Real-Time Intermodal Substitution

Strategy for Airline Recovery from Schedule Perturbation and for Mitigation of Airport Congestion

Yu Zhang and Mark Hansen

A strategy for airline recovery from schedule perturbation caused by adverse weather or other temporary events is proposed. Capacity reduction at a hub airport caused by these temporary events is reflected as reduced slots (or later controlled time of arrivals) for airlines from ground delay programs. It is suggested that under severe circumstances legacy airlines with hub-and-spoke networks implement ground transportation into their network to substitute for flights. This strategy is called real-time intermodal substitution (RTIMS). RTIMS is different from air–rail cooperation practiced in Europe in that it is triggered only by a severe demand and supply imbalance at major hub airports and it performs operational integration of airside and ground transportation. Mathematical programming is presented to help airlines make decisions about if and how to delay, canceling flights, or substituting flights with buses. An approximation algorithm is proposed to obtain solutions, avoiding substantial computation time required to solve large-scale nonlinear integer programming. As an evaluation, experimental data are used to compare scenarios with and without tactical intermodal substitution.

Shortages of airline resources and airspace capacities are two major causes of airline operation disruption, which costs the airline industry billions of dollars in annual losses. Shortages of airline resources stem from a variety of causes, such as crew sick leave, aircraft maintenance outages, and lack of ground personnel or gates. Airspace resources include en route (center) capacities and terminal (airport) capacities. Shortages of airspace capacities are caused by adverse weather, air traffic control facility malfunctions, security threats, and so forth. Bratu and Barnhart give a detailed description of airline disruptions (*1*).

Under the current system of collaborative decision making (CDM), once a severe imbalance of demand and supply is detected, a ground delay program (GDP) advisory is suggested, as illustrated in Figure 1. This advisory assigns scheduled flights an estimated arrival time, most of which are later than the original scheduled arrival time. Airlines respond to this advisory by canceling and reordering flights. If the imbalance is resolved by these adjustments, GDP ends; otherwise, the air traffic control system command center issues each remaining scheduled flight an expected departure clearance time and a controlled time of arrival (CTA). Airlines manage their CTAs (or arrival slots) to their best internal business interests. An airline operation center may cancel more flights if necessary, reorder flight sequences, and reassign flights to CTAs to use the arrival slots of delayed or canceled flights. This flexibility to use slots under CDM encourages airlines to cancel and combine flights, thus reduce demand, passenger delay, and airline disruption costs.

Compared to airline resource shortages, disruptions caused by capacity deficiencies have more networkwide effects. Thus an airline's recovery from this kind of schedule perturbation is more critical to its operation and National Airspace System performance. This study deals with this kind of disruption, especially when there is capacity reduction or closure at major hub airports that are considered pivots of the National Airspace System. An option other than current strategies is proposed. Under this new strategy, tactical intermodal substitution (TIMS), when severe capacity deficiencies occur, the airline brings other transportation modes into its hub-and-spoke network and integrates multimodal scheduling if this will help reduce disruption costs. Consequently, airlines have to make the following decisions:

1. Which flight leg should be delayed, canceled, or substituted by other transportation modes?
2. If a flight leg is not canceled or substituted, how long should it be delayed?
3. How should passengers be reassigned to flight legs or pseudo legs carried by other transport modes?
4. What are the departure times of other transport vehicles?

These decisions are subject to constraints, such as aircraft capacity restriction, passenger conservation, aircraft availability, and airport capacity limit. To effectively make decisions, managers must have knowledge of predicted airport capacity profile (or assigned slots from the FAA GDP) and passenger itinerary information. Making decisions to recover operations in the aftermath of disruption is a comprehensive process. Currently, decisions related to schedule, aircraft, crews, and passengers are made sequentially. Because of this complication and the limitation of operations research technology, there is no model that successfully integrates all components. Encouragingly, although some assumptions still have to be made to simplify the problem, the research community is moving toward integrating two or more components to pursue a more global optimization. This study focuses on schedule and passenger recovery for a one-stage hub-and-spoke network, considering both arrival and departure capacity constraints.

Airline intermodal experience in disruption recovery is rare, and decisions are made reactively, case by case, and unsystematically. With increase in demand, major airports are vulnerable to capacity reduction caused by adverse weather and other temporary events. There is a need for a procedure to utilize other transport modes for recovery by coordinating airside and groundside operations. Benefits

Y. Zhang, 107A McLaughlin Hall, and M. Hansen, 144 McLaughlin Hall, National Center of Excellence for Aviation Operations Research, University of California, Berkeley, CA 94720. Corresponding author: Y. Zhang, zhangyu@berkeley.edu.

Transportation Research Record: Journal of the Transportation Research Board, No. 2052, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 90–99.
DOI: 10.3141/2052-11

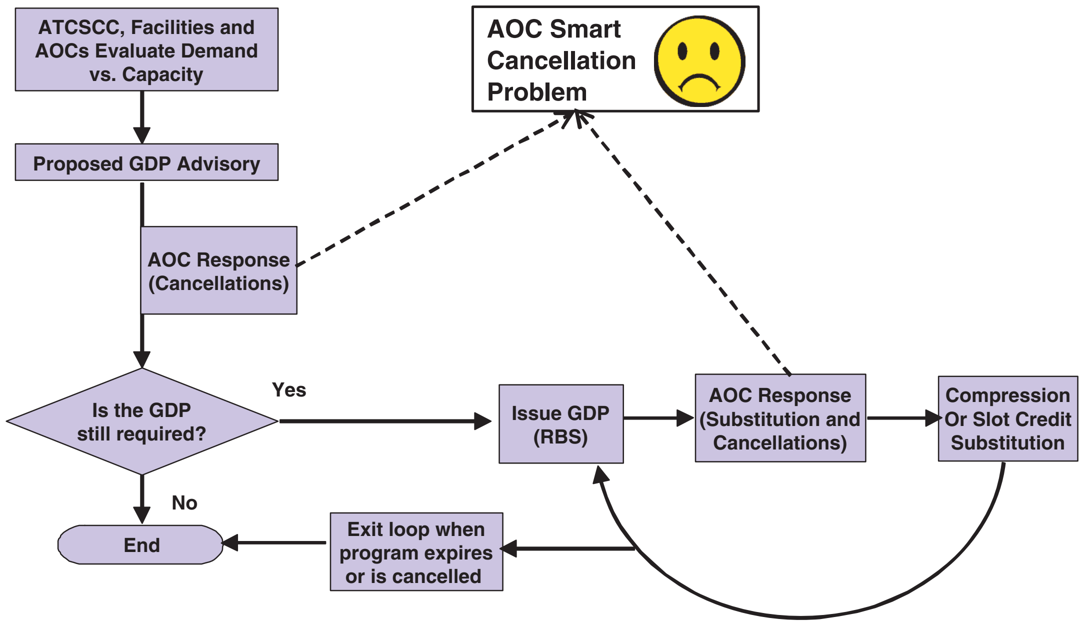

FIGURE 1 Flowchart of collaborative decision making (AOC = airline operation centers; RBS = ration by schedule).

