

## Research Article

# Modelling Stochastic Route Choice Behaviours with a Closed-Form Mixed Logit Model

Xinjun Lai and Jun Li

Guangdong Provincial Key Laboratory of Intelligent Transportation System, School of Engineering,  
Sun Yat-sen University, Guangzhou 510270, China

Correspondence should be addressed to Jun Li; [stsljijun@mail.sysu.edu.cn](mailto:stsljijun@mail.sysu.edu.cn)

Received 5 June 2014; Revised 16 September 2014; Accepted 25 September 2014

Academic Editor: Kang Li

Copyright © 2015 X. Lai and J. Li. 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.

A closed-form mixed Logit approach is proposed to model the stochastic route choice behaviours. It combines both the advantages of Probit and Logit to provide a flexible form in alternatives correlation and a tractable form in expression; besides, the heterogeneity in alternative variance can also be addressed. Paths are compared by pairs where the superiority of the binary Probit can be fully used. The Probit-based aggregation is also used for a nested Logit structure. Case studies on both numerical and empirical examples demonstrate that the new method is valid and practical. This paper thus provides an operational solution to incorporate the normal distribution in route choice with an analytical expression.

## 1. Introduction

Route choice is one of the crucial issues in transportation analysis because it models the travelling behaviours so as to provide predictions for the future demand. Drivers always try to maximize their travelling welfare when choosing a path from a given origin-destination (OD) pair. However not all of them choose the best alternative because of the imperfect knowledge of network. To model this perception error and the stochastic route choice behaviour, Probit and Logit models are two of the most widely used methods. The utility of each alternative is decomposed into a deterministic and a random portion. Assume that there are  $n$  paths between an OD pair and the route choice set is  $K$ ; the utility (welfare) of an alternative path  $i$  can be represented as

$$U_i = V_i + \eta_i, \quad i \in K, \quad (1)$$

where  $U_i$  is the utility of path  $i$ ,  $V_i$  is the deterministic part which is composed by attributes such as length and cost that can be explicitly captured, and  $\eta_i$  is the random term that captures the perception error. A rational traveller would select a path with the maximum utility among the alternatives in  $K$ .

Probit assumes that the random portion is normally distributed; besides, it provides a highly flexible structure for

correlation. However, it is limited due to the computation burden. It does not have a closed-form formula when there are more than two alternatives. Generally, the computation of multinomial Probit requires either Clark's approximation [1, 2], Monte Carlo simulation [3], or numerical integration [4]. Yai et al. [5] used the multinomial Probit model in the context of route choice in the Tokyo rail network, but the maximum number of alternatives is limited to four. On the other hand, Logit is more popular for its analytical tractability. Logit assumes that the error term is type I extreme value (EV) distributed. Moreover, it assumes each of the error terms is independently identically distributed (IID), which leads to a closed-form mathematical structure to simplify the computation in estimation and prediction. As a consequence, Logit has two main disadvantages because of the IID assumption: (1) it cannot represent the path correlation which leads to enlarged probabilities of the overlapped paths, namely, the overlapping problem and (2) it cannot represent the heterogeneity in perception errors which would produce unreasonable results, namely, the scaling problem.

There are several modified methods to address the two drawbacks of Logit in the context of route choice. Regarding the first disadvantage, the overlapping problem, the improved models are classified into two types.

- (i) Modifications of multinomial Logit (MNL), such as path size Logit (PSL) [6–9], C-Logit [10], and implicit availability/perception (IAP) model [11]: in these models, an additional term is introduced in the utility function to capture the correlations of paths, so as to decrease the attraction of the overlapped path. This method maintains the simple form of MNL. Besides, the log-likelihood function of this method is globally concave, so it guarantees a global optimum for parameters estimations. However, the additional terms are convenient approximations. Previous researches show that they might be too sensitive to the composition of the choice set [9, 12].
- (ii) Generalized extreme value (GEV) proposed by McFadden [13]: the most widely used methods of this type in route choice are the link-based crossed nested Logit (CNL) [14, 15] and the paired combinatorial Logit (PCL) [16–19] model. These two models all have a tree structure to represent the link-path relation, where alternatives with shared attributes are classified into the same nest so the correlation can be explicitly captured. The link-based CNL model treats each link as a nest, and each path uses several links which are classified into the corresponding nests. The PCL model compares paths by paired combinations, and each path pair is a nest. The CNL model has a large set of parameters that need to be estimated, so some researches provide approximated formulas [14, 20]. Besides, some researchers suggest that the parameters can be achieved by solving a system of equations of the correlation and constraints [21, 22]. Likewise, the PCL model also requires a parameter to represent the correlation, and the specifications are provided by Gliebe [23] and Prashker and Bekhor [24].

As for the second drawback of Logit, the scaling problem, Pravinongvuth and Chen [18] propose origin-destination specific scaling factor to represent the different scale of diverse networks. Chen et al. [25] examine the scaling effect when applying route choice model in stochastic equilibrium models. Miwa et al. [26] examine how to set the scale parameter (dispersion parameter) and apply a multiclass stochastic user equilibrium (SUE) assignment model to consider differences in drivers' perception errors.

