

Research Article

Bus Route Design with a Bayesian Network Analysis of Bus Service Revenues

Yi Liu , Yuanhua Jia, Xuesong Feng , and Jiang Wu

MOE Key Laboratory for Urban Transportation Complex System Theory and Technology, Beijing Jiaotong University,
No. 3 Shangyuanacun, Haidian District, Beijing 100044, China

Correspondence should be addressed to Yi Liu; 16114209@bjtu.edu.cn

Received 27 May 2017; Revised 31 October 2017; Accepted 13 December 2017; Published 30 January 2018

Academic Editor: Erik Cuevas

Copyright © 2018 Yi Liu 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.

A Bayesian network is used to estimate revenues of bus services in consideration of the effect of bus travel demands, passenger transport distances, and so on. In this research, the area X in Beijing has been selected as the study area because of its relatively high bus travel demand and, on the contrary, unsatisfactory bus services. It is suggested that the proposed Bayesian network approach is able to rationally predict the probabilities of different revenues of various route services, from the perspectives of both satisfying passenger demand and decreasing bus operation cost. This way, the existing bus routes in the studied area can be optimized for their most probable high revenues.

1. Introduction

The bus network design problem has become one of the most common hard-to-solve issues of many cities in the world today. Unsatisfying bus services resulting from irrational bus routes make their revenues decrease. Therefore, a good design of a bus route is essential to satisfy the bus travel demand and improve the bus service revenue. Many methods for rationally designing a bus route or network have been continually developed. Ceder and Israeli [1] combined mathematical programming approaches and decision-making techniques to solve the transit network design problem. Chien and Spasovic [2] studied a grid bus transit system to optimize route spacing, station spacing, headway, and fare with the objectives of maximum total operator profit and social welfare. Lee and Vuchic [3] offered an iterative approach to solve the relationship between the variable transit trip demand and the transit network design, under a given fixed total demand. Wirasinghe and Vandebona [4] considered the express route planning problem based on both grid and nongrid road networks to minimize operating costs, passenger access costs, waiting time, and traveling time. Tirachini et al. [5] introduced a social welfare maximization model with the interplay between congestion and crowding externalities, with the aim of optimizing the design of urban bus routes. Ceder et al.

[6] used a mathematical modeling approach which includes considerations of uneven topography to design the placement of stops along a single bus route.

The majority of researchers tried to minimize the total travel time, or the generalized cost. The genetic algorithm (GA), Tabu search, simulated annealing methods, and so forth have all played important roles in recent research on bus network design. Zhao and Zeng [7] combined GA and simulated annealing to minimize transfers with reasonable route directness while maximizing service coverage. Fan and Machemehl [8] provided a multiobjective model for considering the design of public transportation networks in the case of variable demand, and the solution methodology was based on the Tabu search method. Pacheco et al. [9] proposed an approach based on a local search strategy to solve the route design and bus assignment problem under the effect of different demands. Szeto and Wu [10] combined GA and a neighborhood search heuristic to simultaneously perform the suburban route design. Nayeem et al. [11] developed two versions of GA based metaheuristics to discuss the transit routing problem, with the aim of minimizing the travel time and the number of transfers simultaneously.

Inherently, the bus service revenues are usually sensitive to passengers, operating costs, and so on. Therefore, there is a strong need to develop a new approach from a comprehensive

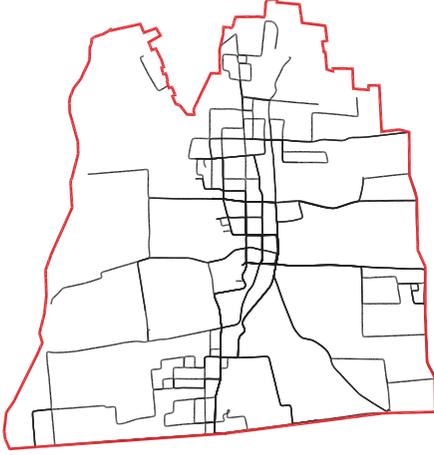


FIGURE 1: The existing layout of bus routes in X .

perspective that takes into account the dependency and uncertainty of the correlated factors of bus routes. A Bayesian network (BN) can be used to discover the overall dependency structure of a large number of variables under uncertainty and incomplete information and to predict future events such as traffic accidents in comparatively accurate manners [12–14]. As a result, this research has newly proposed an approach based on BN to predict the probabilities of different revenues of various route services to satisfy passenger demand and decrease bus operation cost. This way, the existing bus routes in the studied area can be optimized for their most probable high revenues.