that airlines can achieve by implementing such procedures include the following:

- Reserve scarce airspace resources for more important flights and reduce aircraft delay;
- Reduce the pressure of passenger reassignment, especially when there is insufficient protected space in the markets where flight legs are canceled;
- Reduce passenger misconnection and delay costs; and
- Obtain flexibility in crew recovery.

INSUFFICIENT INTERMODAL PRACTICE IN AIRLINE RECOVERY

The following information was obtained mainly from conversations with airline operational personnel.

Because of a severe thunderstorm passing through Dallas–Fort Worth, Texas (DFW), American Airlines (AA) canceled flights from DFW to Oklahoma City, Oklahoma (OKC), Austin, Texas (AUS), and San Antonio, Texas (SAT), airports within a 3- to 4-h drive. Because of the cancellation and high booking rates on flights on consecutive days, according to airline personnel, "buses became the best option." The biggest problem that AA encountered was communication with passengers—who will be reaccommodated on buses and where should they go. AA sat up pseudo flight numbers for the buses, and the passengers were booked according to first landing in, first out. AA admits that the bus operation did not run perfectly, but for passengers it beat spending a night at the airport. Thus AA intends to make it a more formal part of the OSO planning package and provide detailed procedures on handling it. Customer feedback on this experiment was mixed. AA said that some customers were upset, stating that "if they wanted to ride the bus, they would have bought a bus ticket," whereas others were happy just to get out of town.

United Airlines (UA) operates with five hubs. Through conversation with customer managers at Chicago O'Hare, Illinois (ORD) and Los Angeles, California (LAX), it was found that if the flights with destinations close to ORD are canceled, passengers will be reaccommodated on buses so they will not have to stay at the terminal overnight, or longer when snowstorms are severe. The direct cost of hiring buses, as estimated by the customer relation division, is equivalent to the cost of providing discount vouchers for hotel accommodation and reassigning passengers on later flights. At LAX, a thunderstorm passing through may lead to cancellations of flights out of the airport. Occasionally, the station will hire buses to transport passengers to their destination of San Diego, California (SAN). The driving time between LAX and SAN is only about 2 h.

Different from the two airlines, high-speed rails in New York State and along the Northeast Corridor allow Continental Airlines (CO) at Newark, New Jersey (EWR) to cooperate with Amtrak and reduce its flight frequency on several markets. For instance, passengers from Philadelphia to San Francisco, California (SFO) or London Heathrow through EWR can book a combination of train and air tickets through the CO reservation system. This cooperation gives CO a means of passenger recovery if remaining flights in those markets must be canceled. According to the station manager at EWR, more passengers with trip destinations in Europe chose the combination of train and air modes.

Other transport modes can be used in intermodality in American air transportation, such as rail; however, the rail system in the United States does not have direct access to most of the major airports, and the situation cannot be improved in the near future. According to AA and UA, it is fairly easy to obtain buses through contacts at a local tour company. A customer manager at ORD said that buses can be had within 1 h. Randomly picked motor coach charter companies were able to provide very close lead times on similar services. Thus the bus can be used as a substitution mode generically and nationwide.

LITERATURE REVIEW

Airline recovery from disruption is a complex problem because of interactions between multiple resources, network propagation effect, integrality of decision variables, and nonlinear characteristics of the cost function. Exploration of this problem can be traced back to the 1980s. In 1984, Teodorovic and Guberinić developed a network model to minimize overall passenger delay in circumstances in which an aircraft unexpectedly becomes unavailable. They used a branch-and-bound procedure to determine the least expensive set of aircraft routings and schedule plan (*2*). The authors assumed a single fleet and applied their methodology for a small-size network with eight flights. However, the methodology is cumbersome for solving realistic-scale problems. Teodorovic extended his research later and coauthored a paper with Stojkovic in 1990 in which they addressed the airline scheduling and routing problems with heuristic algorithms based on dynamic programming (*3*). They constructed a two-stage optimization, first minimizing the total number of canceled flights and then minimizing the total passenger delay on flights not canceled.

The recovery problem for solving unpredicted aircraft shortage and maintenance requirements attracted more attention in the 1990s. Jarrah et al. presented a decision support framework for airline flight delay or cancellation implemented at UA (*4*). They modeled the schedule recovery as a minimum-cost network flow problem and considered delaying or canceling flights separately. Yan and Yang proposed four models for dealing with temporary unavailability of a single aircraft. The objective functions were to minimize the duration of disturbance and determine the most profitable schedule in the perturbation period (*5*). The four models handle different combinations of delaying, canceling, or ferrying flights. Lagrangian relaxation and subgradient methods are used to find near optimal solutions and good bounds. Their models are realized for a relatively small airline, China Airlines. Cao and Kanafani introduce an integrated delay and cancellation model at multiple airports (*6, 7*). They presented detailed calculation of downstream delay cost in the quadratic 0-1 programming model. Their model can be extended to formulate some special cases, such as ferry of surplus aircraft and replacement of different types of aircraft.

