Schematic Transit Network Design using Minimum Spanning Tree Algorithm

Hyoung-Chul Kim [a], Ikki Kim [b], Jin-Seok Hahn [c]

[a] *Chungnam Institute, 73-26, Yeonsuwon-gil, Gongju-si, Chungcheongnam-do, 314-140, Republic of Korea*
[b] *Department of Transportation Engineering, Hanyang University, 1271, Sa 3 Dong, Sangrok-Gu, Ansan, Gyunggi-Do, 425-791, Republic of Korea*
[c] *Korea Environment Institute, Division of Atmospheric Environment, Environmental Policy research Group, 370, Sicheong-daero, Sejong-si, 339-007, Republic of Korea*

[a] *E-mail: khcchk@gmail.com*
[b] *E-mail: ikkikim@hanyang.ac.kr*
[c] *E-mail: hahn1027@gmail.com*

Abstract: This study is focused on exploring the public transportation routes. Exploration of the alternative routes of the public transportation needs a simultaneous reflection of the locational connectivity, demand for the public transportation, operating costs of the public transportation routes, etc. This study proposes an exploration methodology of the alternative routes of the public transportation using the minimum spanning tree algorithm. Prim's Algorithm was used to form the minimum spanning tree. Also, this study shows, as an example, that the explored alternative routes of the public transportation can change depending on the focus of the policy direction of the public transportation from the supply and demand perspective.

Keywords: Bus Rapid Transit, Decision Making, Minimum Spanning Tree, Prim's Algorithm, Greedy Method

1. INTRODUCTION

1.1 Background and Purpose

Recently, projects planning and evaluating the major express routes, such as BRT (Bus Rapid Transit) are increasing in order to connect the passage between the regions. This is probably due to the advantages possessed by the public transportation methods, such as bus, from the aspects of the construction costs and the efficient utilization of the existing road infrastructure, as compared to the subway system.

Generally, the macro procedure for planning the public transportation routes, such as the subway and the bus routes, is classified into exploration of alternative routes, evaluation of each alternate route, selection of the optimal alternative route, etc., and go through a sequential analysis process. The existing studies have focused on the second and the third stages of evaluation of the alternative routes and selection of the optimal alternative routes. Pattnaik *et al.* (1998) used genetic algorithms (GAs) for design of the urban bus route network. Carlos Lucio Martins and Margarida Vaz Pato (1998) reported on computing solutions for a specific problem (FBDP, Feeder Bus Network Design Problem) arising in public transport system. Giuseppe Bruno *et al.* (2002) proposed a mathematical model for the location of a rapid transit alignment in an urban setting. Quentin K. Wan and Hong K. Lo (2003) suggested a mixed integer formulation for multiple-route transit network design problem (MRTNDP). Bin Yu *et al.* (2005) developed an optimization model for bus transit network based on road network and zonal OD. Hiroshi Shimamoto *et al.* (2010) evaluated an existing bus network from the perspectives of passengers, operators, and overall system efficiency using transportation network optimization model. Gilbert *et al.* (2011) proposed a model for the design of a robust rapid transit network that the effect of disruption on total trip coverage is minimized. These studies have mostly focused on evaluating the alternative routes of the public transportation using quantitative indexes. However, it is difficult to find study cases on the first stage of exploration of alternative routes with schematic design concept.

This study is focused on exploring the public transportation routes. Exploration of the alternative routes of the public transportation needs a simultaneous reflection of the locational connectivity, demand for the public transportation, operating costs of the public transportation routes, etc. Also, this study is proposing, as an example, that the explored alternative routes of the public transportation can change depending on the focus of the policy direction of the public transportation from the supply and demand perspective. This study proposes an exploration methodology of the alternative routes of the public transportation using the minimum spanning tree algorithm. Prim's Algorithm was used to form the minimum spanning tree.