This paper is organized as follows. Section 2 introduces the study area represented as X here. The proposed processes of BN learning and probabilistic inference are given in Section 3. Section 4 presents the application of the methodology to the optimization of bus routes in X . Section 5 summarizes conclusions and comments.

2. Study Area

The proposed BN has been implemented on real bus routes in X which is displayed in Figure 1. Data was obtained from the administrative statistics of the government department. The routes in Figure 1 are mainly concentrated in the north area of X . The uneven distribution of routes causes an uneven distribution of bus passenger demands, and it also causes traffic congestion which makes the traveling speed excessively low in the north area of X . It is found that the ratio of bus routes whose average load factor is below 0.3 [15] reaches 25.00%. Most routes are relatively long and their daily passenger transport volumes are comparatively low, which makes their revenue per day often unable to meet their costs. Merely about 50% of bus routes are very detoured, which causes the average nonlinear coefficient of those routes to exceed 1.4. Because of its relatively high bus travel demand and unsatisfactory bus services, X has been selected as the study area of this research to explain the proposed BN approach.

3. Methods

3.1. Bayesian Network. The model structure of a BN combining the principles from graph theory, probability theory, and statistics is known as a directed acyclic graph G defined by the respective sets of nodes and directed edges [16, 17]. The nodes represent random variables, and the directed edges represent relationships between variables. Associated with each node of a BN is a conditional probability distribution ($CPtable$) quantifying how much a node depends on its parent node(s) [18, 19].

Formally, the directed acyclic graph G encodes conditional independence between the variables. Each variable X_i is independent of its nondescendants in the graph given the state of its parents. Thus, the joint distribution of n variables X_1, \dots, X_n which can be described by (1) is the product of the conditional distributions of each variable given its parent node(s).

$$P(X_1, \dots, X_n) = \prod_{i=1}^n p(x_i | \pi_i), \quad (1)$$

where n represents the number of all variables, π_i is an instantiation of the set of parent node(s) of X_i , and x_i denotes the state of X_i .

The conditional probabilities that appear in the product form are stored in $CPtable$ as the parameters. When the topology and $CPtable$ are completed, Bayes' theorem [20] can be used as a tool for performing probabilistic inference which can be thought of as a message (e.g., the available state of a certain variable) passing process in the BN. The theorem is shown in

$$P(x_i | x_j) = \frac{P(x_i, x_j)}{P(x_j)} = \frac{P(x_j | x_i) P(x_i)}{P(x_j)}, \quad (2)$$

where $P(x_i | x_j)$ is the posterior probability distribution of x_i given x_j , $P(x_j | x_i)$ is a prediction term for x_j given x_i , and $P(x_i)$ and $P(x_j)$ are the prior probabilities of x_i and x_j , respectively. A value that is assigned to a variable will be propagated through the network and will update the marginal posterior distributions of other nodes as explained above.

3.2. Determining Variables. Before assessing the graph topology of a BN and the parameters of the $CPtable$ explained in the BN, different variables represented by the nodes of the BN ought to be decided according to the characteristics of the bus routes in the research area and the availability of the data for the BN analysis. In this study, the revenue per unit distance (RUD) is used as the query variable to measure the operating quality of bus routes. The evaluation calculation is interpreted by

$$RUD = \frac{R_i}{LEN_i}, \quad (3)$$

where R_i denotes bus service revenues of route i (unit: Yuan RMB) and LEN_i represents the length of route i (unit: km).

Besides the RUD, the nodes of the BN developed in this research also represent fitted-out vehicles (FOV), average

daily passengers (ADP), traveling speed (TRS), average load factor (ALF), length (LEN), nonlinear coefficient (NOC), and average station spacing (ASS) for the study area. The FOV are the fitted-out vehicles of route i (unit: vehicles), the ADP represents the average daily passengers of route i (unit: passengers), and the ASS denotes the average station spacing of route i (unit: m). Other variables are explained by

$$\text{TRS} = \frac{\text{LEN}_i}{\text{TT}_i}, \quad (4)$$

$$\text{ALF} = \frac{Q_{id}}{Q_i \times \text{run}_i}, \quad (5)$$

$$\text{NOC} = \frac{\text{LEN}_i}{\text{SL}_i}, \quad (6)$$

where LEN_i represents the length of route i (unit: km), TT_i represents the entire traveling time of route i (unit: h), Q_{id} denotes the average daily passenger transport volumes of route i (unit: passengers), Q_i denotes the rated passenger transport volumes of a bus of route i (unit: passengers), run_i represents the daily schedules of route i (unit: trips), and SL_i indicates the linear distance of route i (unit: km).

