Research Article

Optimizing Bus Frequencies under Uncertain Demand: Case Study of the Transit Network in a Developing City

Zhengfeng Huang,^{1,2} Gang Ren,¹ and Haixu Liu³

¹ School of Transportation, Southeast University, Nanjing 210096, China

² Department of Civil and Environmental Engineering, University of Wisconsin, Madison, WI 53705, USA

³ School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China

Correspondence should be addressed to Gang Ren; rengang@seu.edu.cn

Received 26 March 2013; Revised 28 April 2013; Accepted 28 April 2013

Academic Editor: Valentina E. Balas

Copyright © 2013 Zhengfeng Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Various factors can make predicting bus passenger demand uncertain. In this study, a bilevel programming model for optimizing bus frequencies based on uncertain bus passenger demand is formulated. There are two terms constituting the upper-level objective. The first is transit network cost, consisting of the passengers' expected travel time and operating costs, and the second is transit network robustness performance, indicated by the variance in passenger travel time. The second term reflects the risk aversion of decision maker, and it can make the most uncertain demand be met by the bus operation with the optimal transit frequency. With transit link's proportional flow eigenvalues (mean and covariance) obtained from the lower-level model, the upper-level objective is formulated by the analytical method. In the lower-level model, the above two eigenvalues are calculated by analyzing the propagation of mean transit trips and their variation in the optimal strategy transit assignment process. The genetic algorithm (GA) used to solve the model is tested in an example network. Finally, the model is applied to determining optimal bus frequencies in the city of Liupanshui, China. The total cost of the transit system in Liupanshui can be reduced by about 6% via this method.

1. Introduction

It is important to determine the optimal transit frequencies when public transport issues, such as network planning or operation plan scheduling, are being decided. In these frequency determination models, passengers' costs and operating costs minimization objective is usually constructed. As to the basic data on transit network passenger flow, it is acquired by either the passenger counting method or a model of passenger route choice behavior [1–5]. In these previous studies, the collected data have not taken the uncertain demand into account. However, passenger demand is actually uncertain, due to the effects of many factors, such as socioeconomic characteristics, population development, land use property, changing travel patterns, and emergency traffic incidents. And the complexity of the prediction process can also make passenger demand uncertain. Therefore, the results of transit frequency determining model will be more useful and robust if uncertain passenger demand is considered. Most studies

[6, 7] in road network assumed the uncertain travel demand can be described as a probability distribution function with mean and variance known. So, the primary objective of this study is to optimize bus frequencies with the eigenvalues of uncertain demand known. More specifically, the deducted process of bus frequencies determining objective by analytical method is the main focus of this paper.

Transit operation performance indices based on likelihood measures are more likely to be given attention in studies on transit network design that incorporate uncertain demand. For instance, it is assumed that the demand and running time of transit are stochastic. Then, systemwide travel time reliability, schedule reliability, and direct boarding waiting-time reliability are defined from the perspectives of the community, transit administration, and the operator and passengers separately [8]. However, few researches have focused on other transit problem related to uncertain demand. In the aspect of road network, many traffic network design problems have investigated the uncertain demand.

For example, a genetic algorithm (GA) combined with a simulation technique has been used to solve urban road transportation discrete network design problems by considering travel demand variation and link capacity degradable [9]. Sensitivity-based, scenario-based, and min-max models have been proposed to develop robust optimal improvement schemes that desensitize urban road transportation continuous network system performance to uncertain demands or allow the system to perform better in the face of a worst-case demand scenario [10]. Provided that highways were built and are operated based on the BOT scheme, the optimal road toll price with demand uncertainty can be determined by considering tradeoffs among operators, road users, and the government [11]. Regarding the above network design problems of capacity expansion or congestion pricing, the sampling simulation method is adopted to deal with uncertain demand. However, this method requires too much running time, and it cannot be applied to large networks. Therefore, several analytical methods have been proposed to counter weak solution efficiency. For instance, the point approximation method has been adopted to determine the near-optimal road toll price under uncertain demand [12], while this method has only been shown to provide a good solution in a small test network. Fortunately, an analytical method in the road network design proposed by the literature [13] has the advantage of efficient computation and large network application. Once the demand conforms to Poissons' distribution, the objective form of total travel time can be acquired by his method. Referring to this idea of objective deduction, an analytical method is proposed in the present study to determine the optimal bus frequencies. By the method of this paper, the uncertainty of demand can be propagated to flow according to the transit assignment probability, and then the total cost of transit system can be obtained efficiently with the uncertain flow by the analytical method.