Some researches combine both advantages of Probit and Logit, and the most representative model is the mixed Logit [27], also named as Logit kernel, error component, or hybrid Logit. It incorporates other distributions other than type I EV to provide a flexible and tractable form to represent the correlation across alternatives, the alternative specific variances, and also taste heterogeneity. Frejinger and Bierlaire [8] use the error component to model the subnetwork so as to represent the path overlap in an abstract network. Bekhor et al. [28] estimate an error component model based on the Boston route choice data. However, the mixed Logit does not have a closed-form expression; consequently the estimation and prediction all require the simulation-based method. Researches [20, 29] show that the simulation-based method requires a large number of draws to achieve stable

predictions. Besides, currently there is no efficient path-based SUE traffic assignment for solving the route choice model with the mixed Logit model [18].

To fully use the advantages of Probit and Logit, this paper proposes a mixed Logit method with a closed-form to model the stochastic route choice behaviours. With a closed-form expression, the computation burden in estimation and prediction would be relieved. Moreover, the closed-form formula alleviates the difficulties in the path-based SUE assignment with a mixed Logit model. The paper is organized as follows. Section 2 describes the methodology, including the nested model structure and the Probit-based aggregation. The validation from a numerical example is presented in Section 3. The new method is applied to real data in Section 4. Finally conclusions and discussions for future study are given in Section 5.

## 2. Methodology

*2.1. Model Structure.* The proposed model has a similar structure as the PCL model. Paths are compared by paired combination. Consider  $n$  alternatives in the choice set  $K$  between an OD pair; by paired combination there are totally  $N = n(n - 1)/2$  paths pairs. The new model has a two-level nest structure. Each path pair is a nest; within the nest there is actually a binary choice case. The expected maximum utility [7] of each path pair is used as the utility of the nest. In the upper level, it is a multinomial choice model with  $N$  nests. Consider a three-alternative case, as shown in Figure 1, paths  $i$ ,  $j$ , and  $k$ , and the path pairs are  $(i, j)$ ,  $(i, k)$ , and  $(j, k)$ . The probability that path  $i$  is chosen among three paths is a combination of the marginal probability  $P(i, j)$  of the nest and the conditional probability  $P[i | (i, j)]$  within the nest, which is

$$P(i) = P(i, j) \cdot P[i | (i, j)] + P(i, k) \cdot P[i | (i, k)]. \quad (2)$$

In order to relax constraints of the Logit model, the error part  $\eta_i$  in (1) is decomposed into two parts,  $\varepsilon_i$  and  $\xi_i$ . The first error  $\varepsilon_i$ , which is IID EV distributed, captures the differences of nests; the second error  $\xi_i$ , which obeys normal distribution, captures the differences within the nest. For the first nest that includes paths  $i$  and  $j$ , their utilities are

$$\begin{aligned} U_i &= V_i + \xi_i + \varepsilon_i, \\ U_j &= V_j + \xi_j + \varepsilon_j, \end{aligned} \quad (3)$$

where  $V_i = \beta X_i$  and  $V_j = \beta X_j$ ,  $X_i$  and  $X_j$  are attributes of paths  $i$  and  $j$ ;  $\beta$  is a vector of parameters that to be estimated.  $\varepsilon_i$  and  $\varepsilon_j$  are nest specific, and they capture the unobserved attributes shared by alternatives in the same nest, so consequently they are the same, which is  $\varepsilon_i = \varepsilon_j = \varepsilon_{(i,j)}$ . The variance of nest  $(i, j)$  is  $\sigma_{(i,j)}^2 = \pi^2/6\theta^2$ , where  $\theta$  is the scale parameter.  $\xi_i$  and  $\xi_j$  are alternative-specific, and they capture the unobserved attributes specific to alternatives  $i$  and  $j$ .  $\xi_i$  and  $\xi_j$  are assumed to be normally distributed with the expectation of zero and the variances  $\sigma_i^2$  and  $\sigma_j^2$ , respectively. Sheffi [30] argues that the variance can be assumed to be

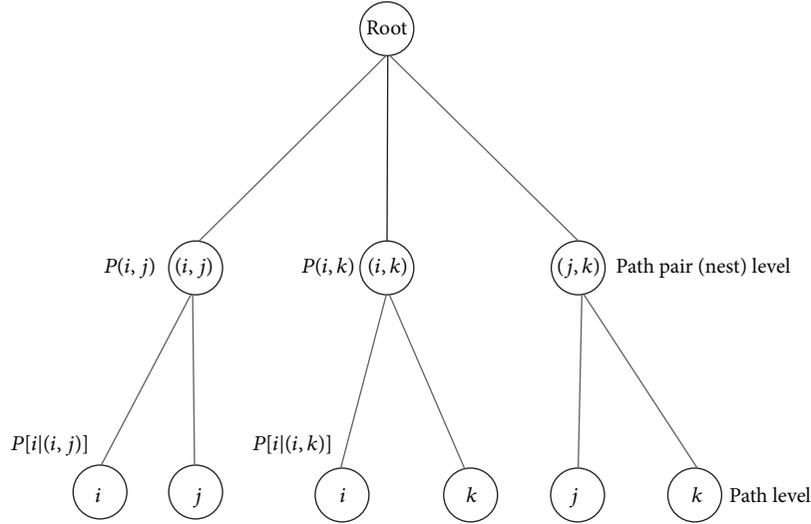


FIGURE 1: Model structure of a three-alternative case.