In this research, the values of all the variables represented by the nodes of the proposed BN for X are calculated from January to May 2016 according to the administrative statistics of the government department. The criteria of various discrete states of different variables are listed in Table 1. The discrete thresholds of the RUD, the ADP, and the FOV are based on the respective average values of the historical data. According to Sun et al. [21], 22.00 km/h and 30.00 km/h are selected as the discrete thresholds of the TRS. The discrete thresholds of the ALF are 0.30, 0.60, and 1.20 [15]. According to CBIP [22], the discrete thresholds of the NOC are 1.10, 1.40, and 2.00 and those of ASS are 500.00 m and 800.00 m. The discrete thresholds of the LEN are 36.23 km and 27.17 km [23].

3.3. Estimating Graph Topology and Parameters. After determining all the variables represented by the nodes of a BN, the structure of the BN (i.e., the dependencies between different variables) and the parameters of the *CPTable* interpreted in the BN (i.e., the strengths of various dependencies as encoded by the entries in *CPTable*) can be estimated according to specific issues. A common and simple approach to the BN structure learning is to rank graph structures via a search-and-score method that measures how well each model structure fits the data to find the global optimization BN structure [24]. Nevertheless, the search-and-score method requires prior information about the ordering of the nodes to reduce the search space [25, 26]. Unfortunately, this prior information about node ordering is not always given in advance. In order to make accurately efficient BN learning based on training data without adequate prior information, a directional dependence analysis (DDA) algorithm based on the dependency analysis of variables is proposed to learn the BN structure [26]. The calculation steps of the DDA algorithm including the sequential steps of building undirected BN, orienting edges, and thinning BN are as follows.

3.3.1. Building Undirected BN. Make a list L to display every pair of distinct nodes (X, Y) with sufficient mutual information $I(X, Y)$ calculated by (7) over a threshold value θ . Sort (X, Y) in L according to the order of $I(X, Y)$ from large to small. Connect a pair of nodes by an edge. Remove the edge e_{XY} with the maximal $I(X, Y)$ from L to list E and then remove the edge e_{XZ} including one of the nodes of the edge e_{XY} from L to E with the conditional mutual information $I(Y, Z | X)$ computed by (8) over a threshold value θ . Iteratively examine the edges in L .

$$I(X, Y) = \sum_{x,y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)} \quad (7)$$

$$I(Y, Z | X) = \sum_{x,y,z} P(y, z, x) \log \frac{P(y, z | x)}{P(y | x)P(z | x)}. \quad (8)$$

Add an edge between a pair of nodes belonging to (X, Y) remaining in L when they cannot be separated by a set of relevant conditional independence tests based on their conditional mutual information $I(X, Y | C)$ with respect to the set of “evidence” variables (i.e., condition set) C .

3.3.2. Orienting Edges. After building the undirected BN, the direction of an undirected edge e_{XY} is oriented using the conditional relative ability to predict CRAP calculated in (9) given the union of domain set $Z = \mathcal{Y}(X) \cup \mathcal{Y}(Y)$, where $\mathcal{Y}(X)$ is the domain of X , $\mathcal{Y}(Y)$ is the domain of Y , and x, y , and z denote an instantiation of X, Y , and Z , respectively.

$$\text{CRAP}(Z, X \rightarrow Y) = \frac{\sum_z P(z) \max [P(x, y | z)]}{\sum_z P(z) \max [P(x | z)]} - 1. \quad (9)$$

According to the causal semantic theory [19], if

$$\text{CRAP}(Z, X \rightarrow Y) - \text{CRAP}(Z, X \leftarrow Y) > \lambda \quad (\lambda > 0) \quad (10)$$

and no loops exist in the causal direction $X \rightarrow Y$, X is chosen as a parent node to Y . Similarly, if

$$\text{CRAP}(Z, X \leftarrow Y) - \text{CRAP}(Z, X \rightarrow Y) > \lambda \quad (\lambda > 0) \quad (11)$$

and no loops exist in the causal direction $X \leftarrow Y$, Y is chosen as a parent node to X . The proposed method based on the causal semantic theory will not be able to orient all the edges in a network because of the weak causal relationship. In this research, the Minimum Description Length (MDL) principle [27, 28] is used to deal with the weak causal semantic of arc $X \rightarrow Y$.