In the remainder of this study, a bilevel model of transit frequency determination is formulated first. In the lower-level model, transit link's proportional flow eigenvalues are acquired based on optimal strategy transit assignment model. In the upper-level model, transit network cost is combined with transit network robustness performance (indicated by the variance in passenger travel time) as the upper-level objective (total cost of transit system), overall size of the bus fleet, and line bus number requirements make up the constraints. Then, the transit frequency optimization problem can be solved with GA. At last, this model is tested with cases of small transit route network and the Liupanshui transit route network in China.

2. Model Development and Analysis

A change in transit frequency can affect not only line capacity but also stop waiting time. Variation in waiting time will affect the passengers' route choice strategy. Therefore, in the previous frequency determination studies by bi-level model, the lower-level model fully considers the effect of frequency

on passengers' route choice strategy. But the selected upper-level objectives (such as passengers' total travel cost [2], the network cost composed of the passengers' total travel cost, and the transit's operating cost [5]) do not accurately reflect the transit network's robustness performance because uncertain demand is not considered. In this study, the lower-level model also considers the effect of frequency. However, the upper-level objective in this study will include transit robustness performance, in addition to transit network cost. We take the passengers' travel time variance, which can reflect the risk aversion of decision maker and make sure that a wide range of demands can be met by the optimal transit frequency, as the transit robustness performance. So, this upper-level objective can contribute to determine the optimal bus frequencies of the operating company under the condition of uncertain demand.

In the bi-level model, the bus frequencies are the output variables of upper-level model and are also the input variables of lower-level model, while transit link's proportional flow eigenvalues are the output variables of lower-level model and are also the input variables of upper-level model.

2.1. Lower-Level Model. In this part, transit link's proportional flow eigenvalues are determined by two steps. First, the choosing probability of each link from each OD pair is acquired by the optimal strategy transit assignment model. Second, the mean transit link's proportional flow (expressed as mean flow on that link from one specific OD pair) and transit link's proportional flow covariance (expressed as the flow covariance between two links from one specific OD pair) is calculated by an analytical form. Then, by comparing with the simulated method in a small transit network case, the validation of the analytical transit link's proportional flow eigenvalues are studied.

2.1.1. Lower-Level Model Solution Process. The optimal strategy model [14], applied for uncongested transit network, is the basic part of the lower-level model. Its solution algorithm is similar to the labeling method of the shortest path problem, which can achieve a high-efficiency solution process. This model should be based on an extended transit network topology, not simply composed of transit lines and bus stops. In the transit network extension process, each transit line segment between two neighboring stops should be divided into waiting, in-vehicle, and alight links. On the one hand, it can prevent the coexistence of more than one link between two neighboring nodes, which is the requirement of algorithm implementation. On the other hand, it can contribute to the construction of the variables and equations in the optimal strategy model. Next, to easily depict the running/waiting time and frequency at the extended links, we will briefly introduce the extension process.

A transit line l composed of two stops can be changed to an extended network with four nodes as shown in Figure 1. In the extended network, the dotted link stands for the waiting link, the solid link stands for the in-vehicle link, and the dashed link stands for the alight link. The variables from left to right in the parentheses represent the bus running time

FIGURE 1: Extended presentation of transit network.

and frequency, respectively. The notation *inf* means that the frequency on the in-vehicle link and alight link is assumed to be infinite. Other notations related to the extended network $G = (I, A)$ are as follows:

- c_a : the consumed time on transit link a ;
- f_a : the frequency on transit link a ;
- I : nodes set, where each node $i \in I$;
- A : transit links set, where each link $a \in A$;
- $A_i^+(A_i^-)$: transit links set, where node i is the links' upstream (downstream) node;
- v_a^r : the assigned flow on link a in the optimal strategy transit assignment model for reaching node r ;
- f_a^r : the total frequency starts from the upstream node of link a , in the optimal strategy for reaching node r ;
- q_{ir} : transit travel demand from node i to destination r . The demand of node r is defined as $q_{rr} = -\sum_{i \neq r} q_{ir}$ to more easily represent the formula;
- w_{ir} : the total waiting time of the passengers at node i whose destination node is r . The expression is the total flow input at node i divided by the total frequencies on the transit links of A_i^+ . These links belong to candidate links reaching destination r .