proportional to the deterministic utility of paths, so as to link the perception error to the paths attributes. Hence the variance-covariance matrix of paths  $i, j$  is

$$\Sigma = \begin{bmatrix} \lambda V_i & \lambda C_{ij} \\ \lambda C_{ij} & \lambda V_j \end{bmatrix}, \quad (4)$$

where  $C_{ij}$  is the correlated utility of paths  $i$  and  $j$ , such as the overlapped links;  $\lambda$  is the scale parameter of the lower-level and it is to be estimated.

The probability that traveller chooses path  $i$  given the nest  $(i, j)$  is chosen as

$$\begin{aligned} P[i | (i, j)] &= \Pr(U_i > U_j) \\ &= \Pr(V_i + \xi_i + \varepsilon_i > V_j + \xi_j + \varepsilon_j) \\ &= \Pr(\xi_i - \xi_j > V_j - V_i). \end{aligned} \quad (5)$$

Because  $\xi_i$  and  $\xi_j$  are both normally distributed with zero means,  $\xi = \xi_i - \xi_j$  is also normally distributed with expectation zero but with variance  $\sigma^2 = \sigma_i^2 + \sigma_j^2 - 2\sigma_i\sigma_j\rho_{ij} = \lambda(V_i + V_j - 2C_{ij})$ , where  $\rho_{ij}$  is the correlation. We have

$$\begin{aligned} P[i | (i, j)] &= \int_{\xi=-\infty}^{V_i-V_j} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi}{\sigma}\right)^2\right] d\xi, \quad \sigma > 0 \\ &= \Phi\left[\frac{V_i - V_j}{\sigma}\right], \end{aligned} \quad (6)$$

where  $\Phi(\cdot)$  is the standard cumulative normal distribution function. Since the Probit model is only used in a binary choice case with a closed-form expression, the advantages of Probit, such as flexible form of correlation and alternative-specific variance, could be fully exploited in our model.

The expected maximum utility of each nest is the aggregation of the paths within the nest; define it as  $V_{(i,j)}$  for nest  $(i, j)$ . The utility of nest  $(i, j)$  is

$$U_{(i,j)} = V_{(i,j)} + \varepsilon_{(i,j)}, \quad (7)$$

where  $\varepsilon_{(i,j)} = \varepsilon_i = \varepsilon_j$ .

The probability that nest  $(i, j)$  is chosen among  $n(n-1)/2$  nests is

$$\begin{aligned} P(i, j) &= \Pr[U_{(i,j)} > U_{(r,s)}, \forall r \neq s; r, s \in K] \\ &= \Pr[V_{(i,j)} + \varepsilon_{(i,j)} > V_{(r,s)} + \varepsilon_{(r,s)}, \forall r \neq s; r, s \in K] \\ &= \Pr[\varepsilon_{(i,j)} - \varepsilon_{(r,s)} > V_{(r,s)} - V_{(i,j)}, \forall r \neq s; r, s \in K], \end{aligned} \quad (8)$$

where  $\varepsilon_{(r,s)} \forall r \neq s; r, s \in K$  is IID EV distributed. According to the properties of EV distributed [7], we have

$$P(i, j) = \frac{\exp[\theta V_{(i,j)}]}{\sum_{r \neq s; r, s \in K} \exp[\theta V_{(r,s)}]}. \quad (9)$$

The probability that traveller chooses path  $i$  in the choice set  $K$  is

$$P(i) = \sum_{i \in K, i \neq j} P(i, j) \cdot P[i | (i, j)]. \quad (10)$$

**2.2. Calculation of the Nest Utility.** The nest utility  $V_{(i,j)}$  in (7) is represented by the expected maximum utility of paths  $i$  and  $j$ . In order to fully use the advantage of the Probit model, this paper incorporates the Probit-based aggregation to represent the utility of the nest.

Clark [1] provided an approximation method with which the expected maximum utility of paths  $i$  and  $j$  is

$$E[\max(U_i, U_j)] = V_i \Phi\left(\frac{V_i - V_j}{\alpha}\right) + V_j \Phi\left(\frac{V_j - V_i}{\alpha}\right) + \alpha \phi\left(\frac{V_i - V_j}{\alpha}\right), \quad (11)$$

where  $\phi(\cdot)$  is the standard normal distribution function and  $\alpha$  is given by

$$\alpha = V_i^2 \sigma_i^2 + V_j^2 \sigma_j^2 - 2\sigma_i \sigma_j \rho_{ij} V_i V_j. \quad (12)$$

Noticing that  $\Phi(-x) = 1 - \Phi(x)$ , the right-hand side of (11) can be simplified as

$$V_j + (V_i - V_j) \Phi\left(\frac{V_i - V_j}{\alpha}\right) + \alpha \phi\left(\frac{V_i - V_j}{\alpha}\right). \quad (13)$$

