Saab Sensis Corp., ATAC Corp., Metron Aviation, "Final Briefing for NASA NRA Characterization of and Concepts for Metroplex Operations," at NASA Langley Research Center, Nov. 2009

The Challenges to an Optimized, De-conflicted, 4 Dimensional Trajectory Solution

- Complex interactions and network impacts
  - Requires integrated planning across airport surface and terminal airspace
- Uncertain future traffic behavior
  - Requires planning under the possibility of multiple different futures
- Competing and nonlinear objectives
  - Requires optimization-algorithms capable of handling complex objective functions

Phase I Research Objectives

- Develop four dimensional trajectory-based traffic management tool called PROCAST by combining technologies from two diverse fields
  - Predictive technology/Data Science: Bayesian Networks (BNs)
  - Optimization technology: Genetic Algorithms (GAs)

- Perform proof-of-concept demonstration by conducting simulation experiments using a test problem—New York metroplex traffic scheduling
  - In Phase I, we focus on a single-airport, arrival-departure-surface scheduling problem
  - Selected John F. Kennedy International Airport (JFK) as the focus site

- Enhance NASA simulation platform to enable terminal airspace traffic simulation and pre-pushback process modeling

PROCAST—Probabilistic Robust Optimization of Complex Aeronautics Systems Technology

Phase I Summary

- PROCAST showed significant benefits in proof-of-concept simulation experiments
  - 3000 hours of metroplex delays saved (assuming 100 days of similar conditions)

- Predictive component by itself (BNs-only) showed benefit by increasing the accuracy of predictions
  - Benefits increased with increasing number of evaluations-over-potential-futures.
  - Speed of computation limited our ability to assess scheduling over a large number of possible futures

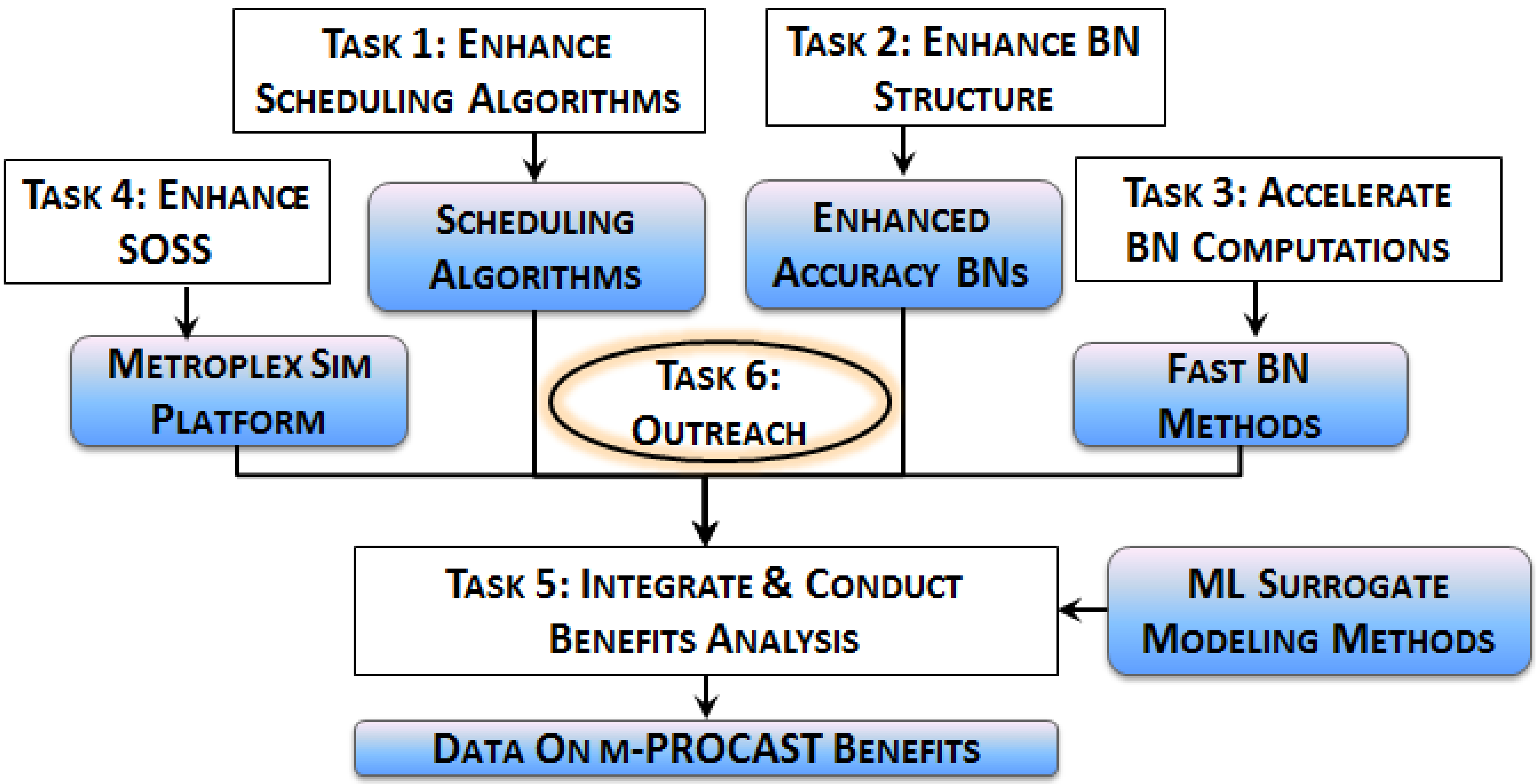
- Optimization-only component (GA-only) added only a small amount of benefit
  - Apparently sensitive to uncertainty in gate pushback readiness times

- Published Phase I work in two papers:
  - Digital Avionics Systems Conference 2014: *“Robust, Integrated Arrival-Departure-Surface Scheduling Based On Bayesian Networks”*
  - AIAA Aviation Technology Integration and Operations Conference 2015: *“A Robust And Practical Decision Support Tool For Integrated Arrival Departure Surface Traffic Management”*

Phase II Objectives

- Extend aircraft trajectory-based traffic scheduling tool
  - Coordinate departures & arrivals for JFK, EWR and LGA airports
  - Utilize Bayesian Networks (BNs) to predict aircraft taxi times and account for uncertainty
- Further investigate Bayesian Networks (BNs) and application to aircraft transit time modeling
  - Revise & enhance Phase I BN model for JFK as needed
  - Develop BN models for EWR and LGA
- Enhance NASA's SOSS simulation platform and perform multi-airport simulation experiments
  - Develop airport & airspace models for EWR and LGA
  - Implement framework for multi-airport simulations
- Collaborate with NASA, FAA, and industry

Overview of Phase II Tasks



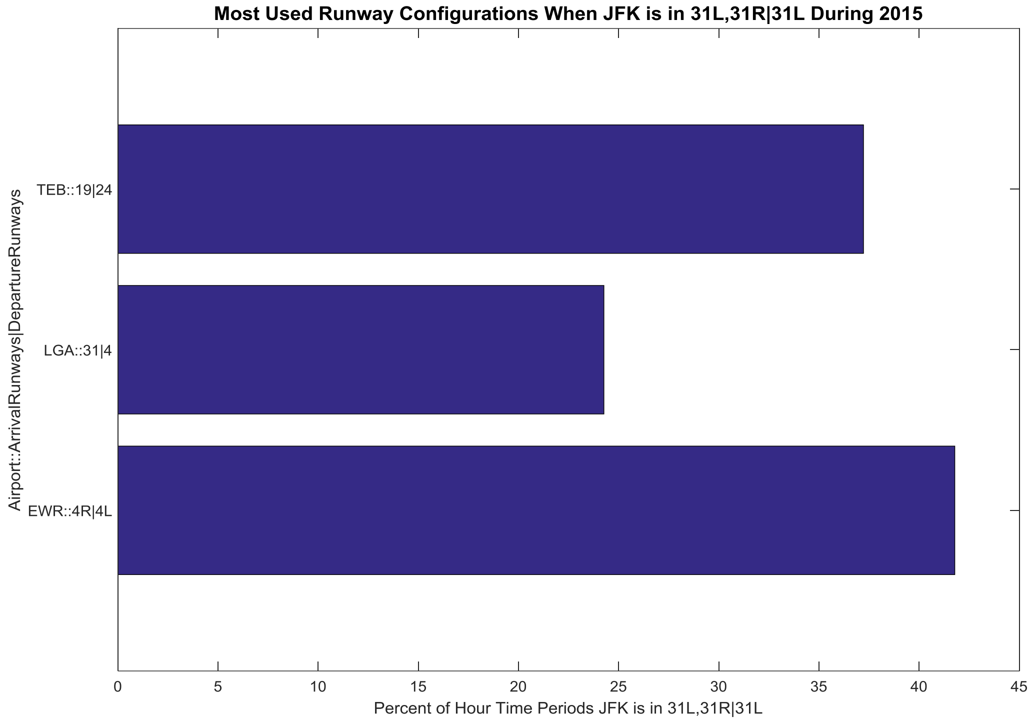
Acknowledgments

- SOSS Adaptation and Traffic Scenarios
  - Valentino Felipe and Sebastian Timar (ATCorp)
  - Robert Windhorst (NASA)
  - Ralph Tamburro (PANYNJ) and Bill Cotton (Cotton Aviation)
- BN Evaluation and Development
  - Ole Mengshoel, Aniruddha Basak, Priya Sundararajan, Erika Menezes, Vinodh Paramesh (CMU)
  - Matt Stillerman (ATCorp)
- Multi-Airport Scheduler
  - Mark Peters (ATCorp)
- Surrogate Modeling and Machine Learning
  - Nikunj Oza (NASA)

SOSS Airport Adaptation

- In Phase II we develop link-node models of the surface and (limited) terminal area for EWR and LGA airports in SOSS.
- The airport model generation process involves the following steps:
  - Selecting the airport runway configuration
  - Generating the required airport adaptation data
    - Surface link-node model, runway geometry, terminal procedures, arrival and departure fixes, surface routes, runway separation requirements
  - Generating traffic scenarios

