Multi-objective Integer Programming Model and Algorithm of the Crew Pairing Problem in a Stochastic Environment

DEYI MOU
Civil Aviation University of China
College of Science
Jinbei Road 2898, 300300, Tianjin
CHINA
dymu@cauc.edu.cn

YINGNAN ZHANG
Civil Aviation University of China
College of Science
Jinbei Road 2898, 300300, Tianjin
CHINA
zhangyingnan1026@126.com

Abstract: Crew scheduling is an important production planning of airlines. Being optimized crew scheduling could make full use of human resources, and reduce operating costs. The traditional airline crew scheduling model is deterministic and does not include potential disruptions due to weather, air traffic control, etc. To take into account of effects of random factors such as weather, air traffic control, passenger demand, etc., we develop a stochastic chance-constrained programming model (SCCPM) for minimizing the crew cost and maximizing the passenger satisfaction. Based on Monte Carlo method, Back Propagation (BP) neural network and genetic algorithm, we develop a hybrid intelligent algorithm to solve the model. To evaluate the robustness of the model, the signal to noise ratio (SNR) method is included in this paper. We present computational results which show the effectiveness of our SCCPM and the hybrid intelligent algorithm.

Key–Words: airline operations, crew scheduling, passenger satisfaction, stochastic chance-constrained programming, hybrid intelligent algorithm.

1 Introduction

Flight scheduling, the optimization process of fleet resources, is a very important production plan of airlines. Crew scheduling is one of the flight scheduling problems, therefore, how to effectively make use of fleet resources is the problem to be solved for various airlines.

In aviation industry crew cost is the second largest after fuel among all the operating costs, which includes employee wages, welfare and crew schedule costing. However, most of crew cost is controllable, which is different from the cost of aviation fuel. Just by optimizing the scheduling can save tens of thousands of expenses for airlines. Even slight improvements can be translated into significant savings. Hence, the study of crew scheduling is very necessary.

For previously described reasons, airline crew scheduling has gained considerable attention. The airline crew scheduling problem has been formulated as a set partitioning problem. In practice, the crew scheduling problem is divided into two separate issues: the first step is the crew pairing; the second step is to assign crew. The legitimate crew pairings must first be generated in accordance with legal provisions and the relevant provisions of airlines, and then search for the optimum pairing according to the requirements of the decision makers, such as the one which minimize the paring cost. In this paper we focus on the pairing optimization model and its algorithm.

In most of the domestic and foreign literature models are considered to be deterministic in a static environment, without taking into account of the operation uncertainties in crew scheduling. Some representative works [1-4] appeared in literatures several years ago. In recent years, crew scheduling problems with stochastic conditions have been studied in this field. For example, Yen et al [5] established stochastic integer programming model for the airline crew scheduling problem and developed a branching algorithm to identify expensive flight connections and branch on multiple variables, finally find alternative solutions. They have demonstrated the effectiveness of branching algorithm. Yan et al [6] developed a stochastic-demand scheduling model considering stochastic disturbances of aviation passengers, and employed arc-based and route-based strategies to develop two heuristic algorithms to solve the model. Tekiner et al [7] proposed a conventional mathematical programming model to solve the crew pairing problem, considering incorporating the set of selected pairings into the model while keeping the increase in the crew cost in an acceptable range. Zhang et al [8] put forward three new crew scheduling models in an uncertain environment and designed a new model solving algorithm. Combined with characteristics of crew scheduling, Zhao [9] put forward a solution of crew scheduling based on adaptive genetic algorithm. He did some research on implementation of genetic algorithm and its application to crew scheduling. Mou et al [10] put forward the concept of probability of the crew delay, presented analytical formula, and constructs the robust mathematical model of the crew pairing based on minimizing the probability of the crew delay. They carried out some computational experiments using Matlab and compared the results of the two models.

From the view of mathematical programming, their models belong to the stochastic expected value model. Because of the diversity of the decision-making in real situations, chance-constrained programming, which is another sort of the stochastic programming, may be more suitable for some problem. To the best of the authors' knowledge, there has been no research on crew scheduling problems where the chance-constrained programming is used as a modeling tool. To remedy this research, we first utilize the chance-constrained programming concept to develop a crew scheduling model, from the basis of the airline's perspective.

