

# Multi-objective airport gate assignment problem in planning and operations

V. Prem Kumar<sup>\*†</sup> and Michel Bierlaire

*Transport and Mobility Laboratory (TRANSP-OR), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland*

## SUMMARY

We consider the assignment of gates to arriving and departing flights at a large hub airport. This problem is highly complex even in planning stage when all flight arrivals and departures are assumed to be known precisely in advance. There are various considerations that are involved while assigning gates to incoming and outgoing flights (such a flight pair for the same aircraft is called a turn) at an airport. Different gates have restrictions, such as adjacency, last-in first-out gates and towing requirements, which are known from the structure and layout of the airport. Some of the cost components in the objective function of the basic assignment model include notional penalty for not being able to assign a gate to an aircraft, penalty for the cost of towing an aircraft with a long layover, and penalty for not assigning preferred gates to certain turns.

One of the major contributions of this paper is to provide mathematical model for all these complex constraints that are observed at a real airport. Further, we study the problem in both planning and operations modes simultaneously, and such an attempt is, perhaps, unique and unprecedented. For planning mode, we sequentially introduce new additional objectives to our gate assignment problem that have not been studied in the literature so far—(i) maximization of passenger connection revenues, (ii) minimization of zone usage costs, and (iii) maximization of gate plan robustness—and include them to the model along with the relevant constraints. For operations mode, the main objectives studied in this paper are recovery of schedule by minimizing schedule variations and maintaining feasibility by minimal retiming in the event of major disruptions. Additionally, the operations mode models must have very, very short run times of the order of a few seconds.

These models are then applied to a functional airline at one of its most congested hubs. Implementation is carried out using Optimization Programming Language, and computational results for actual data sets are reported. For the planning mode, analyst perception of weights for the different objectives in the multi-objective model is used wherever actual dollar value of the objective coefficient is not available. The results are also reported for large, reasonable changes in objective function coefficients. For the operations mode, flight delays are simulated, and the performance of the model is studied. The final results indicate that it is possible to apply this model to even large real-life problems instances to optimality within short run times with clever formulation of conventional continuous time assignment model. Copyright © 2013 John Wiley & Sons, Ltd.

**KEY WORDS:** airport gate assignment; planning; mathematical modeling

## 1. INTRODUCTION

The airline industry has long been a fertile area for applying optimization techniques. This paper describes the airport gate assignment problem (GAP) as experienced by congested hub airports and

<sup>\*</sup>Correspondence to: Prem Kumar Viswanathan, Transport and Mobility Laboratory (TRANSP-OR), School of Architecture, Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Switzerland. E-mail: prem.viswanathan@epfl.ch

<sup>†</sup>A major part of the work was performed while the author was affiliated to Symphony Marketing Solutions, Bangalore, India.

large airline companies. Airport gates are restricted resources and are used by incoming and outgoing flights to park the aircraft, disembark the passengers of the incoming flights, and board the passengers of the outgoing flights.

There are some subtle differences in the airport GAP encountered at different airports across the world. For instance, in many European and Asian airports, it is normal to assign an aircraft to a remote bay (also referred to as apron or stands), far away from the airport terminal, and then the passengers are disembarked and boarded using shuttle busses, in the absence of an available gate. However, in USA, remote bays are not allowed, and all passengers are mandatorily required by law to be disembarked and boarded through an airport gate. This makes the problem highly restricted as well as complex because there have been instances, especially during major disruptions, when the aircrafts are required to wait on the tarmac for several hours to disembark the passengers.

There are some other differing features in the operations of airports across the world. In USA, airport gates are resources that are owned or leased by a particular airline company for a specific period based on a medium-term to long-term contract. In case an airline falls short of gates, it will either have to negotiate and sublease gates from a competitor or downsize its operations at that airport. On the contrary, airport gates are largely managed by the airport authorities in Asia and Europe. Thus, in the USA, the onus of efficient ground operations lies entirely with the concerned airline.

The period that an aircraft spends on the ground between incoming (also referred to as arriving) and outgoing (also referred to as departing) flights is called a turn. In the rest of the paper, we use the term “turn” to refer to a flight combination associated with the same aircraft. For every turn, the aircraft is assigned to a gate, and the same gate is utilized by many aircrafts in the course of a day. Airport operations team develops gate assignment plans by using an optimization model that assigns gates to every turn, while balancing operational constraints, given the fleet and turn information through a station. Each hub airport must have a gate plan based on its geography and layout.

Although the different optimization criteria of the problem considered by us in this paper are explained in detail in section 3, we will now present some features and restrictions of the problem that have been observed at our study airports. These features and constraints are not airport specific but have wider applications to all airports.

- *Adjacency constraints:* Adjacency constraint is described as a situation when two large aircrafts cannot be accommodated in adjacent (near-by) gates because of structural or space limitation. Thus, when gate A is occupied by aircraft type 1, the adjacent gate B does not allow aircraft type 2 and vice versa. An example would be a situation with gate B11 that has a B747 parked on it. At the same time, gate B12 cannot accommodate a wide bodied aircraft, such as B747, B767, B777, or an A340. Please refer to Figure 1 for further details. This constraint is observed at almost all major airports in the world and has been widely studied by the researchers as well.
- *Last-in first-out (LIFO) gates:* LIFO gates are observed in a situation where two gates are one behind the other as shown in Figure 2. Thus, if the gate #1 is occupied by an aircraft, then the aircraft in gate #2 cannot depart, or the gate #2 cannot be used to accommodate an incoming aircraft during occupancy of gate #1, even if it is free. This constraint is neither widely considered nor extensively studied by the researchers.
- *Towing:* Towing means that an aircraft is towed away after it arrives at a gate and the passengers are allowed to disembark. It will then be towed out and brought back to a gate for departure. The departure gate may or may not be the same as the arrival gate. The purpose of towing is to free

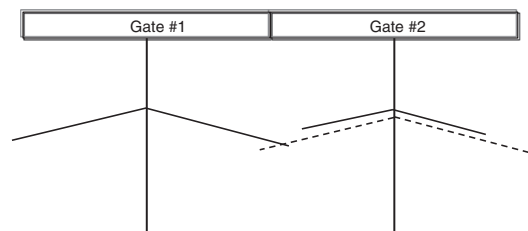


Figure 1. Adjacent gate constraints.

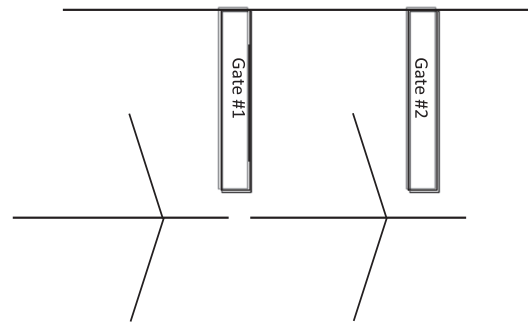


Figure 2. Last-in first-out constrained gates.

up a gate for other turns' use. So, it is only worthwhile to tow turns with a long duration, that is, a long turn time; say more than 2 hours or so. Further, every time an aircraft is towed away or towed into a gate, there is a cost associated with it. Thus, it is imperative to minimize this cost. A towing operation is illustrated in Figure 3.

Conceptually, a long turn that is towed out or towed in is broken down to two separate turns. An arriving flight is combined with a dummy outgoing flight after providing adequate time for passenger disembarkment. Similarly, a departing flight is combined with a dummy incoming flight, providing adequate boarding time.

- *Gate rest:* Gate rest is defined as the duration for which the gate is kept idle between a departing flight and the next arriving flight. The purpose of gate rest is to ensure that the gate plan remains fairly robust in the event of minor delays in the flight schedule. A gate rest of 10 minutes ensures that two successive flights are assigned the same gate if and only if the arrival time of the later flight is scheduled at least 10 minutes after the departure time of the former in the planning phase.
- *Preferred gates:* Although some gates are perceived to be favorable for certain turns, some others may not be perceived to be favorable. Certain pre-determined sets of conveniently located gates are preferred to be assigned to business markets and premium service flights. It is also preferred to assign international flights to international gates and domestic flights to domestic gates, even though it is possible to disembark and board passengers otherwise. Similarly, certain gates may be able to technically handle a particular type of aircraft, without being a preferred assignment. Assignment of turns to preferred gates are maximized.
- *Unassigned turns:* Given that airline companies have peak activities over a small window of period during morning and evening, it is possible that an airline does not have adequate number of gates for all the aircrafts parked on the ground. Under such situations, either some aircrafts are made to wait till some gate is freed or the airline company borrows a gate from its competitors. Both of these are usually not preferred. Whereas the first one impacts the customer satisfaction levels, the second one involves a certain cost and is subject to availability.

We have explained so far some of the main features and restrictions relevant to an airport GAP. This paper is organized in the following manner. In the next section, viz. section 2, we review some of the relevant literature on this topic. Section 3 describes the different objectives considered for our problem.

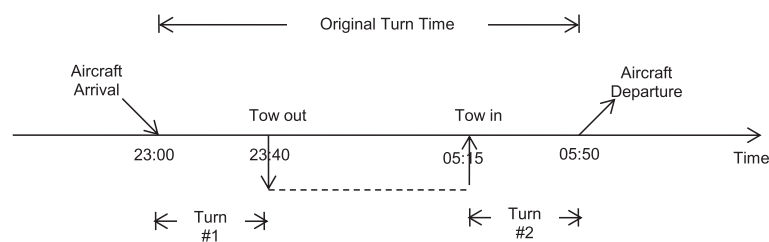


Figure 3. Towing representation.

We outline the existing gaps in the literature on this subject and describe our contribution in plugging those gaps. Section 4 describes the mathematical model developed by us to solve this problem, and section 5 illustrates the results of the model on a real-life example. Section 6 concludes the paper and provides some future directions for research on this topic.