1.2 Scope and Method

Largely, there are two items of consideration when designing the alternative routes of the public transportation. First is the selection of the regions to be connected by the public transportation, and the second is the designing of the public transportation routes connecting the selected regions. The selection of the target regions to be connected by the public transportation is excluded from this study because it can change depending on the public transportation policy directions of the nation or the local autonomous organizations. However, we are proposing a methodology that can generally be considered in the selection of the target region for connection.

Therefore, this study focuses on proposing a methodology for exploring the alternative routes of the public transportation by taking into consideration the demand of the public transportation use and the construction or operations costs of the public transportation routes. Meaning, the maximization of the demand for the public transportation use from the transportation demand perspective and the minimization of the costs for establishing the public transportation routes from the transportation supply perspective are considered concurrently. The assumption that the costs of establishing the public transportation route are proportional to the extension of the public transportation routes is included here.

The exploration of the alternative routes of the public transportation is forming the minimum spanning tree for the given cost function on the simple network composed of node and link to be minimized, and the formed spanning tree is defined as the alternative route of the public transportation. The simple network expresses all regions as a node, each, and the road connected to each node is expressed as a single link. Also, the spanning tree has connectivity by including all nodes. Meaning, the spanning tree formed on the simple network is the public transportation route. Prim's Algorithm was used to compose the minimum spanning tree. Here, the main factors used in connecting the public transportation routes between regions are the demand for the public transportation use between the regions and the distance between the regions. Weighted value was applied to these two factors, and an example network was proposed as the results of the exploration of the alternative routes of the public transportation are diverse depending on the importance of each factors.

2. MINIMUM SPANNING TREE ALGORITHM

The tree that connects all nodes within a given network is called a spanning tree. The minimum spanning tree means the spanning tree with the minimum cost of all links included in the spanning tree when the cost function was given to each link on the network. The minimum spanning tree algorithm can be applied in establishing a connectivity with the minimum costs between the nodes, and primarily used in the optimization method. For example, questions of connecting with the minimum costs, such as natural gas piping, electric wire, road, telephone line, etc., are all minimum spanning tree questions. [1]

This study applied the Prim's Algorithm in creating the minimum spanning tree. The Prim's Algorithm has the characteristic of creating the same minimum spanning tree even if harm is found on a node. Moreover, it is a method of Greedy Method concept by adding the selected link from exploring adjacent link, one by one, to the selected link group, and even if the adjacent links are explored sequentially, it deducts the harm that becomes minimum finally. Dijkstra's Algorithm uses the same type of exploration method, and it is the algorithm for exploring the minimum distance.

The procedure for each stage of the Prim's Algorithm is as follows. [1]

[Stage 0] - Initialization

- Selecting a random nodal point i.
- Include the selected nodal point to the group U.

$U \leftarrow \{i\}$ (1)

- Initialize the cost C_j, from U to each nodal point i.

$C_j \leftarrow C_{ij}$ (2)

[Stage 1] – Select an arc adjacent to the intersection included in U and has the minimum cost

- Select the intersection j that has the minimum C_j among the adjacent intersection included in the intersection U.

$C_j \leftarrow m\dot{n}\{C_k: k \notin U\}$ (3)

- j is included in U, and the selected arc (i, j) is included in T.

$U \leftarrow U \cup \{i\}$ (4)

$T \leftarrow T \cup \{(i, j)\}$ (5)

[Stage 2] – Modified the costs from the intersection included in U to other intersections

- The intersection recently included in U is called j, modify the costs for each intersection k, not included in U.

$$C_k \leftarrow m\dot{n}\{C_k, C_{jk}\} \quad (6)$$

• If there is no nodal point not included in U, then move onto Stage 3. Otherwise, move to Stage 1.

[Stage 3] - Conclusion
• Finish. T is the minimum spanning tree.

3. CONSIDERATION FOR SELECTION METHOD OF THE PUBLIC TRANSPORTATION CONNECTING REGIONS

The public transportation connection region must consider the mobility and the ease of the passenger demand. The public transportation is focused on acquiring mobility of the passenger demand, rather than the function of smooth operation of passenger demand and freight demand as in road. Therefore, the public transportation policy is established to obtain smooth mobility of the passenger demand. Moreover, there are policies that aim appropriate land usage by connecting the traffic and furnish various transfer systems for the ease of utilizing the public transportation, such as TOD (Transit-Oriented Development).