In this paper we assume that no flights are cancelled and each flight operates its original duty. We first present the traditional airline crew pairing model in Section 2. We introduce our stochastic chance-constrained programming model in Section 3. To solve the model, we develop our algorithm in Section 4. In Section 5 we present a method to evaluate the robustness of the model. A real-life example and its computational results appear in Section 6. Finally, we conclude our paper and point out future research directions in Section 7.

2 Traditional Crew Pairing Model

In the traditional airline crew pairing model, the objective is to minimize the cost of paring, the constraint is the basic covering one, that is each flight is covered. The general model is established as a deterministic integer programming, where the costs are taken as fixed value, the decision variables are binary. The crew scheduling problem (CSP) model [5] is as follows:

$$\begin{cases} min \quad \sum_{j=1}^{M_D} c_j x_j \\ s.t. \\ \qquad \sum_{j=1}^{M_D} a_{ij} x_j = 1, \quad \forall i \in F \\ \qquad x_j \in \{0,1\}, \qquad \forall j \in D. \end{cases} \tag{1}$$

Where c_j means the cost of the *j*th pairing, (a_{ij}) is a 0-1 matrix, if flight leg i is covered by pairing j, $(a_{ij}) = 1$, otherwise 0. And M_D is the number of possible pairings, M_F is the number of flight legs in considered. D is denoted as the set of all legal pairings, F means the set of all flight legs. Finally, x_j equals to 1 if pairing *j* is selected, and otherwise 0.

3 Stochastic Chance-Constrained Programming Model

The planned cost was only considered so as to find the optimal objective value in the traditional crew pairing model. In fact, the cost is uncertain until flight has finished. The crew cost is at least as large as its planned cost. Exactly speaking, The crew cost is the sum of the planned cost and the extra cost. Therefore, it's necessary to introduce some stochastic factors into the model in order to ensure its practicality. The most important factor affecting the crew cost is flight time. Flight delays are expensive and lead to loss of time, money and passengers' trust. In the improved model, we will introduce delay costs and passenger satisfaction as two random factors based on the time disruption.

We assumed that ξ_i^a is the flight delay time of flight leg *i*, ϕ_i^a is the ground delay time. For convenience, define the flight delays,

$$\xi_i = \begin{cases} \xi_i^a & \xi_i^a > 0 \\ 0 & \xi_i^a \le 0 \end{cases}, \tag{2}$$

and ground delays,

$$\phi_i = \begin{cases} \phi_i^a & \phi_i^a > 0 \\ 0 & \phi_i^a \le 0 \end{cases}. \tag{3}$$

The delay cost is approximated as a nonlinear function of the delay time: $\tilde{c}(\xi_i, \phi_i)$, $i \in F$. When the delayed flight leg *i* is covered by pairing *j*, we add the delay cost to the planned cost c_j, and the actual cost of pairing *j* is $c_j^* = c_j + \sum_{i=1}^{M_F} \tilde{c}(\xi_i, \phi_i)$. For any crew pairing *j*, let c_j^* be its expected crew cost under delays. Generally, the actual cost of any pairing j can be described as,

$$c_j^* = c_j + \sum_{i=1}^{M_F} a_{ij} \tilde{c}(\xi_i, \phi_i), \ j \in D \tag{4}$$

The airline belongs to service industries. The level of the service is a significant indicator for their image and public praise among the people. Airlines should not only focus on their own profit, but also consider the customers' feeling. In consideration of these reasons, we introduce an indicator named as customer/passenger satisfaction.

Yu [11], Mao et al [12] pointed out that passenger sanctification (PS) was affected by taking off and landing time, flight time, delay time and other factors. We construct linear approximations of the basic satisfaction degree for each pairing and the satisfaction degree under delay situations. Denote the satisfaction of pairing *j* without considering the delay as s_j. We assume that the PS's affected-function by delays is $\tilde{s}(\xi_i, \phi_i),\ i \in F$. Therefore, the PS under delay disruption becomes:

$$s_j^* = \lambda_1 s_j + \lambda_2 \frac{\sum_{i=1}^{M_F} a_{ij}\tilde{s}(\xi_i, \phi_i)}{count^n}, \quad j \in D. \tag{5}$$

Where

$count^n$ - the number of flights included in pairing *j*.

