Dynamic Control of Airport Departures: Algorithm Development and Field Evaluation

Ioannis Simaiakis, Melanie Sandberg, and Hamsa Balakrishnan

Abstract—This paper proposes dynamic programming based algorithms for controlling the departure process at congested airports. These algorithms, called Pushback Rate Control protocols, predict the departure throughput of the airport, and recommend a rate at which to release aircraft from their gates in order to control congestion. The paper describes the design and field-testing of a variant of Pushback Rate Control at Boston airport in 2011, and the development of a decision-support tool for its implementation. The analysis of data from the field trials shows that during 8 four-hour test periods, fuel use was reduced by an estimated 9 US tons (2,650 US gallons), and taxi-out times were reduced by an average of 5.3 min for the 144 flights that were held at the gate.

I. INTRODUCTION

A. Motivation and background

Airport surface congestion contributes significantly to taxi times, fuel burn and emissions at airports. Annually, taxi-out delays at major US airports exceed 32 million minutes, while taxi-in delays exceed 13 million minutes [1]. Recent studies have also shown that low-thrust taxi emissions have significant impacts on the local air quality near major airports [2–4]. The objective of this paper is to develop, implement and evaluate a control policy that can reduce surface congestion.

B. Related work

An airport congestion control strategy in its simplest form would be a state-dependent pushback policy aimed at reducing surface congestion. One such approach is the N-Control strategy. N-Control is an implementation of the virtual queue concept described in the Departure Planner [5], and variants of it have been extensively studied [6–9]. N-Control is based on the typical variation of departure throughput with the number of departures on the surface (denoted \( N \)). As more aircraft pushback from their gates onto the taxiways, the throughput of the departure runway initially increases. However, as the number of taxing departures exceeds a threshold, denoted \( N^* \), the departure runway capacity becomes the limiting factor, and there is no additional increase in throughput. Any additional aircraft that pushback simply incur taxi-out delays [10]. Figure 1 illustrates this behavior for the most frequently used runway configuration at Boston Logan International Airport (BOS) in 2011, using surface surveillance data from a system known as ASDE-X [11].

The N-Control policy is effectively a threshold heuristic: If the total number of departing aircraft on the ground exceeds a certain threshold, \( N_{\text{ctl}} \), where \( N_{\text{ctl}} \geq N^* \), aircraft requesting pushback are held at their gates until the number of aircraft on the ground is less than \( N_{\text{ctl}} \). While the choice of \( N_{\text{ctl}} \) must be large enough to maintain runway utilization, too large a value will be overly conservative and reduce the benefits of the control strategy. A similar heuristic, based on the concept of an Acceptable Level of Traffic (ALOT), is used by Air Traffic Controllers at BOS during extreme congested situations [12]. The N-Control policy is also closely related to the constant work-in-process (CONWIP) policy used in manufacturing systems. The main benefits of CONWIP systems are their simplicity, implementability and controllability [13]. They present an efficient way to control congestion by accepting an adjustable risk of capacity loss.

Optimization-based policies have also been considered recently for surface traffic. Burgain et al. used advanced modeling tools for the characterization of optimal pushback policies. Their control protocol was a generalization of N-Control, where the state of the surface at any time was mapped to an on-off input signal. The solutions were full-state feedback policies, which presented implementation challenges [14].

There has been much prior research on the optimal control of a variety of queuing systems, considering different decision variables and control objectives [15–17]. However, several challenges remain when attempting to apply these results to the control of the airport departure process. Firstly, on-off or event-driven control policies for controlling the pushback process are difficult to implement in practice. Both the air traffic controllers and the airlines prefer a pushback rate that is valid for a predefined time period, after which it can be updated. Controllers prefer such pushback rate recommendations for workload and procedural reasons, and airlines prefer them because of their predictability, which helps with ground
crew planning. Secondly, the control input is applied at the gates during pushback, whereas the main bottleneck is the runway. The control strategy has to accommodate stochastic travel times between the gate and the runway, due to factors such as the pushback process, flight checklists, communication delays and variable taxi speeds.

Several approaches to departure metering have been proposed, including the Ground Metering Program at New York’s JFK airport [18], the field-tests of the Collaborative Departure Queue Management (CDQM) concept at Memphis (MEM) airport [19], the human-in-the-loop simulations of the Spot and Runway Departure Advisor (SARDA) concept at Dallas Fort Worth (DFW) airport [20], the trials of the Departure Manager (DMAN) concept [21] in Athens International airport (ATH) [22], and the field-tests of an N-control based heuristic, Pushback Rate Control (henceforth referred to as PRC) [23]. However, none of these efforts have explicitly estimated the stochasticity of the underlying processes, or developed optimal control policies that account for the uncertainty.

Surface traffic control has also been formulated as an optimization problem minimizing a delay function [24–27]. These formulations are MILP problems that have been shown to be NP-hard [24, 28]. Practical implementations deterministically schedule a small number of operations, typically 20-30 flights at a time, using a rolling horizon. In other words, the solutions are open-loop policies subject to periodic reoptimization. They assume knowledge of the location and intent of all aircraft, no uncertainty, and the ability to instantaneously modify aircraft speeds; they are not applicable in the current operational environment. Receding horizon approaches have also been used for multi-airport capacity management [29].

In this paper, a queuing model is used for the prediction of the takeoff rate to derive two control algorithms, using dynamic programing and approximate dynamic programming (referred to as PRCv2.0 and PRCv2.1, respectively). The paper also presents the design of a decision support tool (DST) for air traffic controllers, and describes the field tests of PRCv2.1 using the new DST at BOS in 2011. Finally, the proposed congestion control algorithm, the takeoff rate predictions, and the results of the field-tests are evaluated.

II. CONTROL STRATEGY DESIGN REQUIREMENTS

The objective of the control strategy is to minimize the number of aircraft taxiing out and thus taxi-out times, while still maintaining runway utilization. It needs to be compatible with currently available information, automation and operational procedures in the airport tower, and have a minimal impact on controller workload. It must also account for uncertainties in the taxi-out process.