## 2. LITERATURE REVIEW

Airport GAP, as a planning problem, has been extensively studied for three decades. Here, we will focus our survey to some of the recent work in this field, without, of course, omitting the pioneering research. Because airport GAP is a special case of generalized assignment problem with specific constraints, its complexity is similar. Mathematical modeling for this problem has also been generally inspired from the modeling techniques for assignment problem. One of the major classifications of the research on airport GAPs is along the lines of modeling methodology, viz. continuous time model and discrete time interval model. Dorndorf *et al.* [13] provide a survey of the state of the art on the airport gate assignment research.

One of the earliest papers reporting the assignment of gates with the objective of minimizing average passenger walking distance (for both departing and arriving flights) is modeled using a continuous time assignment model by Babic *et al.* [1]. The model assigns aircrafts to gates or stands and ensures that larger number of passengers walk less, while ensuring that all flights are assigned to a gate or stand. Mangoubi and Mathaisel [2] also present a continuous time gate assignment model that optimizes the average passenger walking distance into and out of the terminal. Their model, additionally, also looks at aircraft-gate compatibility and the connecting passengers to model the assignment of two flights to gates such that the distance between them is kept within a certain limit. However, the model emerges as a quadratic assignment model, which is linearized in a not very efficient manner. Thus, they employ Linear Programming (LP) relaxation and greedy heuristics to solve the problem.

Haghani and Chen [3] formulate a multiple time slot version of the GAP with the objective of passenger walking and baggage transport distance minimization as an integer program. The problem is formulated as a quadratic assignment problem and is solved using an iterative heuristic. Yan and Chang [4] formulated the airport gate assignment as a multi-commodity network flow problem. The objective of this model is to “flow” all the airplanes in each network, at a minimum cost, which is equivalent to the minimization of total passenger walking distance. An algorithm based on the Lagrangian relaxation, with subgradient methods, accompanied by a shortest path algorithm and a Lagrangian heuristic was developed to solve the problem. The model was tested using data from a Taiwanese airport.

Xu and Bailey [6] propose a tabu search algorithm for a continuous time airport GAP with the objective of minimizing the passenger walking distances, to reach the connecting flights. Given that the problem has a non-linear (quadratic) objective function, a simple tabu search meta-heuristic is used to solve the problem. The algorithm exploits the special properties of different types of neighborhood moves and creates effective candidate list strategies.

Recent research has been concentrating on robust gate assignment plans without focusing on walking time minimization criteria. Bolat [5] provides a model for robust gate assignment, which can be maintained during the real-time operations. This is carried out by maximizing the gap between a departing flight and the next arriving flight. However, the objective function is non-linear, and hence, the problem is solved using a heuristic. Yan *et al.* [8] introduce flexible buffer times to absorb stochastic delays in gate assignment operations. They propose a simulation framework that is not only able to analyze the effects of stochastic flight delays on static gate assignments but can also evaluate flexible buffer times and real-time gate assignment rules. Lim *et al.* [11] consider the more realistic situation where flight arrival and departure times can change. Although the objective is still to minimize walking distances (or travel time), the model considers time slots allotted to aircraft at gates deviate from scheduled slots within a time window. The solution approach uses insert and interval exchange moves together with a time shift algorithm. These neighborhood moves are used within a tabu search framework.

More recent research on this topic focuses on multiple objectives and other special ways of mathematical formulation. There have also been efforts to combine the problem in planning and operations

phase to develop stochastic models. Yan and Huo [7] formulate a dual objective 0–1 integer programming model for the aircraft-gate assignment. The first objective tries to minimize walking times for all passengers, whereas the second objective aims to minimize passenger waiting times in the event of the aircraft not being able to find a free gate. Ding *et al.* [9] consider the over-constrained GAP, which is described as a situation when there are too many flights for the available gates. They propose a 0–1 quadratic program model that minimizes the number of ungated turns and also minimizes the passenger walking distance. They use a greedy algorithm that minimizes the ungated flights, although a neighborhood search technique called the Interval Exchange Move allows flexibility in seeking good solutions within a tabu search framework.

Lim and Wang [10] attempt to accurately build an evaluation criteria for the ability of an aircraft-to-gate assignment to handle uncertainty on aircraft schedule, and to accurately and effectively search the most robust airport gate assignment. They develop a stochastic programming model and transform it into a binary programming model by introducing the unsupervised estimation functions without knowing any information on the real-time arrival and departure time of aircrafts in advance. A partition-based search space encoding, two neighborhood operators for single or multiple aircrafts reassignment, and a hybrid meta-heuristic combining a tabu search and a local search are implemented.

Yan and Tang [15] consider the GAP in the planning mode along with stochastic flight delays that occur in actual operations. They argue that it would be sub-optimal to handle the problem in planning and operations separately, without addressing the inter-relationship between these two stages. They suggest a heuristic approach to solve such a model that includes three components, a stochastic gate assignment model, a real-time assignment rule, and penalty adjustment methods. Diepen *et al.* [12] propose a set partitioning formulation involving modeling on a series of flights that are to be assigned to the same gate. This assignment is called a gate plan. A major advantage of this new formulation is that feasibility can be checked easily during the pre-processing stage. Furthermore, even cost calculation of a gate plan is also pre-processed. This is also one of the few papers that consider the adjacent gate restriction as observed at Schipol airport.

Dorndorf *et al.* [14] propose two methods to incorporate robustness into the gate assignment models through overlap methods and fuzzy sets. Dorndorf *et al.* [16] consider the multiple objectives of maximization of the total assignment preference score, minimization of the number of unassigned flights during overload periods, minimization of the number of tows, and also maximization of the robustness of the resulting schedule with respect to flight delays. However, they present a unique approach involving simple transformation of the flight–gate scheduling problem to a graph problem, that is, the clique partitioning problem. The algorithm used to solve the clique partitioning problem is a heuristic based on the ejection chain algorithm.

Drex1 and Nikulin [17] consider the multiple objectives of minimizing the number of ungated flights and the total passenger walking distances or connection times as well as maximization of the total gate assignment preferences. The problem is formulated as a quadratic assignment formulation and solved by Pareto simulated annealing to obtain a representative approximation for the Pareto front. Hu and Di Paolo [18] employ genetic algorithm to solve the multi-objective airport GAP.

To summarize, airport gate assignment in planning mode is an extensively researched topic over the last few decades. Although the early approaches considered one objective (usually minimization of passenger walking times) and formulated the problem as an integer or quadratic assignment formulation with continuous time slots, the researchers in late 1990s started to look beyond the continuous time formulation and proposed discrete time interval and network formulations. With the advent of 21st century, the problem has been formulated with a fresh perspective, such as set partitioning approach or a clique partitioning graph model, and the focus shifted to several other objectives that are commonly observed for this problem. The objectives considered for the problem have usually been to

- (1) minimize passenger walking times from (or to) the terminal and connecting flight gates,
- (2) minimize the number of ungated turns,
- (3) minimize the number or costs of towing procedures,
- (4) maximize (or minimize) the preference of certain turns to be assigned to favorable (or unfavorable) gates,



- (5) maximize the schedule robustness using different features such as inclusion of stochastic delays in the model, increase the idle time between two turns, or include time windows.

Airport gate assignment in operations mode is studied as a recovery problem in conjunction with aircraft, crew, and passenger recovery problems. However, operational objectives before recovery, such as minimizing gate plan deviations and maintaining schedule feasibility as specific objectives in the operations are not studied.

### 3. PROBLEM FEATURES AND ASSUMPTIONS

The airport GAP studied by us in this paper is inspired from a real-life case study at Chicago O'Hare, even though not all the restrictions reported in section 1 are applicable for this airport. It is indeed an interesting research challenge to study the different ways of formulating the airport GAP in the planning mode. However, it is surprising to note that one common thread that runs across the literature is the fact that the problem has rarely been solved using exact methods to optimality. In fact, we would go on to state that heuristic and meta-heuristic techniques have been "over-employed" to solve the problem. The probable reason is that the scale of the problem is often so large that it is not possible to solve it to reasonable or acceptable levels of optimality with the mathematical formulations themselves. Given that the airline company for which we solve the GAP is among the top 3 in terms of passenger traffic in the world and Chicago O'Hare is one of its busiest hubs, we want to show that simple, continuous time assignment model with a wide range of multiple objectives can produce high quality solutions in reasonable computing times.

Some of the considerations observed at actual airports are not reported in the literature. Some features unique to the GAP studied by us include LIFO gates and towing constraints. These constraints have never been explicitly modeled for either continuous time assignment models or discrete time slot assignment model in any of the papers studied by us. One of the reasons may be that most of the solution procedures have eventually relied on heuristics and they do not see the need for modeling these features as additional constraints. Diepen *et al.* [12] do consider the adjacent gate constraint, but the constraint is pre-processed in their set partitioning formulation. One of our major contributions in this paper is to conceptually model all these physical constraints, viz. adjacent gate constraints, LIFO gate constraints, and towing constraints, into logical mathematical ones.

Walking time minimization is one of the earliest objectives considered for the GAP in the literature. This problem can be easily modeled as a quadratic programming model. Although some of the available literature attempts to linearize the non-linear model, most of the other papers solve the problem using heuristics—where the non-linear objective or constraints really do not matter. We believe that walking time minimization is a realistic objective for the airport GAP; it, however, is not perceived as a priority from the business point of view. Most airline companies like to know how walking time criteria can really impact their bottom line apart from perceived customer satisfaction. In this context, the walking time criteria are really the most important criteria for connecting passengers—with especially short connection times.