Therefore, the main selection criteria for selecting the public transportation connection regions should be to select the points with high passenger traffic. At the same time, it is necessary to select a region that needs to induce the traffic to the public transportation because the share of public transportation method is relatively less than the personal transportation method. This could maximize the impact of developing the public transportation. Moreover, selecting a region with high traffic volume for going to and from work or school occurring repeatedly on a daily basis as a connection region of the public transportation may be applied as well.

4. DESIGNING METHOD FOR ALTERNATIVE OF THE PUBLIC TRANSPORTATION

This study established a cost function of the ring connecting each intersection, and the major factors are the reciprocal value on the size of the traffic volume of each link and the travel distance of the shortest path. Also, the alternative route of the public transportation is explored by composing a minimum spanning tree where the cost function of the link connecting all nodes becomes the minimum.

4.1 Assumptions of the Analysis

There are three assumptions of the analysis. First, it is assumed that there will be more switch to the public transportation method with more traffic volume between each node. This means that there are more potential traffic volumes to switch to the public transportation method. Second, the distance between each node is assumed to be the shortest path, and it is assumed that the costs of construction and operation in establishing the public transportation route is directly proportional to the distance. Third, the method of the public transportation is limited to BRT, surface car, regular buses, etc., using the public streets. This research focuses on the public transportation methods using the roads.

4.2 Analysis Methodology

This study explores the alternative route of the public transportation using the simple network composed of nodes and links. The public transportation connecting regions are expressed with nodes on the simple network, and the streets connecting each region is expressed in links.

The cost function of the link is composed of two variables and one parameter. The first is the variable taking the reciprocal of the size of the traffic volume. This variable has the characteristic of having smaller value as the size of the traffic volume increases. Meaning, the link with large traffic volume is less likely to be included in the link group composing the alternative route of the public transportation. The second is the variable on the travel distance of the shortest route between each node. The link with short shortest route distance also has more possibility of being included in the alternative route of the public transportation group. The last is the parameter that changes the characteristic of the cost function depending on the characteristic of the public transportation route. This parameter has the value between 0 and 1 depending on the characteristic of the public transportation route. This parameter regulates whether to connect the exploring public transportation route with the region with large demand for the public transportation use or to connect with the region that can minimize the costs related to the establishment and operation, etc. of the public transportation route. Meaning, this parameter allows the exploration of the alternative routes of the public transportation from diverse methods, in transportation demand and supply aspects. Therefore, depending on the policy direction of the public transportation, the minimum spanning tree is formed to minimize the total link costs on the network by applying different weighted value on the parameter, and exploring the concerned spanning tree as the alternative route of the public transportation.

On the network composed with many nodes, the traffic volume size of each link is defined as the reciprocal value trp_{ij} after adding the traffic volume moving from i to j (OD_{ij}) and the traffic volume moving from j to i (OD_{ji}), as shown in the Formula (7).

$$trp_{ij} = 1/(OD_{ij} + OD_{ji}) \quad (7)$$

Moreover, the shortest route traffic distance of each link (i,j) is defined as $dist_{ij}$.

The Formulas (8) and (9) shows the definitions of Trp_{ij} and $Dist_{ij}$, each, by standardizing the reciprocal value on the size of the traffic volume of each link (i,j) (trp_{ij}), and the shortest path traffic distance $(dist_{ij})$.

$$Trp_{ij} = trp_{ij}/\overline{trp_{ij}} \quad (8)$$

$$Dist_{ij} = dist_{ij}/\overline{dist_{ij}} \quad (9)$$

Here,

Trp_{ij} : The standardized value by taking reciprocal of the size of the traffic volume of each arc(i,j),

$Dist_{ij}$: The standardize value of the shortest path traffic distance for each arc (i,j),

$\overline{trip_{ij}}$: The average value of $trip_{ij}$ for all arcs(i, j),
$\overline{dist_{ij}}$: The average value of $dist_{ij}$ for all arcs(i, j).