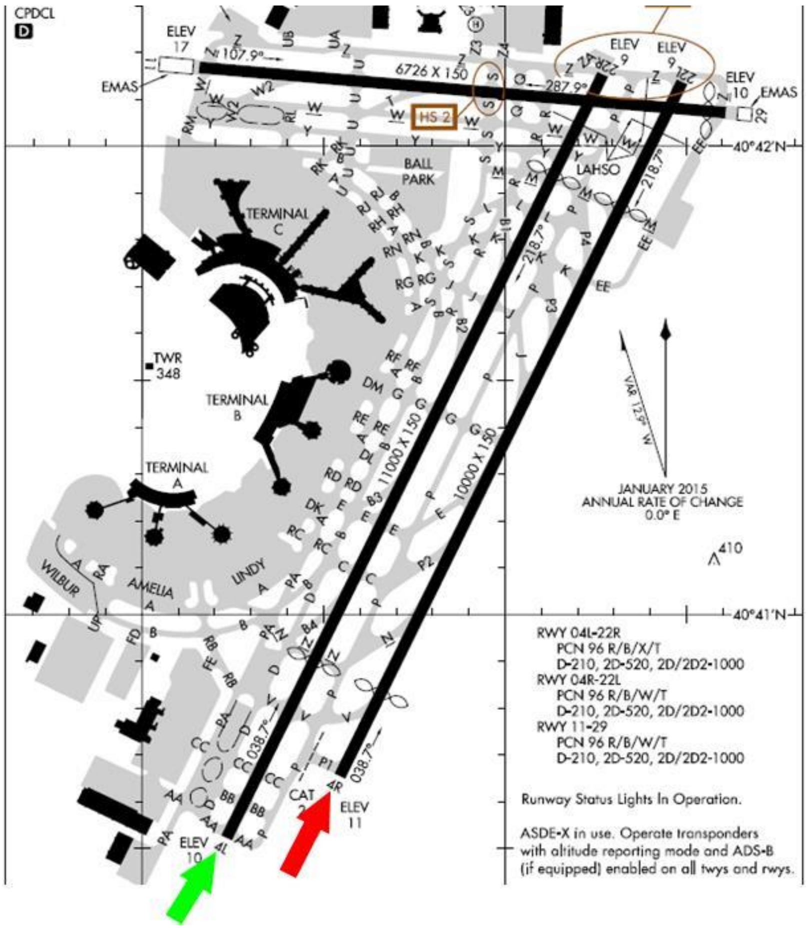
Airport Runway Configurations

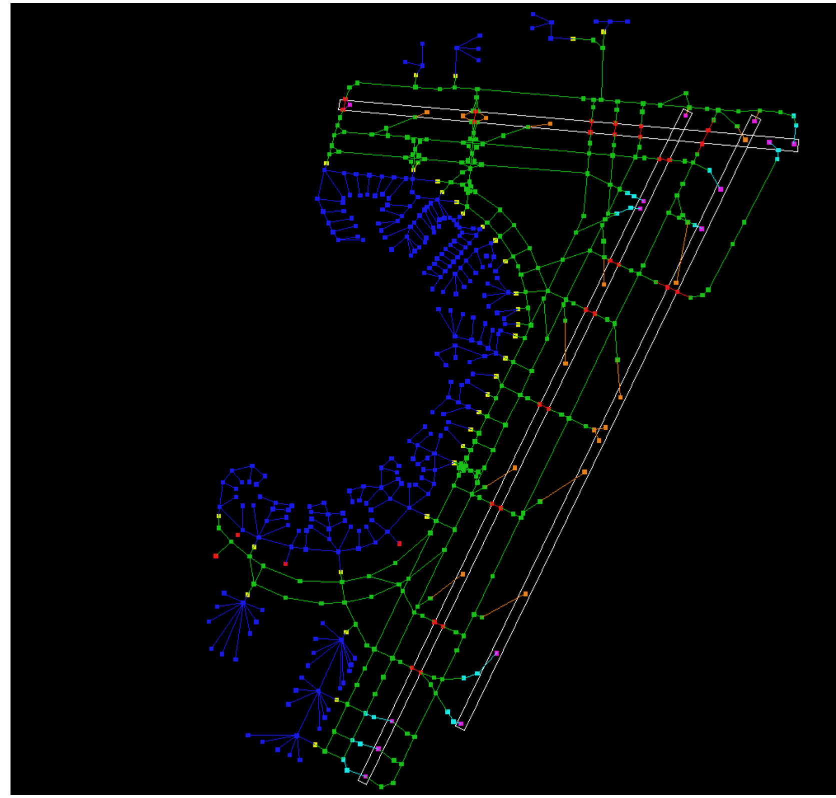


- Runway configurations are based on the JFK runway configuration used in Phase I: arrivals using both runways 31L and 31R, and departures using runway 31L.
- We analyzed FAA ASPM data from 2015 to determine the most commonly used runway configurations for the other New York area airports.
  - We model LGA with arrivals using runway 31 and departures using runway 04 (31|04).
  - We model EWR with arrivals using runway 04R and departures using runway 04L (04R|04L).

SOSS Adaptation for EWR

Runway Configuration: 4R|4L

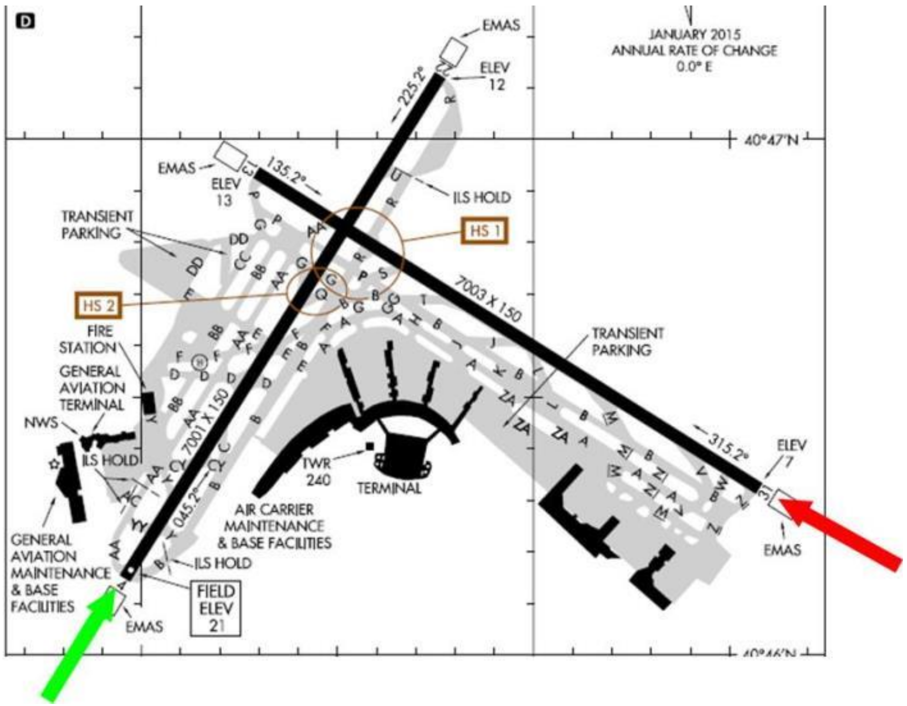


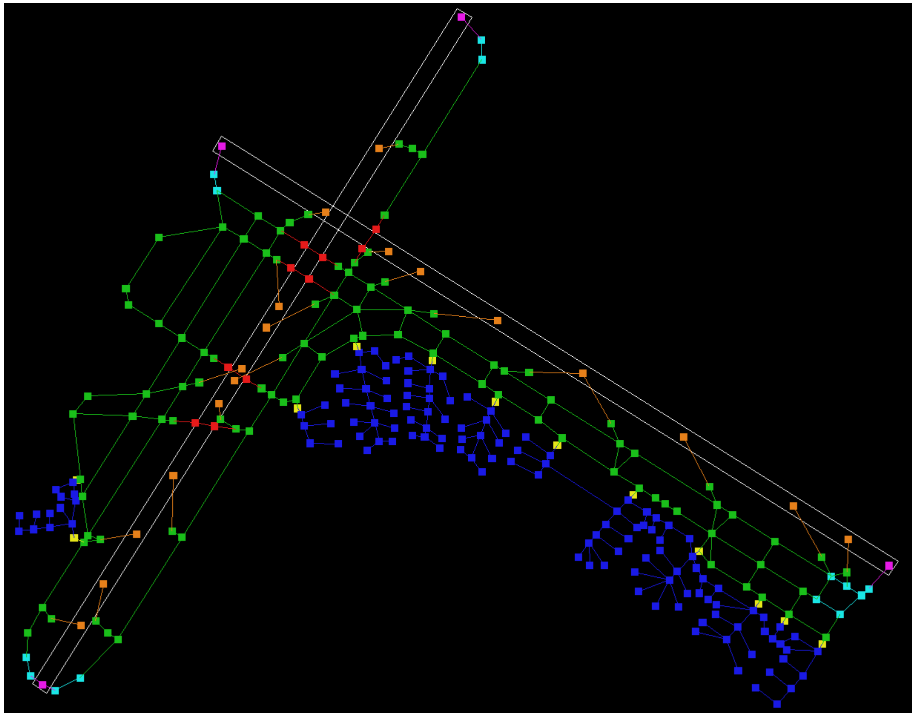


Node Type	Quantity	Link Type	Quantity
Arrival	20	Arrival	19
Departure	14	Departure	2
Queue	20	Queue	157
Runway Crossing	28	Runway Crossing	18
Taxiway	225	Taxiway	356
Spot	32	Spot	1
Ramp	99	Ramp	356
Gate	158	Gate	18

SOSS Adaptation for LGA

Runway Configuration: 31|4





Node Type	Quantity	Link Type	Quantity
Arrival	18	Arrival	19
Departure	4	Departure	5
Queue	14	Queue	9
Runway Crossing	10	Runway Crossing	15
Taxiway	96	Taxiway	141
Spot	12	Spot	1
Ramp	57	Ramp	61
Gate	74	Gate	74

LEARN Phase II Traffic Scenarios

	EWR	JFK	LGA
Validation	9/5/12 18-23 Local 7/25/12 6-11 Local	5/13/12 2-7 Local 3/16/12 13-16 Local	9/5/12 18-23 Local 7/25/12 6-11 Local
Training	5/13/12 0-23 Local 6/11/12 0-23 Local 9/5/12 0-23 Local	5/13/12 0-23 Local 6/11/12 0-23 Local 9/5/12 0-23 Local	5/13/12 0-23 Local 6/11/12 0-23 Local 9/5/12 0-23 Local
Evaluation	7/25/12 8-10 Local	7/25/12 8-10 Local	7/25/12 8-10 Local

- We obtained historical demand schedules generated by the FAA's Air Traffic Organization–Planning (ATO-P) representing 16 days in the NAS during the 2011 – 2012 time period.
- We use a subset of historical demand schedules to generate traffic scenarios for validation, training, and eventual evaluation of our PROCAST implementation.
- For SOSS validation, we developed traffic scenarios for time periods during which the airports were actually operating with the runway configurations used in our SOSS adaptations.
- We then compare SOSS traffic counts and taxi-in/taxi-out times with FAA ASPM data for those time periods of interest.

Validation of SOSS Airport Models

- Compared aircraft taxi times from SOSS simulation with FAA airport operational data at JFK, EWR and LGA
  - Days/time periods where airports operated in runway configurations modeled in our SOSS adaptations
  - Created SOSS input traffic files from FAA flight schedule data for those days/time periods
  - Compared SOSS output aircraft transit times to FAA airport operational data for those days/time periods
- Findings
  - Consistency in transit times in SOSS simulation
  - Some difference in average transit times between FAA ASPM and SOSS simulation data due to modeling approximations

Multi-Airport Scheduling Algorithms

- The LEARN Phase II arrival and departure scheduler is an emulation of the concepts and capabilities of the Departure Metering System (DMS) and Time Based Flow Management (TBFM) decision support tools currently in use at NY area airports.