A passenger with 2-hour connecting time would not bother to walk a little more distance to her gate (which also gives an opportunity for shopping). However, a passenger with 35-minute connection time would be particularly bothered as she has to disembark, walk to the connecting flight gate, and board the flight within this short time. In particular, airline also realizes that if the passenger misses the connection, they may not only have to make alternative arrangements for the passenger but also not realize the complete revenue until the journey is completed. Such connections are identified as "connections at risk." and it is important for the airline to make suitable arrangements—such as making these passengers sit on one of the front rows to enable faster disembarking, assigning the connecting flight to a "nearby" gate, and so on. In this context, our work differs from the other research works in this area—even though the underlying model is conceptually the same. Our first objective is not to minimize the average walking distances but to maximize the connection realizations through our optimal gate assignment model, which we consider as a fresh contribution on this topic.

The second objective considered by us is to limit the number of zone usages to minimum when the hub activity is thin. Although it is difficult to reduce the number of zones during the day times, it is

however possible to limit the number of zones during the night times. This objective has not been considered in the airport gate assignment so far in the literature, and it is indeed one of the contributions of our paper.

Robust scheduling has been adequately addressed in the prior works as one of the objectives of the airport gate assignment model. Robustness, as a measure, can have different definitions. In our paper, we already provide some gate rest between a departing turn and the next arriving turn. This is also referred to as idle time in the literature. The purpose of this gate rest is to absorb small delays in the outgoing or incoming aircrafts. Another purpose of providing gate rest is also to ensure safety by providing a reasonable separation. This would ensure that there is sufficient gap between departing and arriving aircraft to minimize the possibility of any accidents. Usually, standard gate rest is provided for all turns depending on the type of aircraft equipment. We propose to increase the robustness of our planned schedule by increasing the gate rest by accounting for the past history of delays. The method followed is quite simple and intuitive. We note the past delay patterns for every flight. We choose the  $k$ th percentile delay of the historical delay in minutes (say,  $k$  is 95th percentile of past 300 days of delays) for every turn and attempt to add the same to the gate rest corresponding to that turn. Because the delays are themselves calculated at flight levels, we choose the maximum of the two delays-arriving and departing flight of a turn-to calculate the percentile delay. The key word here is “attempt to” because it may not be feasible to provide such a gate rest. For every minute of violation of this additional gate rest, there would be an associated penalty. Although idle time maximization is one of the measures used extensively in the literature to increase the schedule robustness, we feel that our measure is far more effective as it tries to maximize the time between turns to the turns where it is needed the most-instead of adding time to all the turns. We feel that this is another major contribution of this paper.

It is worth highlighting that the recovery procedures considered for the airport GAP in the literature largely focus on the joint schedule recovery for the aircrafts, crew, and passengers. Although it is best to study the recovery models encompassing all functions of the airline business, it is often elaborate and time consuming and helps little at the actual time of “fire-fighting” on the ground when a number of aircrafts arrive with large delays. Before the schedule is recovered, it is imperative that the arriving and departing aircrafts are provided airport gates subject to the given set of operational constraints. This period between the disruption and schedule recovery is handled by ground operations and calls for certain robust and quick gate assignment models. In operations mode, the planning objectives that relate to the profitability are ignored. The primary focus is on schedule feasibility and ensuring that there is minimal further disruption to the flight schedule. The second important feature of operations model is the run time. Small run times are ideal, and it is critical that the operations mode models indeed have very small run times, say, a few seconds. It is not possible to wait, say, even for half an hour, for the model to produce an optimal output because the ground staff literally fights against time while managing disrupted flights.

In this context, it is relevant to note that operations models in gate optimization have two main objectives. The first basic objective is to minimize the deviation from the planned schedule. In the event of delayed arrival and subsequent departure of a large number of turns, it is imperative that the assignment to the originally planned gates may no longer be possible. Given the fresh arrival and departure times for the turns, the first objective aims to minimize the penalty due to reassignment of gates (re-gating). In addition to allowing flight re-gating, the operations model also aims to maintain a similar number of departures and arrivals for every zone for given pre-fixed time intervals. Any deviation in the number of aircrafts in a particular zone on ground with respect to the planned schedule is penalized. All the physical and logical constraints of the airport as mentioned in section 1 would be applicable in the operations problem as well. Further, the operations model with this objective should be capable of handling constraints relating to the following: (i) specific turns that should not be re-gated and/or (ii) specific turns that are allowed to be re-gated.

While running the gate assignment model in the operations mode with the previously mentioned objective, it is quite likely that a scenario emerges when the number of aircrafts on ground is actually more than the available gates. This could be a result of several delayed arrivals or departures piling up just around the hub peak time. Thus, the second objective in the operations mode deals with maintaining the schedule feasibility by retiming some of the flights to a later time. Care is taken to

ensure that the extent of retiming is limited for an individual turn while ensuring that such re-timings are minimal and heavily penalized.

So far, we mentioned different objectives and constraints considered in our version of the airport GAP. We have also highlighted the distinctive features in our model, which may be fresh contributions on this topic. However, there are also some inherent assumptions and limitations in our model that are described as follows:

- Connection revenue is realized only if the passenger is able to disembark, walk between the gates, and board before the departure of connecting flights. This is a fair assumption and a perfect synopsis of the real instance. However, it does not consider any additional costs that would be borne if the passenger misses the connection.
- Disembarking time, walking time between different gates at the airport, and boarding times are provided as point estimate inputs. This is a strong assumption because different passengers may actually take different amounts of time for the same sequence of activities. It is especially true for passengers on wheel chairs and passengers traveling with families. However, the ensuing model would become extremely complicated if pedestrian behavior is to be included in it.
- Connection revenue is provided as point estimate inputs. This is again a strong assumption as the connection revenue might tend to change on a day-to-day basis. The estimates used in the model would be based on average values over a fairly long period because it would not be possible to change gates for turns on a daily basis.
- For schedule robustness, gate rest accounts for minimum gate rest and a certain percentile delay of the flights in the turns. Although this increases the schedule robustness by some extent, it is certainly not the best way. It would be ideal to model the stochastic flight delays into the planning model, but the resulting stochastic mixed integer program would be too complex to handle.
- For the zone minimization objective, it is conveniently assumed that the non-productive time for the employees to travel within the zone would be same for all the gates within the zone and would necessarily be less than the travel time to cross the zone. This is a reasonable assumption based on the actual layout of the airport where the zones are fairly spread out and inter-zonal distances are usually much more than the intra-zonal distances.
- Some of the papers in the literature allow for aircraft waiting in the times of congestion for the planning problem. However, this has not been taken into account for our problem as per the wishes of the airline company. It is fairly reasonable because the gate assignment, at least during the planning stage, should not plan for aircraft waiting. This might eventually happen during the real-time operations (when retiming certain turns beyond their actual arrival), but it would indeed be a bad plan to allow aircraft waiting in the absence of gates. In case a gate plan is highly infeasible, the airport operations team would work with the flight scheduling team to move some of the flights to non-peak periods while negotiating for more gates with the airport authority and the competitors.

We now describe the mathematical model, which is a 0–1 mixed integer program that helps us produce a feasible gate plan in the light of all the previous business constraints.

#### 4. MATHEMATICAL MODEL

In this paper, we first consider the gate assignment for planning mode where cost minimization and revenue maximization are major optimization criteria as opposed to feasibility of solutions or walking times for passengers. In the planning mode, flight schedule and gate plan are used to arrive at a gate assignment schedule while ensuring that the business constraints are satisfied and the objective function obtains an optimal value. We develop a basic 0–1 integer program mathematical model formulation with linear objective function and constraints that would assign one gate to every flight and ensure that all business constraints are satisfied.

We would focus on cost minimization as the objective of this model. The different cost components in the model relate to the cost of providing unfavorable gates to preferred flights, the cost of towing, and the cost of an ungated turn. In this model, we would also describe all business constraints associated with the problem. Additional objectives of connection revenue maximization and minimization of zone usage cost and robust gate plan would be described later.



Although most of the parameters and decision variables in the problem are described with the first objective, specific parameters and decision variables corresponding to the other objectives will be described later.

The following are the data sets for the turn schedule and the airport.

*Sets:*

<i>TURNS</i> :	set of turns to be gated represented as $i$ , or $j$
<i>LTURNS</i> :	set of long turns for which towing is allowed, represented as $t$ , $LTURNS \subset TURNS$
<i>GATES</i> :	set of gates represented as $k$ or $l$
<i>ADJACENT</i> :	set of adjacent gate pairs that have the adjacent gate restriction represented as $(k, l)$
<i>LIFO</i> :	set of last-in first-out gate pairs represented as $(k^F, l^R)$ to distinguish front and rear gates
$(i_1,$	$i_2)$ : new turns arising out of a towed turn $t$
$E_k$ :	set of equipment (aircraft) types that gate $k$ can handle, $k \in GATES$
$E_k^1,$	$E_l^1$ : sets of equipment (aircraft) types in $E_k^1$ is occupying $k$ , no aircraft of any type in $E_l^1$ may use gate $l$ , where $(k, l) \in ADJACENT$ and vice versa.