Optimal Strategy Transit Assignment Model

Objective function (minimal expected travel cost of all the passengers):

$$\min \sum_{a \in A} c_a v_a^r + \sum_{i \in I} w_{ir}. \quad (1)$$

Transit link flow conservation constraint:

$$\sum_{a \in A_i^+} v_a^r - \sum_{a \in A_i^-} v_a^r = q_{ir}, \quad i \in I. \quad (2)$$

Constraint of relation between the flow outputs at node i and the transit link flow of A_i^+ :

$$v_a^r \leq f_a^r w_{ir}, \quad a \in A_i^+, \quad i \in I. \quad (3)$$

Nonnegative transit link flow constraint:

$$v_a^r \geq 0, \quad a \in A. \quad (4)$$

The optimal strategy transit assignment model for reaching node r with the decision variable v_a^r is shown in formulae (1)–(4). Fortunately, the variable v_a^{ir} , assigned flow on link a from the OD pair ir , can also be obtained from the optimal strategy algorithm in the literature of Spiess et al. [14].

The choosing probability of travelling through link a from the OD pair ir is presented as p_a^{ir} . Once the bus frequencies are determined, the choosing probability p_a^{ir} can be a fixed value unrelated to the passenger demands in the optimal strategy model. According to this specific characteristic, the choosing probability p_a^{ir} can be expressed as follows:

$$p_a^{ir} = \frac{v_a^{ir}}{q_{ir}}, \quad i, r \in I, \quad (5)$$

where p_a^{ir} can be acquired easily. We only need to set $q_{ir} = 1$, then v_a^{ir} will be the calculated value of p_a^{ir} .

Using the same method as the road network applied, the mean value u_{ir} and the variance value σ_{ir} of uncertain transit travel demand q_{ir} is obtained by survey in advance. Then, the mean (u_a^{ir}) of v_a^{ir} and the covariance ($\text{Cov}(v_a^{ir}, v_b^{ir})$) between v_a^{ir} and v_b^{ir} can be expressed in the following two equations, according to the properties of mean and covariance:

$$u_a^{ir} = p_a^{ir} u_{ir}, \quad i, r \in I, \quad a \in A, \quad (6)$$

$$\text{Cov}(v_a^{ir}, v_b^{ir}) = p_a^{ir} p_b^{ir} \sigma_{ir}, \quad i, r \in I, \quad a, b \in A. \quad (7)$$

When $a = b$, the result of $\text{Cov}(v_a^{ir}, v_b^{ir})$ is equal to the variance of v_a^{ir} . Then, formula (7) can be denoted as $\sigma_a^{ir} = (p_a^{ir})^2 \sigma_{ir}$.

2.1.2. Validation of Transit Link's Proportional Flow Eigenvalues. The variables u_a^{ir} and $\text{Cov}(v_a^{ir}, v_b^{ir})$ are the input variables of the upper-level model; so we need to test its effectiveness. This subsection expressed the transit link flow eigenvalues as the function of transit link's proportional flow eigenvalues first. Then, its effectiveness can be tested by comparing the analytical transit link flow eigenvalues with the simulated transit link flow eigenvalues.

For the purpose of simple expression, all OD pairs are assumed to be independent of each other, and the passengers on each link come from many OD pairs. Then, according to the central limit theorem, the transit link flow follows a normal distribution, regardless of which distribution the demands follow. Consequently, the mean and variance of the transit link flow can be expressed as follows:

$$u_a = \sum_{i, r \in I} u_a^{ir}, \quad (8)$$

$$\sigma_a = \sum_{i, r \in I} \sigma_a^{ir}. \quad (9)$$

The effectiveness of formulae (6) and (7) can be tested by computing the closeness between the analytical results of (8) and (9) and its corresponding simulated results. A small transit network of six stops was borrowed from the literature of Yu et al. [2] (shown in Figure 2) to validate

TABLE 1: Mean matrix of transit travel demand.

OD	3	4	5	6
1	400	350	700	800
2	500	700	990	1200

Transit line	Line direction using stops	Bus frequency (veh/min)
l_1	1 → 3 → 5	1/5
l_2	1 → 3 → 4 → 6	1/3
l_3	2 → 3 → 4 → 5	1/3
l_4	2 → 3 → 6	1/5

FIGURE 2: Small transit network. Because there are lots of parallel transit lines passing through the same road segment, the extended transit network will certainly create many links and nodes. It is difficult to show large elements in the picture; thus extension figure is not provided herein.

the formulae (6) and (7). The bus running times for the transit links, transit line directions along stops, and bus frequencies are all shown in Figure 2. The independent OD demand follows a normal distribution, with the mean shown in Table 1 and a variance of 50. Several transit link flows are first obtained by the Monte Carlo sampling method. In this method, 1000 OD groups are generated at random; first the flows on transit links are assigned according to each OD group; then the normal distribution function can be fitted according to all the flows of each transit link at last. The variance of these flows, as shown in Table 2, is then determined according to the fitted normal distribution function. These data can be used to verify the correctness of analytical results. In Table 2, the analytical transit link passenger flow, deviation degree (defined as $|\text{simulated variance} - \text{analytical variance}|/\text{simulated variance}$), and the number of OD pairs past the transit link are also listed. The transit link $l_k^{(i \rightarrow j)}$ shown in the first line of the table represents the in-vehicle link from i to j that passed by line l_k .