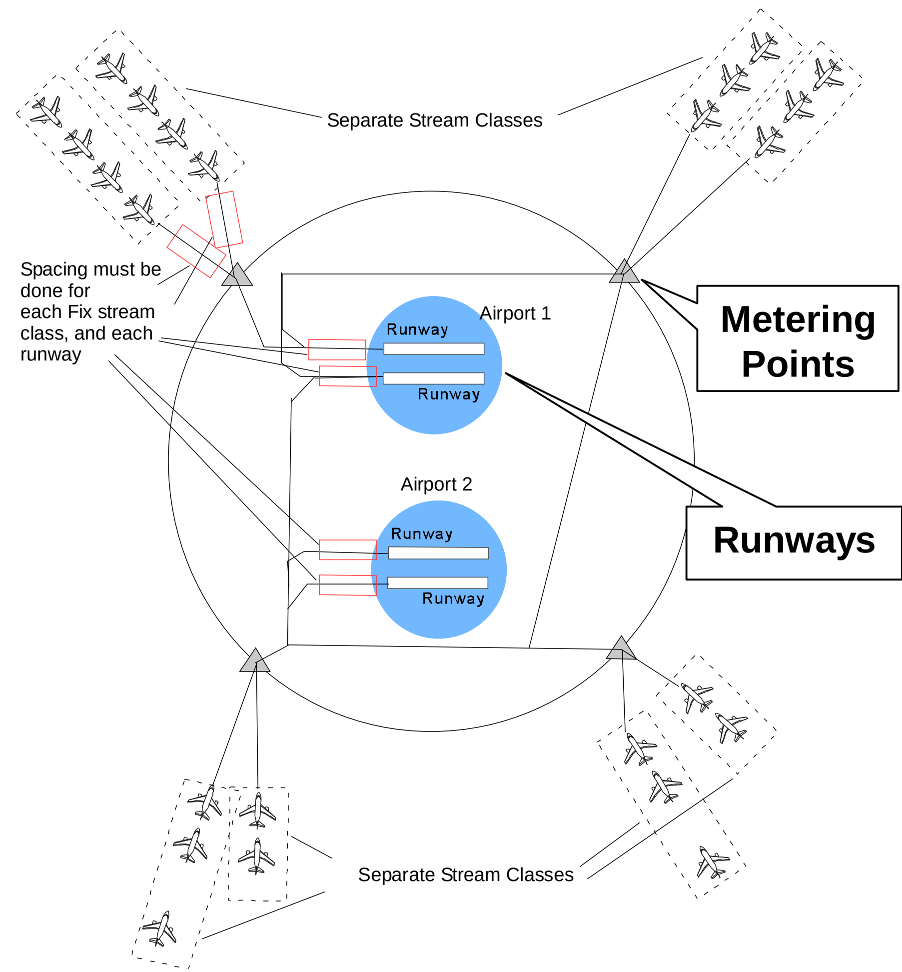
- Departures at JFK are metered by a DMS that computes recommended target movement area entry times (TMATs) for individual flights to keep movement area taxi times and departure queue lengths manageably small.

- Arrivals at JFK are metered by the TBFM decision support tool (DST) that assists the Center Traffic Management Coordinators (TMCs) and center controllers with planning and controlling major-airport arrival traffic flows.

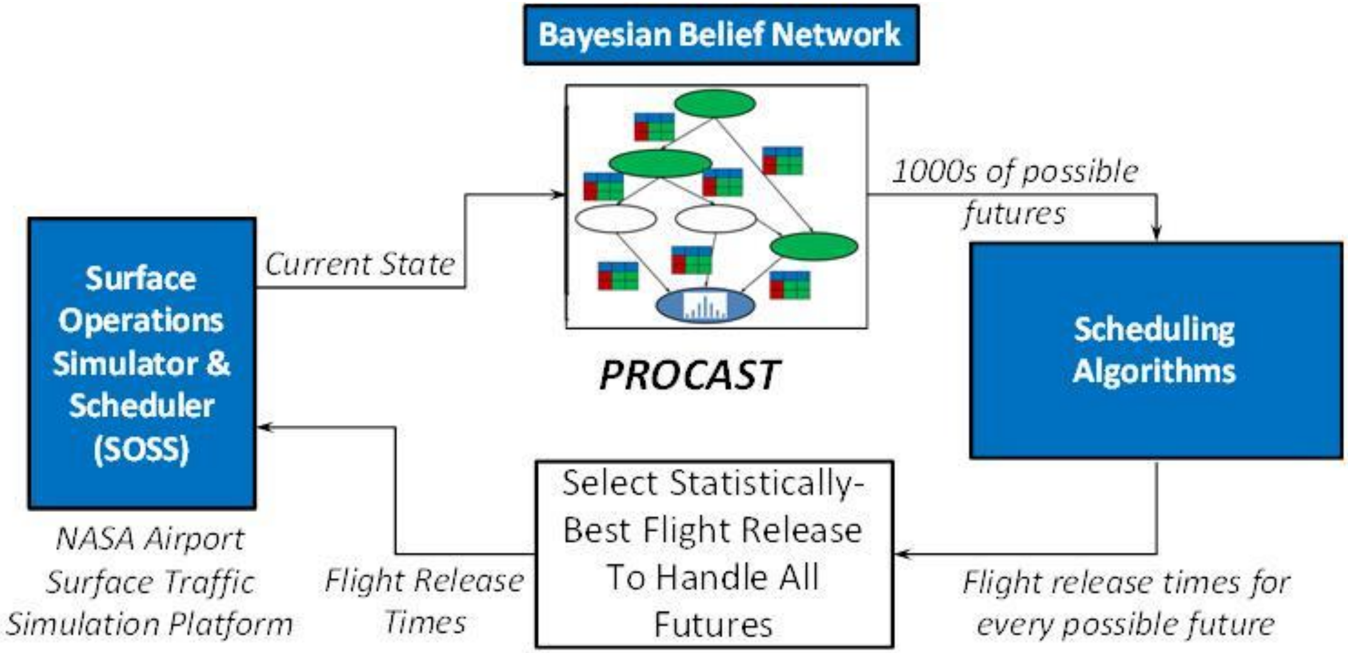
- A key challenge in developing a metroplex integrated arrival and departure scheduler is to manage the use of shared departure fixes across the metroplex airports.

Multi-Airport Scheduling Algorithms

- Sequence & schedule arrival & departure flights
  - Metering point crossing times
  - Airport runway landing and takeoff times
- To satisfy flow constraints
  - Shared waypoints and runways
  - Flow rates to limit traffic levels
  - Aircraft spacing for safety
- Accounting for uncertainty in aircraft transit times
  - Times to metering points
  - Times to runways



Phase II PROCAST Solution Architecture



- In Phase II we focus on developing and investigating the use of BNs to probabilistically model and predict taxi-times on the airport surface within the PROCAST framework.
- The hypothesis is that the use of probabilistic modeling for taxi-time prediction makes the integrated arrival and departure scheduling more robust to uncertainties and therefore more likely to be useful in future Air Traffic Management Decision Support Tools.

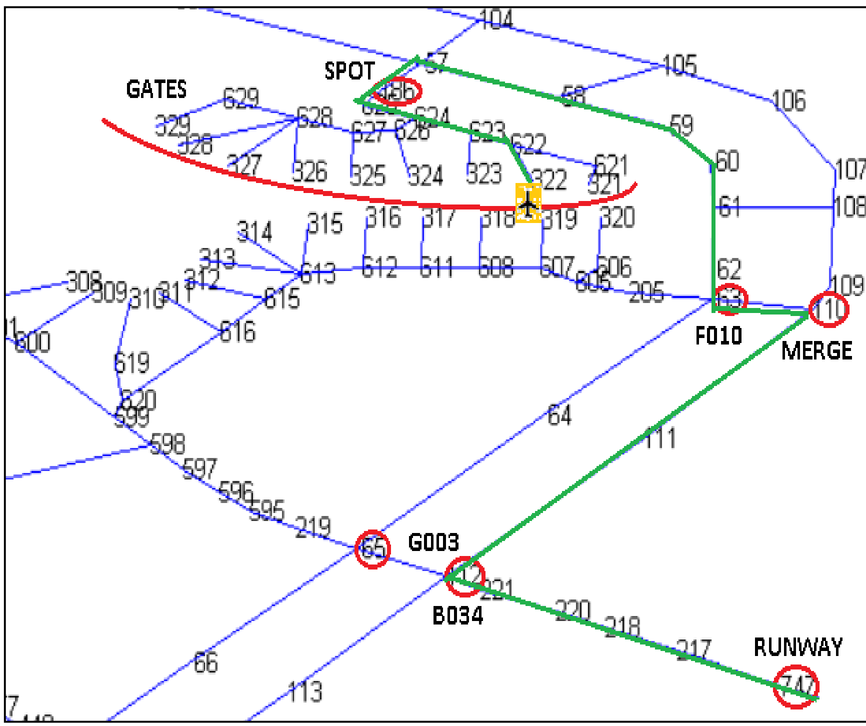
BN Evaluation and Development

- Taxi Time Modeling Using Bayesian Networks
- Challenges
- Contributions
- Improving Phase 1 BN model
- Evaluation framework for Phase 1
- Phase 2 : Extending to multiple airports
- Evaluation framework for Phase 2
- Structure Learning
- Experiments with Phase 2 data
- Conclusion

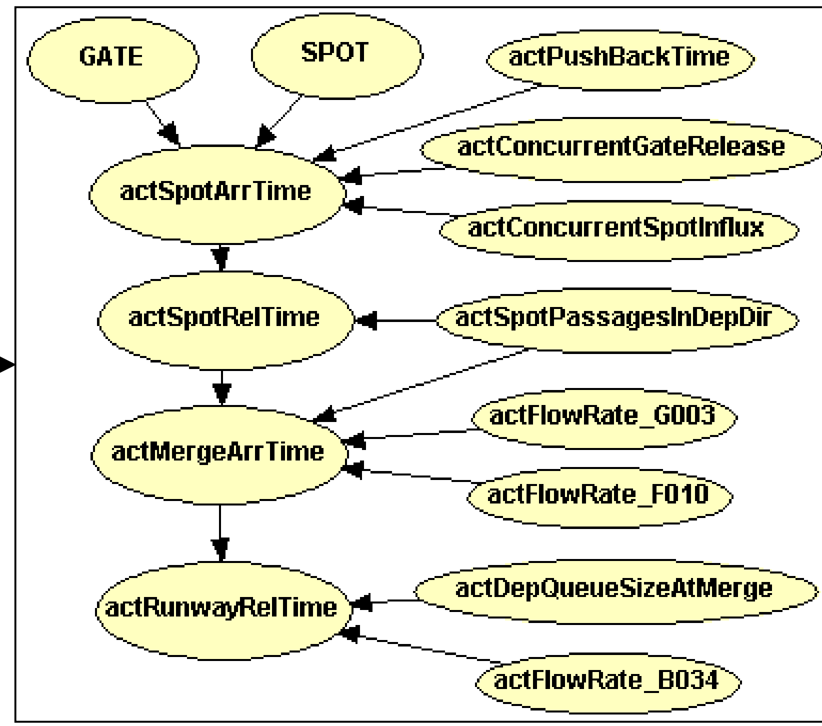
Taxi Time Modeling using Bayesian Networks: Phase I

Created based on expert advice

JFK Airport Snapshot



JFK Airport – Phase 1 BN Structure



Challenges Met in Phase II

- Generalization to arbitrary airports, not rely on subject matter experts:
  - Data driven approach
  - Machine learning
- Feature engineering
- Large and sparse data
- Complex BN model

Improving Phase I Approach

- The BN prediction method in Phase I required a few corrections
  – Debugging Matlab Bayes Net toolbox
  – Fixing the problem of “zero probability values” during inference
  – Investigating the choice of bin size for discretizing transit times
  – Considered both *stagewise* and *posterior* sampling.
- We call the refined method: Enhanced Phase I
- We develop an Evaluation framework
  1. Systematically compare one ML method to another.
  2. Fine tune parameters

Develop Bayesian Networks (BNs) for JFK, EWR & LGA Taxi Times

- Explore BN *Structure Learning* methodologies to discover the any airport transit time model from SOSS simulation data
  - Generalized feature extraction and selection
  - BN structure learning
  - BN parameter learning
- Apply the same methods to model taxi times in JFK, EWR, and LGA