3.3.3. Thinning BN. Find a minimal cut-set $M(X, Y)$ separating X from Y in the BN graph using heuristic approaches [19]. Remove an edge $X \rightarrow Y$ ($X \leftarrow Y$) if $I(X, Y | M(X, Y))$ is less

TABLE 1: The criteria of various states of different variables.

Variables	States
RUD	Low [0, 575.00]; medium (575.00, 1000.00); high (1000.00, $+\infty$)
FOV (vehicles)	Low [0, 8]; high (8, $+\infty$)
ADP (passengers)	Small [0, 2125]; medium (2125, 3718]; large (3718, $+\infty$)
TRS (km/h)	Low [0, 22.00]; reasonable (22.00, 30.00]; high (30.00, $+\infty$)
ALF	Small [0, 0.30]; medium (0.30, 0.60]; large (0.60, 1.20]
LEN (km)	Short [0, 27.17]; reasonable (27.17, 36.23]; long (36.23, $+\infty$)
NOC	Reasonable [1.10, 1.40]; large (1.40, 2.00]; overlarge (2.00, $+\infty$)
ASS (m)	Reasonable [500.00, 800.00]; unreasonable; otherwise

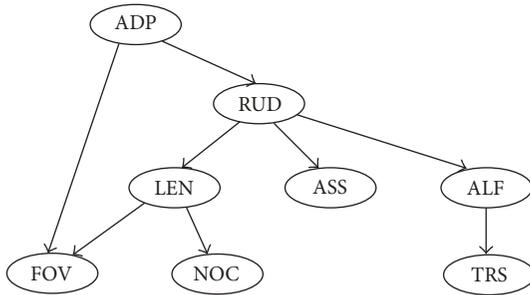


FIGURE 2: The structure of the BN developed for the study area.

than a certain threshold value. Iteratively examine the edges in the BN graph.

Received by virtue of applying the DDA algorithm with entering the statistical data of the variables represented by the nodes of the BN in this research, the structure of the BN newly developed for the estimation of bus service revenues of X is shown in Figure 2. In this study, the mean absolute percentage error (MAPE) is used to validate the accuracy of the developed BN model by calculating the average percentage difference between predicted values and observed ones given validation data [29]. The result of MAPE of each variable in the BN is almost less than 0.10, which demonstrates the good robustness of the BN model.

3.4. Probabilistic Inference. Given the structure of the BN and the parameters of the $CPTable$, inference queries can be evaluated by making full use of conditional independent information between different variables. Predictive inference (also called top-down reasoning) and diagnostic inference (also called bottom-up reasoning) are considered as two main types of probabilistic inference. They are based on the evidence nodes connected to the queried node through its parent and children nodes, respectively. In consideration of the predictive property of the probabilistic inference in this study, the Clique tree (CT) [13, 30], which is one of the major approaches to inference in multiply connected BN, is applied for the top-down reasoning work in this study because of its accuracy inference computation and high computational efficiency.

The CT is an undirected tree. Each tree node consists of a set of nodes from the BN which is called Clique. The

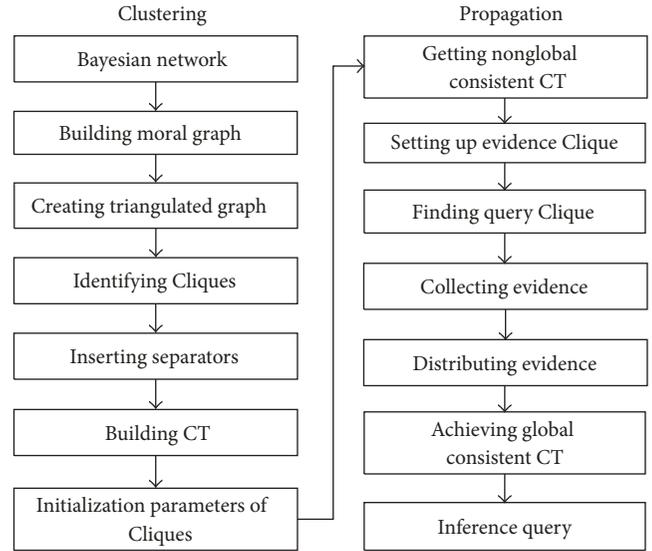


FIGURE 3: The process of CT.

