

스마트카드 데이터를 이용한 심야버스 이용 수요 특성분석 : 강남역을 중심으로

Analysis of User Demand Characteristics of Currently-established Night Bus in Seoul by Using Smart Card Data : Case Study on Gangnam Station

김민주* · 이영인**

* 주저자 : 서울대학교 환경대학원 환경계획학과 교통학전공 석사

** 교신저자 : 서울대학교 환경대학원 환경계획학과 교수

Min ju Kim* · Young ihn Lee**

* Transportation studies Group, Department of Environmental Planning, Seoul National University

** Department of Environmental Planning, Seoul National University

† Corresponding author : Young Ihn Lee, yilee@snu.ac.kr

Vol.16 No.1(2017)

February, 2017

pp.101~116

ISSN 1738-0774(Print)

ISSN 2384-1729(On-line)

[https://doi.org/10.12815/kits.](https://doi.org/10.12815/kits.2017.16.1.101)

2017.16.1.101

요약

본 연구에서는 대중교통 이용객의 대부분이 사용하는 스마트카드 자료를 이용하여 실제 심야시간 통행량을 추정하고, 이를 현재 심야버스 노선과 비교하여 KT통신사에서 심야시간에 발생한 통화량을 기준으로 선정된 심야버스 노선과 실제 심야통행량을 비교하였다. 심야버스와 심야통행량을 비교하기 위해 일치도와 관련된 지표를 제시하고, 현재 얼마나 서비스가 제공되고 있는지 계산하였다. 연구의 특이한 점은 같은 정류장명을 가진 버스정류장이라도 어느 방향에 있느냐에 따라 속한 행정동이 다르기 때문에, 기준이 되는 영향권의 지표를 지하철역으로 잡고, 지하철역과 인접한 행정동을 지하철역의 영향권의 범위로 설정한 것이다. 심야버스가 심야시간 주요 교통수단으로 자리 잡은 만큼, 실제 교통량을 계산하여 노선을 선정한다면, 더 나은 서비스를 제공할 수 있을 것으로 기대된다.

핵심어 : 스마트카드데이터, 빅데이터, 버스노선, 심야버스, 이용자수요

ABSTRACT

This Study estimates the actual night traffic using the smart card data used by most of the public transportation users, and compares it with the current night bus routes by KT Telecom based on the night time call volume. In order to compare the current night bus and night trips evaluated by smart card data, we presented indicators related to the degree of matching, and estimated the volume of service currently provided. The unique approach of the study is that we chose subway station instead of bus stop for the unit of the study. Bus stops has their complexity in a way that stops with same name could belong to different administrative area depending on its direction. For this reason, we decided to use subway station and defined its adjacent administrative district as the scope of influence. Since night bus is the primary means of transportation during the late night, it is anticipated that they will be able to provide better service by calculating the actual traffic and selecting the routes.

Key words : Smart Card Data, Big Data, Bus Routes, Night Bus, User Demand

Received 29 November 2016

Revised 16 December 2016

Accepted 1 February 2017

© 2017. The Korea Institute of Intelligent Transport Systems. All rights reserved.

I. 서론

1. 연구배경

최근 현대인들은 높은 삶의 질을 추구하면서 여가시간을 중요하게 생각하게 되었다. 따라서 퇴근 후 바로 귀가하지 않고 여가통행을 하면서 다양한 심야 통행이 발생하게 되었다. 그렇게 되면서 2013년 처음 도입된 심야버스는 최근 3년간 666만 명이 이용할 정도로 심야시간대의 주요 교통수단으로 자리매김하였다(Yonhop News, 2016). 심야 통행이 증가함에 따라 심야버스나 콜버스(Call-bus) 등과 같이 야간에도 통행 서비스를 제공하는 교통수단이 점점 늘어나고 있다. 하지만, 현재 심야버스노선은 실제 심야시간 통행량이 아닌 KT통신사에서 심야시간에 발생한 통화량을 조사한 자료를 토대로 설정된 노선이므로 이 노선이 실제 통행량에 맞게 제대로 서비스가 제공되고 있는가에 대한 연구가 필요하다.

2. 연구목적 및 필요성

기존 심야버스 노선은 KT통신사에서 심야시간에 발생한 통화량을 조사하여 서울시 공공데이터와 연계하여 분석하였다(News Jelly, 2016). 이는 심야통화량을 기준으로, 심야시간에 인구가 많이 몰려 있는 곳을 대상으로 심야버스의 수요 예측 및 심야버스 노선을 선정하였다. 심야시간 통화량으로 심야시간 유동인구를 추정하고, 이를 통해 심야버스 노선을 결정한 것은 빅데이터(Big data) 분야에서 높게 평가할 만한 업적이다.