*Parameters:*

$\alpha$ :	minimum gate rest
$C_1$ :	the actual cost of towing an aircraft
$C_2$ :	the cost of not assigning a gate to a turn, determined by cost of borrowing a gate from a competing airline
$a_i$ :	planned arrival time of turn $i \in TURNS$
$b_i$ :	planned departure time of turn $i \in TURNS$
$C_{ik}$ :	notional cost of assigning turn $i$ to gate $k$
$e_i$ :	equipment (aircraft) type used by turn $i \in TURNS$

*Decision Variables:*

$x_{ik} \in$	$\{0, 1\}$ : 1 if turn $i$ is assigned to gate $k$ ; 0 otherwise
$y_i \in$	$\{0, 1\}$ : 1 if turn $i$ is not assigned to any gate; 0 otherwise
$w_t \in$	$\{0, 1\}$ : 1 if long turn $t$ is towed; 0 otherwise

**Objective Function:** We start by minimizing all the cost elements in our model—which in this case are the cost of assigning unfavorable gates to certain preferred, premium service turns; cost of towing a turn; and the cost of not assigning a turn to a gate (ungated turn). It may be noted that the costs of assigning a gate are symbolic based on the perceived importance of certain flights; the cost of towing and the cost of ungated turn are on an actual basis.  $C_{ik}$  represents how unfavorable would be the assignment of turn  $i$  to gate  $k$ . This coefficient, usually positive, is affected by a number of business and operational preferences:

- Some predefined sets of conveniently located gates are preferred for turns that contain the top business market flights and premium services flights.
- Some international terminal gates are capable of accommodating domestic arrivals, but they are less preferred than gates at domestic terminals.
- Some gates are less preferred for some fleet types because of gate features.

$$\text{Minimize } \sum_{i \in T} \sum_{k \in K} C_{ik} x_{ik} + C_1 \sum_{t \in L} w_t + C_2 \sum_{i \in T} y_i \quad (1a)$$

*Constraints:* A turn has to be assigned to exactly one gate or none. It is also required that the assigned gate is capable of handling the aircraft associated with the turn. This is modeled as

$$\sum_{k \in GATES: e_i \in E_k} x_{ik} + y_i = 1, \quad i \in TURNS \quad (2)$$

It is possible that the airline company wants a particular turn  $i'$  to be only assigned to a certain gate  $k$ . This can be modeled as

$$\sum_{i' \in TURNS} x_{i'k} = 1, \quad k \in GATES \quad (3)$$

It is possible that the airline does NOT want a particular turn  $j'$  to be assigned to a certain gate  $l$ . This can be modeled as

$$\sum_{j' \in TURNS} x_{j'l} = 0, \quad k \in GATES \quad (4)$$

There cannot be two turns on the same gate at the same time, including the gate rest time after the turn has departed. This is referred to as overlap and can be modeled as

$$x_{ik} + x_{jk} \leq 1, \quad i, j \in TURNS; k \in GATES : a_i < a_j \wedge a_i < b_j + \alpha \wedge a_j < b_i + \alpha, i \neq j \quad (5)$$

Two adjacent gates cannot handle certain types of aircraft types simultaneously. This can be logically modeled as

$$x_{ik} + x_{jl} \leq 1, \quad i, j \in TURNS; k, l \in GATES; (k, l) \in ADJACENT : a_i < a_j \wedge a_i < b_j \\ \wedge a_j < b_i, i \neq j; e_i \in E_k^1; e_j \in E_l^1 \quad (6)$$

The following two constraints ensure that no aircraft can exit the front gate as long as there is an aircraft in the rear gate and that no aircraft can enter the front gate as long as the aircraft in the rear gate has not departed.

$$x_{ik^F} + x_{jl^R} \leq 1, \quad i, j \in TURNS; (k^F, l^R) \in LIFO : a_j \leq a_i \wedge a_i \leq b_j, i \neq j \quad (7)$$

$$x_{ik^F} + x_{jl^R} \leq 1, \quad i, j \in TURNS; (k^F, l^R) \in LIFO : a_j \leq b_i \wedge b_i \leq b_j, i \neq j \quad (8)$$

We now introduce the following constraints to represent the towing of a long turn. Note that long turn  $t$  is broken down into two possible half turns  $i_1$  and  $i_2$ , such that arrival time of  $i_1$  is same as the arrival time of  $t$  and departure time of  $i_2$  is same as the departure time of  $t$ . The first constraint ensures that the long turn is split only if the option of long turn is chosen by paying the towing costs and the two half turns are assigned to different gates. The next constraint ensures that there is no

overlapping for the two half turns arising out of breaking a long turn. The last constraint ensures that there are no adjacent gate limitations for the half turns.

$$x_{i_1k} - x_{i_2k} \leq w_t, \quad i_1, i_2 \in \text{TURNS}; t \in \text{LTURNS}; k \in \text{GATES} : i_1 \neq i_2 \quad (11)$$

$$x_{i_1k} + x_{j\ell} - 1 \leq w_t, \quad i_1, j \in \text{TURNS}; t \in \text{LTURNS}; k \in \text{GATES} : \\ i_1 \neq j \wedge j \neq i_2 \wedge b_{i_1} + \alpha < a_j \wedge b_j + \alpha < a_{i_2} \quad (12)$$

$$x_{i_1k} + x_{j\ell} - 1 \leq w_t, \quad i_1, j \in \text{TURNS}; t \in \text{LTURNS}; (k, \ell) \in \text{ADJACENT} : \\ i_1 \neq j \wedge j \neq i_2 \wedge b_{i_1} + \alpha < a_j \wedge b_j + \alpha < a_{i_2}, e_{i_1} \in E_k^1, e_j \in E_\ell^1 \quad (13)$$

This model with objective as (1a) and constraints (2)–(13) would minimize the overall operational costs while satisfying all business-related constraints. This mathematical formulation solves the gate assignment model for all the complex constraints, although the objective function is kept trivial. Although the model can be enriched with the introduction of additional objectives, the constraint sets (2)–(13) would always be needed to produce feasible solutions. As a result, we will refer to this formulation as base model in the rest of the paper.

#### 4.1. Planning mode

The previous model determines least cost feasible gate assignment that satisfies all the logical and business constraints. However, this feasible assignment is not sufficient in real life where decision-makers are often faced with several additional multiple objectives that include maximization of passenger connection revenue, minimization of zone usage cost, and maximization of schedule robustness. We will now present the different objectives that are considered by us in the planning mode of the airport GAP, starting the objective of revenue maximization.

##### 4.1.1. Maximization of passenger connection revenue

In this section, we will extend the base model to maximize the passenger revenues by optimizing the connection time for “connections at risk.”

Although there are stipulated minimum and maximum connection times, it must be noted that flights gated at distant gates could possibly result in misconnections if the connection time is fairly tight. For a passenger who transfers to a connection flight at the airport, the connection time is readily defined as the walking time required from the arrival gate of her incoming flight to the departure gate of her outgoing flight. For the planning model, the arrival time of the incoming flight and the departure time of the outgoing flight are assumed to be fixed as per the published flight schedule. Each flight must be assigned to exactly one gate, and there should be sufficient time for passengers boarding at the gate.

When building a 0–1 integer program formulation, one of the key issues is the choice of decision variable. We consider the gating plan of an incoming flight connection as well as an outgoing flight connection rolled into one variable. Thus, for a flight schedule with 800 flights and 100 gates, the worst case scenario could result in 6.4bn 0–1 variables. Fortunately, every flight does not always present a connection opportunity to passenger with every other flight. Incidentally, a flight can potentially connect to barely 20 other flights, and, in most cases, the connection time is often more than the longest walking time between two gates at the airport, which means that there are few “connections at risk.”

We now present an extension to the model given previously (1a)–(13) to incorporate this objective in the mathematical formulation. As before, we will first define the additional sets and parameters and then the additional decision variables, objective function, and constraints.

*Sets:*

CNX: set of all revenue connections involving turns  $i$  and  $j$ , that is,  $(i,j)$

*Parameters:*

$REVENUE_{ij}$ : revenue generated by connecting turn  $i$  to turn  $j$ . Note that  $(i,j) \in CNX$

$Walk(k,l)$ : wholesome walking time including boarding, de-boarding, and other components of time to move from gate  $k$  to gate  $l$

*Decision Variables:*

$z_{ijkl}$ : 1 if turn  $i$  is assigned to gate  $k$  and turn  $j$  is assigned to gate  $l$  and  $(i,j) \in CNX$ , 0 otherwise

*Mathematical Model:*

Maximize

$$\sum_{(i,j) \in CNX} \sum_{k \in GATES} \sum_{l \in GATES} REVENUE_{ij} * z_{ijkl} \quad (1b)$$

$$\sum_k z_{ijkl} \leq \sum_k x_{ik} \quad \forall (i,j) \in CNX; i,j \in TURNS; k,l \in GATES \quad (15)$$

$$\sum_l z_{ijkl} \leq \sum_l x_{jl} \quad \forall (i,j) \in CNX; i,j \in TURNS; k,l \in GATES \quad (16)$$

$$x_{ik} + x_{jl} - 1 \leq z_{ijkl} \quad \forall (i,j) \in CNX, i,j \in TURNS, k,l \in GATES : a_i + Walk(k,l) \leq b_j \quad (17)$$

$$z_{ijkl} = 0 \quad \forall (i,j) \in CNX, i,j \in TURNS, k,l \in GATES : a_i + Walk(k,l) > b_j \quad (18)$$

Constraint (1b) is the objective function addition to the existing model to capture the revenue maximization. We maximize the overall revenue by realizing the connection revenue between flights, which are components of turns  $i$  and  $j$ . Revenue computation would be carried out carefully by looking at the possibility of connection from  $i$  to  $j$ .

Constraints (15) and (16) are upper bounds that ensure  $z_{ijkl}$  would not take a value of 1 unless both turns  $i$  and  $j$  are assigned to gates  $k$  and  $l$ , respectively. Note that this is a necessary condition for this variable to take a value 1, but by no means sufficient. Constraint (17) is a lower bound for this variable. This variable  $z_{ijkl}$  is created for only “select” valid connections between turns  $i$  and  $j$ . By using “select” connections, we refer to those connections where the connection time is greater than the minimum time for a passenger to de-board, walk, and board

another connection whereas less than the maximum time for a passenger to de-board, walk, and board another connection for a given hub airport. For instance, the maximum time for a passenger to deplane, walk, and board another flight at this hub is 41 minutes for domestic connections irrespective of the aircraft type. It is not difficult to prove that we do not sacrifice optimality by such an assumption because any connection with connection time less than minimum walking time would never materialize even if the corresponding equipments are gated at closest gates. Similarly, there would be no adverse impact on the revenue by assigning connecting turns to farthest gates (or any other combination of gates) if the total connection time is greater than the maximum walking time between two gates at the airport.