Because the mean flows obtained by the two methods are almost equivalent, only the variances are compared here. It can be observed that the analytical variance and simulated variance are close. So the results of analytical method are validated. In addition, more OD pairs through a given transit link results in a smaller degree of deviation for this link. It indicates that the increase of OD pairs through one link will cause analytical variance being close to the true one, which complies with the central limit theorem. In urban transit networks, each transit link serves many OD pairs, and the analytical method is therefore more suitable for this application.

2.2. Upper-Level Model

2.2.1. Deduction of Upper-Level Objective. The SP survey presented in the literatures [15, 16] indicates that the travel time and its reliability are the two most important factors influencing passengers' travel behavior. Experienced traffic engineers generally use travel time variance as an indication of travel time reliability, which can reflect both transit network robustness performance and risk aversion of decision maker. Besides, incorporating this term as a part of the upper-level model, the passengers' benefits can be reflected. Therefore, the upper-level objective in this study uses all passengers' mean travel time and its corresponding variance to represent the passengers' benefits. To consider the benefit of transit operating company, the objective also includes the term of transit operating costs. The upper-level objective is expressed as follows:

$$\begin{aligned} \min E \left(\sum_{a \in A, r \in I} c_a v_a^r + \sum_{i, r \in I} w_{ir} \right) \\ + \alpha \text{Var} \left(\sum_{a \in A, r \in I} c_a v_a^r + \sum_{i, r \in I} w_{ir} \right) + \beta \sum_l F_l f_l. \end{aligned} \quad (10)$$

The first subterm of (10) represents the expected travel time of the passengers; the second sub-term of (10) represents the total travel time variance; the third sub-term of (10) represents the transit operating costs. The meaning of relevant parameters is shown as follows:

$E(\cdot)$: this notation is used to solve the expectation value;

$\text{Var}(\cdot)$: this notation is used to solve variance value;

F_l : the cost per frequency increases on transit line l ;

α : the coefficient showing the importance degree of uncertain demand;

β : the coefficient showing the importance degree of transit operating costs.

Because the bus frequency is the variable of upper-level model, the following work describes objective (10) as the function of frequency variable only. According to the formulae (6) and (7), the first sub-term of objective is expressed as follows after the analytical deduction:

$$E \left(\sum_{a \in A, r \in I} c_a v_a^r + \sum_{i, r \in I} w_{ir} \right) = \sum_{i, r \in I} \sum_{a \in A} \left(\frac{c_a + 1}{f_a^r} \right) u_a^{ir}, \quad (11)$$

where variable f_a^r represents the total frequencies which belong to the optimal strategy reaching the node r and starting from the upstream node of link a . The mean transit link's proportional flow u_a^{ir} from the lower-level model is known in this equation. So the bus frequency is the only

TABLE 2: Validation test of transit link's proportional flow variance.

Transit link	$l_2^{(1 \rightarrow 3)}$	$l_2^{(3 \rightarrow 4)}$	$l_2^{(4 \rightarrow 6)}$	$l_4^{(2 \rightarrow 3)}$	$l_4^{(3 \rightarrow 6)}$	$l_1^{(1 \rightarrow 3)}$	$l_1^{(3 \rightarrow 5)}$	$l_3^{(2 \rightarrow 3)}$	$l_3^{(3 \rightarrow 4)}$	$l_3^{(4 \rightarrow 5)}$
Simulated variance	76.9	35.5	0	70.7	101.5	27.7	18.6	60.1	78.6	45.7
Analytical variance	78.1	34.8	0	71.1	100	28.1	19.6	58.6	79.3	44.6
Deviation degree	1.6	2	0	0.6	1.5	1.4	5.4	2.5	0.9	2.4
OD pairs	4	2	0	4	2	4	2	3	4	2

variable here. Similarly, the second sub-term can be deduced as follows:

$$\begin{aligned}
 & \text{Var} \left(\sum_{a \in A, r \in I} c_a v_a^r + \sum w_{ir} \right) \\
 &= \sum_{i, r \in I} \left(\sum_{a \in A} \left(\frac{c_a + 1}{f_{a^-}^r} \right)^2 \sigma_a^{ir} \right. \\
 & \quad \left. + 2 \sum_{b, d \in A \cap b \neq d} \left(\frac{c_b + 1}{f_{b^-}^r} \right) \left(\frac{c_d + 1}{f_{d^-}^r} \right) \text{Cov}(v_b^{ir}, v_d^{ir}) \right), \quad (12)
 \end{aligned}$$

where the transit link's proportional flow covariance $\text{Cov}(v_b^{ir}, v_d^{ir})$ is obtained from the lower-level model.