In order to express with ratio scale to have the values between 0 and 1 for the standardized $Trip_{ij}$ and $Dist_{ij}$, each were divided with the respective maximum values, as shown in Formulas (10) and (11).

$Trip\_max = max\{Trip_{ij}\}\ (\forall i, j)$ (10)

$Dist\_max = max\{Dist_{ij}\}\ (\forall i, j)$ (11)

$TR_{ij} = Trip_{ij} / Trip\_max$ (12)

$D_{ij} = Dist_{ij} / Dist\_max$ (13)

Here,

$Trip\_max$: The maximum value of $Trip_{ij}$ for all links (i, j),
$Dist\_max$: The maximum value of $Dist_{ij}$ for all links (i, j),
TR_{ij} : The standardized value between 0 and 1 of the reciprocal value of the size of the traffic volume for each link (i, j),
D_{ij} : The standardized value between 0 and 1 of the shortest path traffic distance for each link (i, j).

The Formula (14) shows the setting the cost function of each arc (i, j) of the given network after applying α and $(1-\alpha)$ to standardized TR_{ij} and D_{ij}, respectively.

$C_{ij} = \{\alpha TR_{ij} + (1-\alpha) D_{ij}\} \times \beta_{ij}\ \ (0 \leq \alpha \leq 1,\ \beta = 1\ or\ \infty)$ (14)

Here,

C_{ij} : Cost function of each arc (i, j),
α : Public transportation policy parameter, set closer to 1 for broader-unit route, and set closer to 0 for connecting short distances within a region,
β_{ij} : Parameter determining whether to reflect the link (i, j), set as 1 if included as part of the alternative route, and set as c otherwise.

The cost function of the link (i, j) is set with the standardized value between 0 and 1 for TR_{ij} and D_{ij}, respectively, and as the parameter for the public transportation policy, α value, is closer to 1, the ratio of TR_{ij}, related to the traffic volume, increases, and the cost function is mostly explained by the reciprocal value of the size of the traffic volume. However, because the proportion of D_{ij}, the traffic distance of the shortest path decreases, the sensitivity and the meaning for the shortest path traffic distance decreases, relatively, on the cost function. These cases can be applied when the policy direction of the public transportation is for designing an inter-regional alternative route, connecting the travel between regions, such as the BRT. Meaning, by setting the value of α, the parameter of the public transportation policy of the cost function, near 1, and composing the minimum spanning tree where the alternative route of the public transportation directly connects a region with large traffic volume, rather than laying over the close-by regions, if possible.