As mentioned before, the preferred form of a congestion control strategy is one that recommends a pushback (release) rate to air traffic controllers [23] at the beginning of each time period and is periodically updated. In general, the length of the time period, \( \Delta \), should equal the lead time of the system, that is, the delay between the application of the control input (setting an arrival rate at the runway server by controlling the pushback rate) and the time that the runway “sees” that rate.

For the departure process, this time delay is given by the travel time from the gates to the departure queue. By choosing a value of \( \Delta \) that is approximately equal to the travel time from the gates to the runway, the flights released from the gate during a given time period are expected to reach the departure queue in the next time period.

Careful monitoring of off-nominal events and constraints is also necessary for implementation at a particular site. Particular concerns are gate conflicts (for example, an arriving aircraft is assigned the same gate as a departure that is being held) and the ability to meet controlled departure times (Expected Departure Clearance Times or EDCTs) and other traffic management constraints. In consultation with the BOS Tower, flights with EDCTs were handled as usual and released First-Come-First-Served. Pushbacks were expedited to accommodate arrivals if needed. Finally, since departures of propellor-driven aircraft (props) were known not to significantly affect jet departures [30], props were exempt from Pushback Rate Control.

III. DEPARTURE PROCESS MODEL

A. State variables

At the beginning of each time-window (called an epoch), the state of the airport system was observed. The takeoff rate in this time-window and the state of the departure process at the beginning of the subsequent time-window were predicted and used to recommend a pushback rate. For the purposes of control, the state was derived from the following inputs:

1) Meteorological conditions and runway configuration, \((MC; RC)\).
2) Number of jet aircraft traveling from the gates to the departure runway, \(G\).
3) Number of jet aircraft in the departure queue, \(D\).
4) Expected number of arrivals in the next 15 min, \(A\).
5) Number of props taxing out, \(P\).

The above quantities are all known in the current tower environment: \(G\) is the number of jet flight strips in the ground controller’s position, \(D\) is the number of jet flight strips in the local controller’s position, \(P\) is known from the same positions, and \(A\) is given by the Traffic Situation Display (TSD).

Given the meteorological conditions and runway configuration \((MC; RC)\) at any time \(t\), the state \(N_t\) of the departure process consists of the number of jet aircraft traveling from the gates to the departure queue \((G_t)\) and the number of aircraft in the departure queue \((D_t)\):

\[
N_t = (G_t, D_t). \quad (1)
\]

Total number of aircraft taxiing out, also known as the total work-in-process of the departure process, \(W_t\), is given by

\[
W_t = G_t + D_t. \quad (2)
\]

B. Selection of time period

The value of \(\Delta\) should equal the lead time of the system, that is, the delay between the application of the control input (setting an arrival rate at the runway by controlling the pushback rate) and the time at which the runway sees that arrival rate. By choosing a time horizon that is approximately
equal to the expected travel time from the gate to the departure queue, flights pushing back during a given time period will reach the departure queue in the next time period.

C. Pushback process

At each epoch (the beginning of each time period), the decision maker chooses a pushback rate (arrival rate into the surface system), \( \lambda \in \Lambda = \{0, 1, \cdots, \lambda_{\text{max}}\} \). \( \lambda \) is expressed as the number of pushbacks per \( \Delta \) minutes. By setting a pushback rate at epoch \( t \), the air traffic controller authorizes \( \lambda \) aircraft to push back in that time period. In other words, \( \lambda \) pushbacks will occur in the time period \( [\tau, \tau + \Delta] \) with probability 1.

D. Runway service process

Each departure runway can be modeled as a single server at which aircraft queue to await takeoff. For the case of BOS, the runway service process is represented with the additional constraint that there are \( R_t \) arrivals during the \( [\tau, \tau + \Delta] \) time interval. Such a system is denoted as \( (M(t)|R_t)/E_k/1 \). Assuming that at epoch \( \tau \) there are \( R_t \) aircraft traveling to the departure runway, the probability density function \( g \) of the \( j^{th} \) arrival at the departure runway at time \( t \in [\tau, \tau + \Delta] \) is:

\[
g(r,t) = \frac{R_t - (r-1)}{(\tau + \Delta) - t}, \quad t \in [\tau, \tau + \Delta], \quad r = 0, 1, \ldots R_t
\]

(3) is derived by considering \( R_t - (r-1) \) uniformly distributed random variables in the interval \( [\tau, \tau + \Delta] \). The probability that one of these lies in \( [t, t + dt] \) is \( (R_t - (r-1))dt/(\tau + \Delta - t) \).

The state of the queuing system at time \( t \) is denoted by \( S_t = (R_t, Q_t) \)

\[
S_t = (R_t, Q_t)
\]

(4) where \( R_t \) is the number of aircraft that were traveling to the departure runway at the start of that epoch but have not reached the departure queue yet, and \( Q_t \in \{0, 1, \ldots, kC\} \) is the state of the embedded chain of the semi-Markov process. An example of the chain for \( k = 2 \) and \( C = 4 \) is shown in Figure 2.

A service completion of an Erlang process with shape \( k \) and rate \( k\mu \) is represented with \( k \) stages of exponentially distributed random variables with rate \( \mu \). Each such stage is known as a stage of work. A state of the Markov chain \( (r,q) \) implies that there are \( r \) aircraft that have been traveling to the runway since the start of that epoch, and there are \( q \) stages of work to be completed at the departure runway server. In other words, there are \( \min(1,q) \) aircraft in service and \( \max([q-k]/k,0) \) aircraft in the departure queue.