하지만, 이는 실제 통행량이 아닌 통화량을 기준으로 결정한 노선이므로, 이 노선이 적절한지에 대한 연구가 필요하다. 따라서 본 연구에서는 심야버스가 시행되기 전의 스마트카드 자료에서 심야 통행량을 추출하는 방법을 구체적으로 제시하고, 이를 이용하여 심야 통행량과 현재 운행 중인 심야버스를 비교하여 심야노선의 서비스 제공률에 대해 분석하고자 한다. 스마트카드로 알 수 있는 심야 통행수요를 현재 운행 중인 심야버스가 얼마나 충족시키는 지를 확인하여 심야버스 이용수요의 특성을 파악하고 앞으로 새로운 심야버스 노선을 계획할 때 고려할 수 있다.

본 연구에서는 집계된 마지막 통행이 비가정기반형(Non-home based trip)통행인 경우, 모두 집으로 돌아가기 위한 심야통행이 발생한다는 가정 하에 심야버스가 운행되기 이전의 자료를 이용하여 현재를 평가하는 연구를 진행한다. 사람의 통행에 대한 분석을 기본으로 하기 때문에, 본 연구에서는 기본적인 가정을 제시하고 이를 토대로 연구를 진행한다.

II. 선행연구 고찰

1. 관련연구

1) 대중교통 심야 운행 계획

Chen et al.(2013)은 택시에 부착된 GPS장치를 이용하여 야간버스의 노선을 계획하였다. 버스가 운행되지 않는 야간에는 택시만이 유일한 교통수단이므로, 택시 GPS정보를 버스 노선으로 전환하였다. 또한 버스정류장을 설정하기 위해서 승하차가 발생하는 비율이 높은 지역을 셀로 나누어 구역을 만들고, 이러한 구역을 연결하는 경로를 생성하는 알고리즘을 구축했다. 가장 짧은 길이의 경로를 탐색하는 것과 같은 기존연구와는

다르게 이 연구에서는 제한된 시간동안 최대의 승객을 이동시킬 수 있는 경로를 탐색하였다.

Yoon et al.(2011)은 현재 경기도 심야버스 운행현황을 제시하고 설문조사를 통하여 실제 이용자의 선호 및 불편 사항을 조사하였다. 또한 대중교통인프라 확대, 사회활동의 변화, 야간 교통사고의 위험성 등과 같은 심야연장 운행의 필요성을 제시함과 동시에 심야연장운행의 문제점으로 수요확보, 장비인력 및 확충, 안전유지 등을 꼽았다. 또한 이러한 문제점에 대한 대안을 제시하면서 광역 전철인 경인선 열차의 심야연장 운행을 제안하였다.

2) 기종점 선정

Munizaga MA. and Palma C.(2012)는 산티아고의 스마트카드 시스템을 통해 얻은 정보로 대중교통의 OD 모형을 추정하고자 하였다. 산티아고에서는 대중교통을 이용 할 때, 스마트카드를 이용하여 요금을 지불한다. 그러므로 이를 통해 이용자의 시공간적 분포를 알 수 있고, 이를 통해 2009년 3월과 2010년 6월의 이용자의 통행패턴과 기종점을 측정하였다.

Ahn and Lee(2007)는 서울시 스마트카드 자료를 이용하여 정류장을 기반으로 하는 대중교통 통합 OD(Origin-Destination) 자료와 네트워크를 구축하고자 하였다. 기존의 행정동 기반의 대중교통 계획의 문제점을 들면서, 기존 수요 추정시 신뢰성이 떨어진다고 하였다. 그러면서 정류장기반 대중교통 OD 및 네트워크 구축의 필요성을 언급하였다. 각 정류장 별 위치좌표와 버스노선자료를 이용하여 노드와 링크를 구축하였고, 정류장 기반 대중교통 OD와 네트워크를 EMME2를 이용하여 분석해 연구에서 구축한 OD 및 네트워크의 장점을 확인하였다.

Chang et al.(2011)은 대중교통의 이동성 지표와 대중교통 잠재수요에 대한 개념을 정립하여 이를 통해 직결노선 OD를 선정하고자 하였다. 잠재수요와 이동성 지표를 이용하여 통행시간을 최대로 줄일 수 있는 모형과 알고리즘을 구축하였다. 또한 국가교통DB(data base)에 대중교통 노선자료가 구축 되어있지 않은 수도권 권을 제외한 지방지역을 사례로, 각 지방의 이동성 사각지대와 잠재수요 발생지역을 분석하고 이를 통해 기종점을 선정하였다.

3) 버스 노선 평가지표

Bae(2003)는 특정 기·종점을 연결하는 지선노선 선정으로 범위를 제한하여 연구를 하였다. 1단계로 일정 기종점을 연결하는 무수한 버스노선경로 중에서, 종점까지의 소요시간이 짧은 도로 구간을 노선 설치 도로망으로 한정하는 검토대상 도로망 한정과 2단계 버스 노선의 일반적인 제약 조건 중 물리적, 시간적 길이를 나타낼 수 있는 요소만을 고려하여 후보 노선을 선정하는 허용노선 장에 의한 후보노선의 열거 단계를 거쳐 3단계로 버스운행자의 수익성을 최대화하는 노선을 최종노선으로 선정하기 위해 노선 평가치를 구축하여 노선을 선정하였다. 실용성에 비중을 두어 간단한 개념의 노선선정 기법을 구축하였고, 마을버스와 같은 지역 내 단거리 운행 버스노선의 선정 작업에 유효한 기법을 제안하였다.