2.2.2. Upper-Level Model Formulation. After the above deduction of transit network robustness performance, the form of upper-level model can be expressed as follows.

Upper-Level Model

Objective function: It reflects the tradeoff between passengers and the transit operating company. The passengers expect there to be higher frequencies to reduce waiting time and improve travel time robustness, whereas the company tends to provide lower frequencies due to the operating costs:

$$\begin{aligned}
 & \min \sum_{i, r \in I} \sum_{a \in A} \left(\frac{c_a + 1}{f_{a^-}^r} \right) u_a^{ir} \\
 & + \alpha \sum_{i, r \in I} \left(\sum_{a \in A} \left(\frac{c_a + 1}{f_{a^-}^r} \right)^2 \sigma_a^{ir} \right. \\
 & \quad \left. + 2 \sum_{b, d \in A \cap b \neq d} \left(\frac{c_b + 1}{f_{b^-}^r} \right) \left(\frac{c_d + 1}{f_{d^-}^r} \right) \text{Cov}(v_b^{ir}, v_d^{ir}) \right) \\
 & + \beta \sum_l F_l f_l. \quad (13)
 \end{aligned}$$

Constraint: the number of operating buses must not exceed the total number of buses:

$$\sum_l \left(\sum_{a \in l} c_a + e_l \right) f_l \leq T. \quad (14)$$

Constraint: at least one bus is guaranteed to run for each line:

$$\left(\sum_{a \in l} c_a + e_l \right) f_l \geq 1 \quad \forall l. \quad (15)$$

In the above constraints, e_l stands for the sum of the total delay time at stops and the recovery time at the terminal station; T represents the total number of buses in the transit operating company.

3. Solution Algorithm

The upper-level model provides the bus frequencies for the lower-level model, and the feedback data of transit link's proportional flow eigenvalues are given back to the upper-level model. This lead-follower or Stackelberg game will not end until the upper and lower models reach equilibrium.

Until now, the majority of bi-level model solving methods are heuristic. Specifically, when the objection function of the upper-level model is nonlinear, the feasible approaches would be classified into decomposition-based algorithm [17], gradient-based algorithm [18], and intelligent algorithm [19]. A shortcoming is that the algorithm may converge to a local optimum solution, existing in the first and second types. Fortunately, the algorithm from the third type is mainly used to achieve a global optimum solution. Because of the complexity of this bi-level model, extensively used intelligent algorithm—GA (as shown in Figure 3)—is adopted to solve this problem. Its key modules are given in the following subsections.

3.1. Encoding. There are many transit lines in the urban transit network. In this situation, the binary code of frequencies may reduce the efficiency of the searching procedure and also make the result inaccurate; thus it is not selected as the coding mode. Here, a real-coded scheme is selected to represent frequencies, and a population member is therefore described as follows:

$$\{f_1, f_2, \dots, f_n\}, \quad (16)$$

where n represents the number of transit lines.

3.2. Initial Population Generation. To ensure that the frequencies in the initial population satisfy the constraints (14) and (15) and diversification, this study generates an adequate random number of buses per line and then changes them to the initial frequencies. The total number of buses is known as

FIGURE 3: GA for bus frequency determination with uncertain demand.

T in priori, the number of operating buses at line i ($i \in I$) is denoted by T_i ; then the frequency generation process of initial population is described as follows.

We will perform probabilistic method in the following loop to produce a series of random bus number vectors. First, for all the bus line i ($i \in I$), uniformly generate a random number of buses from the feasible integer set $[1, T_i - n + 1]$; second, give up this vector if the total bus number is larger than T . This loop will not end until the required population size is satisfied. After this process, all these numbers are transformed into frequencies. This process can not only satisfy constraints (14) and (15), but also ensure the diversification of initial population P_0 .

3.3. Fitness Computation. In the GA searching method, the total cost of transit system is used as the fitness. If the constraint cannot be met under the corresponding individual, the fitness should be equal to a large number M . Because the minimal fitness of the individual is the objective, this large number can be a penalty for the inappropriate individual. Then, this individual will be more easily eliminated in the genetic operation. If the individual can meet all the constraints, the optimal strategy transit assignment process for each individual should be conducted to determine the corresponding fitness value.

