Estimating the Effect of Subway Network Structure on Travel Characteristic in the Seoul Metropolitan Area

Seong-Man Jang¹, Young-Soo An², Seung-Il Lee³

¹ The doctor’s course, Dept of Urban Planning, University of Seoul, 163 Siripdaero, Dongdaemun-gu, Seoul Korea, jangsm@uos.ac.kr

² Research professor, Dept of Urban Planning, University of Seoul, 163 Siripdaero, Dongdaemun-gu, Seoul Korea, ysan@uos.ac.kr

³ Professor, Dept of Urban Planning, University of Seoul, 163 Siripdaero, Dongdaemun-gu, Seoul Korea, silee@uos.ac.kr

ABSTRACT:

With increased awareness of the seriousness of environmental problems, sustainable development has been a key issue worldwide. In particular, in order to encourage the use of public transportation, new subway links are continually being constructed. It is highly important to make an efficient use of the existing links, especially because it requires significant capital and time to build new links on the subway. In order to efficiently utilize the existing links, it is necessary to conduct a study to analyze the characteristics of the subway. Due to the increasing complexity of the subway network, it has become important to examine the network centrality in addition to simply analyzing the regional characteristics in order to understand the travel characteristics.

The purpose of this study was to examine the effect of the structural characteristics of the subway network of Seoul Metropolitan Area on the user-behavior of passengers. For this purpose, the network centrality was determined and its effects on the variables such as the number of users and moving distance of each station were analyzed using a polynomial regression model. The results of the analyses performed in this study can be divided into three types according to the indicator of network centrality. First, the higher the reach centrality, the more choices passengers had in selecting a neighboring station as the start point or the end point of their travel, which resulted in the total traffic. Also, the proportion of short-distance travel increased because passengers using the station concerned had more choices in terms of the stations that they could choose as the start point or the end point of their travel. Second, the betweenness centrality value had no significance with respect to the short- and long-distance travel ratios, which means that it has no relation to travel distances. However, it showed a positive correlation to the amount of traffic at stations. Lastly, an increase in closeness centrality led to a decrease in the proportion of long-distance travel and an increase in short- and mid-distance travel.

This study verified that the subway network centrality has an impact on the user-behavior of passengers, which proves that network centrality is an important factor to be considered by the government when establishing policies related to subway.

Keywords: Social network analysis, Network centrality, Subway network structure, Travel characteristic.
1. Introduction

With increased awareness of the seriousness of environmental problems, sustainable development has been a key issue worldwide. In the same context, there has been an active discussion of New Urbanism, a movement to create an urban environment that centers on humans and public transportation (Kim Seong-Hee et al. 2001). The Korean government has also made efforts with a policy target of increasing the modal share rate of public transportation by 50% from 2011 to 2016 (1). In particular, in order to encourage the use of public transportation, new subway links are continually being constructed. As a result, at present, the subway network of Seoul Metropolitan Area spans about 883km in total distance and is comprised of 15 different routes that are intricately connected (2).

It is highly important to make an efficient use of the existing links, especially because it requires significant capital and time to build new links on the subway (Lee Seung-Il, 2004). In order to efficiently utilize the existing links, it is necessary to conduct a study to analyze the characteristics of the subway. Accordingly, there have been many related studies carried out at home and abroad. In particular, majority of the studies have reported the results of the analyses of the characteristics of land use around the subway station areas that affect the number of users and travel distance to the station. However, due to the increasing complexity of the subway network, it has become important to examine the network centrality in addition to simply analyzing the regional characteristics in order to understand the travel characteristics. In other words, an analysis of the subway network centrality on the use-behavior of passengers can contribute to the enhancement of the railway network efficiency and predict the changes resulting from the establishment of new links.

The purpose of this study is to analyze the structural characteristics of the subway links in the Seoul Metropolitan Area on the use-behavior of passengers. For this purpose, the subway network centrality was calculated and its effect on the travel characteristics (Number of users and travel distance) was analyzed using the polynomial regression model. The spatial scope of this study was Seoul Metropolitan Area, and the temporal scope is the year 2011.