λ_k - the weights of the basic satisfaction and the satisfaction effected by delays for *k*=1,2.

Remarks on formula (5): If a selected pairing j could cover 5 flight legs, $\sum_{i=1}^{M_F} a_{ij}\tilde{s}(\xi_i, \phi_i)$ means the sum of the satisfaction degree of 5 flight legs, now we should calculate its average value to ensure the degree bands in (0,1).

In reality, companies not only consider their own benefits but also passengers'. This paper is based on the stochastic chance-constrained programming which initialed by Liu et al [13]. We take delay costs and passenger satisfaction into account, and establish the following model,

$$\begin{cases} min \quad P_1 d_1^+ + P_2 d_2^- \\ s.t. \\ \qquad Pr\left\{ \sum_{j=1}^{M_D} c_j^* x_j - b_1 \le d_1^+ \right\} \ge \alpha_1 \\ \qquad Pr\left\{ -\frac{\sum_{j=1}^{M_D} s_j^* x_j}{count^d} + b_2 \le d_2^- \right\} \ge \alpha_2 \\ \qquad \sum_{j=1}^{M_D} a_{ij} x_j = 1 \qquad\qquad \forall i \in F \\ \qquad x_j = \{0, 1\} \qquad\qquad \forall j \in D \end{cases} \tag{6}$$

Where

$count^d$ - the number of pairings included in the decision variables.

$d_1^+,\ d_2^-$ - the positive and negative deviation variables, are both greater than or equal to 0, which on representative of decision value over the target part.

P_k - the priority factor of d_k.

$\alpha_k \in (0, 1]$ - the pre-given confidence level by the decision makers, $k = 1, 2$. α_k. Usually it takes a larger value.

4 Hybrid Intelligent Algorithm for This Model

The traditional method of solving the chance-constrained programming is to convert the chance constrains into their determinate equivalent forms, and then to solve the determinate model. However, this method is only applied to some special circumstances. There are two stochastic variables involved in our model, thus its highly constrained. Neither the classical algorithm nor the converting method above can be effective. In this paper, we employ a intelligent algorithm which was mentioned by Liu [13] in his monograph. This algorithm can solve the general chance-constrained programming. The main thought of this algorithm is to convert the stochastic model to the certain one. First, we adopt stochastic simulation technologies to produce the delay time length as the input data, and then calculate the output data according to a formulation we have designed. These inputs and outputs are used as the training samples for BP neural network. Second, we train a BP neural network as the fitting function in the next stage. Finally, we use the improved-genetic algorithm to find the best solution. The flowchart for this algorithm is as follows:

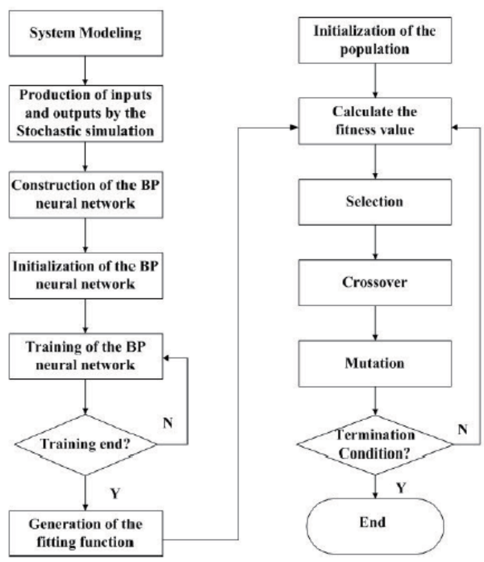

Figure 1: The flowchart for Hybrid Intelligent Algorithm

4.1 Stochastic Simulation Technologies

The SCCPM is different from the traditional model which contains some random variables. The probabilities of random events are taken as constrains. Thus, these probabilities should be obtained first. We use stochastic simulation method to compute them.

4.1.1 The Principle of Monte Carlo Technology Approaching Probability

We consider $Pr\{f(x,\xi) \leq b\}$, where $f(x,\xi)$ is a function of stochastic vector. First of all, *N* independent random vectors $\xi_n(n = 1, 2, ..., N)$ are generated from the probability distribution $\varphi(\xi)$ of random vector ξ.

