Independent Study
Novel, Multidisciplinary Global Optimization under Uncertainty, Phase II

Pantelis Takos
Masters in Software Management
Carnegie Mellon University
Moffett Field, CA 94035, USA
pantelis.takos@sv.cmu.edu

Dr. Ole J. Mengshoel
ECE Department
Carnegie Mellon University
Moffett Field, CA 94035, USA
ole.mengshoel@sv.cmu.edu

ABSTRACT
The purpose of this Independent was to develop a robust Bayesian Network to improve the predictive technology of an airspace traffic management system. For the purposes of our research we studied the *Probabilitic Graphical Model for Departure Traffic Prediction from One Terminal-Gate Area to One Departure Runway* [1]. It is important to note that for the purposes of our research we differentiate our Bayesian network approach from the one described in the aforementioned PGM, by removing the dependecy between two of the nodes.

1. INTRODUCTION
The Bayesian Network (BN) we developed consists of three node types. First of all, the observable nodes, *Pre-pushback Process Completion State, Departure Gate, Pushback Rates from Adjacent Gates, Adjacent Spot Rates, Opposite Direction Traffic Rates* and *Departure Runway Queue Length at Spot Arrival Time*, with respective node names in the BNs, *t, dg, pag, as, odt, drg.* Secondly, the intermediate nodes, *Pushback time* and *Spot Arrival Time*, with respective node names in the BNs, *pt, sat.* Finally, our target node *Runway Time of Arrval*, represented as *rta* in the BNs.

We are going to present and discuss the results we got by using two different tools, Hugin Lite 8.2 [2] and Netica "5.18" [3]. Since we were using evaluation versions of these software tools, one of the main challenges was the allowed number of nodes and states we were permitted to introduce to our BN. Although, we had to reduce the possible values of each state to meet license requirements, the provided Network displays all possible factors affecting the target node.

[1]Aditya Saraf, Kris Ramamoorthy, Steven Stroiney, Saab Sensis Corporation, Campbell, CA Bruce Sawhill and Jim Herriot, NextGen Aerosciences LLC, Williamsburg, VA. ROBUST, INTEGRATED ARRIVAL-DEPARTURE-SURFACE SCHEDULING BASED ON BAYESIAN NETWORKS, Figure 4

[2]API: HUGIN 8.2 (x64), ©1995-2015 Hugin Expert A/S All rights reserved

[3]Norsys Software Corp. ©1992-2014

2. THE BAYESIAN NETWORK
As we shortly described in the introduction we build this Bayesian Network using nine different nodes. Due to uncontrolled factors, we did not receive in time the necessary data, on which our research would be based, and therefore the values presented are continuous real values we chose using common sense. Below, we will explain each of the nodes individually along with their probability tables and calculating functions (where applicable).

2.1 Pre-pushback Process Completion State node
This node represents the delay in minutes an aircraft might have during its preparation activities. We classify it in the following five states:

1. -30 - 0
2. 0 - 30
3. 30 - 60
4. 60 - 90
5. 90 - INF

This node's dependencies (parent nodes) consist of the following activities:

1. Pre De-boarding activities
2. De-boarding process
3. Pre Fueling activities
4. Fueling process
5. Pre Boarding activities
6. Boarding process

Due to software restrictions and to avoid network complexity, at this preliminary stage, this BN is provided separately.

As it becomes evident, this node's dependencies, are prone to human error, and therefore we decided that they should follow a uniform distribution. Additional research is required, based on the data that will be provided, to adjust the probability table of this node.

2.2 Pushback time node

This node represents the time needed for an aircraft to be moved from a passenger terminal to a runway or taxiway. A basic assumption here is that a pushback tug is always assigned to a gate, and no delay should be considered for a pushback tug to arrive. Thus, the time needed depends only on the spot at which the aircraft needs to be pushed to. The node states are:

1. 0 - 2
2. 2 - 4
3. 4 - 6
4. 6 - 8
5. 8 - 10

Due to function expression requirements, in Hugin, equation lower limit has to be -infinity and higher limit +infinity(infinity), in the probability table there are presented two more states which should be ignored. Again, because of the absence of data, we have decided to use a normal distribution with mean $\mu = 5$ minutes and variance $\sigma^2 = 1$ minute.

Hugin expression: *Normal (5, 1)*
Netica equation: *P(pt) = NormalDist (pt, 5, 1)*

2.3 Departure Gate node

In this node we reflect the delay in minutes, that might occur in a departure gate, due to prior delayed or in advance departures from the gate. The node states are:

1. -15 - 0
2. 0 - 15
3. 15 - inf

Our fundamental assumption for the probabilities presented in this node is that they might only slightly affect the intermediate and target nodes. Hence, we have set the probability between 0 and 15 minutes to dominate (92%) the possible deparure gate delay.

2.4 Pushback Rates from Adjacent Gates node

This node reflects possible delays from adjacent gates. The node states are:

1. -15 - 0
2. 0 - 15
3. 15 - inf

Following the similar logic with the previous node (low influence on intermediate and target nodes), we have set the probability over and around 0 to dominate the probability distribution.