Constraint (18) would ensure that the decision variable  $z_{ijkl}$  takes a value of 1 if and only if walking times between flights involved in turns  $i$  and  $j$  are less than the difference between departure and arrival flights in turns  $j$  and  $i$ , respectively.

Note that these features and constraints are only addition to the existing model. There is no change in the basic framework of the old model, and the fresh addition only introduces several new variables and constraints that would help us solve the objective of maximizing connection revenue.

#### 4.1.2. Minimization of zone usage cost

In this section, we consider the next objective of minimization of zone usage. There are certain time intervals during the day when the number of flights on the ground at a station is very less. At such instances, it is preferable to ensure that all flights to be serviced are restricted to a limited number of zones so that employees do not waste time walking between different zones. This can be handled as an extension to the model given previously. Let us again start with introduction additional sets, parameters, decision variables, objective function, and constraints to the model.

##### Sets:

ZONE: set of all zones,  $n$ , such that  $k \in n$

SHIFT: set of all possible shifts  $s$  with a given shift begin,  $s_b$  and end time,  $s_e$ —for example, 00:00–4:00 AM, 01:00–5:00 AM, . . . , 11:00 PM–3:00 AM, assuming a minimum shift of 4 hours.

This model does not assume any predefined shift timings and would also help the airline company realign the shift timings of the employees to extract maximum benefits.

##### Parameters:

$Z\_PENALTY$ : notional penalty for using a zone

$Big\_M$ : a large number (a number larger than the number of gates in all zones)

##### Decision Variables:

$m_{ns}$ : 0–1 binary variable to indicate if a zone  $n$  is utilized in shift  $s$

##### Mathematical Model:

Minimize

$$Z\_PENALTY \sum_n \sum_s m_{ns} \quad (1c)$$



Subject to

$$\sum_{k \in n} \sum_i x_{ik} \leq BIG_M \times m_{ns} \quad \forall s, n : (s_b \leq b_i \leq s_e \vee s_b \leq a_i \leq s_e) \quad (19)$$

This objective ensures that the zone utilization over any shift time is penalized whereas the constraint checks if a turn is scheduled in a zone for a particular shift. Needless to say, the zone minimization models would optimize objective function (1c) with constraint (19) in addition to constraints (2)–(13). The zone minimization model can also be run in conjunction with revenue maximization model, in which case, all the constraints (2)–(19) will be part of the optimization model. The objective function would minimize (1a)–(1b)+(1c) in the joint model. Let us now consider the third objective of maximization of gate plan robustness.

#### 4.1.3. Maximization of gate plan robustness

Gate rest is a concept that is utilized to improve gate plan robustness. The gate assignment plan should consider different minimum gate rest characteristics for different types of aircrafts and flight sectors on which the turns operate. It also dynamically considers gate rest, given expected inbound arrival delays, or other operational characteristics. It should be capable of handling, without excessive disruption, the propagation involved in a typical “out of service” aircraft problem. By incorporating these criteria in the GAP, a robust gate plan can be handed off to airport staff to better manage the gate plan during operations.

We use the same existing model to optimize gating robustness with a minor change. We create a delay variable, *delay*, for a particular turn to include the extent of flight delay in both arrival and departure. *delay* for a turn *i* is computed using historical arrival delay and departure delay data as

$$d_i = \text{Max}(\text{kth percentile Arrival Delay}_i, \text{kth percentile Departure Delay}_i)$$

We use the following additional variable to the existing model, which captures the amount of gate rest violation from the desired amount, and try to minimize the same.

*Parameters:*

$d_i$ :	delay factor for turn <i>i</i>
$\beta$ :	desired gate rest ( $\beta \geq \alpha$ )
$GR\_PENALTY$ :	per minute penalty (notional) for violating gate rest

*Decision Variable:*

$g_i$ : gate rest violation in minutes for turn *i*

*Mathematical Model:*

Minimize

$$GR_{PENALTY} \sum_{i \in TURNS} g_i \quad (1d)$$

Subject to

$$(x_{ik} + x_{jk} - 1) \cdot (b_i + \beta + d_i - a_j) \leq g_i \quad \forall i, j \in Turns \quad (20)$$

The gate rest maximization model would optimize objective function (1d) with constraint (20) in addition to constraints (2)–(13). This completes the mathematical model of all the three objectives described for the planning mode in this paper.

#### 4.2. Operations mode

Now let us consider the mathematical models for the operations mode. As opposed to the planning mode, the primary optimization criterion in the operations mode is to maintain feasibility and minimize the impact of the disruption. This is achieved by minimizing the extent of schedule deviations from the planned gating solution and then retiming a limited number of turns that could not be gated after the application of the first objective.

##### 4.2.1. Minimization of schedule deviations

Let us now focus on the model to minimize schedule deviations. The primary objective of this mathematical model is to provide real-time gating solution based on the actual arrival and departure time of the turns. Some delays cannot be absorbed by gate rest provided in the planning model. If there are several large delays, it could potentially result in piling up of flights during a peak hour at the hub airport. Gating of delayed flights in addition to the planned schedule puts a strain on the resources. To accommodate these delays, some flights are re-gated to ensure that minimal flights remain without a gate at the end of the exercise. In addition to re-gating, this model should aim to minimize any zone-wise deviations in terms of the number of arrivals and departures based on the planned schedule. We propose the following mathematical model to take care of this problem:

*Parameters:*

$plan\_gate_{ik}$ :	1 if turn $i$ is assigned to gate $k$ at planning stage, 0 otherwise
$reg\_status_i$ :	0 if turn $i$ is not allowed to be re-gated, 1 otherwise
$REGATE\_P$ :	penalty for re-gating a turn $i$
$ZONE\_P$ :	penalty for one variation in the sum total turns serviced in a zone
$sch\_a_i$ :	scheduled arrival time of turn $i$
$sch\_b_i$ :	scheduled departure time of turn $i$

*Decision Variables:*

The decision variables  $x_{ik}$  and  $y_i$  have the same relevance as used in the planning stage, but operations problem is re-run for new/actual arrival and departure times for the turns.

$x_{ik} \in$	$\{0,1\}$ : 1 if turn $i$ is assigned to gate $k$ ; 0 otherwise
$y_i \in$	$\{0,1\}$ : 1 if turn $i$ is not assigned to any gate; 0 otherwise
$y_i \in$	$\{0,1\}$ : 1 if turn $i$ is re-assigned to a different gate; 0 otherwise
$z_{ns}$ :	an unrestricted integer for all zones, $n$ , and all shift periods, $s$
$u_{ns}$ :	non-negative integer for all zones, $n$ , and all shift periods, $s$
$v_{ns}$ :	non-negative integer for all zones, $n$ , and all shift periods, $s$

*Objective Function:*

Operations mode would have its objective function, and as mentioned before, the stress is really on operational feasibility, rather than optimization. The first two terms in the objective function aim to minimize the deviation of the new schedule with respect to the planned schedule. The notional costs

corresponding to re-gating and zone deviation penalties have little significance, as long as these two terms of non-negative reflect the relative importance of the two deviations for the operations personnel. The third term in the objective has the same purpose as in the planning, that is, to ensure as many turns are gated as possible.

$$\text{Minimize } REGATE_P \sum_i y_i + ZONE_P \sum_n \sum_s (u'_{ns} + v'_{ns}) + C_2 \cdot \sum_i y_i \quad (21)$$

*Constraints:*

The logical constraints among the set (2)–(13) would be equally applicable for the problem in the operations mode and hence would not be ignored. Some constraints, such as (3) and (4) enforced by the planners may be ignored, if the airport operations manager so decides. For these constraints, the arrival and departure times of the turns,  $a_i$  and  $b_i$ , would be replaced by the actual arrival and departure times,  $sch\_a_i$  and  $sch\_b_i$ . In addition, the following constraints would also be required.

$$x_{ik} + y_i = 1 - y_i \quad \forall i \in TURNS, k \in GATES : plan_{gat}e_{ik} = 1 \wedge reg_{statu}s_i = 1 \quad (22)$$

This constraint ensures that either a turn is re-gated or a turn remains ungated.

$$x_{ik} = plan_{gat}e_{ij} \quad \forall i \in TURNS, k \in GATES : reg_{statu}s_i = 0 \quad (23)$$

This constraint ensures that planned gate schedule cannot be changed, if not allowed.

$$y_i = 0 \quad \forall i \in TURNS : reg_{statu}s_i = 0 \quad (24)$$

This constraint ensures that if a turn is assigned to its planned gate, it cannot be re-gated.

$$\sum_i \sum_k (plan_{gat}e_{ik} - x_{ik}) = z'_{ns} \quad \forall n, k \in n; i, s : (s_b \leq sch_{b_i} \leq s_e \vee s_b \leq sch_{a_i} \leq s_e) \quad (25)$$

This constraint captures the difference between planned gating and the new gating for every zone, by every shift. Each of the  $z'_{ns}$  variable can be negative or positive.

$$z'_{ns} + u'_{ns} \geq 0 \quad \forall n, s \quad (26)$$

$$z'_{ns} \leq v'_{ns} \quad \forall n, s \quad (27)$$

These two constraints ensure that the positive value of  $z'_{ns}$  variable is captured in one of them, whereas the other one remains 0—because of the minimizing objective.

The model with objective (21) and constraints (2)–(13) and (22)–(27) completes the gate assignment model in operations mode. Given that the basic model, without any specific profitability objectives, (1b), (1c), or (1d) is solved within a few seconds, it is expected that this model too would be solved reasonably fast. This is confirmed in the results in section 5.