Denote the number of f_n which satisfy the condition $f(x,\xi) \leq b$ as *M*. According to the Law of Large Numbers, when N is sufficiently large, $p = \dfrac{M}{N}$ is used as an approximation of $Pr\{f(x,\xi) \leq b\}$.

We know that the minimum *d* which satisfies the inequality:

$$Pr\left\{\sum_{j=1}^{M_D} c_j^* x_j \leq d\right\} \geq \alpha_1,$$

should be achieved at the equal sign, that is

$$Pr\left\{\sum_{j=1}^{M_D} c_j^* x_j \leq d\right\} = \alpha_1.$$

When comes to the inequality

$$Pr\left\{\frac{\sum\limits_{j=1}^{M_D} s_j^* x_j}{count^d} \geq d\right\} \geq \alpha_2,$$

the maximum *d* is the one we needed. It should be achieved at the equal sign also,

$$Pr\left\{\frac{\sum\limits_{j=1}^{M_D} s_j^* x_j}{count^d} \geq d\right\} = \alpha_2.$$

Here, we denote $f(\xi) = \sum\limits_{j=1}^{M_D} c_j^* x_j$, the sequence

$$f(\xi^1), f(\xi^2), ..., f(\xi^N)$$

is produced after *N* simulations.

Define $h(\xi^k)$,

$$h(\xi^k) = \begin{cases} 1 & f(\xi^k) \leq d \\ 0 & otherwise \end{cases}, \quad k = 1, 2, ..., N.$$

According to the law of large number, when *N* tends to infinity the value of

$$\frac{\sum\limits_{k=1}^{N} h(\xi^k)}{N}$$

tends to α_1 which presents the probability of $f(\xi) \leq d$. On this basis, when the sequence

$$f(\xi^1), f(\xi^2), ..., f(\xi^N)$$

is sorted in ascending order, the *N*'th element in the arrangement gives the required value, where *N'* is the integer part of $\alpha_1 * N$. Similarly, the maximum *d* for

$$Pr\left\{\frac{\sum\limits_{j=1}^{M_D} s_j^* x_j}{count^d} \geq d\right\} = \alpha_2$$

should be found at the *N*"th location of a descending order sequence, where *N"* is the integer part of $\alpha_2 * N$.

4.1.2 Monte Carlo Technology to Produce the Simulation Data

Now, define the following function,

$$U : \boldsymbol{x} \rightarrow (U_1(\boldsymbol{x}), U_2(\boldsymbol{x})) \tag{7}$$

$$\boldsymbol{x} = \{x_1, x_2, ..., x_{M_D}\}^T \tag{8}$$

Where,

$$U_1(\boldsymbol{x}) = min\left\{d \middle| Pr\left\{\sum_{j=1}^{M_D} c_j^* x_j \leq d\right\} \geq \alpha_1\right\}, \tag{9}$$

$$U_2(\boldsymbol{x}) = max\left\{d \middle| Pr\left\{\frac{\sum\limits_{j=1}^{M_D} s_j^* x_j}{count^d} \geq d\right\} \geq \alpha_2\right\}. \tag{10}$$

Then d_1^+ in model (6) is equal to $[U_1(\boldsymbol{x}) - b_1] \vee 0$, d_2^- is equal to $[b_2 - U_2(\boldsymbol{x})] \vee 0$. Next the input and output data of the function (7) should be generated. The steps for simulating $U_1(\boldsymbol{x})$ are as follows:

Step 1. Initialize the parameters: number of simulation *N*, confidential level α_1 and α_2, the number of pairings M_D, the number of flights M_F.

Step 2. Generate M_D-dimensional column vector randomly, that is $\boldsymbol{x}$.

Step 3. Generate ξ, ϕ , randomly. Note that the value generated by the random function should be treated, so as to ensure non-negative.

Step 4. Calculate the value of $\sum_{j=1}^{M_D} c_j^* x_j$.

Step 5. Repeat steps 2-4 *N* times and then get *N* values of $\sum_{j=1}^{M_D} c_j^* x_j$. Sort these values in ascending order, the *N'*th is the minimum d needed. Where *N'* is the integer part of $\alpha_1 * N$.

Thus, give an $\boldsymbol{x}$, there is a $U_1(\boldsymbol{x})$ (the minimum *d*) through *N* times' simulation. Repeat *M* times like this, we obtain *M* training samples. By these samples, train a neural network (4.2) which is used to approximate the uncertain function *U*.