Lee and Park(2003)은 도시 내 버스 전체 노선 체계에 대한 합리적이고 형평성 있는 평가를 위해 평가기준과 정량적 지표를 설정하고 평가사례를 들어 제시하였다. 3가지 평가 기준으로 버스이용의 편의성, 운행의 생산성, 사회적 비용의 최소화를 제시하였고, 이를 바탕으로 버스이용자, 버스 운영자의 관점으로 대상별로 나누어 버스 이용자는 접근성, 승차안락성, 환승률, 버스 운영자는 운행생산성, 노선직결성, 그리고 사회적 관점에서 지역적 형평성, 차량 소요대수 등 7개의 정량적 지표를 설정하였다. 또한 적용사례로 경기도 시흥시를 대상으로 기존 노선을 포함하여 4개의 대안을 제시하여 평가지표별로 대안을 평가하였다. 평가지표별

로 유리한 대안이 각각 달라 종합적인 평가를 위해 판단기준표준화 방법으로 선형정규화법을 사용하였으며, 그 결과 대안3인 기존노선을 부분 개선하는 방안과 대안4인 노선을 전면 개편하는 방안이 바람직한 대안이라는 결과가 도출되었다.

Ha and Cheon(2011)은 버스의 노선변경에 따른 노선의 적정성 평가 방법에 대한 연구를 하였다. 변경된 버스노선이 직접적으로 영향을 미치는 노선 운영자와 이용자 측면으로 구분하여 평가 지표를 설정하고, 실제 사례를 통해 평가하였다. 6514번 버스의 노선이 변경되는 것을 사례로 하여, 운영자 측면에서는 버스 한 대당 승객 수를 산출하고, 이용자 측면에서는 통행시간 변화를 평가지표로 제시하였다. 그 결과, 운영자 측면에서는 하루 총 승객수가 807명 증가하였으므로 버스 노선 변경이 적정하다고 판단할 수 있지만, 이용자 측면에서는 구간에 따라 통행시간이 단축된 이용자와 서비스를 받지 못하는 이용자로 나뉘기 때문에 정확한 분석을 하기는 어렵다. 이에 대한 분석을 위해서는 기종점 버스 정류장을 보행권역으로 군집화 시켜 분석할 필요성이 있다고 하였다.

4) 휴대폰 자료를 이용한 수요추정

Demissie et al.(2016)은 휴대폰 자료를 이용하여 개발도상국의 통행 수요를 추정하였다. 이를 이용하여 효율적인 교통 노선과 정류장을 구축하고자 하였다. 항상 양방향 통행이 발생하기 때문에, 집에 있는 전화타워의 위치를 이용하여 심야시간(10PM-7AM)에 빈번하게 발생하는 지점을 심야통행이 발생하는 곳으로 보았다. 이렇게 심야통행이 발생하는 전화타워들을 군집화하고, 이를 연결하여 노선을 선정하는 방법을 제시하였다.

2. 시사점 도출

기존 연구를 통해 대중교통의 심야운행과 스마트카드 자료를 이용하여 기종점을 선정하는 방법, 버스 노선을 선정하는 경우 고려해야할 분석 지표를 살펴보았다. 그리고 우리나라의 심야버스 노선 선정 방식과 유사한 휴대폰 자료를 이용한 심야통행 수요 추정하는 연구도 볼 수 있다.

대중교통의 심야운행은 사람들이 다양한 활동을 하기 시작하였고, 대중교통의 영향권이 점점 넓어지고 있기 때문에 필요하다. 하지만, 심야 수요에 대한 확보가 필수적이기 때문에 스마트카드 데이터 자료를 이용하여 심야통행량을 추정할 수 있다. 추정된 수요를 가지고 버스 노선을 선정하는 데에 있어서는 기존 연구는 이용자와 운영자의 측면으로 나눠서 분석 지표를 설정하여 노선을 선정하였다. 또한 이러한 분석 지표 설정은 다양한 지표 중 분석가의 주관적인 판단에 의해 설정된 지표이므로 분석가에 따라 다양한 결론이 나올 수 있다. 그리고 각 도시별, 노선별로 분석 지표가 다르게 적용되어야 하므로 사례에 따라 유동적으로 적용시켜야 한다.

본 연구에서는 서울시 스마트카드 자료를 이용하여 실제 심야통행량을 추출하여 OD를 구축하고, 노선 선정에 있어서 다양한 지표를 고려해야 하지만, 심야시간대 통화량과 실 통행량을 비교하여 심야버스 이용수요 특성에 대해 분석하고, 이를 통화량을 기준으로 계획한 심야 버스 노선과 비교하고자한다.