CT takes the mechanism of information propagation with the steps of distributing evidence and collecting evidence. According to the difference of the information propagation programmes, the CT method can be further categorised into the Lauritzen–Spiegelhalter method, the HUGIN method, and the Shenoy–Shafer method. The HUGIN algorithm [31] is employed in this study because of its computational efficiency. The process of the HUGIN algorithm, which includes the sequential steps of clustering and propagation, is shown in Figure 3.

3.4.1. Clustering. First, an initial moral graph G_M is constructed by making an undirected copy of the BN and then augmenting it as follows. Let X systematically range over all nodes in G_M . For each node X , HUGIN adds to G_M an edge between each pair of nodes in parent nodes of X if no such edge already exists in G_M . Second, HUGIN triangulates the moral graph G_M , creating a triangulated graph G_M^* . The achieved triangulated graphs are proven to be the Cliques by Kjærulff [32]. Third, a CT is created from the Cliques by inserting separators. Last, the $CPTable$ of the BN translates into potential functions of the CT by initialization.

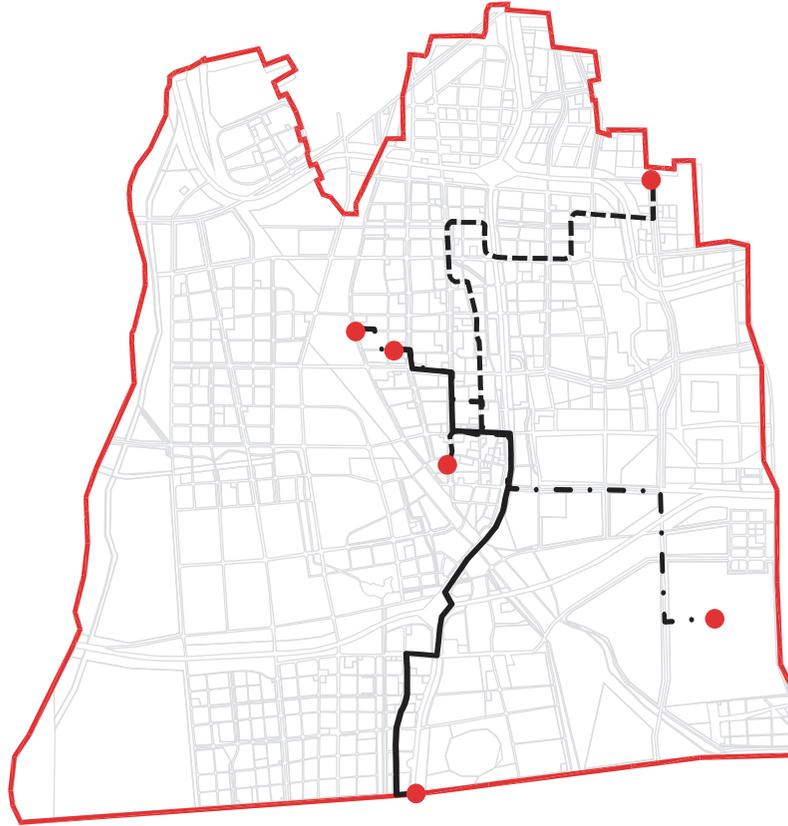


FIGURE 4: The existing layouts of route 26, route 28, and route 36 in X.

3.4.2. *Propagation.* The evidence information is introduced into potential functions of the CT. Find a query Clique C_Q which includes the query variable (i.e., the revenue per unit distance) as the root cluster. In the collecting evidence step, information is propagated from the furthest Clique to the query Clique C_Q until all the information reaches C_Q . On the contrary, in the distributing evidence step, information is propagated from the query Clique C_Q to the furthest Clique until all the information reaches each Clique. This way, the CT built on the basis of the BN achieves global consistency. Thereafter, the posterior probability distributions of the variables represented by the nodes of the BN are available from the prior $CPtable$ explained in each Clique of the CT. Then, the conditional probability of the query variable X_Q given the values of the evidence variables X_E in the BN is equal to

$$P(X_Q | X_E) = \frac{P(x_E, x_Q)}{\sum_{x_Q} P(x_E, x_Q)}, \quad (12)$$

where $\sum_{x_Q} P(x_E, x_Q)$ is the marginal probability of X_E .