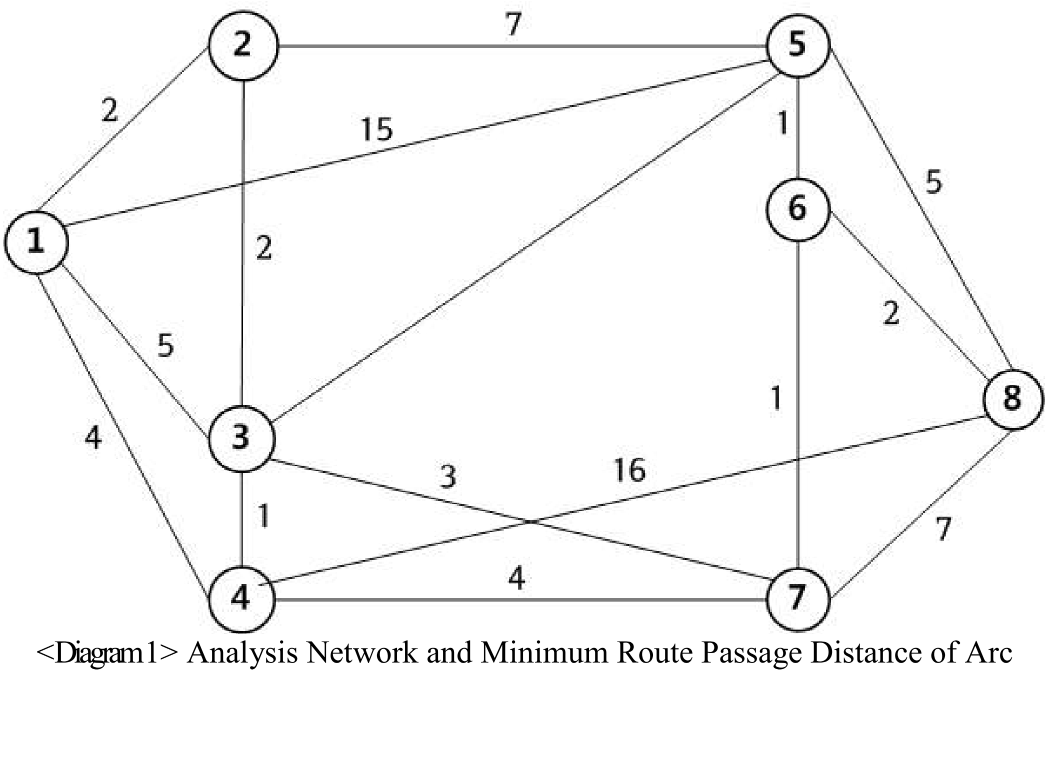
As the value of α, the parameter of the public transportation policy, becomes closer to 0, the proportion of D_{ij}, the traffic distance of the shortest path, increases, and the cost function is mostly explained with the traffic distance of the shortest path. Similarly, the proportion of traffic volume related TR_{ij} decreases, and the sensitivity or the meaning related to the traffice volume decreases, relatively. These cases can be applied when the policy direction of the public transportation aims at connecting the near-by regions, such as local bus or shuttle bus and increasing the accessibility within a reagion. Meaning, the alternative route of the public transportation can be explored by setting the value of α, the public transportation policy parameter of the cost function, near 0, and composing the minimum spanning tree to connect by minimizing the running distance and the costs, rather than connecting the regions with relatively heavy traffic volume.

The parameter β_{ij}, which determines the reflection of the link, is set at 1, when including the link (i,j) as the alternative route of the public transportation, and set as ∞ when intending to exclude the overlapping public transportation routes in a region with the existing subway line passing through according to the policy direction. In composing the minimum spanning tree, link (i,j) with the cost of ∞ is not selected, so depending on the analysis purpose, β_{ij} can be utilized to explore the alternative route of the public transportation.

5. INTRODUCTION OF EXAMPLE NETWORK AND ANALYSIS RESULTS

5.1 Introduction of the Example Network

The example network is composed of 8 nodes, and the minimum route passage distance of the link is seen in the <Diagram 1>.



<Diagram 1> Analysis Network and Minimum Route Passage Distance of Arc

It is assumed that the shortest path traffic distance for each link is same without regards to the directionality.

In the example network, the link not connected between each node is set with ∞ for the β_{ij} of Formula (14), and the link is not included in the minimum spanning tree. Therefore, the example network also excluded these in its expression.

The cases where analysis was performed by setting β_{ij} as ∞ are when intending to reflect the installation of the existing public transportation routes in reality, such as subway, main express route, regular route bus, etc. In other words, unless a dual line is newly installed for lack of capacity, it is very rare to explore alternative routes with the same routes when establishing a new public transportation route. Therefore, depending on the analysis purpose, the dual routes may be not reflected.

The variables composing the cost function are calculated similar to <Table 1>. At this time, $\overline{Trip}$ and $\overline{Dist}$ are 0.005 and 4.938, respectively, and $Trip\_max$ and $Dist\_max$ of Formulas (10) and (11) are 2.089 and 3.241, respectively.