2. Review Theory and Previous Research

2.1 Theories Related to Social Network Analysis

In a study by Wegener (1996), accessibility was analyzed as a factor that provides opportunities for spatial exchanges and has an important influence on people’s travel behavior (Lee Seung-Il, 2000). Accessibility is a concept that includes various meanings such as proximity, spatial distance, ease of spatial interaction and potential for diverse activities, and cannot be defined by one of these factors (Lee Jong-Yong, 1997). Thus, the definition of accessibility varies from study to study as each study defines the term to suit its purpose.

Transportation Geography regards a transportation network as being comprised of nodes and links, and defines the concept of accessibility as a question of how well the nodes are connected to one another (Won Kwang-Hee, 2003). One of the ways to calculate accessibility in Transportation Geography is an analysis methodology in which social network analysts review the structural characteristics of the network and understand the position or location of the actors in the network (Han Ju-Seong, 2010). The social network analysis based concept is theoretically grounded in insights from Simmel and Durkheim and methodologically grounded in classical graph theory (James Moody and Douglas R. White, 2003).

Social network analysis views social relationships in terms of network theory consisting of
nodes and ties (also called edges, links, or connections). Nodes are the individual actors within the networks, and ties are the relationships between the actors. The resulting graph-based structures are often very complex. There can be many kinds of ties between the nodes. Research in a number of academic fields has shown that social networks operate on many levels, from families up to the level of nations, and play a critical role in determining the way problems are solved, organizations are run, and the degree to which individuals succeed in achieving their goals (David L. Passmore, 2011).

Social network analysts themselves have addressed the broader use of network analysis in the social sciences and increasingly, the physical sciences. With rapidly growing interest in the potential uses of Social network analysts, it is both appropriate and timely to review its applications and potential within the field of urban planning (C. Scott Dempwolf & L. Ward Lyles, 2011). Social network analysis is performed to derive a wide range of indices and in the field of urban planning, the major indices include reach centrality, gravity centrality, betweenness centrality, closeness centrality and straightness.

The Reach centrality measure (Sevtsuk, 2010) captures how many surrounding nodes each node Reach centralityes within a given Search Radius on the network. Whereas the Reach centrality measure simply counts the number of destinations around each node within a given Search Radius (optionally weighted by node attributes), the Gravity measure additionally factors in the spatial impedance required to travel to each of the destinations. The Betweenness centrality of a node is defined as the fraction of shortest paths between pairs of other nodes in the network that pass by node (Freeman 1977). The Closeness centrality of an Input node is defined as the inverse of cumulative distance required to Reach centrality from that node to all other nodes in the system that fall within the Search Radius along the shortest paths (Sabidussi 1966). The Straightness metric (Vragovic, Louis, et al. 2005) illustrates the extent to which the shortest paths from a node of interest to all other nodes in the system resemble straight Euclidian paths (Quoted in Andres Sevtsuk and Michael Mekonnen, 2011).

2.2 Review Previous Research

Previous studies were reviewed to examine the possibility of social network analysis in the field or urban planning.

C. Scott Dempwolf and L. Ward Lyles (2011) are studied that what unique value does social network analysis offer compared to other approaches and methods commonly used in urban planning, and How has social network analysis been applied in urban planning research or practice and what contributions have these applications made. They conclude that SNA has the potential to advance and operationalize certain aspects of communicative planning theory, which have been criticized due to their reliance on tacit knowledge in the form of judgment and experience. In particular, SNA has the potential to help planners visualize, measure, and document sources of communicative distortion and to identify specific steps to address these situations. Choi, Ja-Eun (2012) analyzed the relation between the visitor mobility and urban leisure space structure, utilizing the social network analysis. The result of this study explains the social network theory that visitor mobility affects the leisure space structure, then the leisure space structure again exercise influence on the visitor mobility. Kim Hyo-jin (2008) derived the in-degree centrality using social network analysis, and studied how this indicator affected the commuter traffic in the Seoul Metropolitan Area. Through this, the regions and network clusters in the Seoul Metropolitan Area with high in-degree centrality by the hour were derived.
A review of precedent studies related to social network analysis revealed that it has been proven to be a useful tool in the field of urban planning and used frequently in research but it has rarely been used to analyze the user-behavior of passengers. The following are studies that analyzed user-behavior of passengers.

Cervero, R. and Kockelman, K. (1997) analyzed the patterns of passengers based on the density and diversity of the area and design elements. The findings showed that higher density and more diverse use of the land resulted in a decrease in car traffic. Also, there was increased use of public transportation if the design elements of the city were more pedestrian-friendly.