4. Results and Discussions

Three types of route services in the study area X are considered as examples here by the inference queries. The dashed

line, the solid line, and the dot chain line shown in Figure 4 stand for route 26, route 28, and route 36, respectively. The red nodes represent the location of terminal stations of bus routes in X. In this research, the locations of terminal stations in X are assumed to be fixed. The proposed BN enables us to achieve the maximum predicted probabilities of high bus service revenues corresponding to the certain states of the variables of route 26, route 28, and route 36, respectively. According to these states, the optimized layouts of the three routes can be designed, shown in Figure 5. The dashed line, the solid line, and the dot chain line stand for the optimized layouts of route 26, route 28, and route 36, respectively. The red nodes represent the locations of terminal stations of bus routes in X.

Table 2 provides the maximum predicted probabilities of high bus service revenues and the values of variables for the existing routes and the optimized routes. The LEN, the NOC, and the ASS are calculated directly by the configuration of the optimized routes. The ADP is obtained according to the split flow of passenger transport demands of the existing bus routes which partly overlap the optimized routes. Then, the FOV is given based on the passenger transport demands of the optimized routes. The ALF is calculated by simultaneously considering the passenger demands and the fitted-out vehicles. The TRS is calculated by means of the traveling speed of the existing bus routes which partly overlap the optimized routes.

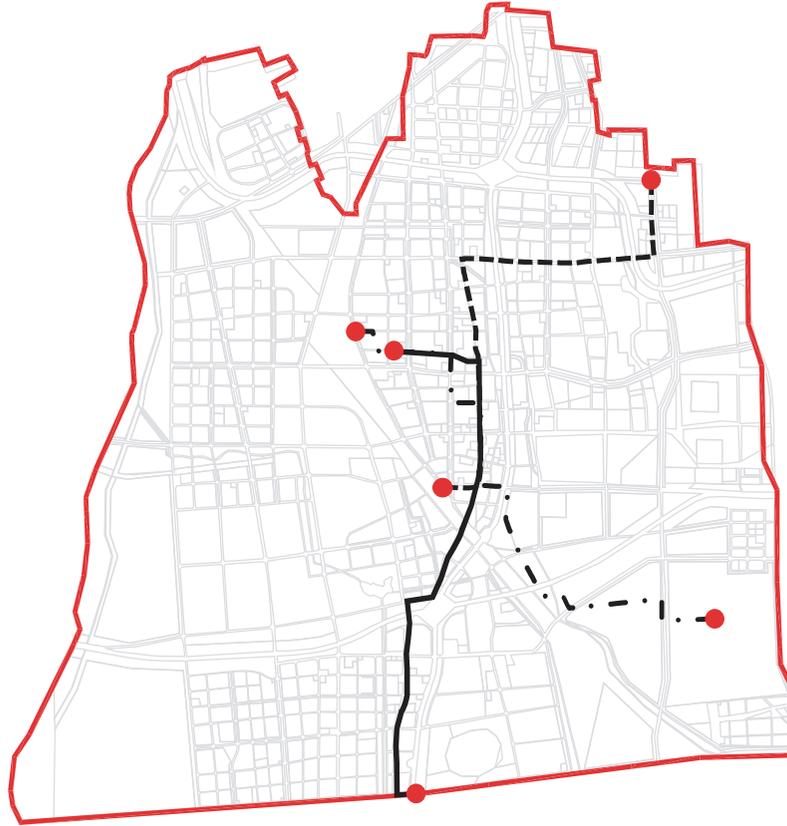


FIGURE 5: The optimized layouts of route 26, route 28, and route 36 in X.

TABLE 2: Maximum probabilities of bus service revenues.