The Probit-based aggregation which is used in (9) is

$$V_{(i,j)} = E[\max(U_i, U_j)]. \quad (14)$$

### 3. Numerical Example

The new model is tested in a three-path network with one independent path and two correlated paths, as shown in Figure 2. Three paths have the same lengths of 10. The value of  $x$  varies from 0 to 10. When  $x = 0$ , the middle and lower paths are completely correlated, so the choosing probabilities for the upper, middle, and lower paths are 1/2, 1/4, and 1/4, respectively. When  $x = 10$ , three paths are all independent, so the probabilities are 1/3 for all of them. When  $0 < x < 10$ , the probabilities for the middle and lower paths should monotonically increase from 1/4 to 1/3, while the probability for the upper route should decrease from 1/2 to 1/3. The purposes of this numerical example are firstly to check whether the models have reasonable outcomes in extreme cases, which are when  $x = 0$  and when  $x = 10$ , and secondly to check whether the shapes of the models are as the same as expected when  $0 < x < 10$ , which means whether they are concave or convex and monotonously increase or decrease.

The proposed model is compared with the Probit, PCL, CNL, MNL, C-Logit, PSL, and Logit kernel (LK) models. The results of Probit, calculated by Clark's approximation, are assumed to be the standard. The new model is actually an improved version of the PCL model with mixed distribution. The LK model is compared because it is also a mixed Logit model but without a closed-form. The CNL is compared because it is also a nested structure as the proposed one. The PSL and C-Logit models are the two most widely used models for their simplicity. The MNL, without addressing the overlapping or scaling issues, is supposed to be the worst. Assume that length is the only attribute in the deterministic utility and its parameter is set to be one. The choosing probability of the middle path in MNL, namely,  $P(2)$ , is

$$P(2) = \frac{\exp(-\theta L_2)}{\sum_{i=1,2,3} \exp(-\theta L_i)}, \quad (15)$$

where  $L_i$  is the length of path  $i$ . The scale parameter  $\theta$  is normalized to one.

The link-based CNL model [14] is chosen because it is the most widely used CNL model in the route choice field because it systematically captures the route-link relations in the road network, where each link corresponds to a nest and the paths that share the same link belongs to the same nest. Therefore we choose the formula as follows:

$$P(2) = \frac{\exp(-\theta L_2 + \ln G_2)}{\sum_{i=1,2,3} \exp(-\theta L_i + \ln G_i)},$$

$$\ln G_i = \ln \sum_{m=1}^4 \left[ \mu I_{im} e^{-L_i(\theta_m - 1)} \left( \sum_{j \in C} I_{jm} e^{-\theta_m L_j} \right)^{(\theta - \theta_m)/\theta_m} \right], \quad (16)$$

where  $G$  is the GEV generating function,  $G_i$  is its partial derivative with respect to  $\exp(-\theta L_i)$ , and  $I_{im}$  is the inclusive parameter and is defined as  $I_{im} = l_m / L_i$ , where  $l_m$  is the length of link  $m$  and  $L_i$  is the length of path  $i$ .  $\theta$  is the root scale parameter and is normalized to 1;  $\theta_m$  is the scale parameter for nest  $m$  and is set as 1.5 in our case. Since this case is just illustrative, a larger or a smaller setting of the parameter  $\theta_m$  would not change the major properties of the model, which means that it would not change its value when  $x = 0$  or when  $x = 10$  and it would not change the concave, convex, or monotonic properties. Without loss of generality, in this case we set the nest parameter  $\theta_m = 1.5$ .

The choice probability of the middle path in the LK model is calculated by

$$P(2) = \exp\left(-\theta L_2 - \lambda_{LK} \sqrt{l_d} \xi\right) \times \left( \exp(-\theta L_1) + \exp\left(-\theta L_2 - \lambda_{LK} \sqrt{l_d} \xi\right) + \exp\left(-\theta L_3 - \lambda_{LK} \sqrt{l_d} \xi\right) \right)^{-1}, \quad (17)$$

where  $l_d$  is the length of link  $d$ ,  $\lambda_{LK}$  is the scale parameter and is set to be 1 in this case just for illustration, and  $\xi$  is a random number from the standard normal distribution. Simulation-based method [28] is used with one million draws.

The formulas of PSL, C-Logit, and PCL are corresponded to the researches of Ben-Akiva and Bierlaire [6], Cascetta et al. [10], and Koppelman and Wen [17], respectively. Probit employs the scale parameter  $\lambda$ , whereas the MNL, PCL, CNL, PSL, and C-Logit models employ the parameter  $\theta$ . To ensure consistency among different models, we assume that perception errors are the same, which is  $\pi^2/6\theta^2 = \lambda L_j$ . In this case all the paths lengths are the same so the variances are the same.

The choice probabilities for the middle route calculated by different models are provided in Figure 3. When  $x = 0$  and  $x = 10$ , the new model has the expected results, so do the PCL, PSL, and C-Logit models; however both the CNL and the LK models fail to produce reasonable value when  $x = 0$ . When  $0 < x < 10$ , the curve of the new model is close to the Probit and they both show a downward concave

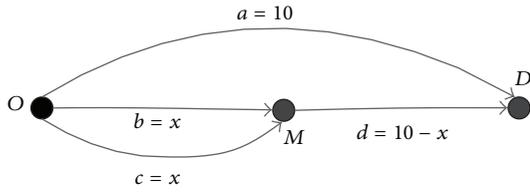


FIGURE 2: A three-path network.