#### 4.2.2. Maintain schedule feasibility by minimum retiming

Clearly, the previous model suffices as long as all the delayed turns are capable of being assigned to a gate. However, there is one situation when this model would not be adequate. If the disruption causes the schedule to pile up too many turns at the airport during the peak hours, it would be impossible to gate all these turns. If there are ungated turns (nonzero  $y_i$ ) in planning stage, it is handled by either

rescheduling some flights or leasing out new gates at the airport. In operations stage, the only possible solution is to delay the assignment of gate to the turn beyond the peak hour, while making it wait at the tarmac. Thus, the arrival and departure times of certain flights would be retimed (delayed) to ensure that all turns are eventually assigned to some gate. We now describe the associated model.

*Sets:*

**TURNS:** Here, it does not refer to all turns as in the previous models but only the set of turns that are candidates for retiming. We choose all turns that could not be assigned a gate on the basis of the previous operations model and all other turns, which are on ground after the arrival or before the departure of these turns.

*Parameters:*

$\alpha$ : minimum gate rest  
**RETIME\_P:** penalty for retiming a turn  $i$  by 1 minute (separately applied for both arrival and departure)  
**MGT<sub>i</sub>:** minimum ground time (or the time that a turn needs to spend on the ground for the passengers to disembark and perform aircraft cleaning and boarding)

*Decision Variables:*

The decision variables  $x_{ik}$  and  $y_i$  have the same relevance as used in the planning stage, but operations problem is re-run for new/scheduled arrival and departure times for the turns.

$pa_i$ : extent by which turn  $i$  is made to wait for gate after its scheduled arrival  
 $pb_i$ : extent of departure delay of turn  $i$  beyond its scheduled departure

It may be noted that the scheduled departure or arrival can only be delayed but not advanced. Thus,  $pa_i$  and  $pb_i$  would be non-negative.

$act\_a_i$ : actual arrival time of turn  $i$  ( $= sch\_a_i + pa_i$ )  
 $act\_b_i$ : actual departure time of turn  $i$  ( $= sch\_b_i + pb_i$ )

Note that the constraints (2)–(13) would no longer be applicable in the same format as it is in the planning model or the previous operations model because the arrival and departure times themselves are variables in this new model. We would not have the variable  $y_i$  as all turns have to be assigned to a gate through retiming. We also need several auxiliary variables to capture overlap, adjacency, and LIFO conditions.

$x_{ik} \in \{0,1\}$ : 1 if turn  $i$  is assigned to gate  $k$ ; 0 otherwise  
 $olp_{ij} \in \{0,1\}$ : an auxiliary variable to track overlaps between turns  $i$  and  $j$   
 $olp_i \in \{0,1\}$ : an auxiliary variable to track overlaps between turns  $i$  and  $j$   
 $adj_{ij} \in \{0,1\}$ : an auxiliary variable to track adjacency between turns  $i$  and  $j$   
 $adj_i \in \{0,1\}$ : an auxiliary variable to track adjacency between turns  $i$  and  $j$   
 $lfi_{ij} \in \{0,1\}$ : an auxiliary variable to track LIFO constraint between turns  $i$  and  $j$   
 $lfi_i \in \{0,1\}$ : an auxiliary variable to track LIFO constraint between turns  $i$  and  $j$

*Objective Function:*

The primary objective of the retiming model is to minimize the extent of retiming beyond the scheduled times of arrival and departures. Scheduled times are indicative of the already delayed arrival and departure times due to disruption.

$$\text{Minimize } RETIME_P \sum_i (pa_i + pb_i) \quad (28)$$

*Constraints:*

As already mentioned, constraint sets (2)–(13) would no longer be applicable. Let us reformulate the constraint for overlap in this model.

$$act_{a_i} \leq act_{b_j} + M \cdot (1 - olp_{ij}) + \alpha \quad \forall i, j \in TURNS \quad (29)$$

$$act_{a_i} \geq act_{b_j} - M \cdot olp_{ij} + \alpha \quad \forall i, j \in TURNS \quad (30)$$

$$act_{b_i} + \alpha \geq act_{a_j} - M \cdot (1 - olp'_{ij}) \quad \forall i, j \in TURNS \quad (31)$$

$$act_{b_i} + \alpha \leq act_{a_j} + M \cdot olp'_{ij} \quad \forall i, j \in TURNS \quad (32)$$

$$\sum_{k \in GATES: e_i \in E_k} x_{ik} = 1 \quad i \in TURNS \quad (33)$$

$$x_{ik} + x_{jk} + olp_{ij} + olp'_{ij} \leq 3 \quad \forall i, j \in TURNS; k \in GATES \quad (34)$$

$$act_{b_i} = sch_{b_i} + pb_i \quad \forall i \in TURNS \quad (35)$$

$$act_{a_i} = sch_{a_i} + pa_i \quad \forall i \in TURNS \quad (36)$$

$$act_{b_i} - act_{a_i} \geq MGT_i \quad \forall i \in TURNS \quad (37)$$

These constraints, (29) to (34), ensure that no two turns can be assigned the same gate if there is an overlap. The first four constraints ensure that at least one of the  $olp$  and  $olp'$  variables is 1, with  $olp = 1$  implying that turn  $i$  departs before turn  $j$  and  $olp' = 1$  implying that turn  $j$  departs before turn  $i$ . When both  $olp$  and  $olp'$  are 1, the turns  $i$  and  $j$  cannot be assigned the same gate.

Constraints (35) and (36) ensure that actual arrival and departure can only be delayed beyond the scheduled arrival or departure. Constraint (37) ensures that a delayed arrival stays on the ground at least to satisfy its minimum ground time requirement.

Now, we would model the adjacency constraints, which are logically written in the same way as overlap constraint previously.

$$act_{a_i} \leq act_{b_j} + M \cdot (1 - adj_{ij}) \quad \forall i, j \in TURNS \quad (38)$$



$$act_{a_i} \geq act_{b_j} - M \cdot adj_{ij} \quad \forall i, j \in TURNS \quad (39)$$

$$act_{b_i} \geq act_{a_j} - M \cdot (1 - adj'_{ij}) \quad \forall i, j \in TURNS \quad (40)$$

$$act_{b_i} \leq act_{a_j} + M \cdot adj'_{ij} \quad \forall i, j \in TURNS \quad (41)$$

$$x_{ik} + x_{jl} + adj_{ij} + adj'_{ij} \leq 3 \quad \forall i, j \in TURNS; k, l \in GATES; (k, l) \in ADJACENT : i \neq j; e_i \in E_k^1; e_j \in E_l^1 \quad (42)$$

Constraints (38)–(42) ensure that at least one of  $adj$  and  $adj'$  is 1. If the two turns are indeed assigned to adjacent gates,  $adj = 1$  implies that turn  $i$  departs before turn  $j$ , and  $adj' = 1$  implies that turn  $j$  departs before turn  $i$ . When both these variables are 1, the turns cannot be assigned to adjacent gate pair  $(k, l)$ .

Next, we try to model the LIFO constraints.

$$act_{a_i} \leq act_{a_j} - M \cdot (1 - lf_{ij}) \quad \forall i, j \in TURNS \quad (43)$$

$$act_{a_i} \geq act_{b_j} + M \cdot lf_{ij} \quad \forall i, j \in TURNS \quad (44)$$

$$act_{b_i} \geq act_{b_j} + M \cdot (1 - lf'_{ij}) \quad \forall i, j \in TURNS \quad (45)$$

$$act_{b_i} \leq act_{a_i} - M \cdot lf'_{ij} \quad \forall i, j \in TURNS \quad (46)$$

$$x_{ik} + x_{jl} + lf_{ij} + lf'_{ij} \leq 3 \quad \forall i, j \in TURNS; k, l \in GATES; (k^F, l^R) \in LIFO : i \neq j \quad (47)$$

Constraints (43) and (44) ensure that  $lf$  variable takes value 1 when a turn occupies the front gate, whereas (45) and (46) ensure that  $lf'$  variable takes value 1 when the front gate is empty. Constraint (47) ensures the feasibility of only one scenario. We will now see the implementation and performance of these models in the next section.

## 5. IMPLEMENTATION AND RESULTS

As we mentioned earlier, the purpose of this entire exercise was to actually optimize the gate assignment plan for a large airline at one of its busiest hubs—Chicago O'Hare. The models described in section 4 are implemented in Optimization Programming Language and run on a SUSE-Linux server with 4 GB RAM. The sample data set flown in the past by a major airline from its Chicago O'Hare hub is used by us and has 619 turns (including long turns), 25 724 flight connection opportunities accounting for revenue of over \$5.3m per day and 73 gates available at the hub airport under consideration.

Input data to the model in the planning mode included the following:

- Turns data (comprising turn id, incoming flight id, outgoing flight id, departure and arrival times of both incoming and outgoing flights, origin, via and destination stations)
- Flights data (comprising flight ids, flight numbers, departure and arrival times, origin and destination stations, origin and destination stations, reference day)

- International routes data (comprising international stations data)
- Itinerary (connection) data (comprising itinerary id, incoming flight id, outgoing flight id, min and max connection times, number of pax, revenue per pax)
- Station gate data (comprising gate information, list of equipments allowed for a gate and other miscellaneous restrictions such as adjacency, LIFO or push back)
- Walking time data (comprising gate to gate walking time data, equipment-wise deplane data, boarding time data, load bridge loading time)

We also implemented the following pre-processing strategy to the planning model objectives to make the model simpler and eliminate large number of superfluous variables.

- Only those connections that are potentially feasible are considered.
- All connections where the connection time is less than walk time between nearest gates + boarding time + deplane time + load bridge time are ignored and assumed to be lost revenue in any case.
- All connections whose connection time is greater than walk time between farthest gates + boarding time + deplane time + load bridge time are ignored and assumed to be realized revenue in any case.
- Minimum gate rest for long haul and short haul are provided for all cases.
- The other two levels (gate rest for long haul, short haul, and delays) are given same incentive.