At epoch \( 0 \), the Markov chain is at the bottom level of Figure 2 in state \( (R_0, Q_0) \), namely, \( R_0 \) aircraft traveling to the departure runway and \( Q_0 \) stages of work to be completed. By the end of epoch \( \Delta \), all \( R_0 \) aircraft will have reached the departure queue, and the Markov chain will be at the top level (0 aircraft traveling). Let \( P_{r,q}(t) \) denote the probability that the queuing system is in state \( (r,q) \) at time \( t \), where \( 0 \leq t \leq \Delta \). The state probabilities \( P_{0,0}(\Delta), P_{0,1}(\Delta), \cdots, P_{kC,\Delta}(\Delta) \) describe the state of the queuing system at the end of the time interval \( \Delta \), and can be determined by considering the possible transitions of the Markov chain (for example, in Figure 2).

The resultant set of equations are known as the Chapman-Kolmogorov equations [33]. For \( 0 \leq t \leq \Delta \), and \( 1 \leq r \leq R_0 \):

\[
\frac{dP_{0,0}}{dt} = k\mu P_{0,1}
\]

(5)

\[
\frac{dP_{0,q}}{dt} = k\mu P_{0,q+1} - k\mu P_{0,q}, \quad 1 \leq q < k
\]

(6)

\[
\frac{dP_{0,q}}{dt} = k\mu P_{0,q+1} + \frac{1}{\Delta - t} P_{q,q-k} - k\mu P_{0,q}, \quad k \leq q < kC
\]

(7)

\[
\frac{dP_{0,kC}}{dt} = \frac{1}{\Delta - t} P_{kC,(kC-1)} - k\mu P_{0,kC}
\]

(8)

\[
\frac{dP_{r,0}}{dt} = k\mu P_{r+1,0} - \frac{r}{\Delta - t} P_{r,0}
\]

(9)

\[
\frac{dP_{r,q}}{dt} = k\mu P_{r+1,q} - k\mu P_{r,q} - \frac{r}{\Delta - t} P_{r,q}, \quad 1 \leq q < k
\]

(10)

\[
\frac{dP_{r,q}}{dt} = k\mu P_{r+1,q} + \frac{r+1}{\Delta - t} P_{r+1,q-k} - k\mu P_{r,q} - \frac{r}{\Delta - t} P_{r,q}, \quad k \leq q \leq k(C-1)
\]

(11)

\[
\frac{dP_{r,kC}}{dt} = \frac{r+1}{\Delta - t} P_{r+1,kC} - k\mu P_{r,kC}, \quad k(C-1) < q < kC
\]

(12)

\[
\frac{dP_{kC,0}}{dt} = \frac{1}{\Delta - t} P_{kC,(kC-1)} - k\mu P_{kC,0}
\]

(13)

\[
\frac{dP_{R_0,0}}{dt} = k\mu P_{R_0+1,0} - \frac{R_0}{\Delta - t} P_{R_0,0}
\]

(14)

\[
\frac{dP_{R_0,q}}{dt} = k\mu P_{R_0+1,q} - \left( \frac{R_0}{\Delta - t} - k\mu \right) P_{R_0,q}, \quad 1 \leq q \leq k(C-1)
\]

(15)
\[ \frac{dP_{0,q}}{dt} = k\mu P_{0,q+1} - k\mu P_{0,q}, k(C-1) < q < kC \] (16)
\[ \frac{dP_{0,kC}}{dt} = -k\mu P_{0,kC} \] (17)

The state probabilities \( P_{i}(\Delta) \), \( i = 0, 1, \ldots, kC \), are obtained by numerically solving (5)-(17) for \( \tau = \Delta \), with initial value \( (R_0, Q_0) \). The probability of the queuing system state at time \( \Delta \) being \( i = Q_\Delta \) is given by

\[ Q_\Delta = \{ f(R_0, Q_0) \] with \( p_{q(i)}(R_0, Q_0) = P_{0,i}(\Delta) \) for \( 0 \leq i \leq kC \)

\[ \implies p_{q}(R_0, Q_0) = P_{0}(\Delta) \triangleq [P_{0,0}(\Delta), P_{0,1}(\Delta), \ldots, P_{0,kC}(\Delta)]'. \]

E. System dynamics

Suppose that at epoch \( \tau \), there are \( R_\tau \) aircraft traveling to the runway, \( Q_\tau \) stages of work left in the queue, and the decision maker selects a pushback rate \( \lambda_\tau \). At \( \tau + \Delta \), \( R_\tau \) aircraft will have reached the departure queue, \( \lambda_\tau \) aircraft will be traveling, and \( Q_{\tau + \Delta} = f(R_\tau, Q_\tau) \) stages of work will remain in the queue. The queuing system therefore evolves according to the following equation:

\[ (R_{\tau + \Delta}, Q_{\tau + \Delta}) = (\lambda_\tau, f(R_\tau, Q_\tau)) \] (18)

Given that the chain is in state \( (r,q) \) at the epoch \( \tau \) and the pushback rate \( \lambda \) is chosen, the probability that the chain is in state \( (i,j) \) at the next epoch \( \tau + \Delta \) is:

\[ \Pr_{(r,q)\rightarrow(i,j)}(\lambda) = \begin{cases} p_{q(i)}(r,q) & \text{if } i = \lambda \\ 0 & \text{otherwise} \end{cases} \] (19)

The state \( S_t \) of the queuing system maps to the state \( N_t \) of the departure process as follows:

\[ N_t = \begin{cases} (\lambda_{t-1, \max}([Q_t - k]/k, 0)), & t \in \{0, \Delta, \ldots\} \\ (V_t + R_{t-1, \max}([Q_t - k]/k, 0)), & \text{otherwise} \end{cases} \] (20)

where \( V_t \) is the number of aircraft that pushed back before the start of the time period in which \( t \) lies, and \( \tau \) is the time of sampling the system every \( \Delta \) minutes, the departure process is decoupled into the pushback and the runway service processes, which are independent of each other within a time period.