GDPs are a key component of air traffic flow management. If traffic demand is expected to exceed capacity at airports (or sometimes en route airspace), a GDP is launched to hold an arrival aircraft at its original airport unless there is reasonable assurance that, after departure, it will be able to proceed to its destination with a minimum amount of delay in the air (*8*). In 1998, Luo and Yu addressed airline recovery from schedule perturbation stemming from GDP. Their first paper is theoretical and methodological, investigating the landing assignment problem with different objectives by assuming aircraft and crew are nonsplitable (*9*). In their second paper, the nonsplitable assumption is relaxed, which leads to a much more complex optimization problem for the different objectives listed in the first paper. The problem becomes NP-hard if the objective is to minimize the maximum delay of outbound flights. The authors transformed the problem and proposed a heuristic so that a real-life problem could be solved in a very short time (*10*). Milner was the first to consider the connection dependencies of hub operations in the decision support model for resolving airline schedule disruption (*11*). He proposed an optimization model to minimize the costs of bank spread and flight cancellation for arrival operations at a hub airport. Carlson argued that the Milner model did not use the scarce arrival slot efficiently (*12*). He refined Milner's model and provided a different formulation to test the computational efficiency for a large-scale problem. He assumed that the CDM procedure is in place and addressed a scenario in which the arrival capacity at an airline's hub airport is significantly reduced and the hubbing airline must make tactical decisions with limited slots assigned by FAA. Neither the Milner nor the Carlson model considers outbound operations or the connection between arrival and departure flights. Most recently, Barnhart and her students have been working on understanding passenger delays in legacy airline hub-and-spoke networks. They proposed a passenger-centric model for schedule recovery (*1*). First, they formulated the number of crews needed because of a crew shortage resulting from canceled or delayed flights. Then they constructed two models, one using disrupted passenger cost only and the other including all passenger costs. Two other papers review the literature on airline disruption management (*13, 14*). Nevertheless, there is no published literature discussing the integration of other modes into the air transportation network for airline recovery during disruptions caused by resource shortages.

MATHEMATICAL FORMULATION

By introducing ground transportation modes, an airline's hub-and-spoke network becomes a network with two sets of links, one indicating flying between spoke airports and the hub airport with short transportation time but possible delays caused by the insufficient capacity at the hub airport, and the other indicating driving passengers between spoke airports and the hub with longer transportation time but no delays because that mode does not require airside scarce resource. In Figure 2, the straight lines indicate links with short transportation time but possible delays, and the curved lines indicate links with longer transportation times but not limited by airside capacities. The enlarged view on the right of the figure highlights some spoke airports within a certain distance from the hub airport that could be substituted by motor coach services because of the relative small differences between flying times and ground transportation times.

This section constructs a mathematical programming model to help airlines answer the questions posed earlier.

Notation

Time of day is divided into a finite set of periods of equal duration, denoted by $t \in \Gamma = \{1, \dots, T\}$. For instance, Γ might be a set of 96 periods of 15 min each, summing to a planning horizon of 24 h. Arrival and departure slots assigned to an airline are denoted as AHCap_t and DHCap_t for $t \in \Gamma$, respectively. Airlines have arrival flights $a \in \text{A} = \{1, \dots, A\}$ with the scheduled arrival time denoted by AT_a and departure flights $f \in \Phi = \{1, \dots, F\}$ with the scheduled departure time denoted by DT_f. En route flying times of these flights are assumed to be constant and recorded in two vectors, AAT_a and ADT_f, respectively. Spoke airports in this hub-and-spoke network are represented by a set $s \in \Theta = \{1, \dots, S\}$. SpokeA and SpokeF trace arrival flight origins and departure flight destinations, respectively. If $\text{SpokeA}_{as} = 1$, arrival flight a comes from spoke airport s, 0 otherwise. Aircraft type is denoted by $k \in \text{K} = \{1, \dots, K\}$. TypeA and TypeF record the types of arrival and departure flights. If $\text{TypeA}_{ak} = 1$, arrival flight a uses a type k aircraft, 0 otherwise. The number of

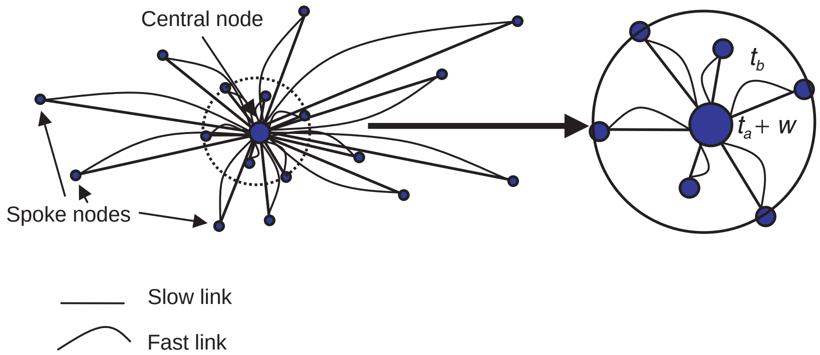

FIGURE 2 Network with node capacity constraint ($t_a + w$ = transport time on fast link and delay; t_b = transport time on slow link).

different types of aircraft available at the beginning of the schedule perturbation period is indicated as a vector TypeO. The number of different types of aircraft required (or scheduled) to be available at the end of the perturbation period is indicated as a vector TypeT.

Passenger itinerary information for this model is listed in Table 1. Passengers with the same itinerary are considered identical, that is, their conditional costs are the same, and only aggregated numbers are important in passenger reassignment. Nevertheless, when airlines implement this strategy, they may give priorities to different passengers, on the basis of passenger loyalty, value to airlines, or other criteria. In Table 1, the first column is arrival flights and the first row departure flights. AHPax_a in the second column is the number of passengers on an arrival flight a with destinations at the hub airport. DHPax_f in the second row is the number of passengers on a departure flight f originated from the hub airport. Pax_{af} denotes the number of passengers whose itinerary involves a transfer from an arrival flight a to a departure flight f. In the last column and last row, the total passengers on arrival flights and departure flights are presented by APax_a and DPax_f, respectively.

It is assumed that the ground transportation time for substituting each flight (arrival or departure) is estimated from historical data and real-time traffic conditions. The time is then recorded into two vectors, BAT_a and BDT_f, for $a \in$ A and $f \in \Phi$. Two other notations needed for this formulation are minimum turnaround times of flights and minimum transfer time of passengers at the hub airport. For simplification, a constant value of 45 min is assumed for aircraft turnaround and 30 min for passenger transfer in this model, denoted by maircraft and mpax, respectively. The values of these two variables, however, can be changed if more information is obtained.

TABLE 1 Input Data: Passenger Itinerary Information

		F_1	F_2	F_3	. . .	
		DHPax_1	DHPax_2	DHPax_3	. . .	
A_1	AHPax_1	Pax_{11}	Pax_{12}	Pax_{13}	. . .	APax_1
A_2	AHPax_2	Pax_{21}	Pax_{22}	Pax_{23}	. . .	APax_2
A_3	AHPax_3	Pax_{31}	Pax_{32}	Pax_{33}	. . .	APax_3
. . .	. . .	. . .	. . .	. . .	. . .	. . .
		DPax_1	DPax_2	DPax_3	. . .	