4.2 Training for Neural Network

Train a neural network (the number of input neurons is M_D, that of hidden layer neurons is *y*, and the number of output neurons is 2) by the input and output data to approximate the uncertain function. The uncertain function is used as the evaluation function in the next layer algorithms. Where y is determined by the formula,

$$y < \sqrt{m+n} + a. \tag{11}$$

Where m, n represent the number of input neurons and output neurons respectively, a is a constant, $1 < a < 10$ [14].

4.3 Genetic Algorithm to Solve the Model

Firstly, the structure of solutions and the initial chromosome should be considered. The solution of the problem is 0-1 variables, therefore, the coding of the solution should also comply with this rule. For example, a chromosome structure is as follow,

1	0	0	1	0	...	1	1

Secondly, the initial chromosome must satisfy the covering constraint,

$$\sum_{j=1}^{M_D} a_{ij} x_j = 1 \quad \forall i \in F \tag{12}$$

The basic steps of genetic algorithm:

Step1. Initialize the chromosomes, compute the fitness of chromosomes by the trained neural network.

Step2. Cross, mutate, select, and compute the fitness of the offspring chromosomes by the trained neural network.

Step3. Pick up chromosomes by roulette wheel selection.

Step4. Repeat steps until the completion of the given number of cycles.

Step5. Obtain the best chromosome, adopt this chromosome as the optimal solution.

5 Robustness Analysis for the Solutions

Because of the randomness of the model, the model solution is not unique. In order to make the obtained solution more robust, we do several experiments to get some satisfactory solutions, and then select a best one among them. In this paper, we adopt a SNR approach for evaluation of the production robustness. This method was first proposed by Dr. Taguchi Gen'ichi [15].

The SNR is the ratio of signal to noise, a greater SNR indicates the products are more robust. It was divided into definite purposed characteristic, smaller-the-better characteristic, larger-the-better characteristic. Definite purposed characteristic means we hope it's better if the deviation between the product's quality and its target value is smaller. It's expected that the better the value of $\frac{\mu}{\sigma}$ greater when the relative error is considered. Where μ is the average value of the data and σ is the standard deviation. Smaller-the-better characteristic means it's hoped that the value of product's quality characteristic is the smaller the better. The quality characteristic is required to have a small average value, but also a lower fluctuation degree. Larger-the-better characteristic can be understood as the opposite of the former one, that is it's expected to have a greater quality characteristic value which is both required a larger average value and a lower fluctuation degree.

The objective of the stochastic chance-constrained programming model in this paper is to obtain the smallest deviation, therefore, we adopt the smaller-the-better characteristic of the SNR to analysis the robustness of solutions. The calculation formula is

$$-10 \lg(\frac{1}{n} \sum_{i=1}^{n} y_i^2)$$

where y_i is the objective value of the *i*th experiment, n is the experiment times. Zhang [8] used the standard deviation as the index for robustness analysis.

Actually, the standard deviation can only measure the fluctuation degree of the data. It contains both the fluctuation degree and the average value of the data if we adopt the SNR method to measure product's robustness. The results could be more comprehensive and accurate.

6 Example and Its Analysis

In order to verify the correctness of the model and demonstrate how to use the proposed algorithm to solve the model, we assume the legal crew parings as in Table 1.

The parings above demand a scheduling cycle no more than two days, the symbol '–' means paring's scheduling cycle is one day. Since the two-day parings involve overnight fee, the cost for two-day parings is more than the one-day parings' under the same flight time.

For convenience, we sort the flight number from small to big which the Flight 105 corresponds NO. 1, the Flight 110 corresponds NO. 2. The details are in Table 2.