<Table 1> Variables of Cost Function for each link (i, j)

link (i,j)	Standardization of the Traffic Volume						Standardization of Shortest Path's Travel Length		
	OD_{ij}	OD_{ji}	$OD_{ij}+OD_{ji}$	$trip_{ij}$	$Trip_{ij}$	TR_{ij}	$dist_{ij}$	$Dist_{ij}$	D_{ij}
(1,2)	100	80	180	0.006	1.044	0.500	2	0.405	0.125
(1,3)	200	40	240	0.004	0.783	0.375	5	1.013	0.313
(1,4)	200	70	270	0.004	0.696	0.333	4	0.810	0.250
(1,5)	500	300	800	0.001	0.235	0.113	15	3.038	0.938
(2,3)	80	90	170	0.006	1.106	0.529	2	0.405	0.125
(2,5)	40	80	120	0.008	1.567	0.750	7	1.418	0.438
(3,4)	170	160	330	0.003	0.570	0.273	1	0.203	0.063
(3,5)	40	50	90	0.011	2.089	1.000	4	0.810	0.250
(3,7)	60	30	90	0.011	2.089	1.000	3	0.608	0.188
(4,7)	70	30	100	0.010	1.880	0.900	4	0.810	0.250
(4,8)	300	200	500	0.002	0.376	0.180	16	3.241	1.000
(5,6)	120	150	270	0.004	0.696	0.333	1	0.203	0.063
(5,8)	50	500	550	0.002	0.342	0.164	5	1.013	0.313
(6,7)	150	50	200	0.005	0.940	0.450	1	0.203	0.063
(6,8)	100	100	200	0.005	0.940	0.450	2	0.405	0.125
(7,8)	90	200	290	0.003	0.648	0.310	7	1.418	0.438

The calculation results for the cost function (C_{ij}) on each link(i, j), when the level of α is 1, 0.5, and 0 respectively, are shown in the <Table 2>.

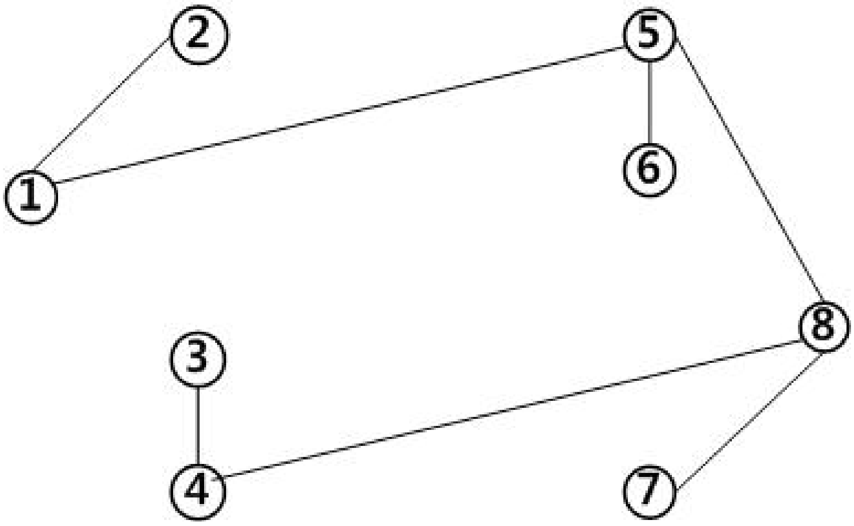
<Table 2> Results of the Cost Function for each link (i, j)

Cost	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 0$
C_{12}, C_{21}	0.500	0.313	0.125
C_{13}, C_{31}	0.375	0.344	0.313
C_{14}, C_{41}	0.333	0.292	0.250
C_{15}, C_{51}	0.113	0.525	0.938
C_{23}, C_{32}	0.529	0.327	0.125
C_{25}, C_{52}	0.750	0.594	0.438
C_{34}, C_{43}	0.273	0.168	0.063
C_{35}, C_{53}	1.000	0.625	0.250
C_{37}, C_{73}	1.000	0.594	0.188
C_{47}, C_{74}	0.900	0.575	0.250
C_{48}, C_{84}	0.180	0.590	1.000
C_{56}, C_{65}	0.333	0.198	0.063
C_{58}, C_{85}	0.164	0.238	0.313
C_{67}, C_{76}	0.450	0.256	0.063
C_{68}, C_{86}	0.450	0.288	0.125
C_{78}, C_{87}	0.310	0.374	0.438