SOSS simulation

- Feature Extraction

taxi times, generalized features

- Feature Selection

Training Data: taxi times, generalized important features

- BN Learning

Test Data: generalized important features

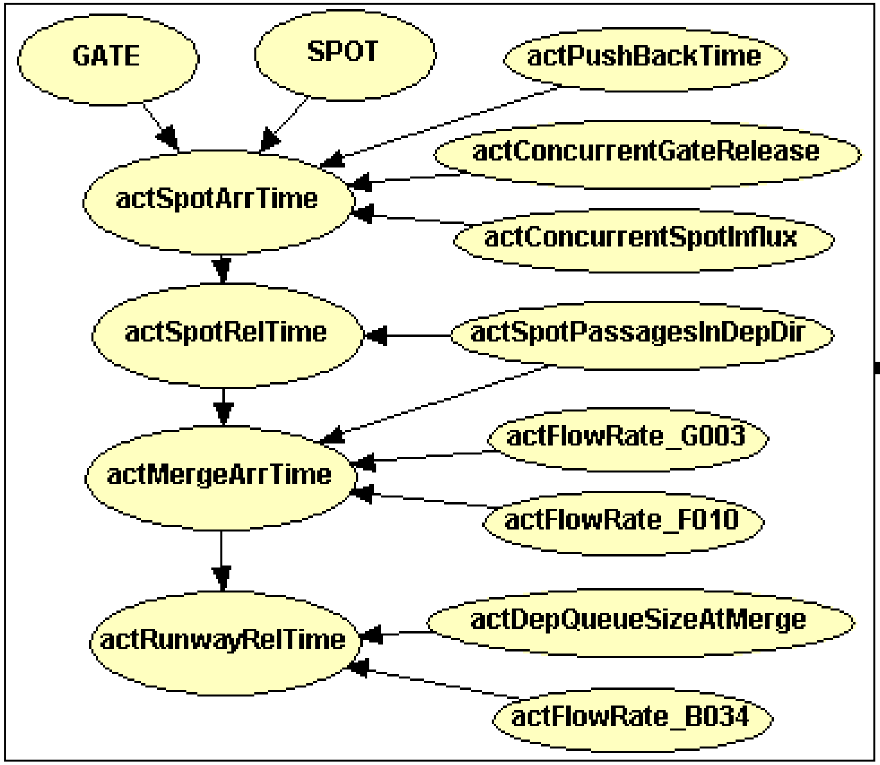
- Taxi Time Predictions from trained BN model

Taxi Time Modeling using Bayesian Networks: Phase II

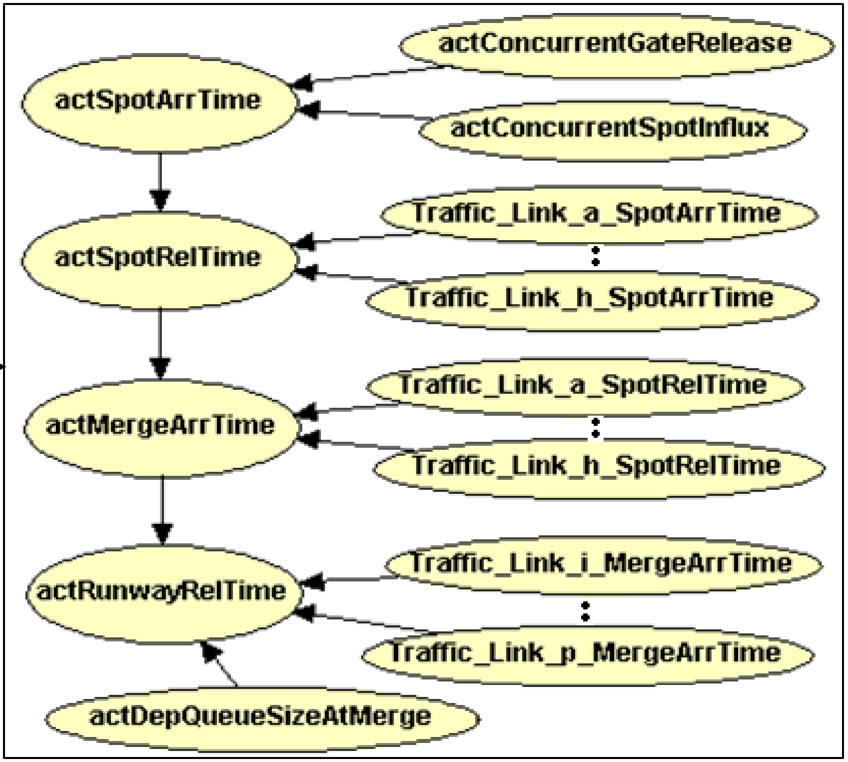
Created based on expert advice

Data driven approach

JFK Airport – Phase I BN Structure



Phase II: Generalized to any airport



Revised Phase I (No Structure Learning): Evaluation Framework

Split flights fairly into training and testing subsets

Digest the simulation data.

Compute time-varying traffic levels in each simulation.

Result is **one row per flight**, with flights from all simulations lumped together.

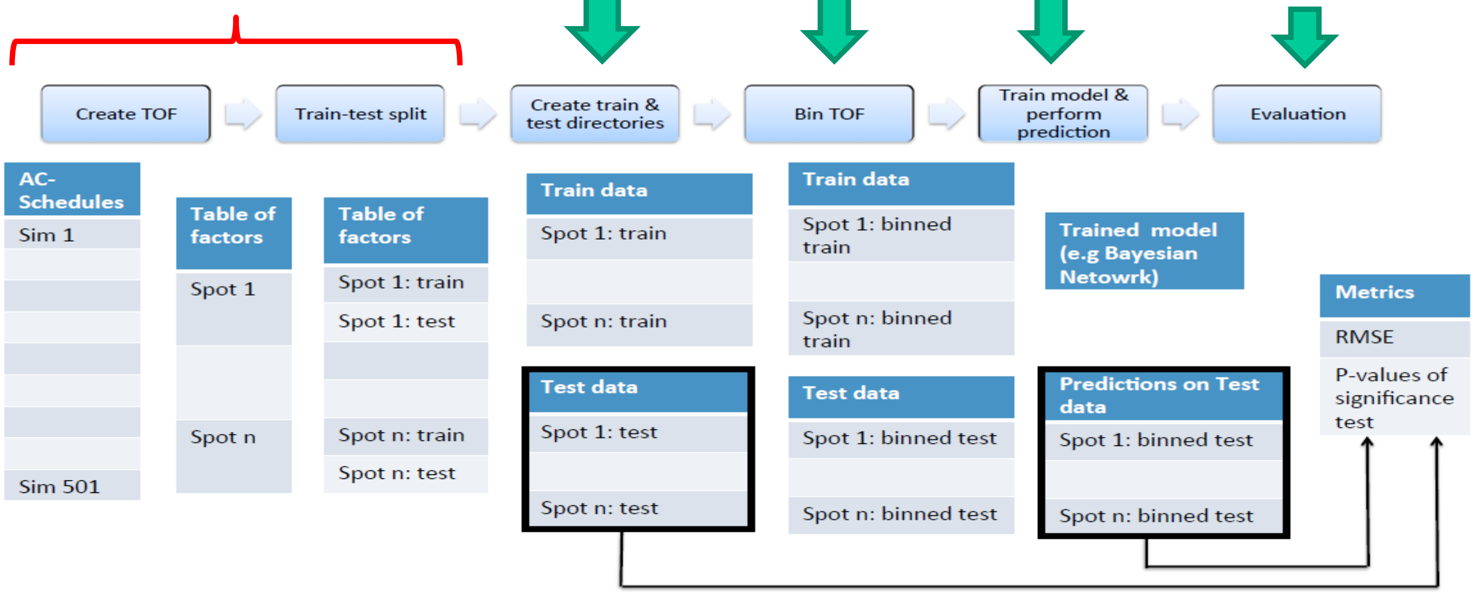
Bin data to create discrete data (optional)

Any ML model can be used.

Train the model.

Predict the test data

Compute metrics by combining predictions with test data



Zero Probability Values

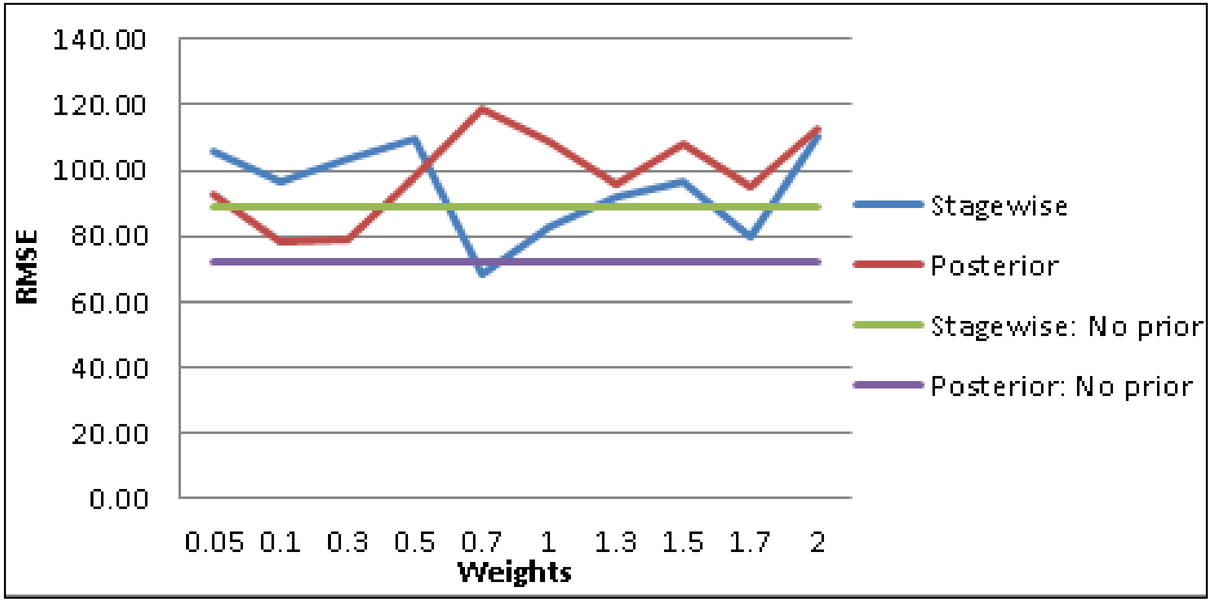
- When we encounter conditions not seen in training data
  - Trained model may not be able to make a good prediction due to zero probabilities

- Investigated approaches
  - Supply a default value (Phase I approach).
  - Create a Bayesian network with a prior distribution.
  - Use larger bin size (and fewer bins) to reduce sparseness of training data.
  - **Use less evidence (fewer inputs) when predicting.**
  - **Use continuous variables in the Bayesian network.**

Bayesian Network with Dirichlet Prior

- Using the Phase I BN, but train the network with a Dirichlet prior distribution.
- Explored *a range of weights* for the distribution, and tried *both forms of sampling*, but focused on departures taxiing through one *spot* for simplicity.
- Transit time prediction results are not encouraging, not convincingly better than the baseline with no prior.