Then the matrix$(a_{i,j})$ mentioned in Section 2 becomes a 12×28 0-1 matrix.

$$a_{i,j} = \begin{pmatrix} 0&1&0&1&1&0&1&0&0&0&0&0&0&0&0&1&0&1&0&0&0&0&1&0&0&0&0&0 \\ 0&0&1&0&0&1&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&0&0&1&1&0&0&0 \\ 0&0&0&0&0&0&0&1&0&1&1&0&1&0&0&0&0&0&1&0&1&0&0&0&0&1&0&1 \\ 0&0&0&0&0&0&0&0&1&0&0&1&0&0&0&0&0&0&0&1&0&0&0&0&0&0&1&0 \\ 0&0&0&1&0&0&1&0&0&1&0&0&1&1&0&0&0&1&0&0&1&1&1&1&1&1&1&1 \\ 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&1&1&1&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&1&1&0&0&0&0&0&0&0 \\ 0&0&0&0&1&1&1&0&0&0&0&0&0&0&0&1&1&1&0&0&0&0&1&1&1&0&0&0 \\ 0&0&0&0&0&0&0&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&1&1&0&0&0&0&0&1&1&1&0&0&0&0&1&1&1 \\ 0&0&0&1&0&0&1&0&0&1&0&0&1&1&0&0&0&1&0&0&1&1&1&1&1&1&1&1 \end{pmatrix}$$

The initial feasible solution is D1-D5-D9-D21 which meets the covering constrain. According to the study by Barnhart et al [1] in 2003, we suppose that the distribution of the flying delay time and the ground delay time is:

$$\xi \sim [-24.5 + \gamma(5.28, 5.07)],$$

and

$$\phi \sim [-0.001 + 146 * \beta(0.61, 23.6)].$$

respectively.

The delay cost function can be defined as

$$\tilde{c}(\xi_i, \phi_i) = (\xi_i^2 - 0.01\xi_i) + 0.1\phi_i,$$

the passenger satisfaction

$$\tilde{s}(\xi_i, \phi_i)$$

is defined as

$$1 - (\xi + \phi_i)/5.$$

Thus, when an $\boldsymbol{x}$ is given, a $U_1(\boldsymbol{x})$ namely d is obtained after 5000 times simulations. Repeat that process 3000 times, we get 3000 training samples which are used to train a neural network so as to approximate the uncertain function U. We get the optimal solution for the deterministic model (1) in Section 2 by the LINGO software. The selected parings are D1-D9-D11-D15-D23, and the lowest cost is 167 units, the maximum satisfaction is 93.25 %. We use these results as our targets in the SCCPM (formula (6)), that is $b_1 = 167$, $b_2 = 0.9325$, so as to searching the minimum deviation.

Since the input data is randomly generated, the format and range of input data is essential for the function fitting. We solved the deterministic model with software and obtained the best optimum solution and several optimal solutions. We found that the number of zeros in the solution, that is the number of the selected pairing, is generally not more than 5. Thus the selection of input data greatly affect degree of function fitting, in this case, the generated input data have been processed, so that in each string of data in the proportion of the number 1 is about 5/28.

The parameter and part of the program in Section 4.2 are as follows:

net =newff(minmax(Y),[10,2],'tansig','purelin','trainrp');
net.trainParam.show=50;
net.trainParam.lr=0.05;
net.trainParam.epochs=1000;
net.trainParam.goal=1e-3;
$[net, tr] = train(net, Y, DCS)$.

We use MATLAB to programme and calculate the procedures. The results can be seen in Table 3.

From Table 3, we can see paring D1-D5-D10-D20 disappears 11 times in 20 experiments, and paring D1-D5-D8-D15-D27 does 6 times. They are the high-frequency solutions. We compare their SNR values and select a better solution.

According to the formula

$$-10\lg(\frac{1}{n}\sum_{i=1}^{n} y_i^2),$$

the smaller-the-better characteristic value of paring D1-D5-D10-D20 is -17.7639, the value of paring D1-D5-D8-D15-D27 is -21.255, that is to say paring D1-D5-D10-D20 is more robust than paring D1-D5-D8-D15-D27. As we all know, the SNR value is expected to be greater. Thus, we choose the paring D1-D5-D10-D20 as the optimal solution. The average value of the 11 results is used to be the optimal value, that is the optimal cost of crew is 175.8412 and the optimal satisfaction is 0.901358. The results of the analysis for the high-frequency solutions are in Tables 4 and 5.