Ⅲ. 스마트카드 자료를 이용한 노선 이용특성 분석 방법론

기존의 심야버스는 심야시간대 통화량을 기준으로 심야시간대의 수요를 추정하여 버스 노선을 선정하였

다. 본 연구에서는 이러한 방식으로 추정한 수요대신 스마트카드 자료를 이용하여 실제 통행량을 구하여, 기존 심야버스 노선이 실제 통행량에 부합하는 서비스를 제공하는지를 분석하고자 한다.

이는 스마트카드에 의해 귀가통행이 발생하지 않은 통행자가 심야버스 서비스가 없는 지역에서는 타 수단을 이용하여 귀가하였다는 가정아래, 이를 심야버스 노선이 신설되면, 심야 버스를 이용할 잠재적인 수요자로 규정한다는 전제 하에 분석된다.

본 연구에서는 스마트카드 자료 내에서 행정동 별로 심야통행이 얼마나 발생하는지를 추출한다. 다음으로 이렇게 추출한 통행량과 기존 심야버스 노선이 지나는 행정동과 얼마나 일치하는지와, 서비스가 제대로 제공되고 있는지를 평가지표를 통해 분석한다.

1. 행정동 별 심야통행 비율

심야시간대에 발생하는 전체 통행량 중 행정동 별 수요를 이용하여 수요 비율을 계산한다.

$$D_{i,j} = \frac{N_{i,j}}{\sum_{j=1}^k N_{i,j}} \quad (1)$$

$D_{i,j}$: 행정동 i 에서 발생하는 전체 심야 통행 중 행정동 j 로 가는 심야통행 비율(%)

$N_{i,j}$: 행정동 i 에서 행정동 j 로 가는 심야통행량

k : 전체 행정동 수

2. 노선영향권과 서비스 제공률

심야버스 노선은 실제통행량을 기준으로 조성된 것이 아니라, 통화량을 기준으로 노선을 선정한 것이기 때문에 실제 통행량에 얼마나 현재 심야버스가 서비스를 제공하는지에 대한 평가가 필요하다. 모든 행정동 전체로 비교를 했을 실제 심야통행량의 행정동과 기존 심야버스가 지나는 행정동을 비교하여 심야버스 이용 수요와 서비스 제공률을 평가하고자 한다.

1) 실제통행과 심야버스 노선 영향권의 일치도

스마트카드 데이터 자료를 이용하여 심야시간 실제통행량을 알 수 있다. 심야시간 실제 통행량의 행정동과 기존 심야버스 노선이 지나는 행정동을 비교한다. 이 때, 기존 심야버스 노선에 인접한 지하철역이 포함하는 영향권을 심야버스가 영향을 미치는 행정동으로 기준을 잡는다. 두 데이터 모두 행정동을 기준으로 구축하였으므로, 두 데이터의 행정동을 비교하여 두 데이터의 일치도(CD : Conformity Degree) 지표를 개발하였다.

$$CD_{i,j} = \frac{S_{i,j}}{A_j} \quad (2)$$

$CD_{i,j}$: 행정동 i 에서 행정동 j 로의 심야통행과 기존 심야버스 노선의 일치도

$S_{i,j}$: 행정동*i*에서 행정동*j*로 발생하는 실제 심야통행

A_j : 기존 심야버스 노선이 지나는 영향권에 속하는 행정동*j*

이 때, 실제 심야통행에 기존 심야버스 노선이 지나가면 $CD_{i,j} = 1$, 그렇지 않으면 $CD_{i,j} \neq 1$ 가 나온다.

2) 해당지역 실제 서비스 제공률

각 행정동에 발생하는 심야 통행량과 그 지역을 지나는 버스 노선을 비교하여 해당지역에 실제로 제공되는 서비스율을 알아보고자 한다. 이 때의 해당지역 실제 서비스 제공 인원(RSP : Real Service People)을 산정하는 식은 다음과 같다.

$$RSP_{i,j} = \sum_{n=1}^m (n \times p) \times c \quad (3)$$

$RSP_{i,j}$: 심야버스로 행정동*i*에서 행정동*j*로 서비스를 제공할 수 있는 인원 수(명)

n : 행정동*i*와 행정동*j*를 지나는 심야버스 노선 수

p : n 노선 운행시간동안의 총 버스 운행 대수(대)

c : 서비스 수준에 따른 탑승인원(명)

여기서, 버스 1대당 탑승인원은 <Table 1>의 입석형 버스의 서비스수준을 따른다. 위에서 구한 $RSP_{i,j}$ 와 실제 행정동*i*에서 행정동*j*로의 심야통행량인 $N_{i,j}$ 을 비교하여 행정동*i*에서 j 로의 실제 서비스 제공률(RS : Real Service rate)을 산정한다.

$$RS_{i,j} = \frac{RSP_{i,j}}{N_{i,j}} = \frac{\sum_{n=1}^m (n \times p) \times c}{N_{i,j}} \quad (4)$$

$RS_{i,j}$: 행정동*i*에서 행정동*j*로 제공하는 서비스율

이 때, 실제 심야 통행량보다 심야버스로 제공할 수 있는 인원이 더 많은 경우 $RS_{i,j} \geq 1$, 실제 심야 통행이 심야버스로 제공할 수 있는 인원보다 많은 경우는 $RS_{i,j} < 1$ 의 값을 가진다.