O’ullivan, S. and J. Morrall. (1996) analyzed the walking distance to light rail stations in the CBD and suburban areas. The results of the analysis showed that the walking distance to the stations differed depending on the conditions of land use. Also, that study emphasized that pedestrian access around suburban Light-Rail Transit stations should be reviewed and improved to reduce circuity.

Jeon Myeong-jin. (1997) analyzed the relationship between the land use patterns and the choice of travel means in the Seoul Metropolitan Area. An analysis using the multiple logit model confirmed that an increase in density of office and residential areas resulted in people relying more heavily on the means of public transportation. Also, the usage rate for personal vehicles was higher than that of the bus in office-concentrated areas, proving that the high density policy signifies a movement to an increased use of public transportation.

Previous studies that analyzed the use characteristics of subway were mainly based on the conditions of using the land in the surrounding area. However, due to the increasing complexity of the subway network, it has become more important to examine the network centrality in addition to the regional characteristics in order to understand the travel characteristics. Accordingly, in this study, the centrality of the Seoul Metropolitan Area subway network and its effects on the number of users per station and distance to the stations were determined to suggest their implications.

3. Analysis
3.1 Developing a Model and Setting the Parameters

To meet the purpose of this study, a polynomial regression model was used with network centrality as the independent variable and user-behavior of passengers as the dependent variable.

\[ TC = \beta_0 + \sum_{i=1}^{n} \beta_i X_i \]  

\( TC = \text{Travel Characteristic} \)
\( X = \text{Centrality index} \)
\( \beta_0, \beta_i = \text{Parameters} \)
\( n = \text{Number of Centrality} \)

In order to determine the use-behavior of passengers, O/D data on each of the stations were derived using the report on the monthly passenger transport performance of the subway for the year 2011 provided by Korea Railroad (KORAIL). The average of users of the Seoul Metropolitan Subway in 2011 were 5,409,417 and the median for the walking distance of users was approximately 11.4 km, while the first quartile standard value was approximately 6.1 km and third quartile standard value was 18.4 km (3).
Fig 1. Scatter plot for Number of Passenger

In order to analyze the user-behavior of passengers, the number of people embarking on or disembarking from the trains was added to calculate the total traffic. The station with the highest traffic was Gangnam Station, whereas the lowest traffic was at Sinwon Station. A travel for less than 1 minute was defined as a short-distance travel, while a travel for 1 to 3 minutes was defined as a mid-distance travel and over 3 minutes a long-distance travel. After, the ratio of short-, mid- and long-distance travels was calculated for each station. The results showed that there was a high percentage of short-distance travels at Yongdap Station and Bupyeongsamgeori Station, and high percentage of long-distance travels at Incheon Int'l Airport Cargo Terminal Station and Incheon International Airport Station.

Network centrality was calculated using social network analysis, which can be used to calculate a wide range of network centralities. In this study, the reach centrality, betweenness centrality and closeness centrality were chosen as they are indicators that can best explain an subway network. The calculation method for each of the indicators is as follows:
3.1.1. Reach centrality
The reach centrality, $C_{R}[i]$, of a node $i$ in a graph $G$ at a search radius $r$, describes the number of other nodes in $G$ that are reachable from $i$ at a shortest path distance of at most $r$. It is defined as follows:

$$C_{R}[i] = ||\{j \in G - \{i\}: d[i, j] \leq r\}||$$

(2)

Where $d[i, j]$ is the shortest path distance between nodes $i$ and $j$ in $G$, and $||S||$ is the cardinality of the set $S$. If the nodes in $G$ are weighted, then reach is defined as follows:

$$C_{R}[i] = \sum_{j \in G - \{i\}, d[i, j] \leq r} W[j]$$

(3)

Where $W[j]$ is the weight of node $j$ (Sevtsuk 2010).