Decision Variables

The decision variables in the model are a set of binary variables that determine if one flight is released during period t or is canceled and substituted by motor coach. If neither of these options happens, the flight is canceled without any substitution. Another decision variable is the number of passengers on a takeoff flight. A departure flight at the hub airport can be held to wait for connecting passengers. Of course, the cost of doing this is to sacrifice the time of passengers and crews who have been ready for departures. The decision variables are described in detail as follows:

$$xf_a^t = \begin{cases} 1 & \text{if arrival flight } a \text{ is planned to arrive at the hub airport during time period } t \\ 0 & \text{otherwise} \end{cases} \qquad a \in A$$

$$xs_a^t = \begin{cases} 1 & \text{if arrival flight } a \text{ is substituted with ground transportation and arrives at the hub airport during time period } t \\ 0 & \text{otherwise} \end{cases} \qquad a \in A$$

$$yf_f^t = \begin{cases} 1 & \text{if departure flight } f \text{ is planned to leave the hub airport during time period } t \\ 0 & \text{otherwise} \end{cases} \qquad f \in \Phi$$

$$ys_f^t = \begin{cases} 1 & \text{if departure flight } f \text{ is substituted with ground transportation and planned to leave the hub airport during time period } t \\ 0 & \text{otherwise} \end{cases} \qquad f \in \Phi$$

$P_f \geq 0$ integer, the number of passengers on departure flight f $\qquad f \in \Phi$

Objective Function

The objective function that airlines want to minimize is the disrupted costs that include not only flight delay and cancellation costs but also passenger costs due to flight delay, cancellation, or substitution. The cost of motor coach service should also be taken into consideration. The formulated airline disruption cost can be expressed as follows:

$$\min\left(\sum_a \mathrm{AD}_a \cdot \mathrm{AHPax}_a + \sum_s \sum_t \left(\mathrm{IPax}_s^t - \mathrm{OPax}_s^t\right) + \sum_f \mathrm{DD}_f \cdot \mathrm{CPax}_f\right)$$
$$\cdot \mathrm{CostP} + \mathrm{DP} \cdot \mathrm{CostD} + \left(\sum_a \mathrm{AD}_a + \sum_f \mathrm{DD}_f\right) \cdot \mathrm{CostF} + \mathrm{BO} \qquad (1)$$

where the first component of the objective function is the passenger delay cost including the delays of arrival passengers whose destination is the hub airport, passenger waiting time at the hub terminals, and delays of departure passengers. The second component is the disrupted passenger cost. Disrupted passenger is defined as the passengers who miss their originally connection and cannot be reassigned to later flights before the end of the planning horizon. The third component is flight delay cost, and the last is ground transportation operating cost. The arrival and departure delays in the objective function are calculated as follows:

$$\mathrm{AD}_a = \sum_t \left(t - \mathrm{AT}_a\right) \cdot \left(xf_a^t + xs_a^t\right) + \left(1 - \sum_t \left(xf_a^t + xs_a^t\right)\right) \cdot \mathrm{ACanT}_a \qquad (2)$$

$$\mathrm{DD}_f = \sum_t \left(t - \mathrm{DT}_a\right) \cdot \left(yf_f^t + ys_f^t\right) + \left(1 - \sum_t \left(yf_f^t + ys_f^t\right)\right) \cdot \mathrm{DCanT}_f \qquad (3)$$

where $A\mathrm{Can}T_a$ and $D\mathrm{Can}T_f$ are estimated delay hours if flights are canceled. According to a previous study, the values are between 6 and 7 h. Use of these numbers in this formulation is an approximate way to count flight cancellation cost into airline disruption cost.

Disrupted passengers are inbound passengers whose arrival flights are canceled and transfer passengers who miss connections at the hub airports. Passengers who miss their original itinerary but are successfully reassigned to other flights or substitution ground transportation, however, are excluded. The number of disrupted passengers is as follow:

$$\mathrm{DP} = \sum_a \left(1 - \sum_t \left(xf_a^t + xs_a^t\right)\right) \cdot A\mathrm{Pax}_a + \sum_s \left(\mathrm{IPax}_s^T - \mathrm{OPax}_s^T\right) \qquad (4)$$

IPax_s^T and OPax_s^T are total inbound and outbound transfer passengers to a destination *s*. Calculation of these numbers is elaborated later. The operating cost of a bus includes fixed cost and valuable cost that is proportional to passenger times for those who are reassigned to ground transportation. It is calculated with the following formula:

$$\mathrm{BO} = \left(\sum_a \mathrm{APax}_a \cdot \mathrm{ABT}_a \cdot \sum_t xs_a^t + \sum_f P_f \cdot \mathrm{DBT}_f \cdot \sum_t ys_f^t\right)$$
$$\cdot \mathrm{CostBP} + \left(\sum_a \sum_t xs_a^t + \sum_f \sum_t ys_f^t\right) \cdot \mathrm{CostB} \qquad (5)$$

Constraints

The first set of constraints is arrival and departure capacity constraints at the hub airport:

$$\sum_a xf_a^t \le \mathrm{AHCap}_t \qquad \forall t \in \Gamma \qquad (6)$$

$$\sum_f yf_f^t \le \mathrm{DHCap}_t \qquad \forall t \in \Gamma \qquad (7)$$

Besides being delayed or substituted, flights can also be canceled. This is presented by the following set of constraints:

$$\sum_t \left(xf_a^t + xs_a^t\right) \le 1 \qquad \forall a \in A \qquad (8)$$

$$\sum_t \left(yf_f^t + ys_f^t\right) \le 1 \qquad \forall f \in \Phi \qquad (9)$$