<Table 1> Service Level of Standing Bus

Level Of Service	Number of Passengers	Notes
A	≤15	<Urban bus> ·Vehicel Area 26.37m ² ·Standing Area 7.5m ² ·Based on 31 Seats
B	≤31	
C	≤40	
D	≤50	
E	≤62	
F	>62	

*KHCM(Korea Highway Capacity Manual), (2013)

IV. 적용 및 검증

1. 스마트카드 자료 구조

스마트카드는 대중교통 이용정보가 시스템을 통해 기록되므로 사람이 직접 조사하는 방법보다 더 정확하고, 1년 365일 24시간 내내 수도권을 포함 대도시 어느 지역에서 이용 가능하므로 지역별, 시간별 자료 수집에 용이하고 이미 시스템을 통해 자료가 수집되고 있으므로 이용 면에서도 효율적이다.

따라서 본 연구는 이러한 대중교통카드 자료를 활용하여 기존 심야버스 노선 선정에 이용되었던 통화량으로 수요를 추정하지 않고 실제 통행 수요량을 추정하는 것이 가능하기에 스마트카드 자료를 이용해서 분석하고자 한다. 이용하는 스마트카드 자료는 <Table 2>와 같이 47개의 정보를 제공하며, 이용자가 승하차를 하며 단말기에 카드를 태그할 때마다 47개의 정보가 새롭게 기록된다.

<Table 2> Structure of Smart Card Data

No	Data	No	Data	No	Data	No	Data	No	Data
1	vCard_No	2	Transportation_Cd	3	Transaction_Id	4	Busroute_Id	5	Busroute_Name
6	Transportoperators_Id	7	Transportoperatos_Name	8	Vehicle_Id	9	VehicleNumber	10	SstartTime
11	SendTime	12	RidingTime	13	RideStation_Id	14	RideStation_Name	15	GetOffStation_Time
16	GetOffStation_Id	17	GetOffStation_Name	18	TransferCount	19	UserclassCod	20	UserGroup
21	Seats	22	RidePay	23	GetOffPay	24	RideViolationPay	25	GetOffViolationPay
26	TotalDistance	27	TotalHour	28	Local_Code	29	Gcnt	30	Sent
31	Ccnt	32	Ecnt	33	Time_Code	34	Tdate	35	Rdate
36	Rtime	37	Gdate	38	Gtime	39	Means	40	Station_Code
41	ZoneCode	42	Station_Name	43	Dong_Code	44	Classification_Code	45	SiDo
46	SiGunGu	47	EupMyeonDong						

본 연구에서는 이 중 <Table 3>의 9개 자료, vCardNo(가상카드번호), RidingTime(승차시간), RideStationId(승차정류장ID), GetOffStationTime(하차시간), GetOffStationId(하차정류장ID), StationCode(역코드), StationName(역명), DongCode(행정동코드), EupMyeonDong(읍면동)만을 이용한다.

이 중에서 가장 중요한 지표는 GetOffStationTime(하차시간)과 RideStationId, GetOffStationId(승·하차 정류장ID)이다. 스마트카드 자료는 승·하차 정류장ID의 행정동 코드 및 전수화 존 코드를 제공하여, 각각의 정류장이 어느 행정동과 존에 위치하고 있는지 알 수 있다.

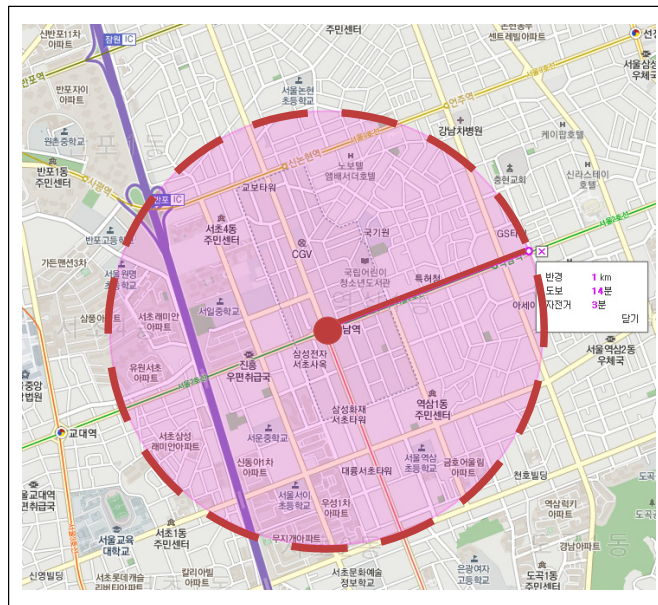
<Table 3> List of Extracted Data

No	Data	No	Data	No	Data	No	Data	No	Data
1	vCard_No	12	RidingTime	13	RideStation_Id	15	GetOffStation_Time	16	GetOffStation_Id
40	Station_Code	42	Station_Name	43	Dong_Code	47	EupMyeonDong		