3.1.2. Betweenness centrality
The betweenness centrality, $C_{B}[i]$, of a node $i$ in graph $G$ estimates the number of times $i$ lies on shortest paths between pairs of other reachable nodes in $G$ that lie within the network radius $r$ (Freeman 1977). If more than one shortest path is found between two nodes, as is frequently the case in a rectangular grid of streets, then each of the equidistant paths is given equal weight such that the weights sum to unity. It is defined as follows:

$$C_{B}[i] = \sum_{j, k \in G - \{i\}, d[j, k] \leq r} \frac{n_{jk}[i]}{n_{jk}}$$

(4)

Where $n_{jk}$ is the number of shortest paths from node $j$ to node $k$ in $G$ and $n_{jk}[i]$ is the number of shortest paths from $j$ to $k$ that pass through $i$, with $j$ and $k$ lying within the network.
radius \( r \) from \( i \). If the nodes in \( G \) are weighted, then betweenness is defined as follows:

\[
C_B^r[i] = \sum_{j,k \in G \setminus \{i\}, d([j,k] \leq r)} \frac{n_{jk}[i]}{n_{jk}} \cdot W[j]
\]  

(5)

The Betweenness measure may be used to estimate the potential of passersby at different locations of the network. If the analysis is weighted by demographics of a certain type in the surrounding nodes for instance, then Betweenness centrality can capture the potential of passersby of that particular demographic at node \( i \).

### 3.1.3. Closeness centrality

The closeness centrality, \( C_C^r[i] \), of a node \( i \) in a graph \( G \) is the inverse of the total distance from \( i \) to all other nodes that are reachable in \( G \) within radius \( r \) along shortest paths (Sabidussi 1966). It is defined as follows:

\[
C_C^r[i] = \frac{1}{\sum_{j \in G \setminus \{i\}, \ d([i,j] \leq r)} d([i,j])}
\]  

(6)

Whereas Betweenness centrality estimates the potential traffic passing by each location in the graph, the Closeness measure indicates how close each of these locations is to all other surrounding locations within a given distance threshold. If weights are applied, then Closeness centrality is defined as follows:

\[
C_C^r[i] = \frac{1}{\sum_{j \in G \setminus \{i\}, d([i,j] \leq r)} d([i,j]) \cdot W[j]}
\]  

(7)

Network centrality was determined based on the subway network of Seoul Metropolitan Area in 2009. The network centrality for each station was calculated using the Urban Network Analysis Tool of GIS 10 [Fig. 3] [Fig. 4] [Fig. 5]. The radius of analysis for reach centrality and betweenness centrality was limited to 10 km, while the radius for closeness centrality was not limited for analysis because limiting the radius would cause a difference in the number of nodes under analysis, thereby affecting the value. The results of the calculation showed that Yongmun Station and Sinchang Station, located in the outer edge of the network, each had low network centrality, while Wangsimni Station and Dongdaemun History & Culture Park Station, connecting various networks, were shown to have high network centrality.
Fig 3. Reach centrality

Fig 4. Betweenness centrality
3.2 Analysis Results

Polynomial regression analysis was performed to analyze the effects of network centrality on the total traffic flow of each of the stations. Based on the analysis, adj. R square value of 0.243 was obtained. All the variables in the model had a significance level of 1%. The total traffic decreased by 0.329 when reach centrality was one unit higher. On the other hand, a one-unit increase in betweenness centrality and closeness centrality led to an increase in total traffic by 0.424 and 0.445, respectively (Table 1).

Table 1. Result of Polynomial Regression Analysis for Number of Trip

<table>
<thead>
<tr>
<th>Number of Trip</th>
<th>Adj.R.squ</th>
<th>Collinearity statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>St. Coef.</td>
<td>t</td>
</tr>
<tr>
<td>Reach</td>
<td>-.329</td>
<td>-3.042</td>
</tr>
<tr>
<td>Betweenness</td>
<td>.424</td>
<td>6.623</td>
</tr>
<tr>
<td>Closeness</td>
<td>.445</td>
<td>5.001</td>
</tr>
</tbody>
</table>

In order to examine the effects of network centrality on the travel distance to each of the stations, polynomial regression analysis was performed 3 times after classifying the dependent variable as short-, mid- and long-distance travel ratios. The results of the analysis showed that the R-square of each model was 0.422, 0.570 and 0.662, respectively. The significance of the
independent variables in each model was examined. In the model analyzing the short-distance travel ratio, reach centrality and closeness centrality had a significance level of 1%. In the model analyzing the mid-distance travel ratio, closeness centrality had a significance level of 1%, which between centrality had a significance level of 10%. In the model analyzing the long-distance travel ratio, closeness centrality had a significance level of 1%.