The implementation of the previously mentioned pre-processing strategies reduced the number of critical (or at risk) connections substantially to just 335. It must be noted that although some of the costs and passenger connection revenue numbers are directly used in the model based on the actual data, some other costs such as the cost of assigning a turn to an unfavorable gate or the cost of using a zone or the cost of violating the desired additional gate rest to maximize schedule robustness are notional and based on the perceptions of the business analyst.

For benchmarking purposes, in our implementation, the first model of cost minimization also contains all perceivable business constraints. This model is considered as base model, and the three additional objectives are subsequently added to this base. So, the model is solved for all the combination of objectives:

- (1) Base, with only cost minimization objective,
- (2) Base plus connection revenue maximization,
- (3) Base plus zone usage minimization,
- (4) Base plus gate rest maximization,
- (5) Base plus connection revenue maximization plus zone minimization,
- (6) Base plus connection revenue maximization plus gate rest maximization,
- (7) Base plus zone minimization plus gate rest maximization,
- (8) All the objectives.

Although clearly the number of objectives under consideration has a direct correlation with the number of variables in the model and thus the problem complexity and run time, other factors, such as minimum and desired gate rest times, too have a significant role in the complexity. In our example, we noticed that increasing the minimum gate rest by even 5 minutes could explode the run time. We present the results in Table I to highlight the relationship between run time and minimum gate rest.

Table I. Results of the mixed integer program model on base case.

Objective function value (\$)	Minimum gate rest (minutes)	Run time (seconds)	Optimality gap (%)	Ungated turns
7427	5–10	10	0	—
7393	10–15	29	0	—
7131	15–20	177	0	—
6914	20–25	≥1000	4	3

As shown in the previous table, the base model runs within 3 minutes when the minimum gate rest is reasonable and the model finds adequate opportunities to schedule all the turns. For the rest of the study, we use a minimum gate rest of 10 minutes for turbo propeller and regional jet aircrafts and 15 minutes for all others. Similarly, the desired gate rest is a 0–5 minutes add-on to the minimum gate rest and thus maximum of 15 minutes for propeller and regional jet aircrafts and 20 min for other larger aircrafts. Note that there is also a historical delay component beyond the desired gate rest, both of which can only be violated with a penalty. Table II presents the results of the different scenarios, their run time, and also the performance of the different objectives. Note that the solution provided by the model when a particular objective is not specifically chosen can still be better than the worst possible. For example, the connection revenue for the base model need not be \$0 as the output of the model can still manage to realize some revenue when two connecting flights are assigned to gates such that walking time criteria will ensure that the passengers will not miss these flights.

Table II elaborates the results of eight distinct cases for the three objectives in all combinations of their selection. The first row refers to the base case, where none of the three objectives are selected and the mathematical model is mainly solved for feasibility with an implicit cost minimization objective. Even in the base model, connection revenue of \$4.154m is achieved, even though the zone usage and gate rest violation are quite high. In rows 2, 3, and 4, connection revenue, zone usage, and gate rest violation are individually optimized, respectively. Thus, the connection revenue for row 2 is highest at \$5.31m, the zone usage is lowest at 3 for row 3, and the gate rest violation for row 4 is lowest at 0.2m. In rows 5, 6, and 7, combinations of two objectives are considered sequentially. Thus, the connection revenue remains higher at \$5.31m, and zone usage remains low at 3 when those particular objectives are chosen. The gate rest violation penalty, when it is chosen with connection revenue in row 6, is 14.4 minutes and is 0.5 minute when it is chosen with zone minimization in row 7. In comparison, the gate rest violation is more than 1200 minutes when it is not chosen explicitly in the objective function (row 5). When all the three objectives are chosen together, connection revenue decreases by \$225 from the best possible value of \$5.31m, zone usage increases to 4 from the best possible value of 3, and gate rest violation increases to 14.6 minutes from the best possible value of 0.2 minute.

It is clear from Table II that the results show improvements when a particular objective is actually chosen as a criterion for optimization separately and deteriorates slightly when more objectives are chosen. For example, the minimum number of zones rise from 3 to 4 or total gate rest violation from 0.2 to 14.4 or 14.6 minutes when the concerned objective is chosen in conjunction with another objective. In any case, the gains from choosing a particular objective in the mathematical model are fairly obvious—with connection revenue going up by more than 25% (rows 2, 5, 6, and 8), zone usage falling to one third (rows 3, 5, 7, and 8) and gate rest violation being reduced to a mere 0.02% to 1.1% from the original violations (rows 4, 6, 7, and 8).

Table II. Results of the mixed integer program model on the different combination of objectives.

Row #	Case	Objective function (\$)	Run time (seconds)	Optimality gap (%)	Connection		
					Revenue (million \$)	Zone usage: 10:00 PM–6:00 AM	Total gate rest violation (minutes)
1	Base	7393	29	0	4.154	9	1258.6
2	Connection revenue max	5317671	258	0	5.310	9	1270.0
3	Zone min	6738	321	0	4.155	3	1258.6
4	Gate rest (GR) max	7388	8	0	4.154	9	0.2
5	Connection rev + zone	5 316 716	221	<0.01	5.310	3	1234.7
6	Connection rev + GR	5 317 207	254	0.00	5.310	9	14.4
7	Zone + GR	6643	>3600	1.62	4.155	3	0.5
8	Connection rev + zone + GR	5 316 452	1891	<0.01	5.310	4	14.6

max, maximization; min, minimization.

From Table II, it is also amply clear that the dollar contribution due to connection revenues have a very high, disproportionate impact in the objective function. From the analyst perspective, it is explained by the fact that the connection revenue is a realized revenue whereas the cost of gate rest violation and zone usage are notional—which is aimed to be optimized to a point beyond the impact of the revenue in itself. The notional cost chosen for 1 minute of gate rest violation per turn is assumed to be \$25. The cost for using one zone for one shift period is assumed to be \$50. The third notional cost of assigning a preferred gate to certain flights varies from \$0 to \$50, depending on the importance of the markets on which these flights operate. However, this cost coefficient does not conflict with any of the three main objectives strongly.

To understand the sensitivity of these notional cost coefficients in the multi-objective function, let us vary these notional costs to evaluate the impact on the corresponding objective functions. We choose values from one-hundredth to hundred times the notional costs to assess the impact. Thus, the gate rest violation per minute is allowed to vary from 25¢ to \$2500, and zone penalization is allowed to vary from 50¢ to \$5000. Unsurprisingly, there is no difference in connection revenue, and the gate rest violation decreases with the increase in penalty. Run time for \$2500 penalty is 570 seconds, and 1120 seconds for \$0.25 penalty. Similar results are noticed for varying the penalization on run time. The results are shown in Figures 4 and 5. It is amply clear from this analysis that the model is fairly robust with respect to the changes in the notional costs of the different elements. The changes in actual gate rest violation or zone usage are much smaller for even very large variances, compared with the base scenario when these objectives are not part of the model, implying that there is no further need for detailed Pareto analysis of the different notional costs.

Now let us shift our attention to operations models. One of the possibilities to validate the models in operations mode is to create (or simulate) a real disruption to the schedule. Typically, the impact of disruptions during the morning or evening peaks creates more infeasibilities than disruptions at other periods. A cursory glance at the data reveals that morning peaks are bigger than the evening peaks.

For example, the maximum number of turns on the ground in the morning during an hour is 91 between 08:00–09:00 AM. On the contrary, the maximum number of turns on the ground in the evening during any hour is 83 between 20:00–21:00 (8:00–9:00 PM). The number of available gates

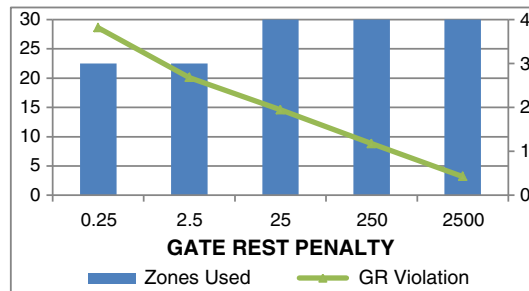


Figure 4. Gate rest violation and zone usage change with changes in gate rest penalty.

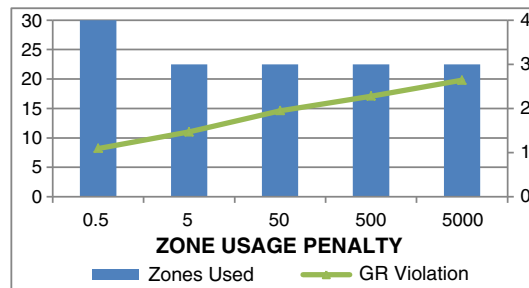


Figure 5. Gate rest violation and zone usage change with changes in zone usage penalty.

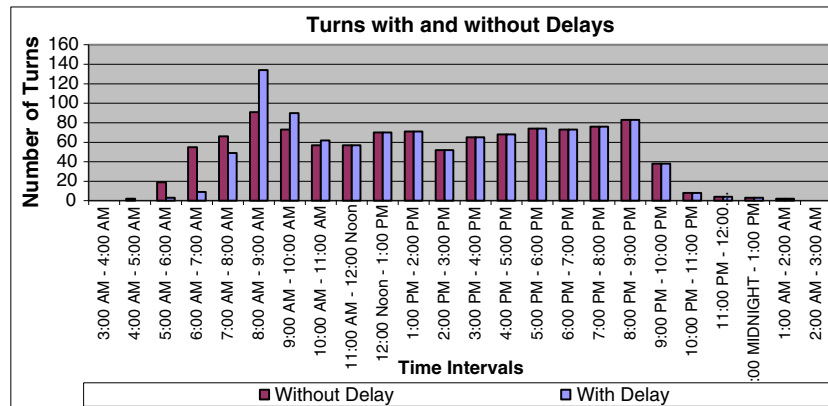


Figure 6. Number of turns by hour on the ground—with and without delays.