2. 적용범위

본 연구에서는 2013년 5월 21일 화요일에 강남역 주변 지역에서 발생한 통행을 대상으로 분석하고자 한다. 대상 지역인 강남역 주변은 심야시간대(오후11시~오후1시) 택시 승·하차 비율이 가장 높은 지점인 강남구 역삼1동(Ahm, 2015)에 속해있으며, 역삼1동 중 가장 유동인구가 많은 지하철역이므로 중심지역으로 설정하였다.

또한 기본 가정에서 지하철역의 공간적 영향권을 반경 1km로 설정하였기 때문에 중심지역의 범위는 강남역을 중심으로 <Fig. 1>과 같이 반경 1km로 볼 수 있다.



<Fig. 1> Central Area

3. 기본 가정

본 연구에서는 스마트카드 데이터를 이용하여 이용자의 통행 행태를 분석하기 위해 몇 가지 가정을 제시한다.

첫째, 마지막 통행이 귀가 통행이 아닌 경우, 그 곳에서 버스나 지하철을 이용한 통행은 끝났지만, 이 후 대중교통 이외의 수단을 이용하여 귀가했다고 가정한다. 이는 스마트카드에 집계된 모든 통행의 마지막 통행이 귀가통행이 아닌 경우, 심야버스 노선이 신설되면 심야버스를 이용할 잠재적인 수요자로 집계할 수 있다는 것을 의미하며, 이를 기반으로 심야버스 노선의 서비스 제공률을 확인할 수 있다.

둘째, 심야 시간에 버스나 지하철을 이용할 수 있다면, 이용객은 보통 지하철 한 정거장의 길이인 1km 정도는 걸어갈 의향이 있다고 가정한다.

현재 국내에서는 일반적으로 역세권의 범위를 역사 중심에서 반경 500m 이내로 보고 있다(Kim, 2012). 하지만 본 연구에서는 심야시간이라는 점을 고려하여 역세권 범위의 2배 거리, 즉 1km를 이용자의 도보허용거리로 가정한다.

셋째, 공간적 기준이 되는 지점은 노선이 지나는 지점과 가장 인접한 지하철역으로 보고, 지하철역의 공간적 영향권은 반경 1km로 설정한다. 도보허용거리를 1km로 가정하였으므로 이용자에게 영향을 미칠 수 있는 범위 또한 역사로부터의 1km로 가정한다.

4. 분석 DB구축

방대한 양의 데이터 자료인 스마트카드 자료를 이용하기 위해서는 분석에 필요한 데이터만을 추출하여 데이터양을 줄이는 작업이 필요하다.

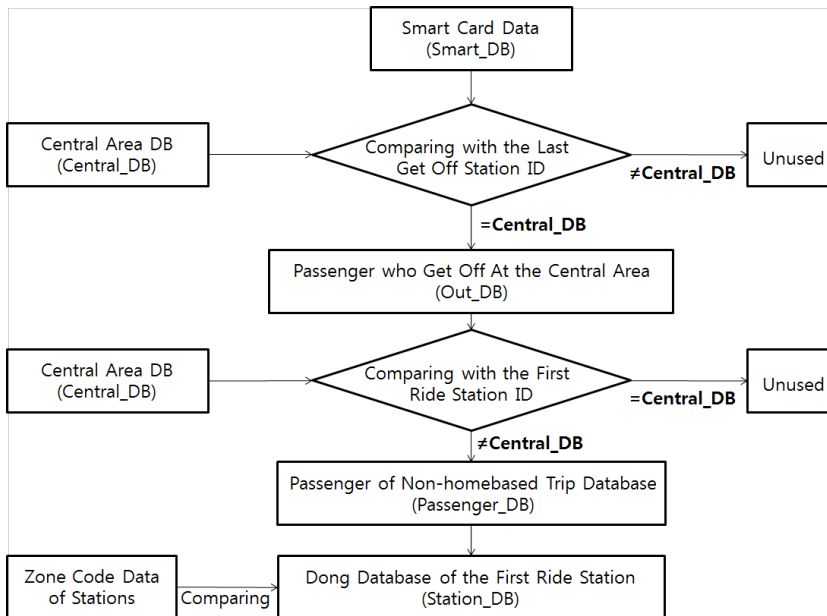
본 연구에서는 방대한 양의 스마트카드 자료에서 분석에 필요한 부분만을 추출하는 방법을 설명하고, 추출된 데이터를 이용하여 심야버스 이용수요를 확인하고자 한다.