5.2 Analysis Results

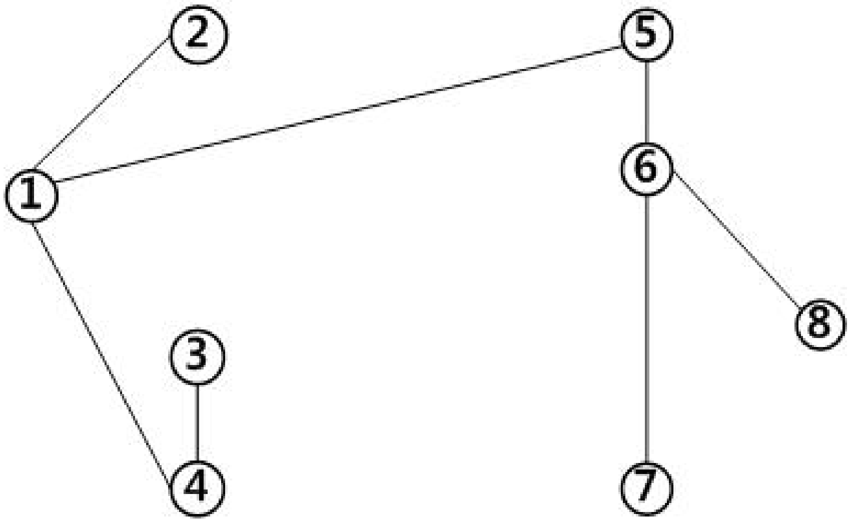
When $\alpha = 1$ in designing the alternative route of the public transportation, the minimum spanning tree is composed by only considering the traffic volume between each node, and is shown in <Diagram 2>.

When looking at the analysis results, an alternative route of the public transportation that is connecting the links (i, j) with high traffic volume while connecting all nodes is generated.



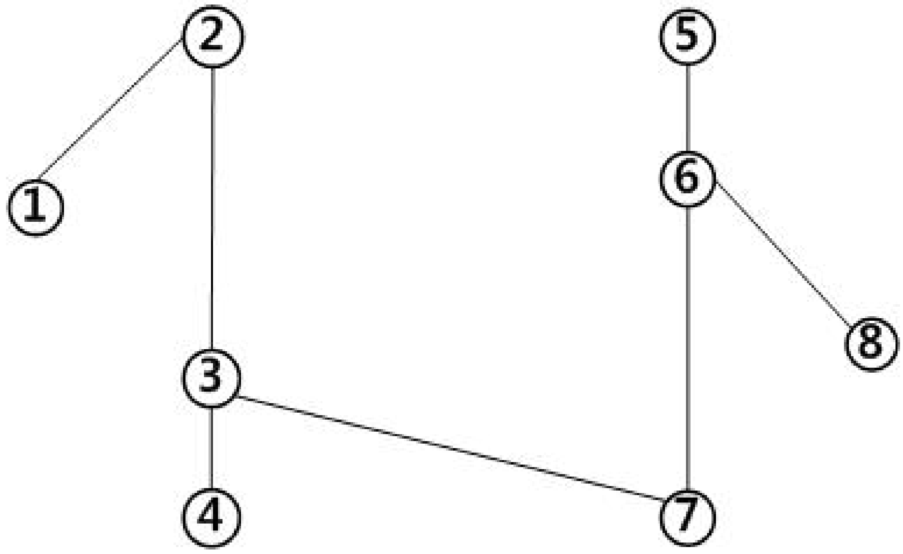
<Diagram 2> Alternative Route of the Public Transportation when $\alpha = 1$

When $\alpha = 0.5$, same weighted value was given to the size of the traffic volume between each node and the shortest path traffic distance, and an alternative route of the public transportation, as shown in <Diagram 3> is generated.