Table 1: The feasible crew pairings

NO.	Flight in the first day	Flight in the second day	Flight time
D1	125	105	11
D2	131	110	5
D3	131	111	5
D4	131	110-138-118	8
D5	133	110	5
D6	133	111	5
D7	133	110-138-118	8
D8	135	113	6
D9	135	114	6
D10	135	113-138-118	9
D11	136	113	6
D12	136	114	6
D13	136	113-138-118	9
D14	138	118	3
D15	131-111	–	5
D16	131-111-113	110	10
D17	131-111-133	111	10
D18	131-111-133	110-138-118	13
D19	131-111-136	113	11
D20	131-111-136	114	11
D21	131-111-136	113-138-118	14
D22	138-118	–	3
D23	138-118-133	110	8
D24	138-118-133	111	8
D25	138-118-133	110-138-118	11
D26	138-118-136	113	9
D27	138-118-136	114	9
D28	138-118-136	113-138-118	12

Table 2: The corresponding table for flight numbers and their codes in this paper

Flight No. 105	110	111	113	114	118	125	131	133	135	136	138
Code No. 1	2	3	4	5	6	7	8	9	10	11	12

Table 3: Experimental results

No.	d_1^+	d_2^-	DEVIATION	PARINGS	COST	P.S.
1	19.9195	1.1616	14.2921	1,5,10,20	186.9195	0.920884
2	10.5642	2.44	8.12694	1,5,8,15,27	177.5642	0.9081
3	14.3432	2.0664	10.66016	1,5,9,11,14	181.3432	0.911836
4	4.6613	4.5293	4.6217	1,5,10,20	171.6613	0.8872
5	5.7355	2.7524	4.84057	1,5,10,20	172.7355	0.904976
6	13.3597	4.857	10.80889	1,5,8,15,27	180.3597	0.8839
7	6.6554	7.0524	6.7745	1,8,20,23	173.6554	0.861976
8	11.8837	3.9815	9.51304	1,5,10,20	178.8837	0.892685
9	13.1552	3.7579	10.33601	1,5,10,20	180.1552	0.894921
10	14.5759	3.6619	11.3017	1,5,10,20	181.5759	0.895881
11	1.6208	1.7012	1.6449	1,5,10,20	168.6208	0.915488
12	5.6148	3.144	4.8735	1,5,10,20	172.6148	0.9011
13	8.7435	2.7711	2.5655	1,5,10,20	175.7435	0.9048
14	21.6965	4.9899	16.6845	1,8,16,27	188.6965	0.882601
15	4.9688	3.6244	4.5655	1,5,10,20	171.9688	0.8963
16	15.1097	6.4671	12.5169	1,5,8,15,27	182.1097	0.8678
17	6.3739	3.1792	5.4155	1,5,10,20	173.3739	0.900709
18	17.0626	4.0479	13.1582	1,5,8,15,27	184.0626	0.892
19	16.1959	3.0793	12.26092	1,5,8,15,27	176.3716	0.8976
20	15.0922	3.945	11.74804	1,5,8,15,27	182.0922	0.89305

Table 4: The analysis of robustness for the high-frequency solutions–Paring 1 : D1-D5-D10-D20

NO. (i)	1	2	3	4	5	6	S-T-B SNR:
optimal value (y_i)	14.2921	4.6217	4.8406	9.5130	10.3360	11.3017	
NO. (i)	7	8	9	10	11	-	$-10\lg(\frac{1}{n}\sum_{i=1}^{n} y_i^2) = -17.7639$
optimal value (y_i)	1.6449	4.8735	2.5655	4.5655	5.4155	-	

Table 5: The analysis of robustness for the high-frequency solutions–Paring 2 : D1-D5-D8-D15-D27

NO. (i)	1	2	3	4	5	6	S-T-B SNR:
optimal value (y_i)	8.12694	10.80889	12.5169	13.1582	12.26092	11.74804	$-10\lg(\frac{1}{n}\sum_{i=1}^{n} y_i^2) = -21.255$

7 Conclusion and Future Directions

In this paper, we discuss the matching optimization model for crew scheduling in uncertain environments. Crew cost and passenger satisfaction are the two random variables affected by delay time.

Since the model is different from the traditional deterministic model, a hybrid intelligent algorithm consisted of the Monte Carlo method and the BP neural network and the genetic algorithm are used to solve the model. Firstly, we simulate the large amounts of data as a BP neural network training samples by the Monte Carlo method. Then we utilize these data to train a neural network (in this case the network is a mapping of $\mathbf{x}$ and d), which is used as the fitness function in the next step. Finally, the genetic algorithm cross, mutate and select until the maximum number of obtain iterations to obtain the optimal solution. MATLAB software is used to programme and realize the algorithm.