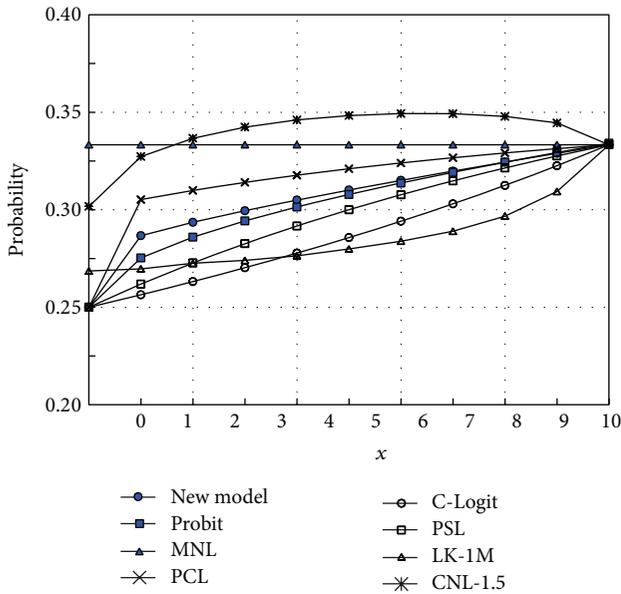


FIGURE 3: Probabilities of selecting the middle route in network 1.

shape. Besides, the result from the new model demonstrates a substantial improvement over the PCL model. The results from PSL and C-Logit also have reasonable performances. However, the results from the CNL model have bizarre behaviours. It overpredicts the choosing probabilities while the paths are partly correlated. The curves of Probit, PSL, PCL, and the new model are upward concave, while the curves for C-Logit and LK are downward concave. The results from MNL are  $1/3$  regardless of the variation of  $x$ . It fails to provide logical prediction while paths are overlapped, as expected. According to the results and the comparisons in this case, the new model is valid and capable to produce reasonable outcomes.

#### 4. Empirical Results

In order to evaluate the performances of the proposed method, we apply the new model to real data. A case study of taxi drivers choosing routes in the city center is presented. The studied city, Guangzhou, is situated in the southern China and it has approximately ten million inhabitants. Only the central business district (CBD), the Tianhe region as highlighted in Figure 4, is studied. The data set for the estimation is from GPS-equipped taxis when they were carrying passengers. The data was collected by



FIGURE 4: Map of the studied region from Guangzhou city.

a management company for monitoring purpose but not for navigation, so the route choice behavior is based on the drivers' own judgment. The vehicles were monitored within a radius of 5 km in the CBD, and 5786 trips from 473 ODs are collected for case study. The information on the studied network is shown in Table 1. Three data sets are collected: the first one is for estimating the parameters of the new model and the compared models, which will be presented in Sections 4.1, 4.2, and 4.3; the last two data sets are for validating forecasting, which is to use the estimated models to predict the choosing probabilities and compare them with the actual choosing behaviors, and the results are shown in Section 4.4.

**4.1. Model Specification.** Three attributes, length, artery road ratio, and the number of signal-controlled intersections, are included into the utility function, as shown in Table 2. Length and time are two highly similar and correlated attributes, so only one of them would be sufficient into the utility function. However a precise actual travelling time is difficult to obtain before departure. When drivers decide which route to choose, they usually process the information which is more stable in their concept, in this case length is relatively more stable than time. Therefore we use length in the utility other than time. The unit of length is kilometer so its magnitude is similar with other attributes for the convenience of the estimation. The artery road ratio is the length of the artery road (major roads and arterial streets) divided by the total length of the trip. We assume that artery roads have a significant and positive impact in the utility, because, compared with minor streets, the artery roads have more lanes, higher capacity, and even less traffic lights in the studied region, which means higher level of service when driving. Therefore the larger value of the artery road ratio is supposed to be more attractive to the drivers and this attribute is included. The number of signal-controlled intersections is expected to have significant and negative impact when driving in the city center. More traffic lights mean more chances of stopping and delay; therefore the more intersections with traffic lights are expected

TABLE 1: Information on the studied network.

Nodes	Unidirectional links	Major roads	Arterial streets	Collector streets	Signal-controlled intersections
208	662	24	34	32	57

TABLE 2: Statistics on routes corresponding to observations.

Attributes	Min.	Average	Max.
Length (km)	0.447	2.84	8.34
Artery road ratio	0	0.797	1
Number of the signal-controlled intersections	0	3.65	14

TABLE 3: Information on the number of actually chosen paths.