3.4. Genetic Operation. The genetic operation adopted in this study is as follows: in the selection part, the Roulette Wheel Selection Method is used; in the crossover part, the two-point crossover method is used; in the mutation part, the Gaussian mutation method is used. All the above

operators are described in detail by the genetic algorithm toolbox developed by Sheffield University. In addition, the termination criterion is that best fitness does not change during 10 successive generations.

4. Test Network

4.1. Small Test Transit Network. The case of Figure 2 is used to illustrate the effectiveness of the algorithm. The running time and OD demand of the aforementioned transit links are adopted here. The cost of each line per unit frequency is assumed to be 3000, which has been transformed to time cost (min), and therefore $\beta = 1$. For each line, the line delay is $e_l = 0$, and total number of buses is 32. According to the literature of [16], the magnitude of travel time variance is investigated to be higher compared to the expected total travel time. Additionally, the survey from [20] showed that the weight-coefficient ratio has an upper bound of 2.25. To be compromised, this study sets $\alpha = 1.5$. For the genetic parameters, the number of initial population is set at 40. To rapidly create some fitness close to optimum fitness in the stochastic searching process, we set the crossover rate at a high value equal to 0.7. To satisfy Holland's schema theorem (denoted as a necessary condition for reaching optimum solution proposed by Holland [21]) that fitter individual can increase exponentially in successive generations, we set the mutation rate at a low value equal to 0.3. The punishment value M is set at 10^7 that is big enough compared to the size of objective value.

According to the above parameter values, the acquired terminal objective is 40,147 min, and the optimal frequencies are $\{0.69, 1.15, 0.58, 1.11\}$ veh/min.

When α and β are changed, the corresponding objective is shown in Table 3.

If α is fixed, an increase in β could lead to a decrease in operating cost and an increase in travel time eigenvalues. It indicates that if decision maker attaches more importance to operating cost, the more passengers' benefits will be removed. The reason is that under low operating cost, fewer bus frequencies would be generated, which could bring increased travel time to the passenger demand.

If β is fixed, an increase in α could lead to the opposite results of the previous situation. It indicates that the transit network robustness performance can be guaranteed at the cost of operating company's benefit. The reason is that the increasing of bus frequencies under higher operating cost could satisfy more uncertain demand. This data development trend could also indicate that varied risk aversion degrees (the amount of unsatisfied demand) may be reflected by giving different values to α .

4.2. Test Transit Network of Liupanshui City. The algorithm was applied in Liupanshui transit network whose peak-hour transit OD demand is acquired through the combination of household travel surveys and on-board surveys. Liupanshui is a medium-size city in western China, with an urban area of 293 km² and a population of 460,000, and its transit network is uncongested. The transit network is composed

TABLE 3: Objective analysis according to different parameter values.

α	β	Frequencies	Mean travel time	Travel time variance	Operating cost	Objective
0	1	{0.35, 0.81, 0.60, 0.60}	23917	6892	7118	31035
	1.5	{0.33, 0.74, 0.44, 0.56}	25072	7655	6176	34336
	2	{0.34, 0.56, 0.31, 0.48}	27055	8922	5084	37222
0.5	1	{0.42, 0.93, 0.61, 0.78}	23013	6210	8957	35075
	1.5	{0.35, 0.77, 0.49, 0.59}	24313	7320	7820	39703
	2	{0.35, 0.62, 0.33, 0.48}	25201	8201	6340	41982
1	1	{0.53, 1.02, 0.57, 0.91}	22140	5802	9358	37300
	1.5	{0.41, 0.79, 0.52, 0.62}	23936	6578	8201	42816
	2	{0.43, 0.71, 0.51, 0.55}	24289	7365	6983	45620
1.5	1	{0.69, 1.15, 0.58, 1.11}	21335	5485	10584	40147
	1.5	{0.59, 0.96, 0.56, 0.73}	22677	6169	8511	44695
	2	{0.46, 0.83, 0.56, 0.58}	23783	6780	7283	48519

TABLE 4: Transformed time cost per frequency and line delay (min).