The next constraint balances the inflow and outflow of passengers, that is, the outbound transfer passengers to a destination *s* leaving at time *t* should be no more than the inbound transfer passengers to the same destination arriving at the hub airport up to t − mpax, where mpax is the minimum passenger turnaround time:

$$\mathrm{IPax}_s^\tau \ge \mathrm{OPax}_s^{\tau-\mathrm{mpax}} \qquad \forall s \in \Theta \qquad \forall \tau \in \{1, \ldots, (T - \mathrm{mpax})\} \qquad (10)$$

where

$$\mathrm{IPax}_s^\tau = \sum_{t=1}^{\tau} \sum_a \sum_f \left(xf_a^t + xs_a^t\right) \cdot \mathrm{Pax}_{af} \cdot \mathrm{SpokeF}_{fs}$$

$$\mathrm{OPax}_s^\tau = \sum_{t=1}^{\tau} \sum_f \left(yf_f^t + ys_f^t\right) \cdot \left(P_f - \mathrm{DHPax}_f\right) \cdot \mathrm{SpokeF}_{fs}$$

Another critical constraint is the aircraft balance. Allow aircraft swapping among flights that utilize the same type of aircraft:

$$\mathrm{TypeO}_k + \mathrm{IAircraft}_k^\tau \ge \mathrm{OAircraft}_k^{\tau+\mathrm{maircraft}} \qquad \forall k \in K$$
$$\forall \tau \in \{1, \ldots, (T - \mathrm{maircraft})\} \qquad (11)$$

$$\mathrm{TypeO}_k + \mathrm{IAircraft}_k^T \ge \mathrm{OAircraft}_k^T + \mathrm{TypeT}_k \qquad \forall k \in K \qquad (12)$$

where

$$\mathrm{IAircraft}_k^\tau = \sum_{t=1}^{\tau} \sum_a xf_a^t \cdot \mathrm{Type\,A}_{ak}$$

$$\mathrm{OAircraft}_k^\tau = \sum_{t=1}^{\tau} \sum_a yf_f^t \cdot \mathrm{TypeD}_{fk}$$

In addition, relative flights are not allowed to be released from their origins earlier than the scheduled time. For an arrival flight from spoke airports, there are two constraints, depending on whether that flight is substituted. The constraints are as follows:

$$\sum_t t \cdot xf_a^t \ge \mathrm{AT}_a \cdot \sum_t xf_a^t \qquad \forall a \in A \qquad (13)$$

$$\sum_t t \cdot xs_a^t \ge \left[\mathrm{AT}_a + \left(\mathrm{BAT}_a - \mathrm{AAT}_a\right)\right] \cdot \sum_t xs_a^t \qquad \forall a \in A \qquad (14)$$

In comparison, there is only one constraint for departure flights at the hub airport:

$$\sum_t t \cdot \left(yf_f^t + ys_f^t\right) \ge \mathrm{DT}_f \cdot \sum_t \left(yf_f^t + ys_f^t\right) \qquad \forall f \in \Phi \qquad (15)$$

The number of passengers on departure flights P_f has to be less than the capacity of aircraft used by flight *f.* Although theoretically if the flight is substituted, there should be no such limit on how many passengers can be transported on the bus, real schedules with relative low frequencies serving a market pair make this a natural upper bound for ground transportation mode as well. If the capacity of departure flights is DCap, then the constraint will be the following:

$$P_f \leq \mathrm{DCap}_f \qquad \forall f \in \Phi \tag{16}$$

COMPLEXITY AND SOLUTION ALGORITHM

A close look at the modeling shows that it is a nonlinear integer programming with all linear constraints but a nonlinear objective function. The number of decision variables is $2 \cdot (A + F) \cdot T + F,$ and the number of constraints is $4 \cdot (A + F) + (S + K + 2) \cdot T,$ where A and F are the number of scheduled arrival and departure flights, respectively, S is the number of spoke airports, K is the number of aircraft types, and T is the number of time periods. A presolving process will eliminate decision variables based on Constraints 13, 14, and 15. However, for a hub-and-spoke network with tens of spoke airports and hundreds of flights, the number of decision variables and constraints for a daily operation can be enormous. Fortunately, airlines usually apply banked scheduling strategies in hub airports to reduce the transfer time (time between connections). Thus it is possible to partition a daily schedule into several time windows within which a set of arrival and departure flights finish their operations. Previous literature (*15*) set a time window as 2 h. However, airlines have smoothed their schedules to reduce operating cost, as a way to wrestle with the cruel financial situations in the aftermath of the terrorist attacks of September 11, 2001. Although the banking structure still exists, time windows should be elongated. The first technique used to get the solution of the mathematical programming for airline daily operations is to cut the time horizon into several overlapped time windows. Considering the depeaking phenomenon, a 4-h time window is used. Solutions are found by running the optimization iteratively for the sequential time windows.

Integer programming problems with a large number of variables are difficult to solve in general. Most successful approaches to solving nonlinear integer problems have involved linear approximation and relaxation techniques. This paper uses the traditional relaxation techniques. An approximation algorithm for this programming is described as follows:

- First, the integrality of decision variables is relaxed and a nonlinear programming is solved.
- Second, with solutions from the first step, adjust the flight decision variables $xf_a^t, xs_a^t, yf_f^t, ys_f^t$ to 1 if the solution value is large than 0.5; otherwise they will equal to 0. With the adjustment in the second step, all constraints other than Constraint 10 remain valid.
- Third, take adjusted flight decision variables as additional inputs to the original programming, and run a simple linear integer programming to find out integer values for number of passengers on departure flights.

The MINLP solver on NEOS Server 4.0 is used to fulfill these tasks. The design and implementation of the NEOS server were discussed by Czyzyk et al. (*16*), Gropp and Moré (*17*), and Dolan (*18*).

EXPERIMENTAL EXAMPLE

Numerical experiments are conducted to explore the features of the mathematical modeling. It is supposed that Airline A is operating in a hub-and-spoke network with a single hub. There are a set of short-haul spoke airports with a same distance to the hub airport and a set of long-haul spoke airports with another same distance to the hub airport. Forty arrivals and 40 departures are scheduled and evenly distributed in a 4-h time window. Among the flights, 25% are short-haul. The matrix of connecting passengers is simplified by assuming the same amount of passenger transfers from a long-haul inbound to a long-haul outbound flight, the same amount of passenger transfers from a short-haul inbound to a long-haul outbound flight and vice versa, and the same amount of passenger transfers from a short-haul inbound to a short-haul outbound flight. It is assumed there are only two types of aircraft. The larger aircraft is used for long-haul markets and the smaller is used for short-haul markets. The capacity of the hub airport drops to half the normal level for 5 h. Consequently, the number of landing and takeoff slots allocated to Airline A is four each for the first 5 h and eight each afterward.