F. Choice of cost function

The control strategy sets the arrival rate of aircraft to the queuing system, namely, the pushback rate, to balance two objectives: Minimize the expected departure queue length, and maximize the runway utilization. The cost function associated with a state \( (r,q) \) of the queuing system is denoted \( c(q) \). This cost is a combination of the queuing cost and the cost of non-utilization of the runway, that is, \( q = 0 \). If \( q \in \{1, 2, \ldots, k\} \), both the queuing and non-utilization costs are zero. For all higher states, \( q > k \), there is a queuing cost which is usually assumed to be a monotonically nondecreasing function of \( q \) with increasing marginal costs [34, 35]. It is assumed to scale quadratically with the state of the queue, because the expected system delay is a quadratic function of queuing state \(((D \cdot (D+1))/2)/\mu)\). A candidate cost function with these properties is:

\[ c(q) = \begin{cases} H, & q = 0 \\ \left(\frac{(q-k)}{k}\right)^2, & q = 1, \ldots, kC \end{cases} \] (21)

where \( H \) is the cost of a loss of runway utilization.

Equations (5)-(17) are solved numerically to calculate

\[ p_{q}(R_0, Q_0, t) = \left[ \sum_{r=0}^{R_0} p_{r,0}(t), \sum_{r=0}^{R_0} p_{r,1}(t), \ldots, \sum_{r=0}^{R_0} p_{r,kC}(t) \right] ' \] (22)

time \( t \). Numerical experiments show that sampling 10 times a minute is sufficient for accurately calculating the expected cost of each state, \( c \), over \( \Delta \) min [32]:

\[ c(R_0, Q_0) = \sum_{i=0}^{10\Delta - 1} \frac{1}{10} p_{q}(R_0, Q_0, i/10) \cdot c(Q_0) \] (23)

IV. Dynamic Programming Formulation

The optimal costs, \( J'(r,q) \), at each state, \( (r,q) \), for the infinite horizon problem with discount factor \( \alpha \) are given by the Bellman equation [36]:

\[ J'(r,q) = \min_{\lambda \in \Lambda} \left\{ c(r,q) + \alpha \sum_{j=0}^{kC} \Pr_{(r,q)\rightarrow(i,j)}(\lambda) J'(i,j) \right\} \]

\[ = \min_{\lambda \in \Lambda} \left\{ c(r,q) + \alpha p_{q}(r,q) \cdot J'(\lambda) \right\} \]

where \( J'(\lambda) = [J'(\lambda, 0), J'(\lambda, 1), \ldots, J'(\lambda, kC)]' \), \( r \in \{0, 1, \ldots, \lambda_{\max}\} \) and \( q \in \{0, 1, \ldots, kC\} \).

The earlier assumptions that the \( R_\tau \) aircraft traveling at epoch \( \tau \) will all reach the queue during the time interval \( (\tau, \tau + \Delta) \) and that the pushback rate \( \lambda_\tau \) set at epoch \( \tau \) will arrive at the runway at \( t > \tau + \Delta \) are now relaxed. For each value of \( \lambda_\tau \) and \( R_\tau \), \( i \) out of the \( \lambda_\tau \) aircraft are assumed to reach the runway during the time interval \( (\tau, \tau + \Delta) \) with probability \( \beta_i \). Similarly, \( i \) out of the \( R_\tau \) aircraft are assumed to reach the runway at \( t > \tau + \Delta \) with probability \( \gamma_i \). Therefore, \( R_\tau \) aircraft reach the runway during the time interval \( (\tau, \tau + \Delta) \), and \( \lambda_\tau \) aircraft at \( t > \tau + \Delta \), with probability \( 1 - \sum \beta_i - \sum \gamma_i \).

Equation (18) therefore becomes:

\[ (R_{\tau + \Delta}, Q_{\tau + \Delta}) = (\lambda_{\tau - i, f(R_{\tau + i, D_{\tau}})}, \text{w.p. } 1 - \sum \beta_i - \sum \gamma_i) \]

\[ (\lambda_{\tau - i, f(R_{\tau + i, D_{\tau}})}, \text{w.p. } \beta_i, i = 1, \ldots, \lambda_{\tau}) \]

\[ (\lambda_{\tau - i, f(R_{\tau + i, D_{\tau}})}, \text{w.p. } \gamma_i, i = 1, \ldots, \lambda_{\tau}) \]

The above equation is seen to maintain the Markov property. For these system dynamics, the Bellman equation for the infinite horizon problem with discount factor \( \alpha \) is:

\[ \min_{\lambda \in \Lambda} \left\{ (1 - \sum \beta_i - \sum \gamma_i)[c(r+i,q) + \alpha p_{q}(r+i,q) \cdot J'(\lambda)] \right\} \]

\[ J'(r,q) = \sum_{i} \beta_i[c(r+i,q) + \alpha p_{q}(r+i,q) \cdot J'(\lambda - i)] \]

\[ + \sum_{i} \gamma_i[c(r+i,q) + \alpha p_{q}(r+i,q) \cdot J'(\lambda + i)] \] (24)

Equation (24) illustrates the tradeoffs involved with the choice of time period, \( \Delta \). A long time period requires less frequent updates of the optimal policy, making implementation easier. It is, however, necessary to predict runway performance and maintain runway utilization over a longer period of time. If \( \Delta \) is significantly less than the lead time, a smaller inventory will be necessary at the departure queue to maintain runway utilization. However, \( \gamma_i \) will be large, and only a fraction of aircraft taxiing will arrive at the runway by the next epoch. As a result, the state at epoch \( \tau \), \( (R_\tau, Q_\tau) \), will not be closely related to the queuing system state at the next epoch, \( Q_{\tau + \Delta} \).
More frequent updates of the optimal policy will also be necessary, increasing air traffic controller workload. This problem can be shown to satisfy the property of weak accessibility. Suppose that at epoch \(0\), the embedded chain is at state \((r_0, q_0)\). At the next epoch, the chain will be at any of the states \((\lambda_0, 0), (\lambda_0, 1), \ldots (\lambda_0, \min(r_0 + q_0, kC))\) with nonzero probability. Suppose that the following control law is applied: For all \((r_0, q_0)\), \(\lambda_0 = \lambda_{\text{max}}\), where \(\lambda_{\text{max}} > \mu\). Then, the queuing system will reach the state \((\lambda_{\text{max}}, kC)\) within a finite number of epochs with nonzero probability. Also, at the next epoch, the state will be in any of the states \((\lambda_{\text{max}}, 0), (\lambda_{\text{max}}, 1), \ldots (\lambda_{\text{max}}, kC)\) with nonzero probability. As before, from any of these states, the chain will reach the state \((\lambda_{\text{max}}, kC)\) within a finite number of epochs with nonzero probability. The state \((\lambda_{\text{max}}, kC)\) is therefore recurrent under this control law, and weak accessibility is satisfied.