Variables	Route 26		Route 28		Route 36	
	Existing	Optimized	Existing	Optimized	Existing	Optimized
	0	96.32%	0	99.86%	13.22%	98.89%
FOV (vehicles)	7	15	8	16	5	10
ADP (passengers)	924	3853	1560	4027	842	2698
TRS (km/h)	28.40	28.40	30.00	24.17	17.16	23.00
ALF	0.50	0.42	0.65	0.65	0.17	0.65
LEN (km)	35.50	33.30	45.00	42.90	14.30	10.70
NOC	1.68	1.58	1.46	1.39	2.07	1.55
ASS (m)	845.20	740.00	1022.70	794.44	621.70	509.52

The passenger demands of the west lines of route 26 are influenced by the split-flow effect, and the trend of the east lines of route 26 is inconsistent with the primary passenger flow. The comparatively low daily passenger transport volumes and the unreasonable layouts of route 26 make the probability of the existing high bus service revenues of route 26 become 0. Route 26 is adjusted by selecting the roads with comparatively few bus routes and stops with relatively large demands in order to increase the passenger transport volumes. Meanwhile, the number of operating vehicles is increased to satisfy the increasing amount of passenger demands. As a result, the revenues of optimized route 26 are the most probable high revenues.

Route 28 is mainly along the expressway which concentrates more than 40 bus routes. The concentrated routes split the passenger transport volumes of route 28. Moreover, route 28 is a relatively long and circuitous route. As a result, the probability of the existing high bus service revenues of route 28 is 0. In order to achieve the highest revenues, route 28 is optimized by selecting the roads with comparatively few bus routes and stops with relatively large demands to ensure larger passenger demands, while slightly improving route directness to ensure lower nonlinear coefficient and increasing the fitted-out vehicles to satisfy the increasing amount of passenger demands and control the average load factor in a reasonable range.

The concentrated routes cause low passenger demands of route 36 by the split-flow effect and make the traveling speed excessively low. Moreover, the route is more detoured, which increases the bus operation cost. The unreasonable layout of route 26 causes the probability of the existing high bus service revenues of route 36 to be 13.22%. Route 36 is modified by selecting the roads with comparatively few bus routes to increase the passenger demands and improve traveling speed, while increasing the number of operating vehicles to satisfy the increasing amount of passenger demands and control the average load factor in a reasonable range. Moreover, eliminating the circuitous alignment is going to ensure lower nonlinear coefficient. Then, this modification can successfully provide the most probable high revenues.

5. Conclusions

Using the area X in Beijing as the study area, a BN approach has been newly developed and applied to forecast the probabilities of different revenues of various route services, from the perspectives of both satisfying passenger demand and decreasing bus operation cost. This way, the existing bus routes in the studied area can be optimized for their most probable high revenues. In the future, combining other methodologies with the BN is worthy of the exploration for more rational estimation of the bus operating revenues. And extending the proposed solution methodology to solve other transport route design problems can be another future research direction.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research is supported by the National Natural Science Foundation of China (71571011) and the Fundamental Research Funds for the Central Universities (2016JBM028, 2017YJS110).