Table 2 shows the results of the numerical example. The Without Substitution column lists the results for the scenario in which there is no ground transportation substitution for canceled flights. For

TABLE 2 Optimization Results for Numerical Experiment

		With Substitution		
Objective Function	Without Substitution	Approximation		Lower bound
Total cost ($)	411,046	221,575	220,188	216,250
Total arrivals and departures	80	80	80	
Inbound cancellation		12	14	
Substitution		10	10	
Outbound cancellation	2	14	16	
Substitution		10	10	
Longest delay (hours)	5.5	1.5	1.3	
Total passengers	7,360	7,360	7,360	
Disrupted passengers	90	14	16	
Computation time (seconds)	$\sim 10^3$	10^1–10^2	$\sim 10^3$	10^1–10^2

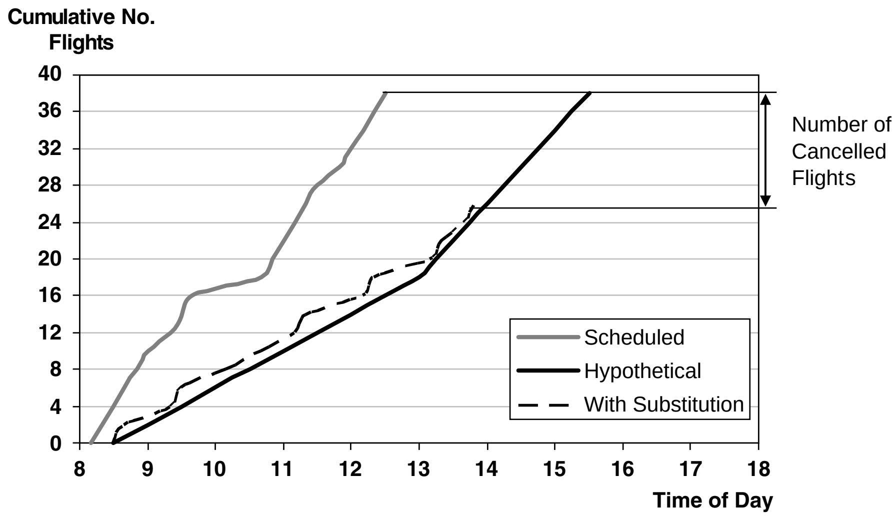

FIGURE 3 Cumulative number of arrivals.

the scenario with substitution, Table 2 shows the results from the approximation algorithm and from optimization and a lower bound without any binary or integer constraints. The optimization algorithm takes a much longer computation time than the approximation algorithm. It can be seen that for the scenario without substitution, there is no cancellation for inbound flights, and only two of 40 departure flights are canceled. Ninety passengers are left behind as disrupted. The longest flight delay under this scenario is up to 5.5 h. On the other hand, if there is ground substitution for canceled flights, the number of flight cancellations increases, and most of the cancellations will be substituted by motor coach. The reduction of the objective function value in the With Substitution scenario is mainly due to the decrease of flight delay and number of disrupted passengers.

Assuming there is no TIMS option for the airlines, with the same cost coefficients and formulation, airline total cost goes up to $411,046. Optimization from relaxing the integrality of decision variables provides a lower-bound of the problem, which is $216,250. With the solution algorithm described earlier, fractional decision variables are adjusted to binary variables for flight decision variables. Given adjusted flight decision variables as inputs and by retaining the integrality of the passenger decision variable, a set of new passenger numbers on departure flights is obtained, and the result for the objective function total cost $221,575. This number is about 3% higher than the lower bound. The exact result from the nonlinear integer programming falls into the range of the lower bound and the result of the approximation algorithm. With intermodal substitution, the number of disrupted passengers decreases from 90, about 1.2% of total passengers, to 14, about 0.1%.

Flight delay savings with RTIMS are illustrated in Figures 3 through 6. In all figures, the horizontal axis is the time of day and the vertical axis is the cumulated number of flights. In each figure, the solid gray curve depicts the cumulated flights arriving at or taking off from the airport following their original schedules. The solid

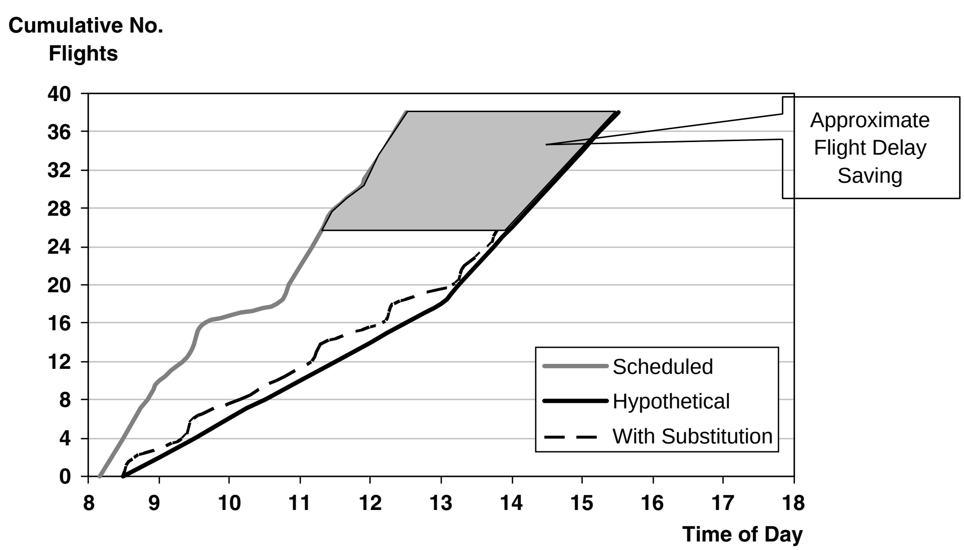

FIGURE 4 Arrival flight delay saving.