<Diagram 3> Alternative Route of the Public Transportation when $\alpha = 0.5$

When $\alpha = 0$, a minimum spanning tree is composed by only considering the shortest path traffic distance between each node, and is shown in <Diagram 4>. Therefore, an alternative route of the public transportation selected to minimize the distance of all connected arc (i, j) by connecting all intersection is generated.



<Diagram 4> Alternative Route of Public Transportation when $\alpha = 0$

6. CONCLUSION

This study proposed a methodology for designing the alternative routes of the public transportation using the minimum spanning tree, and utilized the Prim's Algorithm. Moreover, I have proposed a methodology for exploring and designing variety of alternative routes by the decision-makers or public transportation policy analysts using the public transportation policy parameters of α and β_{ij} .

This study proposed a measure for selecting a regional alternative route, where the focus of the policy evaluation index is on the size of the traffic volume between regions, a measure for selecting an alternative route focusing on the distance that effects the construction cost, maintenance cost, operating cost, etc., and a measure for selecting the alternative routes of the public transportation by reflecting all impact factors from the above two.

The designing methodology for the alternative routes of the public transportation proposed by this study can be useful in exploring and designing alternative routes with many regions to be connected by the public transportation or from diverse policy evaluation indexes. Especially, the developing counties of the East Asia, with insufficient public transportation system, have many opportunities to newly plan the public transportation routes. Therefore, it is expected that the methodology of exploring public transportation routes proposed by this study can be applied to coincide with the characteristics of the public transportation routes.

It is necessary to conduct studies on the methodology to rationally determine the connecting regions of the public transportation and on the policy evaluation index in selecting the alternative routes of the public transportation between regions in the future.

REFERENCES

Bin Yu, Zhongzhen Yang, Chuntian Cheng, and Chong Liu, "Optimizing bus transit network with parallel ant colony algorithm," *Proceedings of the Eastern Asia Society for Transportation Studies*, Vol. 5, pp. 374-389, 2005.

By S. B. Pattnaik, S. Mohan, and V. M. Tom, "Urban Bus Transit Route Network Design Using Genetic Algorithm," *Journal of Transportation Engineering*, pp. 368-375,

August 1998.
Carlos Lucio Martins and Margarida Vaz Pato, “Search strategies for the feeder bus network design problem,” *European Journal of Operational Research*, pp. 425-440, 1998.
C. W. Kirkwood, Strategic Decision Making, Duxbury Press, Belmont, 1997.
G. Zhou and M. Gen, Genetic algorithm approach on multi-criteria minimum spanning tree problem, *European Journal of Operational Research*, 1999.
Giuseppe Bruno, Michel Gendreau, and Gilbert Laporte, “A heuristic for the location of a rapid transit line,” *Computers & Operations Research*, pp. 1-12, 2002.
Gilbert Laporte, Angel Marin, Juan A. Mesa, and Federico Perea, “Designing robust rapid transit networks with alternative routes,” *Journal of advanced transportation*, pp. 54-65, 2011.
Hiroshi Shimamoto, Naoki Murayama, Akimasa Fujiwara, and Junyi Zhang, “Evaluation of an existing bus network using a transit network optimization model: a case study of the Hiroshima City Bus network,” *Transportation*, pp. 801-823, 2010.
J. R. Evans and E. Minieka, Optimization Algorithm for Networks and Graphs, 2nd ed., rev. and expanded, Marcel Dekker, NY, 1992.
M. K. Kang, Network and Algorithm, Bak-young-sa, Seoul, 1991.
Quentin K. Wan and Hong K. Lo, “A mixed integer formulation for multiple-route transit network design,” *Journal of Mathematical Modelling and Algorithms*, pp. 299-308, 2003.
R. L. Graham and Pavol Hell, “On the History of the Minimum Spanning Tree Problem,” *Annals of the History of Computing*, Vol. 7, No. 1, January 1985.