또한, 심야버스 이용수요의 특성을 파악하여 현재 운행 중인 심야버스가 얼마나 이용자에게 서비스를 잘 제공하고 있는지를 비교한다.

<Fig. 2>은 스마트카드 내의 자료를 1)부터 5)까지의 과정을 통해 심야통행을 추출하는 과정을 도식화 한 알고리즘으로, <Fig. 2>을 따라 추출한 데이터를 이용하여 실제 심야버스 통행량을 계산할 수 있고, 이를 이용하여 심야버스 노선 서비스 제공률 등의 평가지표를 계산할 수 있다.

본 연구에서는 하나의 카드번호에서 발생한 통행 중 환승을 고려하지 않고, 단지 최초 통행과 최종 통행만을 이용하여 심야통행을 조사한다. 구체적으로 최초 통행의 승차정류장ID와 최종 통행의 하차정류장ID를 이용한다. 또한, 연구에 이용하기 위해 공간적 범위를 지하철역을 중심으로 하여 반경 1km로 공간적 범위를 설정하고 각 정류장에 따른 승객DB를 구축한다.

<Fig. 2>에서 데이터를 추출하는 작업을 자세히 살펴보면 다음과 같다.



<Fig. 2> The Algorithm of late night Passing Data

1) 스마트카드 DB구축(Smart_DB)

스마트카드에서 제공하는 데이터 중에서 연구에 필요한 부분만을 추출하여 스마트카드DB(Smart_DB)를 구축한다.

2) 중심지역 DB구축(Central_DB)

버스정류장의 경우는 같은 역명의 정류장이라도 승차정류장ID와 하차정류장ID는 같지 않다. 또한 정류장이 영향을 미치는 공간적인 범위를 구체적으로 지정할 필요가 있다.

이에 여기서는, 기준점을 중심으로 심야시간 도보권으로 가정한 1km 반경 안에 들어가는 모든 버스정류장과 지하철역을 하나의 중심지역으로 묶어 중심지역DB(Central_DB)를 구축한다.

3) 중심지역 하차승객 DB구축(Out_DB)

중심지역 하차승객을 추출하기 위해 최종 통행의 하차정류장ID와 2)의 Central_DB 정류장ID를 비교한다.

이 때, 하차정류장ID가 Central_DB의 정류장ID와 일치한다면, 중심지역에서 통행이 끝났다고 할 수 있다. 하차정류장ID와 Central_DB의 정류장ID가 일치하는 데이터만 수집하여 중심지역 하차승객DB(Out_DB)를 구축한다.

- 하차정류장ID=Central_DB의 정류장ID ⇒ 추출(Out_DB)
- 하차정류장ID≠Central_DB의 정류장ID ⇒ 버림

4) 귀가통행이 아닌 하차승객 DB구축(Passenger_DB)

중심지역에서 통행이 끝난 승객에는 귀가통행을 한 승객도 포함된다. 따라서 귀가통행이 아닌 승객 데이터를 추출하기 위해 다음을 비교하고자 한다.

먼저, 3)의 Out_DB에서 이들의 최초 통행 승차정류장ID와 2)에서 구축한 Central_DB의 정류장ID를 비교한다. 귀가 통행의 경우 승차와 하차는 같은 지역에서 일어 날 것이므로 최초 통행의 승차정류장ID가 Central_DB의 정류장ID와 다른 데이터만을 수집한다. 이렇게 수집된 데이터는 최종 통행이 중심지역으로의 귀가통행이 아닌 비가정기반형통행(Non-home based trip)이다. 이를 통해 우리는 귀가통행이 아닌 중심지역 하차승객DB(Passenger_DB)를 구축한다.

Out_DB에서

- 최초 통행 승차정류장ID=Central_DB의 정류장ID ⇒ 버림
- 최초 통행 승차정류장ID≠Central_DB의 정류장ID ⇒ 추출(Passenger_DB)

5) 최초 통행의 승차정류장 행정동 DB구축(Station_DB)

가정에 따르면, 4)에서 추출한 최종 Passenger_DB의 승객은 기본가정에 따라 버스나 지하철의 운행시간이 지난 후, 귀가통행을 할 것이다. 이 승객은 심야버스가 신설될 경우, 심야버스를 이용할 잠재적인 수요로 볼 수 있다.

야간통행 수요 및 공간적 분포를 알기 위해 최초 통행의 승차정류장ID와 정류장별 존 코드자료를 비교하여 Passenger_DB의 행정동 DB(Station_DB)를 구축한다.

6) 기존 심야버스 노선의 영향권 DB구축(Area_DB)

현재 심야버스의 노선별 정류장을 조사하여, 그 노선 각각의 정류장과 가장 인접한 지하철역을 기준으로 하여 각 지하철 영향권에 속하는 행정동DB(Area_DB)를 구축한다.

V. 결 과

1. 분석 결과

1) 분석 DB구축

중심 지역인 강남역 반경 1km내의 모든 정류장을 조사해 보면, 역 수는 지하철역 4개, 버스정류장 240개로 총 244개의 정류장이 있다. 총 244개의 정류장에 해당하는 정류장ID로 구성된 Central_DB를 구축한다.