2.5 Spot Arrival Time node

This is the second intermediate node we use, where we basically aggregate the delays from the three aforementioned nodes. The node states are:

1. -70 - -55
2. -55 - -40
3. -40 - -25
4. -25 - -10
5. -10 - 5
6. 5 - 20
7. 20 - 35
8. 35 - 50
9. 50 - 65
10. 65 - 80
11. 80 - 95
12. 95 - 110
13. 110 - 125
14. 125 - inf

The value of this node would be significantly greater if the use of a tool without restrictions was available, where the volume of node states would reflect with great accuracy the most probable scenarios. The functions that will combine the node states in each tool are:

Hugin expression: *dg + pt + pag*
Netica equation: *sat (dg, pt, pag) = (dg + pt + pag)*

2.6 Adjacent Spot Rates node

This node represents the delay in minutes because of traffic priority to adjacent gates. Normally, we expect the probability of two gates being ready at the same time highly unlikely. Hence, we are using only three states to describe this node:

1. -10 - 0
2. 0 - 10
3. 10 - inf

In the first node we have assigned a 90% probability, which basically means no delay due to adjacent gates. In the second node, we have aasigned a 9% probability, which basically interprets in a relatively small probability of delay being caused due to traffic in nearby gates. Finally, we consider the probability of an aircraft being halted more than ten minutes, technically more than the time required for one pushback time, extremely improbable and therefore assigned it to a 1% probability.

2.7 Opposite Direction Traffic Rates node

Here we reflect the probability of delay in minutes, due to opposite direction traffic(mainly landing). We considered the scenario of such a delay extremely plaussible, especially in airports with high volume of flights. For this reason we decided to provide 4 states to describe this node, and slightly reduce the predominance of the node describing the state over and around 0 to 80%. Below are the node states:

1. -10 - 0
2. 0 - 10
3. 10 - 20
4. 20 - inf

2.8 Departure Runway Queue length at Spot Arrival Time node

In this node we follow the exact same concept with the previous node, but in this case we describe the probability of a delay caused by a queue on the runaway. Node states are as follows:

1. -10 - 0
2. 0 - 10
3. 10 - 20
4. 20 - inf

2.9 Runway Time of Arrival node

This is our target node, where basically all dealys summarize. This node represents the last position an aircraft will have in an airport surface. Its node states are:

1. -100 - -80
2. -80 - -60
3. -60 - -40
4. -40 - -20
5. -20 - 0
6. 0 - 20
7. 20 - 40
8. 40 - 60
9. 60 - 80
10. 80 - 100
11. 100 - 120
12. 120 - 140
13. 140 - 160
14. 160 - inf

and the functions for our software tools are:

Hugin expression: *as + odt + sat + drq*
Netica equation: *rta (as, odt, drq, sat) = (as + odt + sat + drq)*

3. SOFTWARE TOOLS

To research and improve the predictability of the Network, we started by taking little to no consideration of any hints about the nodes' states. We studied their Conditional Probability Distributions building the belief networks by using two tools; Hugin Lite and Netica.

3.1 Hugin Lite

In the former phase of our research, we focused on building the network using Hugin. During the first steps, and while we were still experimenting with creating the nodes, messages constsantly warned us regarding the version's restrictions. After a couple of iterations, which enabled us to comprehend better the vesrion's limits, we started synthesizing our belief network.

Hugin has a painless user interface, which made setting up and connecting the network nodes very simple. The node window was generally easy to follow, and providing a state-tab proved to be really useful many times. The greatest benefit of using Hugin however, was the availability of node's Probability Distributions table along with the related expression. This allowed us not only to observe the possible value of the parent node but also the likelihood of an observation to belong in more than one node states.

3.2 Netica

While we were looking for the next tool to evaluate our network, Netica became our obvious choice, because in its description it was stated that networks built using Hugin are supported with Netica. The restrictions on the number of states and nodes were more flexible than in Hugin, allowing us to explore the BN deeper. However, once we tried to import the .net files we ran into errors. Even after removing the lines that were causing the import to fail, Netica still failed to import the files. We believe that the issue is caused because of compatibility; specifically, in the description it is stated that the files built using Hugin 4.*, 5.* and 6.* may be read, but we built our network using the latest version 8.2.

On one hand, Netica had a certain level of complexity when it came to building the belief network. The node states are not as obvious as in Hugin, and separating the table equation from the probability table is generally complicated. Yet, the biggest disadvantage was the explicit use of integers and, in particular, multiples of 10 to express the Conditional Probability Tables. What was more ambivalent about this approach, is that it would sometimes omit outcomes adjacent to states with maximum likelihood probability.

On the other hand, Netica has two strategic advantages in comparison with Hugin. First one was the abiliity to easily reset the probability distributions. By simply dragging the column that represented the probability of a node state, the table probabilities automatically balance to their new most likely probabilities and would also recalculate the new probabilities of the child nodes. The second and most important advantage of Netica, was the ability to easily build an interactive website, also known as a HED system (Human-Electronic Dialog). By converting all observable nodes into questions for the user to provide input, this system allows users to have a real time reflection of the state of the target node.

Conclusions

We need to add the dependencies of the *Pre-pushback Process Completion State* node into the BN described above. Also, we need to investigate the available data, to adjust the probability table of this node. Furthermore, we expect that there will be a certain dependency on the activities (parent nodes) and each aircraft type, which might lead in building different BNs, depending on aircraft type, but all of which will follow the pattern we described in this report.

Another aspect that has to be closely examined is the correlation between *Pre-pushback completion state*, *Pushback time* and *Spot Arrival time.* Since currently the inputs for *Spot Arrival time node* are only *Departure Gate, Pushback time* and *Pushback Rates from Adjacent Gates*, the most probable values appear to be between 5 and 35 minutes of delay. However, the probability of *Pre-pushback completion state* is more than 30 minutes in almost 60% of our plots, and therefore the PGM needs to be investigated deeper.

The observable nodes *Opposite Direction Traffic Rates* and *Departure Runaway Queue Length at Spot Arrival Time* are directly proportional to the size and traffic of an airport. Consequently, part of future research should be revisiting node state limits to match the size of the airport being investigated.