In the end, we give an example. Due to the uncertainty of the model, multiple runs of the solution are not unique. But even in this uncertain environment, the solution still has strong robustness. Smaller-the-better characteristic SNR is used to measure the calculated results in this paper. The measurement results show that the algorithm have strong robustness, also verify the feasibility of this model which can be applied to practical problems.

However, the airline customer is usually divided to two classes, business and economy. Different factors of service quality and other influences which are important according to the customer class. In this paper we only consider one situation. Due to that the large number of samples are demanded, the Monte Carlo simulation process is time-consuming (20 minutes each time) which in future studies needs to be improved. How to improve the efficiency of the algorithm to solve large-scale crew scheduling problems will be the next focus of our studies.

Acknowledgements: This research was supported by the Fundamental Research Funds for the Central Universities (GRANT: ZXH2011008) and Research Foundation of Civil Aviation University of China (2011Kys03).

References:

[1] Barnhart C, Belobaba P, Odoni A R. Applications of operations research in the air transport industry. *Transportation Science, Vol.* 37(4), 2003, pp. 368-391.

[2] Yan S, Tu Y P. A network model for airline cabin crew scheduling. *European Journal of Operational Research, Vol.* 140, 2002, pp. 531-540.

[3] Chu H D, Gelman E, Johnson E L. Sovling large scale crew scheduling problems. *European Journal of Operational Research, Vol.* 97, 1997, pp. 260-268.

[4] Klabjan D, Johnson E L, Nemhauser G L *et al.* Airline crew scheduling with time windows and place-count constraints. *Transportation Science, Vol.* 36(3), 2002, pp. 337-348.

[5] Yen J W, Bride J R. A Stochastic Programming Approach to the Airline Crew Scheduling Problem. *Transportation Science, Vol.* 40(1), 2006, pp. 3-14.

[6] Yan S, Tang C, Fu T. An airline scheduling model and solution algorithms under stochastic demands. *European Journal of Operational Research, Vol.*190, 2008, pp. 22-39.

[7] Tekiner H, Birbil S et al. Robust crew pairing for managing extra flights. *Computers & Operations Research, Vol.*36,2009, pp. 2031-2048.

[8] Zhang J, Zhou J, Zheng F. New Stochastic Set Partitioning Models for Airline Crew Scheduling Problem under Uncertainty. *Proceedings of the Sixth International Conference on Information and Management Sciences*. Lhasa, China, 2007, pp. 372-379.

[9] Zhao H, *Flight Crew Scheduling Algorithm and Its Application*. Nanjing University of Aeronautics and Astronautics,M. D thesis, 2007.(in Chinese)

[10] Mou D, Wang Z, Xia Q. Robust Crew Pairing Problem Based on Probability of Crew Delay.*Journal of Systems & Management, Vol.* 20(2), 2011, pp. 207-212.(in Chinese)

[11] Yu J. Research of Airlines Customer Satisfaction Based on Structural Equation Model. *Journal of Civil Aviation University of China, Vol.* 256, 2007, pp. 38:45.(in Chinese)

[12] Mao M, Zhu J.Airlines Customer Satisfaction Testing Model And Case Study. *Journal of Transportation Systems Engineering and Information Technology, Vol.* 105, 2010, pp. 201-207.(in Chinese)

[13] Liu B, Zhao R, Wang G. *Uncertain Programming with Applications*, Tsinghua University Press, 2003.(in Chinese)

[14] LI Xiao-yan. *Study on Optimized Grey Neural Network Prediction Models*. Wuhan University of TechnologyM. D thesis ,2009.(in Chinese)

[15] Liu W. *Design of Experiments*, Tsinghua University Press, 2005.(in Chinese)

[16] Massoud B. *Airline Operations and Scheduling*, Ashgate Press, 2010.

[17] Schaefer A J, Johnson E L, Kleywegt A J, et al. Airline Crew Scheduling Under Uncertainty. *Transportation Science, Vol.* 39(3), 2005, pp. 210-223.

[18] Myungsook An, Yonghwi Noh. Airline customer satisfaction and loyalty: impact of inflight service quality. *Service Business, Vol.* 3, 2009, pp. 293-307.