Average	Max.	Min.
4	12	2

to have a lower utility. The deterministic utility is shown in

$$V_i = \beta_L \times \text{Length}_i + \beta_{\text{ARR}} \times \text{Artery Road Ratio}_i + \beta_S \times \text{Signal}_i. \quad (18)$$

**4.2. Route Choice Set.** According to the researches from Ramming [20] and Frejinger and Bierlaire [8], it is difficult to generate a choice set including all the actually chosen paths. At best 84% of the observed paths are found by combining all the choice set generation algorithms that Ramming [20] had tested. In order to avoid generating a choice set that misses the important chosen paths, we employ a data mining method to build the choice set for each individual. Assume that we observe the trips between a given OD pair for a long enough time period  $T$ , and if there are totally  $n$  paths that are actually chosen by the travellers, we can conclude that these  $n$  paths are the choice set  $K$  of this OD. Since the GPS data is large enough and is continuously provided, it is possible to find out the choice sets for all the ODs. The advantage of this method is that it would not miss the important and actually chosen paths, but the shortcoming is that it may require a very long observation time  $T$  to determine a stable choice set  $K$ .

According to 5786 trips from 473 ODs collected by taxi drivers in a period of one week, this paper analyzes the size of the choice set, denoted by  $n$ . As shown in Table 3, the number of actual chosen paths between any OD pair are not larger than 12, and the average is just 4 paths. It suggests that it is rational to use the paired combination model in route choice, because the magnitude of  $N = n(n-1)/2$  does not lead to a heavy computation burden.

The objective of this paper is purely illustrative. It does not provide a full analysis to determine how long the observation time  $T$  should be or how large the number of the observed trips should be. More tests on this subject would be desirable.

**4.3. Model Estimation.** This paper uses the maximum likelihood estimation method to calibrate the parameters. Five

models are estimated and compared in this section: the proposed model; the MNL, expected to have a poor result; the two most widely used models with tree structure, the PCL and CNL; the LK model with 100 draws estimated by a simulated-based method [31]. Ramming [20] points out that the estimation of the Logit family either is normalizing the scale parameter as  $\theta = 1$  or alternatively is actually the jointly estimates of  $\theta\beta$ . To facilitate the comparison among different models, a scaled parameter estimate is also provided. The scaling,  $\theta\beta_L$ , is based on the estimated length parameter in the MNL model. The magnitude of the scaled estimate for the parameters is consequently the same among the models.

The signs of all the estimated parameters are as expected, as shown in Table 4. The positive sign of the parameter  $\beta_{\text{ARR}}$  suggests that taxi drivers tend to travel on the artery roads. The scaled estimates of  $\beta_{\text{ARR}}$  and  $\beta_{\text{Signal}}$  in models MNL and PCL and the new model have the same magnitude, and the magnitude of these parameters is approximately ten times smaller. The magnitude of the scaled and nonscaled estimates is the same for models PCL and CNL, but not for the new model. Actually, the log-likelihood function of the new model is not globally concave; when searching for the optimum it is easily “trapped” in a local point. The selection of the initial values, upper and lower bounds, is important in estimation. It would be a reasonable approach to estimate a MNL model first in order to gain adequate information on the parameters. The parameters estimates of all the models but LK are all significantly different from zero. Two parameters of LK,  $\beta_{\text{ARR}}$  and  $\beta_{\text{Signal}}$ , are not significantly estimated. The reason may be that the draws are not enough (only 100 draws). The estimation time for MNL, PCL, CNL, and the new model is within five minutes thanks to the closed-form expression. On the other hand, the LK model uses 19 hours, with only 100 draws, and still cannot attain good estimates.

Table 5 provides the likelihood ratio test between the models. It is asymptotically distributed as  $\chi^2$  with  $F$  degrees of freedom, where  $F$  is the difference of the numbers of estimated parameters between two models. Results from this test show that the new model and LK and CNL models are significantly better than the PCL and MNL within the 95% confidence interval. Since the number of estimated parameters for CNL and the new model is the same, it is possible to compare the goodness-of-fit by the final log-likelihood. The data shows that the new model is better than LK, but the CNL model has a better model fit than the new model.

**4.4. Forecasting.** Route choice models are important to predict individual behavior; therefore the comparison of models should not just focus on model fit, but also on the forecasting results. Two out-of-sample data sets are chosen to validate the forecasting ability of models. “Out-of-sample” means these

TABLE 4: Estimation results.

Parameters	MNL	PCL	CNL	LK	New model
$\beta_L$	-0.999	-0.812	-0.448	-0.723	-0.0184
Scaled estimates	-0.999	-0.999	-0.999	-0.999	-0.999
Std. Err.	0.0281	0.0245	0.0248	0.227	0.00218
$t$ -test 0	35.5	33.1	18.1	3.19	88.5
$\beta_{ARR}$	2.34	2.05	0.571	0.277	0.0278
Scaled estimates	2.34	1.99	0.411	0.5126	1.12
Std. Err.	0.0407	0.0329	0.0453	0.189	0.00287
$t$ -test 0	57.7	62.3	12.6	1.47	101
$\beta_{Signal}$	-0.197	-0.168	-0.0828	-0.277	-0.00273
Scaled estimates	-0.197	-0.154	-0.0279	-0.0322	-0.149
Std. Err.	0.0103	0.00860	0.0108	0.478	0.00125
$t$ -test 0	19.3	19.5	7.64	0.579	23.1
$\lambda$	—	—	—	0.537	0.0484
Scaled estimates	—	—	—	0.1356	0.559
Std. Err.	—	—	—	0.193	0.00355
$t$ -test 0	—	—	—	2.79	139
$\theta_m$	—	—	4.94	—	—
Scaled estimates	—	—	4.06	—	—
Std. Err.	—	—	0.125	—	—
$t$ -test 0	—	—	39.4	—	—
Final log-likelihood	-8220	-8130	-6930	-8710	-7990
Estimation time*	60.2 s	85.1 s	79.6 s	18.4 h	199 s

\*Computed with MATLAB on Intel i5 with 4 GB RAM, one processor.