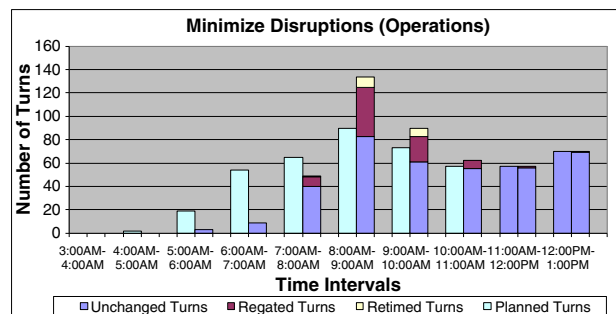


Figure 7. Results of the operations model with normal penalties for re-gating and retiming.

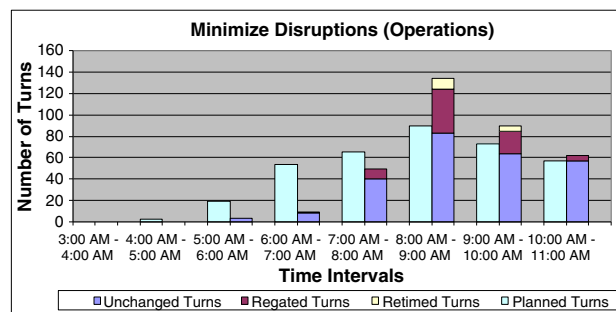


Figure 8. Operations model with higher penalty for re-gating non-delayed turns.

is 73, and there are instances of time in the morning when 71 gates are occupied, although the maximum occupation in the evening is only 64.

Thus, to evaluate the performance of our operations models for the worst case scenario, we generate exponentially distributed random delays with  $\lambda = 30$  minutes to all morning flights till 8:00 AM and apply it to departure time. The extent of delay is kept at 30 minutes and not more because we intend to mimic the worst case scenario when the number of turns on the ground is very large. This extent of delay is expected to move a large number of flights from the high activity non-peak hour of 07:00–08:00 AM. Any delay distribution with  $\lambda$  in excess of 30 minutes tends to smoothen the peaks by moving the flights further later in the day and, hence, not desirable for validation. After the simulation of delays, the number of turns during 08:00–09:00 AM increases to 134, indicating a 50% jump in the morning peak. The impact of delay stabilizes after 11:00 AM. The output of the delay simulation is presented in Figure 6.



For this delayed schedule, we now run the operations model with objective (21) and constraints (2)–(13) and (22)–(27). The run time is within a minute. The penalty of not gating a turn is kept very high, whereas the penalties for re-gating and deviations within a zone are kept relatively lower sequentially. After the runs, it is found that a large number of turns continue to be assigned to their planned gates. However, the schedule stabilized only after 1:00 PM even though there are only two turns beyond 11:00 AM that are re-gated. About 25% turns till 11:00 AM are re-gated, and 6% turns are retimed. All the other turns are assigned to the planned gates. All turns are eventually assigned to a gate, and the average extent of retiming was 42 minutes for the fifteen turns that are retimed. Unfortunately, it is noticed in the first model that several non-delayed turns (arriving after 8:00 AM) were re-gated. To avoid it, we propose double penalization for re-gating turns that are not already delayed. The schedule stabilizes earlier, that is, by 11:00 AM, even though the extent of retiming now increases to 54 min. The results of the two models are shown in Figures 7 and 8, respectively.

## 6. CONCLUSION

We have presented the actual airline gating problem as observed by the industry. The problem considered in this paper is actually implemented for a very large airline at Chicago O'Hare, which is also one of the busiest hubs for this airline. One of the major contributions of this paper relates to the mathematical representation of different types of business and operational constraints for this problem, which have not been carried out before.

Another important contribution of this paper is to show that simple continuous time assignment modeling formulation works fine even for very large, real-life problem instances. This is also one of the few papers where optimal or near-optimal solutions are reported for the airport gate assignment model. In this paper, we have shown that it is possible to model and solve three distinct objectives for the problem including selecting the objectives in any combination. The three objectives are as follows:

- Maximization of connection revenue for particularly those connections that are at risk,
- Minimization of zone usage during sparse operations,
- Maximization of gate plan robustness by providing adequate gate rest based on the historical pattern of flight delays.

The scale of the problem instance involved over 1200 flights and 70 gates. The total connection revenue was in the range for \$5.3m per day, whereas the overall revenue from all flights at the airport is \$40m per day for the particular season for which we obtained the data. The largest problem involving all the three objectives taken together was solved to near optimality with an optimality gap of less than 0.01% in about 30 minutes. Other versions of problems with fewer objectives are solved more quickly to optimality, usually, within 5 minutes, except one instance.

One of the major findings of this work is that even simple mathematical formulations can perform well on large problem instances if the pre-processing is carried out cleverly. This work will find immense applications not only in other similar problems relating to infrastructure assignment in transportation such as assignment of platforms to trains at a railway station or assignment of berth terminals to vessels at the ports but also in a host of other scheduling problems in different contexts. This model can be easily extended to different job-shop scheduling problems with varied objectives and different side constraints.

There are several directions in which the future research on this topic can proceed. One of the most obvious research directions would be in the area of coming with a Pareto frontier for the different notional parameters, which impact the objective. Although some of the cost and revenue coefficients in the objective are actual dollar figures, there are some others such as the coefficient for unfavorable gate assignments, coefficient for zone usage, and coefficient for additional gate rest violation per minute, which are simply notional and based on the perception of the business analysts. Apart from the coefficients provided to us, we also tried to estimate the most optimal solution for zone minimization and schedule robustness by suggesting a very high penalty for those corresponding coefficients in the model. So, in our results, we could present the best and the trade-off solutions for the different

objectives, although we could not build a Pareto frontier for every objective. This may be a potential area for further research.

Operations models have primarily been tried on a single simulation of disruption. More varied types of disruptions, as observed from the real data, must be tried. Additionally, it must be noted that if the disruption is very high and large number of flights need to be retimed, the retiming model is unlikely to run into a few minutes. It may be required to try out some heuristics for fixing the timings of certain flights so that the number of variables in the second model is not too high.

Another interesting direction would be to explore a stochastic flight delay model into the planning model to see if there can be further improvements. Additionally, there is scope for a simple extension of this work to concentrate on variability in the rate of walking of the different types of passengers. A more advanced study can also include pedestrian behavior models within the modeling framework.

#### ACKNOWLEDGEMENTS

Authors are indebted to M. B. Pradeep Kumar of Symphony Marketing Solutions, Bangalore, India, for his valuable inputs in mathematical models and implementation.

#### REFERENCES

1. Babic O, Teodorovic D, Tosic V. Aircraft stand assignment to minimize walking. *Journal of Transportation Engineering* 1984; **110**:55–66.
2. Mangoubi R, Mathaisel D. Optimizing gate assignments at airport terminals. *Transportation Science* 1985; **19**:173–188.
3. Haghani A, Chen M. Optimizing gate assignments at airport terminals. *Transportation Research – A* 1998; **32**:437–454.
4. Yan S, Chang C. A network model for gate assignment. *Journal of Advanced Transportation* 1998; **32**:176–189.
5. Bolat A. Procedures for providing robust gate assignments for arriving aircraft. *European Journal of Operations Research* 2000; **120**:63–80.
6. Xu J, Bailey G. The airport gate assignment problem: mathematical model and a tabu search algorithm. *Proceedings of the 34th Hawaii International Conference on System Sciences* 2001; Hawaii, **3**:10–19.
7. Yan S and Huo C. Optimization of multiple objective gate assignments. *Transportation Research – A* 2001; **35**:413–432.
8. Yan S, Shieh C, Chen M. A simulation framework for evaluating airport gate assignments. *Transportation Research – A* 2002; **36**:885–898.
9. Ding H, Lim, A, Rodrigues B, Zhu, Y. New heuristics for the over-constrained airport gate assignment problem. *Journal of the Operational Research Society* 2005; **32**(7):1867–1880.
10. Lim A, Wang F. Robust airport gate assignment. *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2005)*, Hong Kong 2005; 74–81.
11. Lim A, Rodrigues B, Zhu Y. Airport gate scheduling with time windows. *Artificial Intelligence Review* 2005; **24**:5–31.
12. Diepen G, van den Akker JM, Hoogeveen JA, Smeltink, JW. Using column generation for gate planning at Amsterdam Airport Schiphol. Technical Report UU-CS-2007-018, Department of Information and Computer Sciences, Utrecht University, The Netherlands, 2007.
13. Dorndorf U, Drexl A, Nikulin Y, Pesch E. Flight gate scheduling: state-of-the-art and recent developments. *Omega* 2007a; **35**:326–334.
14. Dorndorf U, Jaehn F, Lin C, Ma H, Pesch, E. Disruption management in flight gate scheduling. *Statistica Neerlandica* 2007b; **61**:92–114.
15. Yan S, Tang CH. A heuristic approach for airport gate assignments for stochastic flight delays. *European Journal of Operational Research* 2007; **180**:547–569.
16. Dorndorf U, Jaehn F, Pesch E. Modelling robust flight-gate scheduling as a clique partitioning problem. *Transportation Science* 2008; **42**:292–301.
17. Drexl A, Nikulin Y. Multicriteria airport gate assignment and Pareto simulated annealing. *IIE Transactions* 2008; **40**:385–397.
18. Hu XB, Di Paolo E. An efficient genetic algorithm with uniform crossover for the multi-objective airport gate assignment problem. *Multi-Objective Memetic Algorithms* eds C-K Goh, Y-S Ong and KC Tan, 71–89. Springer: Heidelberg, 2009.