As reach centrality increased by one unit, the short-distance travel ratio increased by 0.256, while a one-unit increase in betweenness centrality resulted in the mid-distance travel ratio decreasing by 0.80. When the closeness centrality increased by one unit, the short-distance and mid-distance travel ratio increased by 0.371 and 0.817, respectively, and the long-distance travel ratio decreased by 0.725.

Table 2. Result of Polynomial Regression Analysis for Each Distance Ratio

<table>
<thead>
<tr>
<th>Each Distance Ratio</th>
<th>Short_ratio</th>
<th>Adj.R.squ</th>
<th>.422</th>
<th>Mid_ratio</th>
<th>Adj.R.squ</th>
<th>.570</th>
<th>Long_ratio</th>
<th>Adj.R.squ</th>
<th>.662</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reach</td>
<td>.256</td>
<td>2.712</td>
<td>.007</td>
<td>-.016</td>
<td>.191</td>
<td>.489</td>
<td>-.117</td>
<td>.613</td>
<td>.107</td>
</tr>
<tr>
<td>Betweenness</td>
<td>.059</td>
<td>1.054</td>
<td>.292</td>
<td>-.080</td>
<td>-1.649</td>
<td>.100</td>
<td>.024</td>
<td>.553</td>
<td>.581</td>
</tr>
<tr>
<td>Closeness</td>
<td>.371</td>
<td>4.774</td>
<td>.000</td>
<td>.817</td>
<td>12.179</td>
<td>.000</td>
<td>-.725</td>
<td>-12.210</td>
<td>.000</td>
</tr>
</tbody>
</table>

4. Conclusion

The purpose of this study was to examine the effect of the structural characteristics of the subway network of Seoul Metropolitan Area on the user-behavior of passengers. For this purpose, the network centrality was determined and its effects on the variables such as the number of users and moving distance of each station were analyzed using a polynomial regression model. The results of the analyses performed in this study can be divided into three types according to the indicator of network centrality.

First, they are results related to reach centrality. Reach centrality is a result of simply counting how many stations are nearby and the higher it is, the more neighboring stations there are. The higher the reach centrality, the more choices passengers had in selecting a neighboring station as the start point or the end point of their travel, which resulted in the total traffic. Also, the proportion of short-distance travel increased because passengers using the station concerned had more choices in terms of the stations that they could choose as the start point or the end point of their travel.

Second, they are results related to betweenness centrality. Betweenness centrality signifies the traffic must pass through a station in order for passengers to move between two regions using the shortest link. In other words, the higher the betweenness centrality, the more important traffic center it is. The analysis showed that the betweenness centrality value had no significance with respect to the short- and long-distance travel ratios, which means that it has no relation to travel distances. However, it showed a positive correlation to the amount of traffic at stations.

Lastly, they are results related to closeness centrality. Closeness centrality is an indicator that shows how close a station is to all the other stations based on a sum of the shortest distances to the other stations. The higher the sum of the shortest distances to the other stations, the longer
the distance one must travel between stations. Closeness centrality is the reciprocal of the sum; the higher the closeness centrality, the shorter the travel distance between stations. The results of the analysis showed that an increase in closeness centrality led to a decrease in the proportion of long-distance travel and an increase in short- and mid-distance travel.

Despite these results, there were limitations in the previous studies as other factors that influence the user-behavior of passengers were not controlled. There are plans to perform additional studies in the future and derive results while controlling the land use conditions in each of the stations.

This study verified that the subway network centrality has an impact on the user-behavior of passengers, which proves that network centrality is an important factor to be considered by the government when establishing policies related to subway.

Notes:
(1) Build a fast and convenient public transport system, Enhanced transportation demand management, Build the foundation for a green public transport, Building the transport system based on the basic services, Strengthen the competitiveness of public transport industry(Kang Hui-up at al. 2011).
(2) On the basis of 2009(Korea Transport Database)
(3) Scatterplot for number of passenger is a right-skewed curve. These case, The Median is better than Mean for representative value (David S. Moore, 2010).

References:
7) Han Ju-Seong. (2010), “Geography of Transportation”, seoul, Hanul Publishing Group

13) Korea Transport Database : http://www.ktdb.go.kr/