Using a discount factor as in (24) may not be appropriate, since the cost of an unutilized runway remains constant in time. An alternate formulation is to determine the policies that minimize the average optimal cost per stage, \(c^*\):

\[
c^* + h^*(r, q) = \min_{\lambda \in \Lambda} \left\{ (1 - \sum \beta_i - \sum \gamma_i) \left[ \bar{c}(r, q) + \bar{p}_q(r, q) \cdot h^*(\lambda) \right] + \sum \beta_i [\bar{c}(r + i, q) + \bar{p}_q(r + i, q) \cdot h^*(\lambda - i)] + \sum \gamma_i [\bar{c}(r - i, q) + \bar{p}_q(r - i, q) \cdot h^*(\lambda + i)] \right\}
\]

(25)

V. A PUSHBACK RATE CONTROL POLICY FOR BOS

This section describes the application of PRC_v2.0, given by Equation (25), to the departure process at BOS. The focus here is on runway configuration 22L, 27 | 22L, 22R in VMC during the peak evening departure push, although other configurations were also studied. As explained in Section II, props at BOS were exempt from pushback control.

A. Selection of time period

The average unimpeded taxi-out time at BOS is 12.6 minutes under VMC [10]. There is an added delay due to taxiway congestion, which is 1-2 minutes under moderate traffic conditions. 15 minutes is therefore a suitable choice of time-window at BOS. Furthermore, due to a lack of accurate measurements, it is assumed that \(\beta_i = \gamma_i = 0\) for all \(i\). Equation (25) therefore becomes:

\[
c^* + h^*(r, q) = \min_{\lambda \in \Lambda} \left\{ (1 - \sum \beta_i - \sum \gamma_i) [\bar{c}(r, q) + \bar{p}_q(r, q) \cdot h^*(\lambda)] \right\}
\]

(26)

B. Estimation of runway service process parameters

The parameters of the runway service process of BOS during peak evening times were extracted using ASDE-X data from Nov 2010-Jun 2011. An Erlang distribution was fitted using the approximate Method of Moments. The mean service time was 1.54 min, and the variance was 0.47 min\(^2\) [32].

C. Maximum pushback rate and cost function

The set of permissible policies is defined as \(\Lambda = \{0, 1, \ldots, \lambda_{\text{max}}\}\). At most airports, there is a natural threshold for the maximum admissible rate of arrivals into the departure process (pushbacks). \(\lambda_{\text{max}}\) is estimated to be 15 aircraft/15 min, and the queuing system capacity \(C\) is estimated to be 30 at BOS. The cost of underutilizing the runway, \(c(0)\), is chosen to be equal to the cost of a queue of 25 departures, reflecting the fact that at BOS, a very long queue can lead to surface gridlock and non-utilization of the runway [23].

D. Calculation of optimal policies

Given the service time distribution \((k, k\mu)\), the time period \(\Delta\), the queuing space \(C\), the set \(\Lambda\) and the costs \(c\), the optimal pushback policies can be obtained by solving (26). It can be solved efficiently using policy iteration with a suitable choice of initial policy. The policy iteration algorithm converges in fewer than 10 iterations. The optimal policies \(\lambda^*\) are a function of the state of the embedded chain \((r, q)\), which is not observable. However, each state of the chain is mapped to an observed quantity, \(N\), through (20). For \(0 \leq G \leq \lambda_{\text{max}}\), the optimal pushback rate is approximated by:

\[
\tilde{\lambda}(G, 0) = \left[ \frac{\sum_{j=0}^{k} \lambda^*(G, j)}{k + 1} + 0.5 \right]
\]

(27)

\[
\tilde{\lambda}(G, D) = \left[ \frac{\sum_{j=0}^{(D-1)k} \lambda^*(G, j)}{k} + 0.5 \right] \quad \text{for } 1 \leq D < C
\]

(28)

Figure 3 shows the contours of the optimal pushback policy \(\tilde{\lambda}\) as a function of the number of aircraft in the departure queue \(D\) and the number of aircraft traveling to the runway \(G\). As expected, the optimal pushback rates decrease for increasing \(D\) and \(G\). The optimal policies can also be characterized by the expected work-in-process at the next epoch, \(\bar{W}_{T+\Delta}\), as a function of the current state using Equation (18), as shown in Figure 4. When \(W_T \geq 23\), the optimal pushback rate is 0, but it is not sufficient to reduce \(\bar{W}_{T+\Delta}\) to 13. By contrast, when \(W_T \leq 13\), the optimal pushback policy increases \(\bar{W}_{T+\Delta}\) to values higher than 13.
that the optimal pushback policy at time $\tau$ is a function of the expected queue length at time $\tau + \Delta$.

Figure 5 shows the scatterplot of the optimal pushback rate $\dot{\lambda}_\tau(G_t, D_t)$ as a function of the expected $D_{\tau+\Delta}(G_t, D_t)$ for all $0 \leq G \leq \lambda_{\text{max}}$ and $0 \leq D \leq C$, along with a fitted convex non-increasing function that minimizes absolute deviations from the calculated points. The equivalent PRC_v1.0 strategy which aims at keeping $W_{\tau+\Delta}$ at 13 aircraft is also shown [23]. For the most part, the two strategies are the same after rounding to the closest integer. However, when the expected queue length at $\tau + \Delta$ is less than 4, PRC_v2.0 increases $W_{\tau+\Delta}$ to 14 or 15 in order to better account for the risk of runway non-utilization.