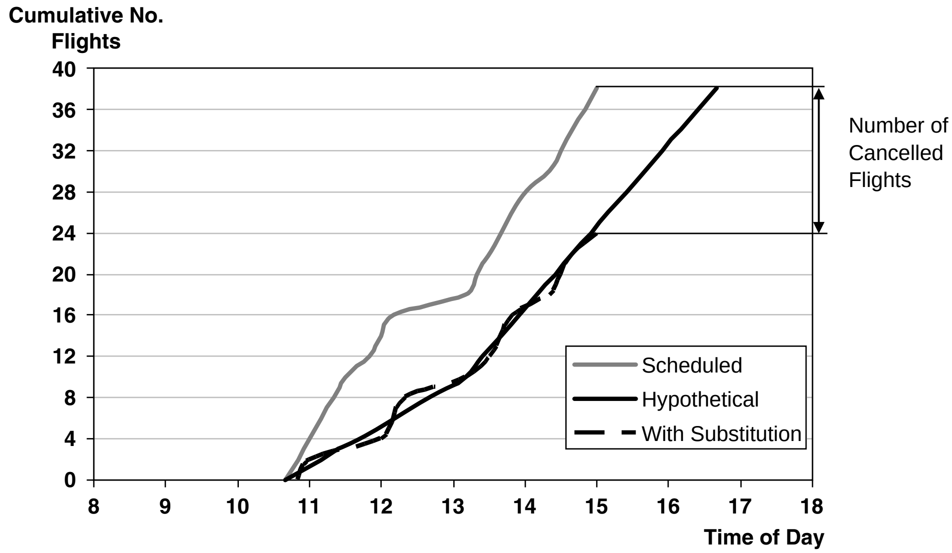

FIGURE 5 Cumulative number of departures.

black curve is the hypothetical arrival or departure curve constrained by the airport capacity. The dash-line curve shows the result of cancellation and substitution. The vertical differences between the gray and dash-line curves are the numbers of canceled flights, and the shaded area represents flight delay savings.

The following sections evaluate how cost coefficients and load factors affect the solutions of the programming.

SENSITIVITY ANALYSIS

Two hypotheses are tested to validate the model:

Hypothesis 1. When the passenger's value of time increases, airline disruption cost increases and the number of disrupted passengers also increases.

Hypothesis 2. When the load factor increases, airline disruption cost and the number of disrupted passengers increase.

Passengers' Value of Time

Assuming other coefficients are the same, passengers' value of time is increased to 1.5 times and two times the current value. Results from optimization show that the total disruption cost increases from $220,188 to $303,898 and $348,269, respectively. The number of disrupted passengers increases from 16 to 90 and finally 180. This test demonstrates, as shown in Table 3, the trade-off between undisrupted and disrupted passengers. With a higher value of time, disrupted passengers, a relative small percentage of total passengers, are weighted less in airline disruption recovery consideration. Hence, Hypothesis 1 is accepted.

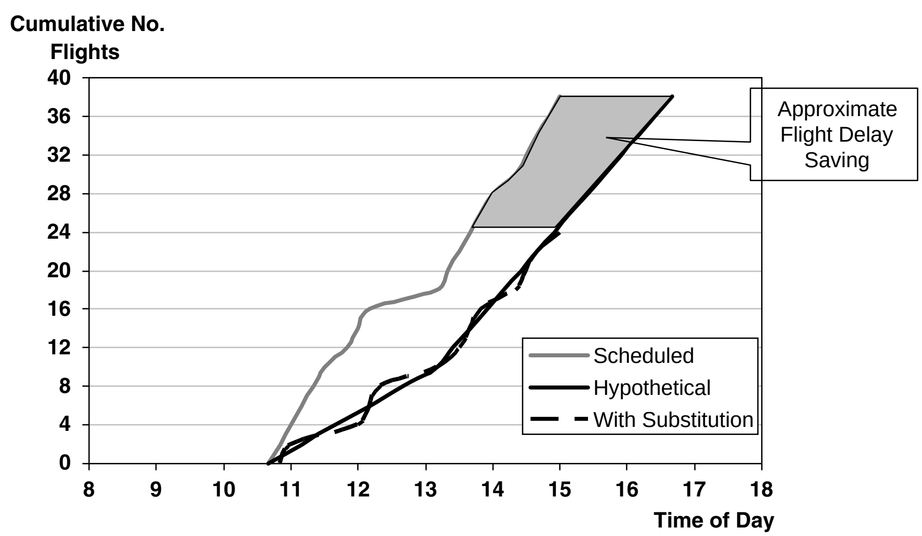

FIGURE 6 Departure flight delay saving.

TABLE 3 Sensitivity Analysis, Passengers' Value of Time

	Passenger Delay Cost		
Objective Function	CostP	1.5 CostP	2 CostP
Total cost ($)	220,188	303,898	348,269
Total arrivals and departures	80	80	80
Inbound cancellation	14	12	12
Substitution	10	10	10
Outbound cancellation	16	12	10
Substitution	10	10	10
Total passengers	7,360	7,360	7,360
Disrupted passengers	16	90	180

Load Factor

Higher load factors of flights leave less space for passenger reassignment and lead to more disrupted passengers and longer flight and passenger delay. For comparison, the aircraft size is decreased or increased without changing the total passenger numbers to represent the changes of load factors. It is assumed that load factors on short-haul flights are larger than those on long-haul flights. The results, as shown in Table 4, indicate that with the increase of load factors, there are fewer cancellations but more disrupted passengers. If load factors decrease, intuitively there will be fewer disrupted passengers. Nevertheless, cancellations also decrease. This is because with lower load factors, passengers are easier to reassign, so flight delays become smaller, which leads to less cancellation. Hence, Hypothesis 2 is accepted.

CONCLUSIONS

A strategy was proposed for airline recovery from schedule perturbation caused by adverse weather or other temporary events. Capacity reduction at a hub airport caused by these temporary events is reflected as reduced slots for airlines from GDPs. Under severe circumstances, it is suggested that legacy airlines with hub-and-spoke networks implement ground transportation into their networks. This strategy is called real-time intermodal substitution (RTIMS) in this study. RTIMS is different from air–rail cooperation for cutting short-haul flight frequency in daily operations. It is triggered only by severe demand and supply imbalance at major hub airports, and it performs operational integration of airside and ground transportation. Mathematical programming was presented to help airlines make decisions on whether and how to delay flights, cancel flights, or substitute flights with motor coaches. The proposed model has captured passenger reassignment at the hub airport, allowing aircraft to be swapped among flights that are assigned to utilize the same type of aircraft. The model traces the number of disrupted passengers and takes this into account in the total cost. An approximation algorithm was proposed to get a solution while avoiding the substantial computation time required in solving large-scale nonlinear integer programming problems. Experimental data were used and results under scenarios of with-TIMS and without-TIMS were compared. As validations, two hypotheses were tested, and both were accepted.