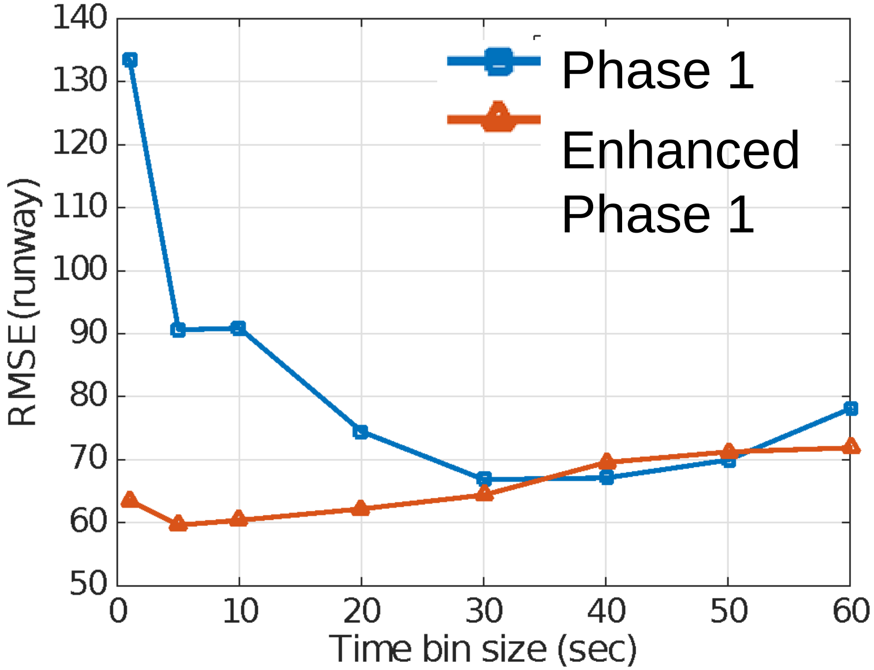
RMSE of Gate to Runway Transit Time Predictions



Effect of Bin Size on BN Predictions (Less Evidence Used)

- Phase 1 predictions are highly sensitive to the choice of bin size
  - Intensified “empty distribution” problem for unfavorable bin sizes
- Enhanced Phase 1
  - **Corrects zero probability values by dynamically reducing the evidence (inputs)**
  - **Other issues fixed**
  - More robust to choice of bin size
  - Achieves smallest prediction error for bin size = 5 sec

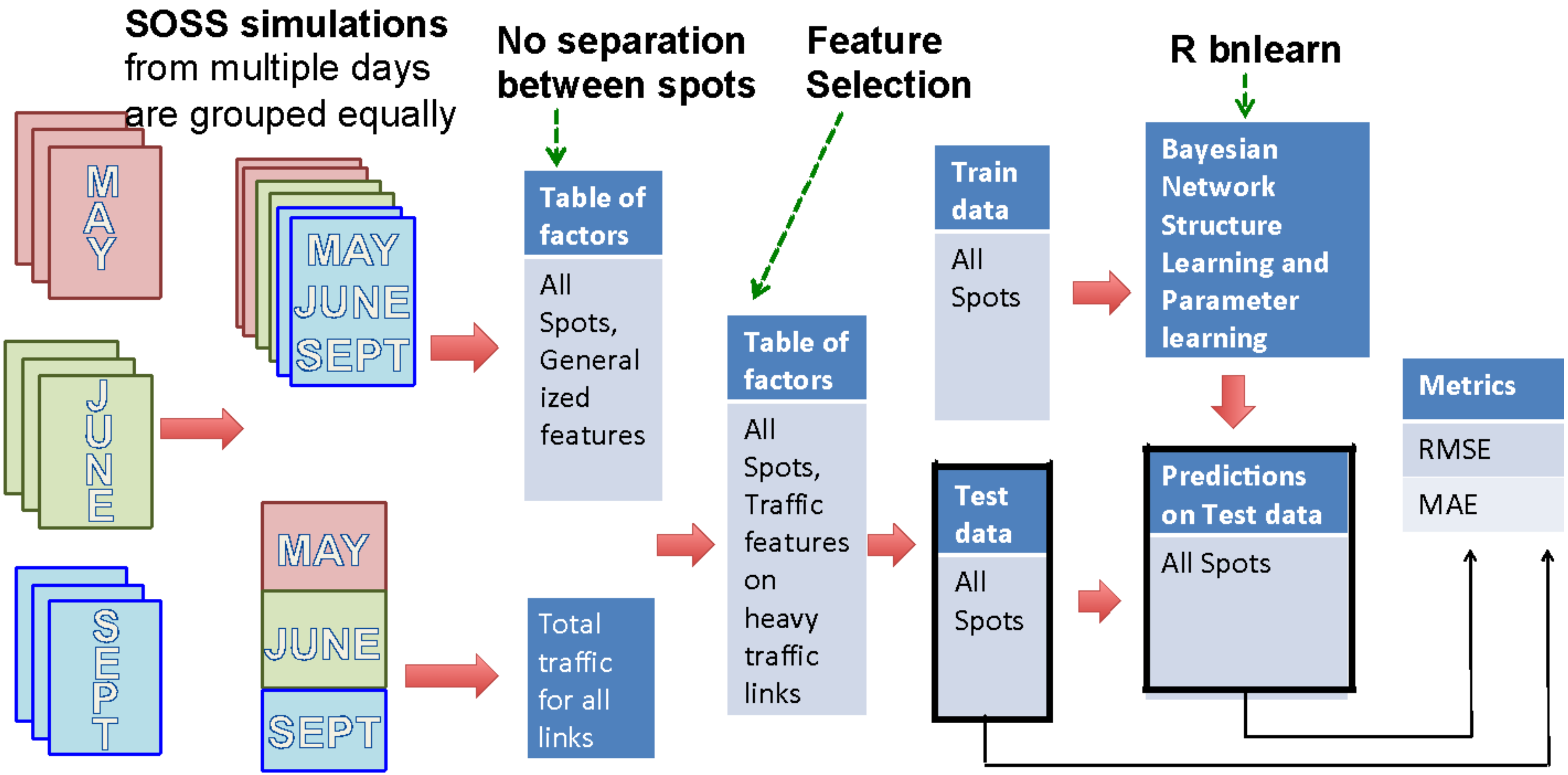
Gate to Runway Transit Time Prediction RMSE (in seconds) as a Function of Time Bin Size



Phase II Evaluation Framework (1)

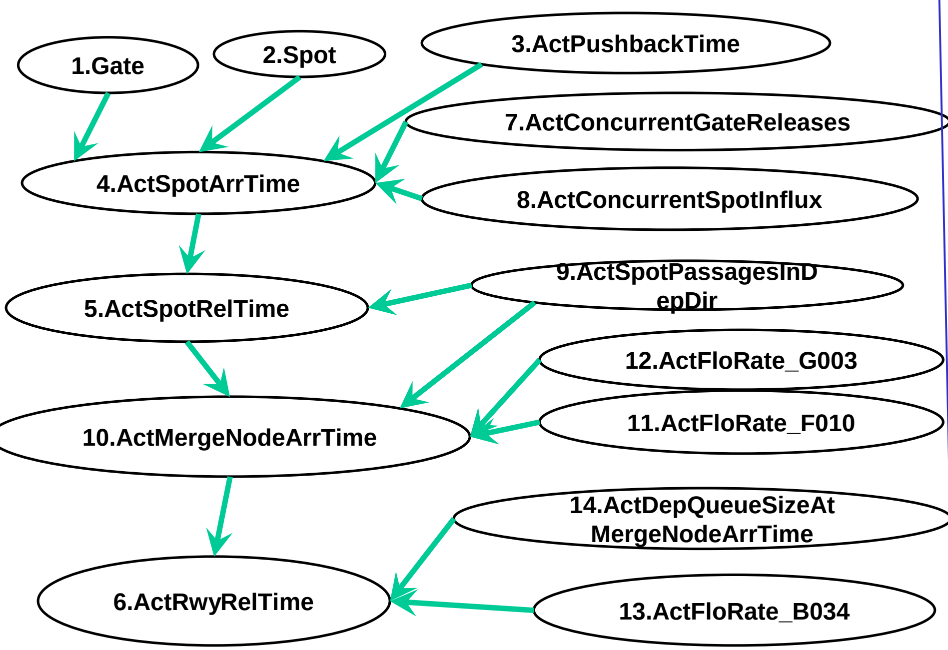
- Challenges
  - Phase II dataset contains simulations from 3 days, 3 airports each
  - Different set of flights in different days
  - Handcrafted JFK-features cannot be used for other airports (EWR, LGA)
- Solution
  - Generalized set of features are computed to replace handcrafted features
  - Feature selection is performed based on total traffic at each link in the airport model
  - BN structure learning on new set of features

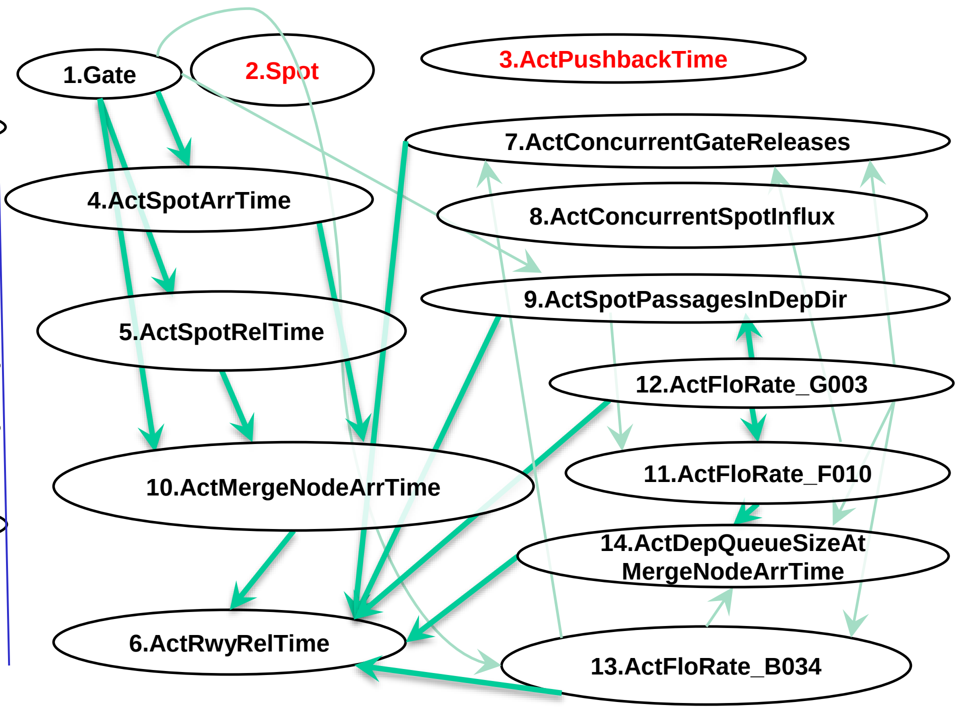
Phase II Evaluation Framework (2)