E. Conditional throughput forecasts

Parameters such as the fleet mix and the expected number of landings can provide a conditional forecast for the runway service time distribution [30, 37]. These parameters explain some of the variance of the departure throughput, and provide a better estimate of the expected departure capacity. For example, the departure throughput in a given 15-min interval for runway configuration 22L, 27 | 22L, 22R in evening periods under visual meteorological conditions, given the arrival throughput and taxiing prop departures, can be estimated from the regression tree shown in Figure 6. This regression tree was validated using 10-fold cross validation.

These conditional forecasts are incorporated into the algorithm as follows:

- At epoch $\tau$, the conditional throughput for the time window $(\tau, \tau + \Delta)$ is predicted from the regression tree, using the expected number of arrivals ($A$) and the number of props taxiing out ($P$).
- The expected takeoff rate in the time window $(\tau, \tau + \Delta)$ and queue length at $\tau + \Delta$ are calculated using a $(M(t)/R_t)/E_k/1$ queuing model with parameters fitted to the throughput forecast from the previous step.
- The PRC_v2.0 curve (Figure 5) is used to calculate the optimal pushback policy for this expected queue length.

The proposed approach, denoted PRC_v2.1, is a heuristic modification of PRC_v2.0 in the spirit of roll-out algorithms [38] to incorporate the conditional forecast. The intuition behind the derivation of PRC_v2.1 is that the conditional forecasts are used only to update the prediction of the queue length. While the reduction in the variance of the capacity distribution could yield a more aggressive control policy, this feature is not exploited. The optimal PRC_v2.1 policy at epoch $\tau$ is denoted $\dot{\lambda}_\tau$, and is a function of the departure queue, the number of aircraft traveling to the runway and the number of props taxiing out at $\tau$, as well as the expected number of landings in $(\tau, \tau + \Delta)$.

F. Rounding of optimal policies

As explained in Section II, the optimal policy needs to be communicated to the controllers as a recommended rate. The optimal pushback rate for each 15-minute time-period is therefore rounded to one of the following: 0 aircraft/15 min (Stop), 1 aircraft/5 min, 1 aircraft/3 min, 2 aircraft/5 min, 1 aircraft/2 min, 3 aircraft/5 min, 2 aircraft/3 min, 4 aircraft/5 min or 1 aircraft/min [23].

VI. DESIGN OF A DECISION SUPPORT TOOL

A Decision Support Tool (DST) was designed in order to implement Pushback Rate Control algorithms, and PRC_v2.1 particular, in the airport tower environment. The device used was a 7" Samsung Galaxy Tab™ tablet computer with the Android™ operating system, which is convenient for application development, while being compact and portable. Two tablet computers were used in the implementation, namely, the rate control transmitter and the rate control receiver. Inputs
were entered into the rate control transmitter, which then determined the optimal pushback rate and communicated it via a Bluetooth wireless link to the rate control receiver. The receiver displayed the recommended rate to the Boston Gate (BG) controller, who authorized aircraft to pushback.

A. Inputs

The inputs to the rate control transmitter were the runway configuration, meteorological conditions, expected number of arrivals in the next 15 minutes, numbers of jets under ground control and local control and number of props taxiing out. The input interface is shown in Figure 7b. The expected takeoff rate and the recommended pushback rate were then calculated using a look-up table for the PRC_v2.1 algorithm, and transmitted to rate control receiver.

B. Outputs

The receiver conveyed the suggested pushback rate to the BG controller through one of two display modes: the rate control and the volume control displays.

1) Rate control display: The output in this mode was a color-coded image of the suggested pushback rate. In this display mode, the BG controller kept track of the time intervals and the number of aircraft that have already pushed back. When the demand for pushbacks exceeded the recommended rate, aircraft were held at the gate until the next time interval. The BG controller kept track of aircraft holds, and released them at the appropriate time.

2) Volume control display: This display mode helped the BG controller keep track of the number of aircraft that had called and had been released. It was an alternative to the handwritten notes that controllers otherwise used to keep track of gate-holds. The volume control mode also provided visual cues of the timeline and upcoming actions.

On the volume control display, a 15-minute time period is broken down into smaller time intervals. For example, if the rate is 3 per 5 minutes, the display shows three rows of three aircraft icons, with each row corresponding to a 5-minute time interval (illustrated in Figure 7a). Aircraft can only be released during an ongoing time interval, indicated by a small black arrow to the left of it. Future positions can only be reserved. Any unused release spots roll over to the next time interval.

C. DST deployment

During the field trials at BOS in 2011, a member of the research team gathered and input data into the rate control transmitter. The rate control receiver was located next to the BG controller, who chose between the rate control and volume control displays. It is expected that in a long-term deployment, the traffic management coordinator (TMC) or the tower supervisor would input the data. For a part of the field-tests, the BG position was merged with another position, either clearance delivery or the TMC to investigate the potential implementation of PRC without requiring an additional controller at the BG position.

VII. FIELD-TESTS AT BOS

PRC_v2.1 was field-tested at BOS during 19 evening (4PM-8PM) demo periods between July 18th and September 11th 2011. However, there was little congestion when the airport operated in its optimal configuration (4L, 4R | 4L, 4R, 9) or when the demand was low. There was enough congestion to warrant gate-holds in only eight of the demo periods. A total of 144 flights were held, with an average gate-hold of 5.3 min. During the most congested periods, up to 44% of flights experienced gate-holds.

This section presents an example describing one test period (July 21, 2011), the methodology used to calculate taxi-out time and fuel burn savings, and a comparison of the predicted and observed takeoff rates during the test periods.