Line	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$F_l(\times 10^4)$	4.48	3.92	4.32	4.20	4.92	3.40	3.84	3.40	3.80	3.40	3.76	3.48	3.96	3.32	4.84	3.32
e_l	12	14	12	8	12	11	10	9	9	11	10	10	11	10	9	11

of 16 lines, 270 stops, and 260 buses, and it resembles a long strip connecting the eastern and western ends of the city, which is shown in Figure 4. It is assumed that the true OD demand follows a normal distribution with the obtained demand value as the mean. According to the literature [9], the standard deviation of demand can be set as the mean value multiplied by uncertain demand coefficient γ (this case sets $\gamma = 0.2$ as Xu and Gao's [9] set). Transformed time cost per frequency F_l and line delay e_l , whose values should refer to the line bus type and line length, are investigated by the bus operating company of Liupanshui and shown in Table 4. Population number and other parameters have the same values as those of the former case.

The computation time of the GA procedure takes around 113 min with processor of Intel Core i7-3520M CPU 2.9 GHz. Their convergence process is shown in Figure 5. The final resulting frequencies are {0.95, 0.21, 0.28, 0.35, 0.28, 0.67, 0.9, 0.64, 0.4, 0.75, 1.07, 0.7, 0.57, 0.51, 0.18, 0.21} veh/min. There are three lines (1, 7, 11) possessing high frequencies. Reasons are listed below.

- (i) The travel demands between east and west districts are high. Both lines 1 and 7 are serving the demands of eastwest direction.
- (ii) Line 1 runs through central business district and line 7 passes the train station.
- (iii) Line 11 is only 5 kilometers long and operates in the central business district.

Furthermore, compared with the current situation (shown in Figure 5) in the city, implementing the optimal solution from this study could decrease the value of the objective by about 6%.

The effect of uncertain demand coefficient γ and fleet size T on objective is shown in Figure 6, where the mean travel demand is assumed to be fixed.

When the fleet size was fixed, the upper-level objective was upward-sloping with the increase of uncertain demand coefficient γ . It showed that the travel demand variance increasing can lead to the objective value increase (how the components of objective affected by the variation of γ will be shown in Figure 7).

When uncertain demand coefficient γ was fixed, the increase of fleet size T can generally reduce the objective value. But in the condition that a very small value was assigned to γ , the objective value was almost unchanged. It indicates that the uncertainty of demand has impact on the use of total buses. If γ is small, only few buses need to be dispatched to improve the objective. In this condition, many buses will sit idle in the operating company. If γ is large, more buses need to be scheduled to meet the uncertain demand. So, according to the value of γ , the adequate fleet size could be determined. If the previous field case is under consideration, the uncertain demand coefficient $\gamma = 0.2$. Then from the Figure 6, we know that $T = 280$ would be the adequate bus number to not only prevent the waste of buses but also satisfy the uncertain demand.

The impact of uncertain demand coefficient γ on objective terms is shown in Figure 7, where T was equal to 260. When uncertain demand variance γ increased, the operating cost and travel time variance would increase simultaneously whereas the mean travel time would decrease. Why operating cost and travel time variance have the same trend of variability with the change of γ ? The reason is that the increased concerns on passengers' benefits would induce more operating cost to be spent at high bus frequencies. The reason of the reverse variability of mean travel time is that the

FIGURE 4: Transit network of Liupanshui city.

FIGURE 5: GA convergence process with crossover and mutation rate equal to 0.7 and 0.3.

improvement of bus frequencies can reduce the total travel time of the mean passengers.

5. Conclusion

- (1) This paper has analyzed optimal bus frequencies determination under uncertain demand by applying a bi-level model. The lower-level model assigns a mean number of passengers to the transit network and

FIGURE 6: The effect of uncertain demand coefficient γ and fleet size T on objective.

acquires the transit link's proportional flow variance, based on the optimal strategy; the upper-level model determines the line frequencies, aiming to minimize the total cost of the transit system, subject to the constraints on the overall size of the bus fleet and line fleet. Uncertain demand, passengers' benefits, and bus operating cost from the company are all taken into account in this model.

- (2) Results from a small test network shows that the GA method is successful in solving the bus frequencies determination problem, and this method was also applied in the uncongested city of Liupanshui in western China. The total cost of the transit system

FIGURE 7: The impact of uncertain demand coefficient γ on objective terms with fleet size $T = 260$.

in Liupanshui can be reduced by about 6% via this method.

- (3) Sensitivity analysis shows the relationship among uncertain demand coefficient, fleet size, and upper-level objective. Particularly at low uncertain demand coefficient, any increase in fleet size causing no change in the upper-level objective. This rule would be beneficial for the determination of bus-buying plan.
- (4) Future studies can be identified as follows. First, basic parameters such as the cost of each line per unit frequency should be determined with more accuracy. Second, scientific survey methods should be developed to acquire the eigenvalues of transit OD demand and the coefficient of total travel time variance. Third, the objective under a congested transit network should be deduced using the analytical method, which will be useful for optimizing bus frequencies of a metropolitan city with demand uncertainty.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (nos. 51078086, 51278101, and 51278429). The authors thank the Jiangsu Provincial Key Laboratory of Transportation Planning and Management of Southeast University, China.