BN Structure Learning: Comparison

JFK BN from *SME Consultation*	JFK BN from *Structure Learning*



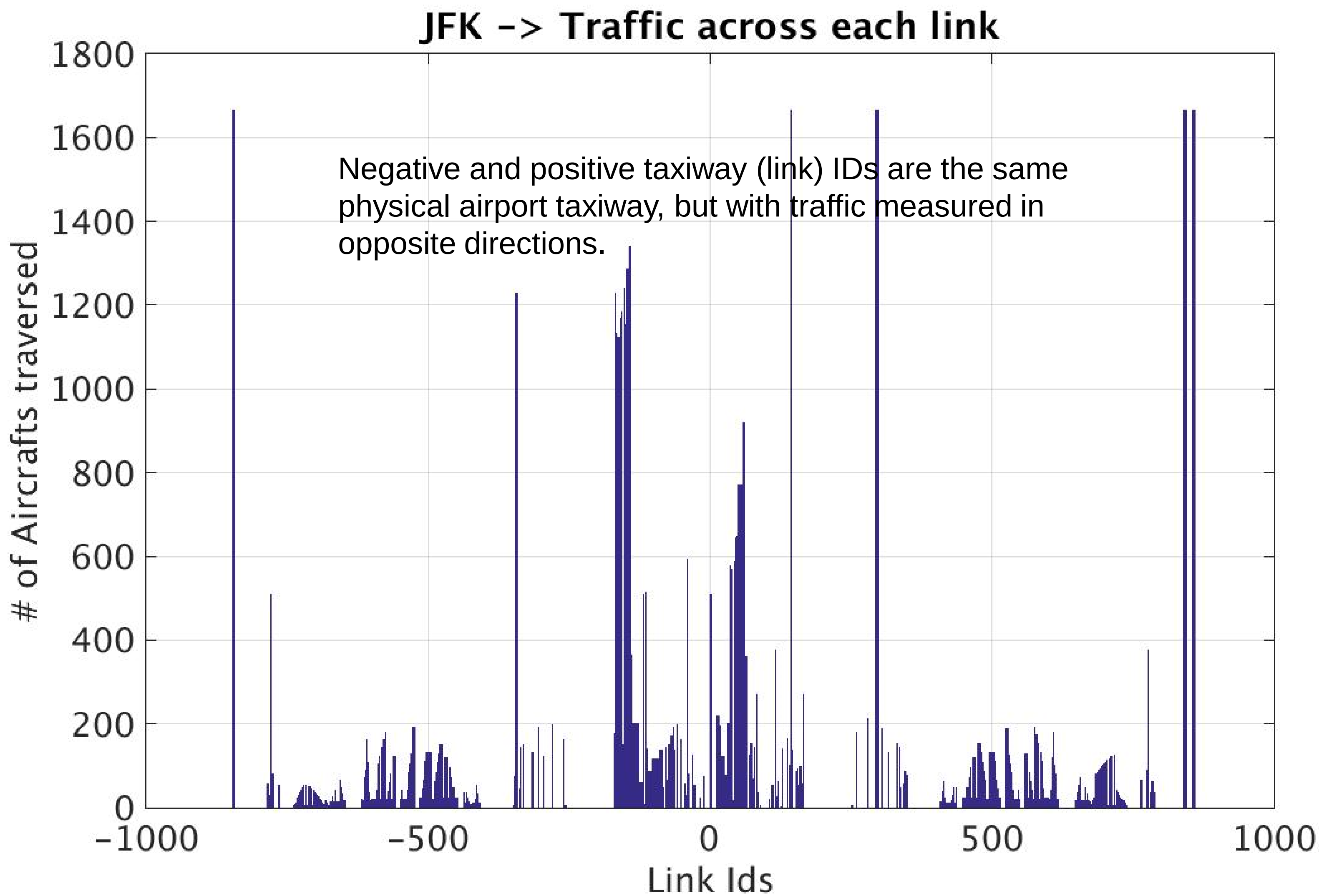


- Using R bnlearn structure learning: Incremental association Markov blanket-based structure learning
- Spot and ActPushbackTime are disconnected form ActRwyRelTime in learnt BN structure: all instances have same value
- ActRwyRelTime has more parents in learnt BN compared to SME
- More edges in learnt BN than SME: not all are intuitive

Phase II Evaluation Framework

- Features:
  - Random variables pertaining to a given departing flight. These are the nodes that will appear in Bayesian Network.
- Why generalize?
  - Some of the Phase 1 features were specific to JFK airport. (e.g ActFlowRate_B034)
  - Designing a set of such features requires subject-matter expert's advice, as well as trial-and-error.
- How to generalize?
  - All of the airport-specific features measure the traffic at some point on the taxiway, in some window of time (relative to the given flight's timeline).
  - Leverage this idea to replace airport specific features with generalized set of features for all high-traffic taxiway links.

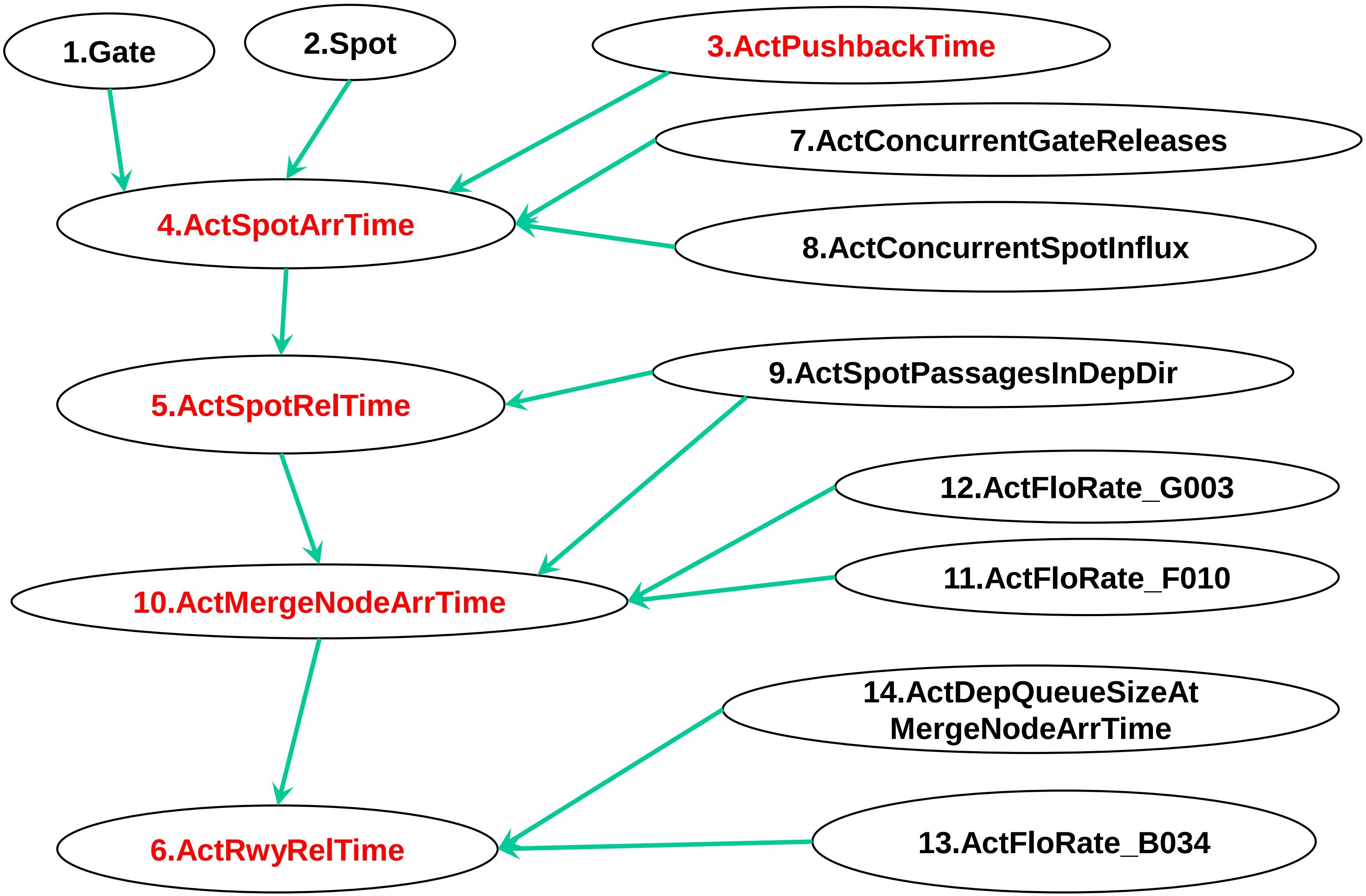
Traffic Statistics for Taxiways at JFK



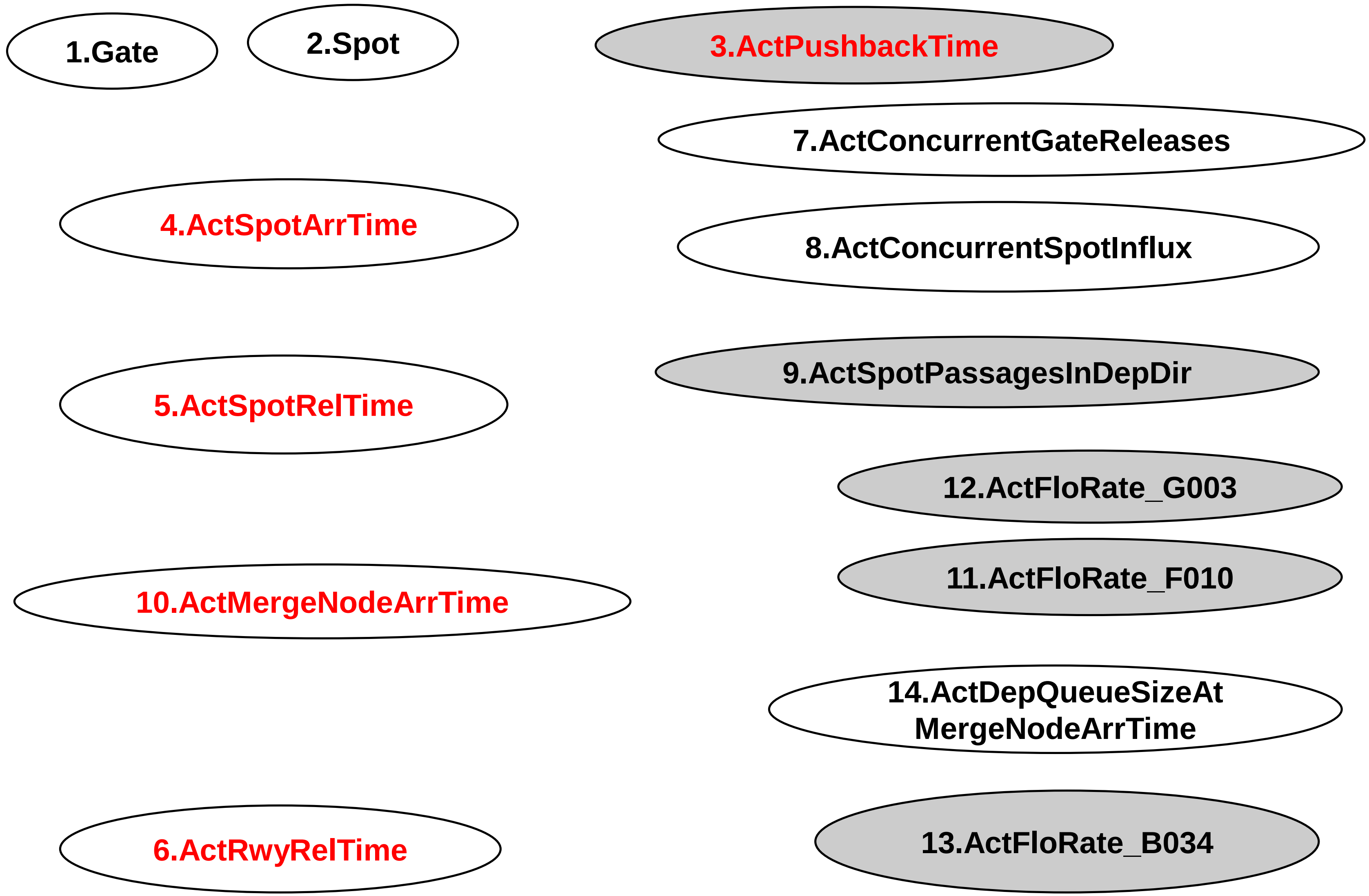
Feature Extraction and Selection

- Procedure
  - Compute traffic statistics per link (node-pair) across all links. Links are directional: each airport link results in two features – we pick the top 100 high-traffic links
  - Key moments in a flight's timeline: Gate release time, Spot arrival time and Merge node arrival time
  - Use time window surrounding these key moments to compute traffic across all high-traffic links in the airport
  - Traffic across each of the high-traffic links is a feature
  - Set of these traffic features capture similar behavior to the node-specific features. So node-specific features can be replaced by these newly computed features