A. An illustrative example from the field-tests

This section illustrates the typical outcomes of the PRC_v2.1 field-tests using a case day with significant gate-holds (July 21, 2011). Figure 8 depicts the events of the demo period, divided into 15-minute windows. The top plot shows the demand for pushbacks (that is, the number of aircraft that called for pushback), the number of pushbacks that were cleared, and the resulting number of jet aircraft actively taxiing out. The middle plot shows the predicted and measured throughput. Finally, the bottom plot shows the average taxi-out times and gate-holding times for aircraft that pushed back in each time interval.

The top plot in Figure 8 shows that as the number of jet aircraft taxiing-out exceeds 14, gate-holds are initiated in order to regulate the traffic to the desired state. For this configuration, the desired state is 13-14 aircraft on the surface. The airport stays in the desired state despite the high variance of the departure throughput (middle plot of Figure 8) and the rounding-off of the recommended pushback rates.

A key objective of the field-test was to maintain pressure on the departure runways, while limiting surface congestion. By maintaining runway utilization, it is reasonable to expect that gate-hold times translate to taxi-out time reduction. Runway utilization was shown to not be adversely impacted by the control strategy by checking that there was always at least one aircraft in the runway queue during the demo periods. This validation was performed both through visual observations and through the analysis of ASDE-X data [39].
The unimpeded taxi-out times of flights are determined using ASDE-X data [32]. Given the pushback clearance time, the sum of the unimpeded taxi-out time and a taxi-out delay due to congestion [32], each flight is propagated to the runway, where it is assigned to the next available takeoff slot in that time period. The difference between this wheels-off time and the pushback clearance time is the expected taxi-out time. The comparison of the actual and predicted runway queuing times helps validate the simulations and assess the impact of the control strategy.

Table I presents detailed simulation results for two days with significant gateholds, Jul 21, 2011 and Jul 22, 2011. The results correspond to flights that were cleared for pushback between 1645 and 2045 hours. The table shows that the model predictions of the mean taxi-out and queuing times match the observed data from the actual operations.

TABLE I: Gate-hold times, mean taxi-times and queuing times from actual data, simulations of actual operations, and simulations of hypothetical operations with no gate-holds.

<table>
<thead>
<tr>
<th>Date</th>
<th># Fls</th>
<th>Gatehold time (min)</th>
<th>Scenario</th>
<th>Avg. taxi-out time (min)</th>
<th>Avg. queue time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/21</td>
<td>121</td>
<td>368</td>
<td>Act. data</td>
<td>16.5</td>
<td>5.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Act. sim.</td>
<td>16.5</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hyp. sim.</td>
<td>19.5</td>
<td>7.9</td>
</tr>
<tr>
<td>7/22</td>
<td>121</td>
<td>279</td>
<td>Act. data</td>
<td>17.9</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Act. sim.</td>
<td>17.9</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Hyp. sim.</td>
<td>20.2</td>
<td>9.2</td>
</tr>
</tbody>
</table>

Figure 9 (top) shows the instantaneous actual and simulated queue on July 21, 2011. The actual queue is seen to be accurately predicted by the simulations. The bottom plot on the same figure shows the actual queue on July 21 with PRC, and the simulated prediction without PRC. The difference between the two queue lengths illustrates the benefit of PRC_v2.1.

The overall estimation of fuel burn savings is conducted by using the simulated taxi-out time savings for all test periods, and models of taxi fuel burn [23, 39]. The total fuel burn reduction estimated to be 2,650 US gallons, which translates to an average savings of about 57 kg per gate-held flight. Table II summarizes the results of the eight demo periods with significant gate-holds.

2) Distribution of benefits: Equity is an important criterion in evaluating potential congestion management strategies. The PRC approach, as implemented here, invokes a First-Come-First-Serve (FCFS) policy in clearing flights for pushback,
TABLE II: Summary of the eight demo periods with significant gate-holds during the PRC_v2.1 field-tests in 2011.

<table>
<thead>
<tr>
<th>Date</th>
<th>Period</th>
<th>Configuration</th>
<th>No. of gate-holds</th>
<th>Total gate-hold time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7/18</td>
<td>4.45-5PM</td>
<td>22L, 22R</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>7/21</td>
<td>5.15-5PM</td>
<td>22L, 22L, 22R</td>
<td>42</td>
<td>384</td>
</tr>
<tr>
<td>7/22</td>
<td>5.15-8.30PM</td>
<td>22L, 22L, 22R</td>
<td>50</td>
<td>290</td>
</tr>
<tr>
<td>7/24</td>
<td>5.15-5PM</td>
<td>4L, 4R, 4L, 4R, 9</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>7/28</td>
<td>5.30-8PM</td>
<td>4L, 4R, 4L, 4R, 9</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>8/11</td>
<td>5.30-8.15PM</td>
<td>22L, 22L, 22R</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>8/14</td>
<td>5.00-6.30PM</td>
<td>22L, 22L, 22R</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9/11</td>
<td>5.30-8.30PM</td>
<td>4L, 4R, 4L, 4R, 9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9/11</td>
<td>6.30-8.15PM</td>
<td>22L, 22L, 22R</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>144</td>
<td>761</td>
</tr>
</tbody>
</table>

Since this is the widely accepted measure of fairness in air traffic control [40]. One would therefore expect no bias toward any airline with regard to gate-holds incurred, and that the number of gate-holds for an airline would be commensurate with its contribution to departure traffic during congested periods. However, in practice, inadvertent resequencing due to pushback procedures, taxi speeds or sudden changes in demand can potentially lead to differences between the gatehold time and taxi-out savings of individual flights. Similarly, the benefit of a gate-hold can also extend to other flights [10]. Figure 10 shows that during the PRC_v2.1 field-tests, the FCFS sequence was mostly maintained, and that the gatehold times were approximately equal to the taxi-time reduction experienced by each airline. It must be noted that the fuel burn benefit to an airline depends on its fleet mix. Figure 10 shows that while the taxi-out time reductions were similar to the gate-hold times, some airlines (for example, Airlines 4, 13, 21 and 27) enjoyed a greater proportion of fuel savings. These airlines were typically those which operated several Heavy aircraft during the evening periods.