TABLE 4 Sensitivity Analysis, Load Factor

	Load Factor (aircraft capacity)		
Objective Function	0.82 0.70	0.88 0.80	0.93 0.89
Total cost ($)	212,141	220,188	228,084
Total arrivals and departures	80	80	80
Inbound cancellation	12	14	12
Substitution	10	10	10
Outbound cancellation	12	16	14
Substitution	10	10	10
Total passengers	7,360	7,360	7,360
Disrupted passengers	0	16	43

NOTATION

$t \in \Gamma$: the set of discrete time periods.
$a \in \mathrm{A}$: arrival flights.
$f \in \Phi$: departure flights.
$s \in \Theta$: spoke airports.
$k \in \mathrm{K}$: aircraft type.
AT_a : scheduled arrival time of arrival flights.
DT_f : scheduled departure time of departure flights.
AAT_a : airborne times of arrival flights.
ADT_f : airborne times of departure flights.
SpokeA_{as} : indicators of origins of arrival flights.
SpokeF_{fs} : indicators of destinations of departure flights.
TypeA_{ak} : indicators of aircraft type of arrival flights.
TypeF_{fk} : indicators of aircraft type of departure flights.
TypeO_k : number of type k aircraft available at the beginning of schedule perturbation.
TypeT_k : number of type k aircraft required to be at the hub airport at the end of schedule perturbation.
AHPax_a : schedule local passengers of arrival flight a.
DHPax_f : scheduled local passengers of departure flight f.
$\mathrm{Pax}_{\mathrm{AF}}$: scheduled transfer passengers from arrival flight a to departure flight f.
APax_a : total scheduled passengers on arrival flight a.
DPax_f : total scheduled passengers on departure flight f.
AHCap_t : arrival capacity at the hub airport in time period t.
DHCap_t : departure capacity at the hub airport in time period t.
BAT_a : bus transportation time if arrival flight a is substituted by ground transportation mode.
BDT_f : bus transportation time if departure flight a is substituted by ground transportation mode.
AD_a : delay of arrival flight a.
DD_f : delay of departure flight f.
DP : number of disrupted passengers.
BO : bus operating cost.
IPax_s^τ : inbound transfer passengers with destination s up to time period τ.
OPax_s^τ : outbound transfer passengers with destination s up to time period τ.
$\mathrm{IAircraft}_k^\tau$: inbound type k aircraft up to time period τ.
$\mathrm{OAircraft}_k^\tau$: outbound type k aircraft up to time period τ.

ACKNOWLEDGMENT

The authors thank the University of California Transportation Center for supporting this research.

REFERENCES

1. Bratu, S., and C. Barnhart. Flight Operations Recovery: New Approaches Considering Passenger Recovery. *Journal of Scheduling,* Vol. 9, No. 3, 2006, pp. 279–298.
2. Teodorović, D., and S. Guberinić. Optimal Dispatching Strategy on an Airline Network After a Schedule Perturbation. *European Journal of Operational Research,* Vol. 15, 1984, pp. 178–182.
3. Teodorović, D., and G. Stojkovic. Model for Operational Daily Airline Scheduling. *Transportation Planning and Technology,* Vol. 14, 1990, pp. 273–285.
4. Jarrah, A., G. Yu, N. Krishñamurthy, and A. Rakshit. A Decision Support Framework for Airline Flight Cancellation and Delays. *Transportation Science,* Vol. 27, No. 3, 1993, pp. 266–280.
5. Yan, S., and D.-H. Yang. A Decision Support Framework for Handling Schedule Perturbation. *Transportation Research, Part B,* Vol. 30, No. 6, 1996, pp. 405–419.
6. Cao, J.-M., and A. Kanafani. Real-Time Decision Support for Integration of Airline Flight Cancellation and Delays, Part I: Mathematical Formulation. *Transportation Planning and Technology,* Vol. 20, 1997, pp. 183–199.
7. Cao, J.-M., and A. Kanafani. Real-Time Decision Support for Integration of Airline Flight Cancellation and Delays, Part II: Algorithm and Computational Experiments. *Transportation Planning and Technology,* Vol. 20, 1997, pp. 201–217.
8. de Neufville, R., and A. Odoni. *Airport Systems: Planning, Design, and Management.* McGraw–Hill, New York, 2003.
9. Luo, S., and G. Yu. Airline Schedule Perturbation Problem: Landing and Takeoff with Nonsplitable Resource for the Ground Delay Program. In *Operations Research in the Airline Industry* (G. Yu, ed.), Kluwer Academic, Boston, Mass., 1998.
10. Luo, S., and G. Yu. Airline Schedule Perturbation Problem: Ground Delay Program with Splitable Resources. In *Operations Research in the Airline Industry* (G. Yu, ed.), Kluwer Academic, Boston, Mass., 1998.
11. Milner, J. M. *Dynamic Slot Allocation with Airline Participation.* PhD dissertation. Massachusetts Institute of Technology, Cambridge, 1995.
12. Carlson, P. Exploiting the Opportunities of Collaborative Decision Making, a Model and Efficient Solution Algorithm for Airline Use. *Transportation Science,* Vol. 34, No. 4, 2000, pp. 381–393.
13. Filar, J. A., M. Manyem, and K. White. How Airlines and Airports Recover from Schedule Perturbation: A Survey. *Annals of Operations Research,* Vol. 108, 2001, pp. 315–333.
14. Kohl, N., A. Larsen, J. Larsen, A. Ross, and S. Tiourine. *Airline Disruption Management: Perspectives, Experiences and Outlook.* Informatics and Mathematical Modelling, Technical University of Denmark, Lyngby, 2004.
15. Karow, M., and J.-P. Clarke. Virtual Hubs: A Schedule Recovery Approach. Presented at AGIFORS Airline Operations Conference, Rome, 2002.
16. Czyzyk, J., M. Mesnier, and J. Moré. The NEOS Server. *IEEE Journal on Computational Science and Engineering,* Vol. 5, No. 3, 1998, pp. 68–75.
17. Gropp, W., and J. Moré. Optimization Environments and the NEOS Server. In *Approximation Theory and Optimization* (M. D. Buhmann and A. Iserles, eds.), Cambridge University Press, London, 1997, pp. 167–182.
18. Dolan, E. *The NEOS Server 4.0 Administrative Guide.* Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill., 2001.

The Aviation Group sponsored publication of this paper.