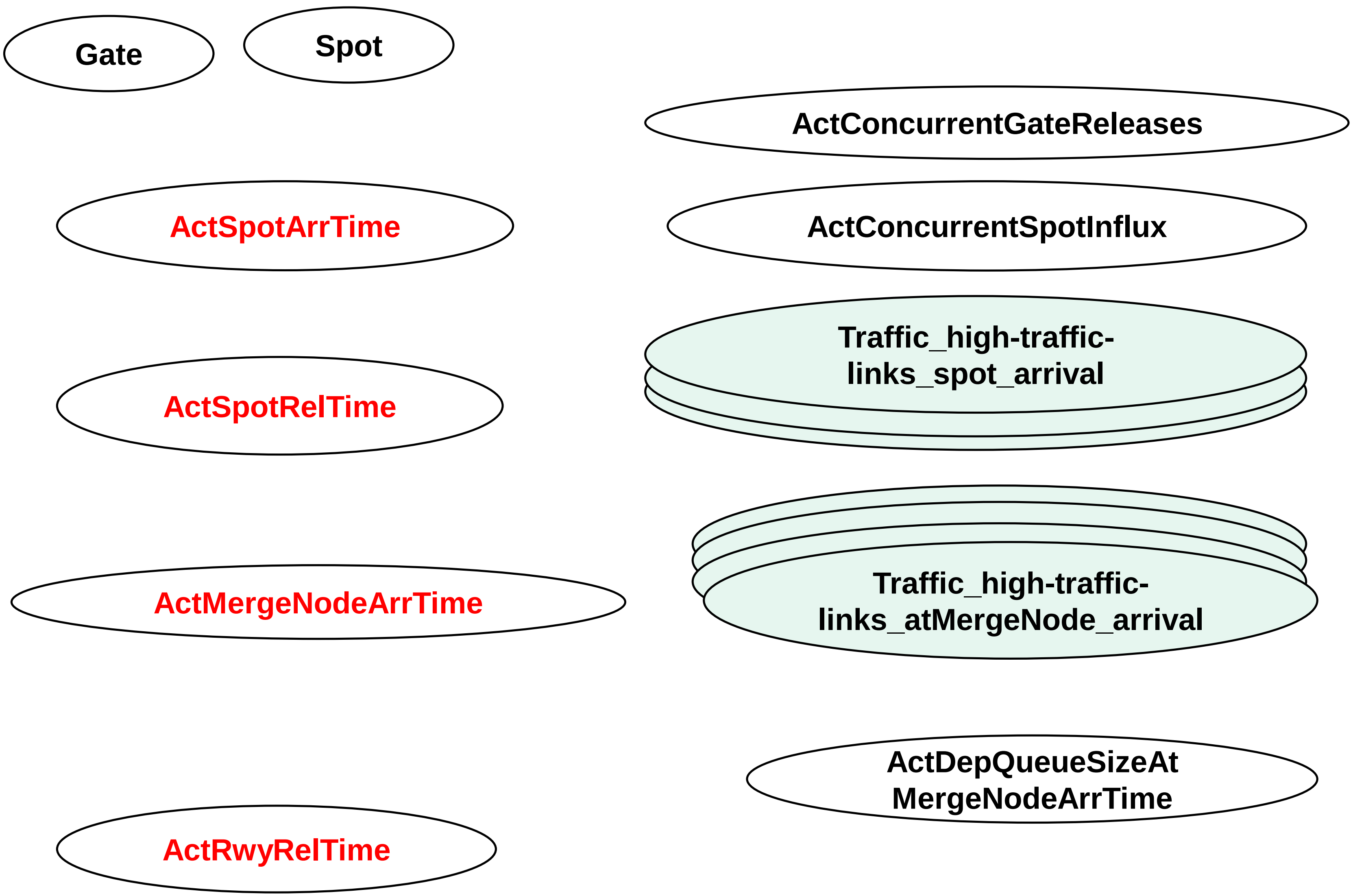
Phase I Bayesian Network



Phase I Features to be Replaced

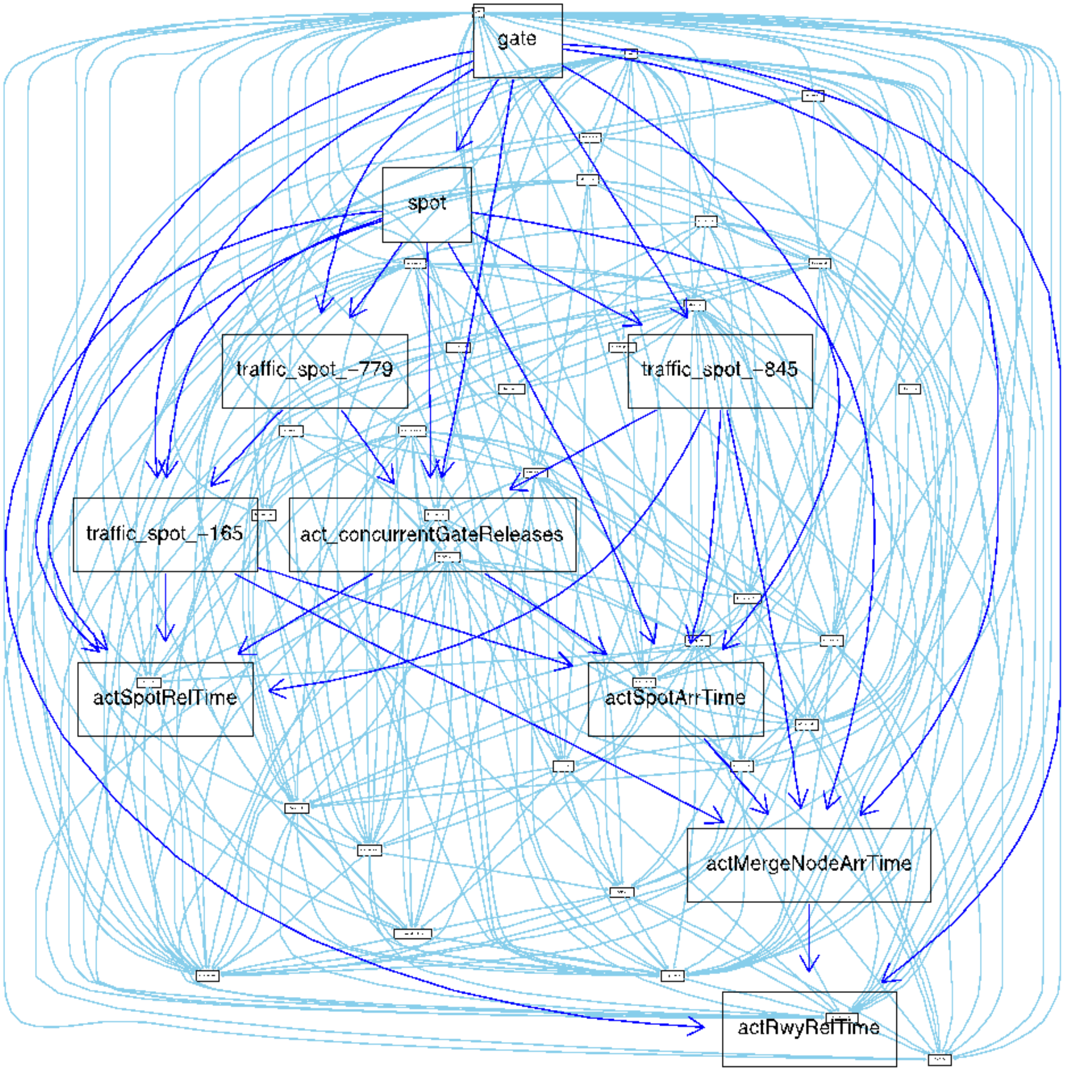


Features Added in Phase II



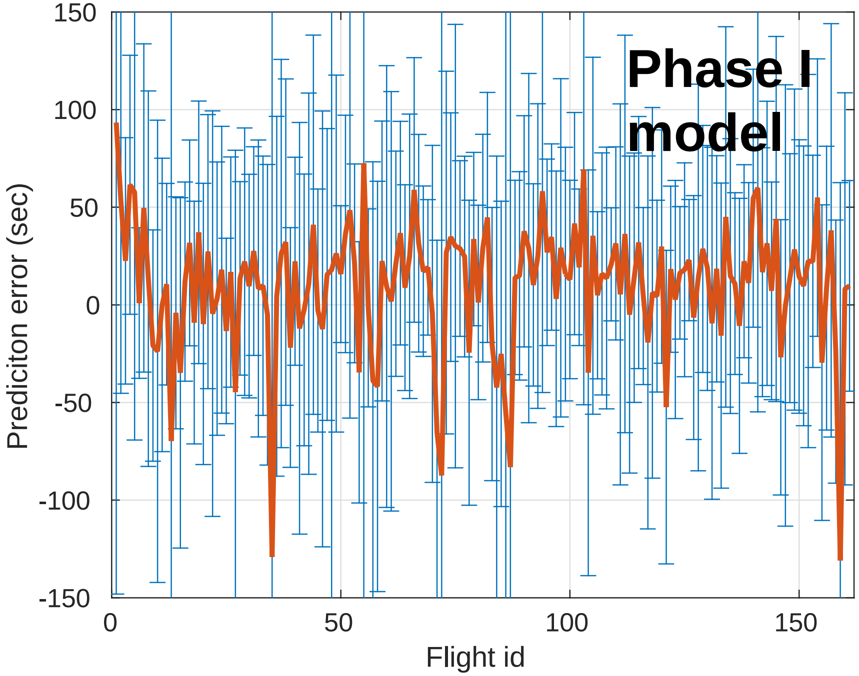
Final BN from Structure Learning

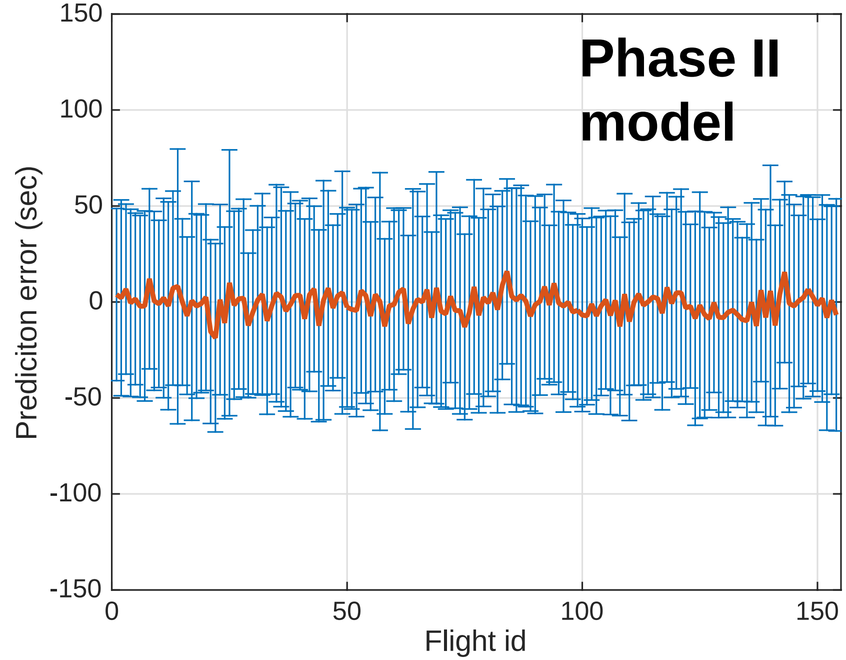
- BN structure for JFK
  - Learnt from Phase II data
  - Nodes are selected features
  - Blacklisted edges provided to training algorithm
    - actRwyRelTime ➔ any node
    - actMergeNodeArrTime ➔ any node except actRwyRelTime
    - …
- Statistics
  - Nodes: 110
  - Edges: 1693