References

- [1] A. Ceder, "Bus frequency determination using passenger count data," *Transportation Research Part A*, vol. 18, no. 5-6, pp. 439-453, 1984.
- [2] B. Yu, Z. Yang, and J. Yao, "Genetic algorithm for bus frequency optimization," *Journal of Transportation Engineering*, vol. 136, no. 6, pp. 576-583, 2010.
- [3] S. I. J. Y. Chien and C. F. M. Tsai, "Optimization of fare structure and service frequency for maximum profitability of transit systems," *Transportation Planning and Technology*, vol. 30, no. 5, pp. 477-500, 2007.
- [4] W. H. K. Lam, J. Zhou, and Z. H. Sheng, "A capacity restraint transit assignment with elastic line frequency," *Transportation Research Part B*, vol. 36, no. 10, pp. 919-938, 2002.
- [5] Z. Gao, H. Sun, and L. L. Shan, "A continuous equilibrium network design model and algorithm for transit systems," *Transportation Research Part B*, vol. 38, no. 3, pp. 235-250, 2004.
- [6] S. V. Ukkusuri, T. V. Mathew, and S. T. Waller, "Robust transportation network design under demand uncertainty," *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 1, pp. 6-18, 2007.
- [7] A. Sumalee, K. Uchida, and W. H. K. Lam, "Stochastic multimodal transport network under demand uncertainties and adverse weather condition," *Transportation Research Part C*, vol. 19, no. 2, pp. 338-350, 2011.
- [8] Y. Yin, W. H. K. Lam, and M. A. Miller, "A simulation-based reliability assessment approach for congested transit network," *Journal of Advanced Transportation*, vol. 38, no. 1, pp. 27-44, 2003.
- [9] L. Xu and Z. Gao, "Bi-objective urban road transportation discrete network design problem under demand and supply uncertainty," in *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL '08)*, pp. 1951-1955, September 2008.
- [10] Y. Yin, S. M. Madanat, and X. Y. Lu, "Robust improvement schemes for road networks under demand uncertainty," *European Journal of Operational Research*, vol. 198, no. 2, pp. 470-479, 2009.
- [11] A. Chen and K. Subprasom, "Analysis of regulation and policy of private toll roads in a build-operate-transfer scheme under demand uncertainty," *Transportation Research Part A*, vol. 41, no. 6, pp. 537-558, 2007.
- [12] L. M. Gardner, A. Unnikrishnan, and S. T. Waller, "Solution methods for robust pricing of transportation networks under uncertain demand," *Transportation Research Part C*, vol. 18, no. 5, pp. 656-667, 2010.
- [13] A. Sumalee, D. P. Watling, and S. Nakayama, "Reliable network design problem: case with uncertain demand and total travel time reliability," *Transportation Research Record*, no. 1964, pp. 81-90, 2006.
- [14] H. Spiess and M. Florian, "Optimal strategies: a new assignment model for transit networks," *Transportation Research Part B*, vol. 23, no. 2, pp. 83-102, 1989.
- [15] TranSystems Corporation, Planners Collaborative, and T. Crikelair, "Elements needed to create high ridership transit systems," TCRP Report 32, Transportation Research Board, 2006.
- [16] T. Uchida and Y. Iida, "Risk assignment: a new traffic assignment model considering the risk travel time variation," in *Proceedings of the 12th International Symposium on Transportation and Traffic Theory*, pp. 89-105, 1993.
- [17] C. Suwansirikul, T. L. Friesz, and R. L. Tobin, "Equilibrium decomposed optimization: a heuristic for the continuous equilibrium network design problem," *Transportation Science*, vol. 21, no. 4, pp. 254-263, 1987.
- [18] T. J. Kim and S. Suh, "Toward developing a national transportation planning model: a bilevel programming approach for Korea," *The Annals of Regional Science*, vol. 22, no. 1, pp. 65-80, 1988.

- [19] Y. Yin, “Genetic-algorithms-based approach for bilevel programming models,” *Journal of Transportation Engineering*, vol. 126, no. 2, pp. 115–119, 2000.
- [20] W. B. Jackson and J. V. Jucker, “An empirical study of travel time variability and travel choice behavior,” *Transportation Science*, vol. 16, no. 4, pp. 460–475, 1982.
- [21] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, MIT Press, 1992.