C. Takeoff rate prediction

As explained in Section V, the PRC_v2.1 algorithm was used to predict the jet takeoff rate. The predictions were validated during shadow testing by means of visual observations, and were subsequently used during the 19 days of the trials. Table III reports the mean error (ME), mean absolute error (MAE) and root mean square error (RMSE) of the predicted jet takeoff rate (relative to the observed value) over the 182 15-min periods of field-tests. The errors for the 93 periods with at least 10 taxiing jet aircraft are also shown, because gate-holds were most likely in these times.

Table III shows that the regression tree-based prediction algorithm used in PRC_v2.1 predicts the takeoff-rate quite well. The mean absolute error is only 1.14 during moderate and high traffic conditions (10 or more jets taxiing-out). Most importantly, the prediction errors are within the level of uncertainty considered in the design of the PRC_v2.1 strategy. For the 22L, 27 or 22L, 22R configuration with at least 10 jet departures taxiing, the takeoff rate was underestimated by at most 2.7, while the PRC_v2.1 algorithm tries to maintain a queue of at least 4 aircraft for these conditions. Similarly, for the 4L, 4R or 4L, 4R, 9 configuration with at least 10 jet departures taxiing, the takeoff rate was underestimated by at most 3.7, while the PRC_v2.1 algorithm tries to maintain a queue of at least 5 aircraft. These observations suggest that the inventory targeted by the algorithm at the queue was sufficient to avoid runway underutilization; a more aggressive congestion control policy may have resulted in an empty runway queue. The importance of maintaining a sufficiently large runway queue has been recognized by other researchers as well [22].

D. Qualitative observations

1) Compatibility with traffic flow management initiatives: An important goal of this effort was to investigate the compatibility of Pushback Rate Control with other traffic flow management initiatives. On two field-test periods, controllers demonstrated that they could handle airspace restrictions such as Minutes-In-Trail (MINIT) programs and target departure times (e.g., EDCTs) while executing the PRC_v2.1 strategy. It was shown that if known in advance, delays due to controlled departure times could be efficiently absorbed as gate-holds.

2) Increased predictability: During the field-tests, once the suggested pushback rate was given to the controller at the start of each time period, the controller communicated the expected release times to all aircraft on hold. The advance notice of expected pushback clearance times improves predictability, and can be useful in planning ground resources.

VIII. EVALUATION OF THE DST

After the field-tests at BOS had been completed, air traffic controllers in the ATCT were surveyed regarding their opinions on the study as a whole, and specifically on the implementation and use of the DST. The survey responses were positive and the controllers liked the DST [41]. Several responses also supported combining BG and another position, removing the need for a dedicated controller during gate-holds. Comments on the best features of the DST included “the ability to touch planes”, “reserve spots”, “[the ability to] count the planes and account for aircraft with long delays”, “the ability to touch planes”, “reserve spots”, “[the ability to]...
“allows me to push and tells me to hold”, and “easy to understand”. Suggestions for improvement included increasing the icon sizes and maintaining more pressure on the runway.

IX. CONCLUSIONS

This paper presented the design and field-testing of a Pushback Rate Control strategy at Boston Logan International Airport (BOS). The proposed approach used historical data to predict the performance of the airport under a set of operating conditions, and used dynamic programming to balance the objectives of maintaining runway utilization and limiting surface congestion. The optimal policy was a recommended rate at which aircraft were cleared for pushback. A decision support interface was designed to display the suggested pushback rate, and to help air traffic controllers keep track of requests for pushback, gate-holds and other metering constraints.

During 8 four-hour tests conducted at BOS during the summer of 2011, fuel use was reduced by an estimated 9 US tons (2,650 US gallons), while carbon dioxide emissions were reduced by an estimated 29 US tons. Aircraft gate pushback times increased by an average of 5.5 minutes for the 144 flights that were held at the gate, but with a corresponding decrease in taxi-out times. Finally, a survey of the air traffic controllers involved in the 2011 demo indicated strong support for the Pushback Rate Control approach, the manner of implementation, and the decision support tools developed for the deployment of such strategies.

ACKNOWLEDGMENTS

This work was supported by the Federal Aviation Administration’s Office of Environment and Energy through MIT Lincoln Laboratory and the Partnership for Air Transportation Noise and Emissions Reduction (PARTNER). The demo was helped make possible by Brendan Reilly, Deborah James, Pat Hennessy, John Ingahurgo, John Melecio, Michael Nelson and Chris Quigley at the BOS Facility; Vincent Cardillo, Flavio Leo and Robert Lynch at Massport; and George Ingram and other airline representatives at the ATA. Vivek Panyam (University of Pennsylvania) volunteered his effort in the development of the Android tablet application. John Hansman, Harshad Khadilkar and Tom Reynolds participated in tower observations before and during the trials; Harshad Khadilkar also analyzed the ASDE-X data. Alex Nakahara helped compute the preliminary fuel burn savings from the gate-hold data. We thank Steve Urllass and Lourdes Maurice at the FAA for their support, and James Kuchar and Jim Eggert of MIT Lincoln Laboratory for their support and help with the ASDE-X data.


Ioannis Simaiakis is a Management Consultant with McKinsey’s Operations Practice in Washington, DC. He received his Ph.D. in Aeronautics and Astronautics from the Massachusetts Institute of Technology. His research focuses on modeling and predicting taxi-out times and airport operations planning under uncertainty.

Melanie Sandberg is associate staff in the Air Traffic Control Division of MIT Lincoln Laboratory. She obtained her Master of Science in Transportation from the Massachusetts Institute of Technology. Her research interests include operational strategies for improvements in air traffic management.

Hamsa Balakrishnan is an Associate Professor of Aeronautics and Astronautics at the Massachusetts Institute of Technology. She received her PhD in Aeronautics and Astronautics from Stanford University. Her research interests include optimization and control algorithms for air transportation systems.