TABLE 5: Likelihood ratio test.

Model 1	Model 2	Test	Threshold (95%)
New model	PCL	276	3.84
New model	MNL	457	3.84
CNL	PCL	2580	3.84
CNL	MNL	2760	3.84
LK	PCL	980	3.84
LK	CNL	1160	3.84

two data sets for validating forecasting have not been used for estimating the models. Each data set includes 200 OD pairs which are randomly selected in the studied region, and all observations associated with these OD pairs are included in the forecasting data sets. 3345 and 3536 observations are in the first and second data set.

To compare the forecasting power of models, we define the out-of-sample error as

$$\epsilon_f^M = \sqrt{\sum_{obs=1}^N (P_{obs,rs}^R(c) - P_{obs,rs}^M(c, \beta^M))^2}, \quad (19)$$

where  $\epsilon_f^M$  is the out-of-sample forecasting error of model  $M$ , and a smaller value would suggest a better forecasting ability in this case;  $P_{obs,rs}^M(c, \beta^M)$  is the predicted probability of the actual chosen path  $c$  computed by model  $M$  with estimated

parameter  $\beta^M$ , between OD pair  $rs$ , and for the observation obs.  $P_{obs,rs}^R(c)$  is the probability of the chosen path in reality, which is 1.  $N$  is the number of observations in the data set.

The out-of-sample forecasting errors of the models are shown in Table 6. The MNL model uses the least time for computation in each data set; however its errors are the second largest among all the compared models. Even though the LK model provides a highly flexible structure by incorporating the multinomial normal distribution, however its errors in both data sets are the largest, and its computation time is also the largest. Although the CNL model has a good model fit in the estimation, it does not provide the most promising forecasting results in this case. The PCL model and the new model, which is also a PCL-based model, provide the least forecasting errors in both data sets. Because of the mixture of the Probit model, the new model uses more computation time than the PCL, but it still uses less time than the traditional mixed Logit model (the LK model) because the new one is closed-form and it does not require the simulation-based method in computation. By summarizing the numerical example, the estimation by real data, and the out-of-sample forecasting results, this paper therefore concludes that the new model is a promising approach in route choice analysis.

## 5. Conclusions

This article proposes incorporating the normal distribution to model the perception error of the travellers in route

TABLE 6: Out-of-sample forecasting errors of compared models.

	MNL	PCL	CNL	LK	New model
Data set 1	48.1	29.9	47.0	54.7	29.1
Elapsed time*	<b>0.13 s</b>	<b>5.27 s</b>	<b>3.92 s</b>	<b>68.2 s</b>	<b>21.5 s</b>
Data set 2	48.7	30.5	48.3	56.1	29.9
Elapsed time*	<b>0.16 s</b>	<b>6.38 s</b>	<b>4.54 s</b>	<b>72.6 s</b>	<b>23.6 s</b>

\*Computed with MATLAB on Intel i5 with 4 GB RAM, one processor.

choice and at the same time retain the superiority of Logit for its closed-form expression. The traditional mixed Logit model requires a simulation-based method for estimation and prediction. This paper provides a new mixed model with an analytical formula, where the computation burden would be relieved. Besides, it would be practical to use the new model in the path-based SUE traffic assignment because of its closed-form expression. Probit model is employed in two places: firstly paths are compared by pairs where the superiority of the binary Probit can be fully used; secondly the Probit-based aggregation is also used for a nested Logit structure. The results on numerical example demonstrate that the proposed approach is valid, and the application with real data shows that the proposed method is practical. The parameter estimates show that the new model is practical, and the likelihood ratio tests present the superiority of the new model compared to the multinomial Logit and paired combinatorial Logit models. Moreover, the out-of-sample forecasting results suggest that the new model has a promising forecast power.

The proposed model provides new perspectives for the mixed Logit model in route choice modelling. It is a reasonable trade-off between the flexibility in correlation and analytical tractability. A natural extension is to loosen the assumption that the scale parameters  $\lambda$  of the normal distribution are all the same, because travellers are supposed to have different perception errors on different path pairs.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This research was funded by the National Natural Science Foundation of China (no. 51178475).