Taxi Time Predictions: JFK

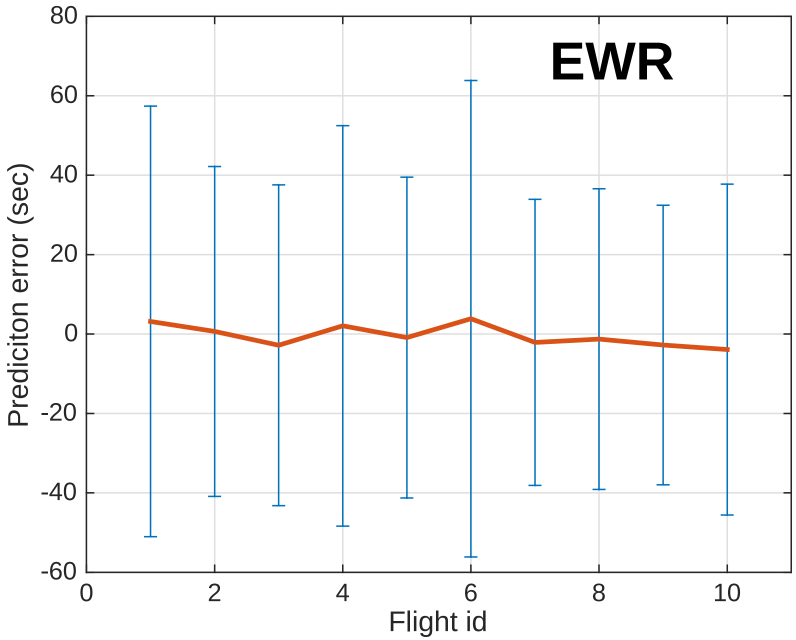
- Predicting gate to runway taxi time

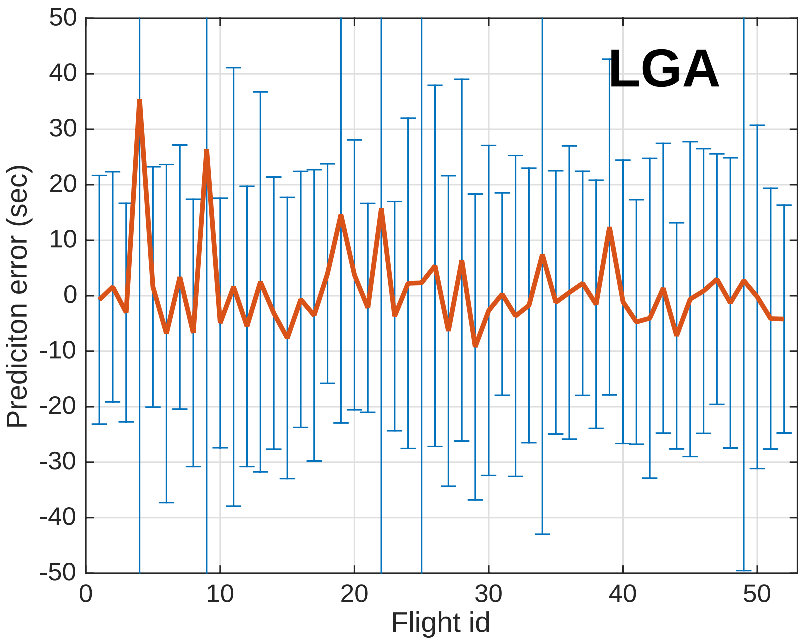




- Brown line and error bars: mean and standard deviation of prediction errors over all test simulations
- Same **Phase II** data in both cases
- **Phase I model**: mean abs. error = 61.5 sec, RMSE = 98.61 sec
- **Phase II model**: mean abs. error = 39.4 sec, RMSE = 51.1 sec

Runway Arrival Time Predictions: EWR, LGA





- Brown line and error bars: mean and standard deviation of prediction errors over all test simulations
- Mean errors in all airports are close to 0
- Error margin < ±1.5 minutes
- Relatively low error in LGA, as expected

Training BN Models with Phase II Data

Challenges

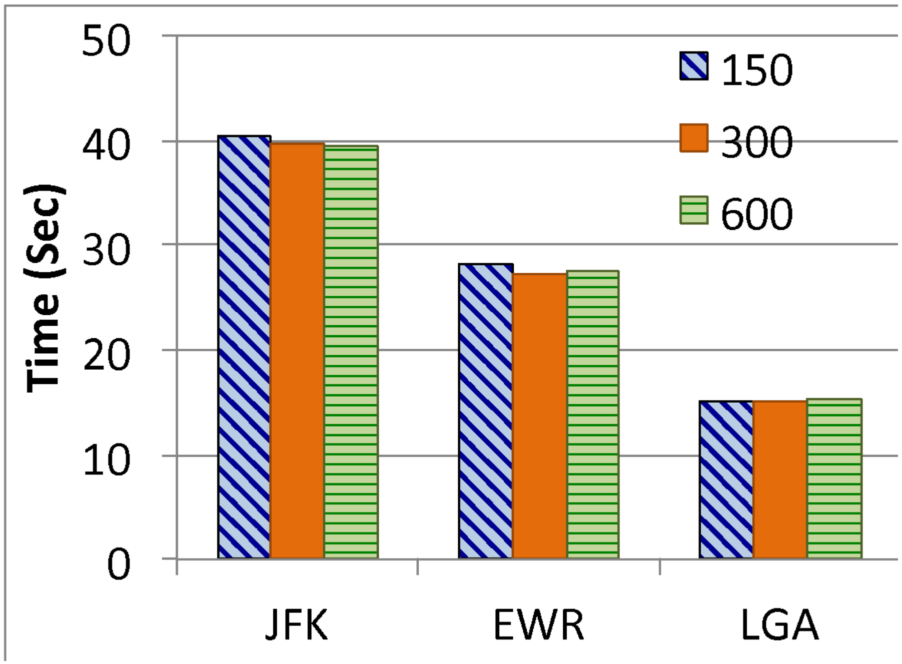
- About 200 times more training data in Phase II
- Each simulation is 24 hours long, compared with 2 hours for the Phase I simulations
- Large computation time required to process all simulation files

Solution

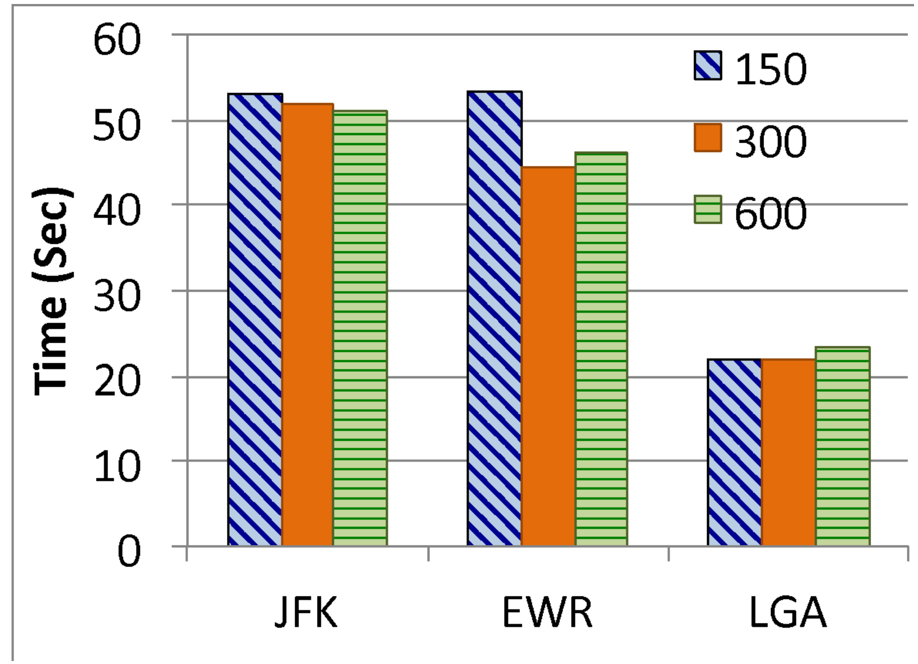
- Setup high-performance server for BN training
- Training set contains a subset of simulations

Research question: How does prediction error vary with sample size (number of files)?

Prediction Errors with Varying Training Sample Size



Mean Absolute Error



Root Mean Squared Error

- Phase II BN models trained with 300 files exhibit low errors
- Adding more training samples: no significant improvement

Comparison of Phase I and Phase II Approaches

	Phase I	Phase II
Features	SME-based judgement and iterative simulation plus trial and error	All plausible generalized features
Variable Types	Discrete, categorical	Continuous
Node Model Type	Table-driven CPD	Linear Gaussian
Structure	Manual, plausible causal links, trial-and-error	BN structure learning
Sampling	Stage-wise	From posterior distribution

LEARN Phase II Evaluation Matrix

- Simulation with ***interacting*** traffic causes sensitivity to initial conditions*, and thus uncertainty in taxi times
- We will compare the effectiveness of the BNs with a simple probability model (distribution)
- We also compare the effectiveness of the probability models based on number of futures

Probability Model	Futures	Notes
None	1	Scheduling based on unimpeded taxi times
Simple	1	Departure taxi time modelled as a normal distribution. Sample, schedule, and down-select.
Simple	10	
Simple	100	
Bayes net	1	Use mean transit time (no down-select)
Bayes net	10	Sample, schedule, and down-select
Bayes net	100	

Traffic Scenario:
7/25/12 - 2 hour period

Preliminary Results

- Table of taxi-in, taxi-out, and total delays at each airport for each evaluation scenario

Discussion of Results

Future Research

- Enhance SOSS airport adaptations
  - Additional airport configurations
  - Additional fidelity of surface and terminal airspace
- Enhance multi-airport scheduling algorithms
  - Add heuristics and optimization features
- Investigate Bayesian networks with hybrid data: continuous + discrete
- Evaluate other BN learning and prediction methods for modeling aircraft taxi time
- Compare different machine learning models (e.g. Gaussian Processes, Support Vector Machines)
- Online update of trained models (using historical batch data) by feedback from previous scheduling cycles
- Integrate taxi time predictions and scheduling in one compute paradigm WHAT DOES THIS MEAN???
  – Constrained optimization with transit time predictions as input parameters

Project Team Contact Information

Name	Organization	Email	Phone
Greg Carr	ATCorp	gcarr@atcorp.com	408-667-0090
Sebastian Timar	ATCorp	stimar@atcorp.com	408-627-2215
Matt Stillerman	ATCorp	mstillerman@atcorp.com	607-257-1975 (x7106)
Mark Peters	ATCorp	mpeters@atcorp.com	757-692-3034
Valentino Felipe	ATCorp	vfelipe@atcorp.com	408-609-7949
Ole Mengshoel	CMU	ole.mengshoel@west.cmu.edu	650-335-2887
Annirudha Basak	CMU	aniruddha.basak@sv.cmu.edu	
Priya Sundararajan	CMU	priya.sundararajan@sv.cmu.edu	
Vinodh Paramesh	CMU	vinodh.paramesh@sv.cmu.edu	
Ralph Tamburro	PANYNJ	rtamburro@panynj.gov	917-828-7741
Bill Cotton	Cotton Aviation	wmcotton@gmail.com	512-261-5119
Nikunj Oza	NASA	nikunj.c.oza@nasa.gov	650-604-2978

NASA Technical Officer: Dr. Robert D. Windhorst 650-604-5020