References

- [1] A. Ceder and Y. Israeli, "User and operator perspectives in transit network design," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1623, pp. 3–7, 1998.
- [2] S. I.-J. Chien and L. N. Spasovic, "Optimization of grid bus transit systems with elastic demand," *Journal of Advanced Transportation*, vol. 36, no. 1, pp. 63–91, 2002.
- [3] Y.-J. Lee and V. R. Vuchic, "Transit network design with variable demand," *Journal of Transportation Engineering*, vol. 131, no. 1, pp. 1–10, 2005.
- [4] S. C. Wirasinghe and U. Vandebona, "Route layout analysis for express buses," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 2, pp. 374–385, 2011.
- [5] A. Tirachini, D. A. Hensher, and J. M. Rose, "Multimodal pricing and optimal design of urban public transport: The interplay between traffic congestion and bus crowding," *Transportation Research Part B: Methodological*, vol. 61, pp. 33–54, 2014.
- [6] A. A. Ceder, M. Butcher, and L. Wang, "Optimization of bus stop placement for routes on uneven topography," *Transportation Research Part B: Methodological*, vol. 74, pp. 40–61, 2015.
- [7] F. Zhao and X. G. Zeng, "Simulated annealing–genetic algorithm for transit network optimization," *Journal of Computing in Civil Engineering*, vol. 20, no. 1, pp. 57–68, 2006.
- [8] W. Fan and R. B. Machemehl, "Tabu search strategies for the public transportation network optimizations with variable transit demand," *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, no. 7, pp. 502–520, 2008.
- [9] J. Pacheco, A. Alvarez, S. Casado, and J. L. González-Velarde, "A tabu search approach to an urban transport problem in northern Spain," *Computers & Operations Research*, vol. 36, no. 3, pp. 967–979, 2009.
- [10] W. Y. Szeto and Y. Wu, "A simultaneous bus route design and frequency setting problem for Tin Shui Wai, Hong Kong," *European Journal of Operational Research*, vol. 209, no. 2, pp. 141–155, 2011.
- [11] M. A. Nayeem, M. K. Rahman, and M. S. Rahman, "Transit network design by genetic algorithm with elitism," *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 30–45, 2014.
- [12] K. Ozbay and N. Noyan, "Estimation of incident clearance times using Bayesian Networks approach," *Accident Analysis & Prevention*, vol. 38, no. 3, pp. 542–555, 2006.
- [13] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, 2014.
- [14] J. Sun and J. Sun, "A dynamic Bayesian network model for real-time crash prediction using traffic speed conditions data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 176–186, 2015.
- [15] X. Kuang, *The Evaluation Research of Urban Public Transport Level of Services*, Jilin University, Changchun, China, 2005.
- [16] F. V. Jensen, *Bayesian Networks and Decision Graphs*, Springer, New York, NY, USA, 2001.
- [17] X. Feng, M. Saito, and Y. Liu, "Improve urban passenger transport management by rationally forecasting traffic congestion probability," *International Journal of Production Research*, vol. 54, no. 12, pp. 3465–3474, 2016.
- [18] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, Calif, USA, 1988.
- [19] J. Cheng, R. Greiner, J. Kelly, D. Bell, and W. Liu, "Learning Bayesian networks from data: an information-theory based approach," *Artificial Intelligence*, vol. 137, no. 1-2, pp. 43–90, 2002.
- [20] F. V. Jensen, *An Introduction to Bayesian Networks*, vol. 210, UCL press, London, UK, 1996.
- [21] F. C. Sun, Z. P. Wang, and J. Wang, "Statistic analysis of the bus average speed of Beijing," *Automotive Engineering*, vol. 25, no. 3, pp. 219–222, 2003.
- [22] *Urban Public Transport Specification*, CBIP (China Building Industry Press), Beijing, China, 1995.
- [23] R. Zhang and H. Yan, *The Theory And Practice of Urban Public Transport Planning*, China Railway Publishing House, Beijing, China, 2007.
- [24] T. Silander and P. Myllymaki, "A simple approach for finding the globally optimal Bayesian network structure," in *Proceedings of the Twenty-second Annual Conference on Uncertainty in*

- Artificial Intelligence (UAI-06)*, R. Dechter and T. Richardson, Eds., pp. 445–452, AUAI Press, 2006.
- [25] D. Heckerman, D. Geiger, and D. M. Chickering, “Learning Bayesian Networks: The Combination of Knowledge and Statistical Data,” *Machine Learning*, vol. 20, no. 3, pp. 197–243, 1995.
- [26] S. C. Wang, *Learning, inference and application of Bayesian Network*, Lixin Accounting Press, Shanghai, 2010.
- [27] J. Suzuki, “A construction of Bayesian networks from databases based on an MDL principle,” in *Proceedings of the Ninth international conference on Uncertainty in artificial intelligence*, pp. 266–273, Morgan Kaufmann Publishers Inc., 1993.
- [28] J. Suzuki, “Learning Bayesian belief networks based on the MDL principle: an efficient algorithm using the branch and bound technique,” 1996.
- [29] F. Zong, H. G. Xu, and H. Y. Zhang, “Prediction for traffic accident severity: comparing the Bayesian network and regression models,” *Mathematical Problems in Engineering*, vol. 2013, Article ID 475194, 9 pages, 2013.
- [30] S. L. Lauritzen and D. J. Spiegelhalter, “Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 50, no. 2, pp. 157–224, 1988.
- [31] S. K. Andersen, K. G. Olesen, F. V. Jensen, and F. Jensen, “HUGIN-A Shell for Building Bayesian Belief Universes for Expert Systems,” *In IJCAI*, pp. 1080–1085, 1989.
- [32] U. Kjærulff, “Triangulation of graphs—algorithms giving small total state space,” 1990.