## References

- [1] C. E. Clark, "The greatest of a finite set of random variables," *Operations Research*, vol. 9, pp. 145–162, 1961.
- [2] M. J. Maher and P. C. Hughes, "A probit-based stochastic user equilibrium assignment model," *Transportation Research B: Methodological*, vol. 31, no. 4, pp. 341–355, 1997.
- [3] Y. Sheffi and W. B. Powell, "An algorithm for the equilibrium assignment problem with random link times," *Networks*, vol. 12, no. 2, pp. 191–207, 1982.
- [4] A. Rosa and M. J. Maher, "Algorithms for solving the probit path-based stochastic user equilibrium traffic assignment problem with one or more user classes," in *Proceedings of the 15th International Symposium on Transportation and Traffic Theory*, pp. 371–392, Adelaide, Australia, 2002.
- [5] T. Yai, S. Iwakura, and S. Morichi, "Multinomial probit with structured covariance for route choice behavior," *Transportation Research Part B: Methodological*, vol. 31, no. 3, pp. 195–207, 1997.
- [6] M. Ben-Akiva and M. Bierlaire, "Discrete choice methods and their applications to short term travel decisions," in *Handbook of Transportation Science*, pp. 5–34, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1999.
- [7] M. Ben-Akiva and S. R. Lerman, *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT press, Cambridge, UK, 1985.
- [8] E. Frejinger and M. Bierlaire, "Capturing correlation with subnetworks in route choice models," *Transportation Research B: Methodological*, vol. 41, no. 3, pp. 363–378, 2007.
- [9] P. H. L. Bovy, S. Bekhor, and C. G. Prato, "The path size factor revisited: an alternative derivation," in *Proceedings of the 87th Transportation Research*, Washington, DC, USA, 2008.
- [10] E. Cascetta, A. Nuzzolo, F. Russo, and A. Vitetta, "A modified logit route choice model overcoming route overlapping problems: specification and some calibration results for interurban networks," in *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France, 1996.
- [11] E. Cascetta, F. Russo, F. A. Viola, and A. Vitetta, "A model of route perception in urban road networks," *Transportation Research Part B*, vol. 36, no. 7, pp. 577–592, 2002.
- [12] J. Li, X. Lai, and Z. Yu, "Paired route impedance correction for multinomial logit model based on equivalent impedance," in *Proceedings of the 92nd Annual Meeting of Transportation Research Board*, Washington, DC, USA, 2013.
- [13] D. McFadden, "Modeling the choice of residential location," in *Spatial Interaction Theory and Residential Location*, pp. 75–96, North Holland, Amsterdam, The Netherlands, 1978.
- [14] P. Vovsha and S. Bekhor, "Link-nested logit model of route choice: overcoming route overlapping problem," *Transportation Research Record*, no. 1645, pp. 133–142, 1998.
- [15] S. Bekhor and J. N. Prashker, "Stochastic user equilibrium formulation for generalized nested logit model," *Transportation Research Record*, no. 1752, pp. 84–90, 2001.
- [16] C. Chu, "A paired combinatorial logit model for travel demand analysis," in *Proceedings of the 5th World Conference on Transportation Research*, Ventura, Calif, USA, 1989.
- [17] F. S. Koppelman and C.-H. Wen, "The paired combinatorial logit model: properties, estimation and application," *Transportation Research Part B: Methodological*, vol. 34, no. 2, pp. 75–89, 2000.

- [18] S. Pravinvongvuth and A. Chen, "Adaption of the paired combinatorial logit model to the route choice problem," *Transportmetrica*, vol. 1, no. 3, pp. 223–240, 2005.
- [19] A. Chen, P. Kasikitwiwat, and Z. Ji, "Solving the overlapping problem in route choice with paired combinatorial logit model," *Transportation Research Record*, no. 1857, pp. 65–73, 2003.
- [20] M. S. Ramming, *Network Knowledge and Route Choice*, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, UK, 2002.
- [21] A. Papola, "Some developments on the cross-nested logit model," *Transportation Research B: Methodological*, vol. 38, no. 9, pp. 833–851, 2004.
- [22] E. Abbe, M. Bierlaire, and T. Toledo, "Normalization and correlation of cross-nested logit models," *Transportation Research Part B*, vol. 41, no. 7, pp. 795–808, 2007.
- [23] J. P. Gliebe, "Route choice using a paired combinatorial logit model," in *Proceedings of the 78th Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 1999.
- [24] J. N. Prashker and S. Bekhor, "Route choice models used in the stochastic user equilibrium problem: a review," *Transport Reviews*, vol. 24, no. 4, pp. 437–463, 2004.
- [25] A. Chen, S. Pravinvongvuth, X. Xu, S. Ryu, and P. Chootinan, "Examining the scaling effect and overlapping problem in logit-based stochastic user equilibrium models," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 8, pp. 1343–1358, 2012.
- [26] T. Miwa, Y. Okada, and T. Morikawa, "Applying a structured dispersion parameter to multiclass stochastic user equilibrium assignment model," *Transportation Research Record*, no. 2196, pp. 142–149, 2010.
- [27] D. McFadden and K. Train, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, vol. 15, no. 5, pp. 447–470, 2000.
- [28] S. Bekhor, M. E. Ben-Akiva, and M. S. Ramming, "Adaptation of logit kernel to route choice situation," *Transportation Research Record*, no. 1805, pp. 78–85, 2002.
- [29] J. Walker, *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables*, The MIT Press, Cambridge, UK, 2001.
- [30] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice Hall, London, UK, 1985.
- [31] K. E. Train, *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge, UK, 2003.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