2013년 5월 21일 화요일에 발생한 통행을 제 4장 4절에서 제시한 방법으로 추출하면, 총 6310개의 Passenger_DB가 구축된다. Passenger_DB는 최종통행의 하차정류장은 중심지역으로 설정한 강남역 주변이지만, 이 지역에 거주하지 않은 사람들의 통행 데이터로 중심지역에서 발생하는 심야 통행이다. 발생한 6310개의 통행을 제 4장 4절 5)의 방법으로 행정동 자료와 비교하면, 강남역에서 발생한 통행을 행정동 별로 분류할 수 있고, 그 결과 Station_DB는 총 683개가 구축된다.

현재 서울시 심야버스 노선은 N13, N15, N16, N26, N30, N37, N61, N62로 총 8개의 노선이 조성되어있다. 이 중에서 중심지역인 강남역 일대를 지나는 3개의 노선은 N13, N37, N61이다. 제 4장 4절 6)에서 3개의 노선 각각의 Area_DB를 구축한다.

2) 행정동 별 심야통행량

위에서 구축한 총 683개의 Station_DB를 식(1)을 이용하여 행정동 별 심야통행량에 따라 순위를 매긴다. 또한 제 4장 4절에 제시한 방법으로 강남역에서 발생한 심야통행량을 구할 수 있고, 특히 제 4장 4절 4)의 Passenger_DB가 <Table 4>의 Traffic을 나타낸다.

이를 보면 241개의 통행이 발생한 서울특별시 서초구 반포1동(No.1)부터 단 하나의 심야통행만이 있는 경기도 연천군 전곡읍(No.683)까지 다양한 지역으로의 심야통행이 일어났음을 알 수 있다.

3) 기존노선과의 일치여부

먼저 실제 통행과 기존 심야버스 노선이 얼마나 일치하는지 보기 위해 제 4장 4절 5)에서 구축한 Station_DB와 6)에서 구축한 노선별 Area_DB를 사용한다.

<Table 4>는 강남역에서 발생한 전체 심야통행을 식(2)을 이용하여 각각의 동과 강남역을 지나는 세 개 노선 각각의 Area_DB를 계산한 표이다. 계산의 결과값이 $CD_{i,j} = 1$ 인 경우에는 1, $CD_{i,j} \neq 1$ 인 경우에는 0으로 표시하고, 각 행정동을 지나는 노선 수만큼 노선의 합계를 표시하였다.

중심지역에서 발생한 심야 통행은 683개의 동에서 발생하였으며, 세 개의 노선이 모두 지나는 동은 8개, 두 개의 노선이 지나는 동은 33개, 한 개의 노선만 지나는 동은 107개이고, 강남역을 지나는 심야버스의 서비스를 받지 못하는 동은 535개의 동으로 나타났다. 이는 전체 683개의 동에서 83.9%의 비율로 굉장히 높은 비율을 보이고 있음을 알 수 있다.

〈Table 4〉 Number of Matches Between Actual Late Night Traffic Volume and Current Night Bus Routes by Administrative District

No.	Dong	Traffic	N13	N37	N61	Σ
1	Banpo1	241	1	1	0	2
2	Seocho3	216	0	0	1	1
3	Banpo4	184	0	0	0	0
4	Yeoksam1	176	1	1	1	3
5	Nonhyun1	170	1	1	0	2
6	Yeoksam2	158	1	1	1	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
320	Dobong1	3	0	0	0	0
321	Sanggye5	3	1	0	1	2
322	Sanggye3·4	3	0	0	0	0
323	Bulgwang1	3	0	1	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
680	Gwangnam	1	0	0	0	0
681	Jangheung-myeon	1	0	0	0	0
682	Gasan-myeon	1	0	0	0	0
683	Jeongok-eup	1	0	0	0	0
Total		6310	70	51	76	197

*Σ = N13+N37+N61

〈Table 5〉 Descending by Route Total

No.	Dong	Traffic	N13	N37	N61	Σ
4	Yeoksam1	176	1	1	1	3
6	Yeoksam2	158	1	1	1	3
14	Seocho2	70	1	1	1	3
23	Seocho4	51	1	1	1	3
39	Daechi4	34	1	1	1	3
49	Daechi2	27	1	1	1	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮
537	Sanggye6·7	1	1	0	1	2
2	Seocho3	216	0	0	1	1
10	Seocho1	97	0	0	1	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
559	Sinwon	1	0	0	1	1
560	Jowon	1	0	0	1	1
633	Cheolsan1	1	0	0	1	1
3	Banpo4	184	0	0	0	0
8	Ogeum	135	0	0	0	0
9	Yangjae1	117	0	0	0	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
680	Gwangnam	1	0	0	0	0
681	Jangheung-myeon	1	0	0	0	0
682	Gasan-myeon	1	0	0	0	0
683	Jeongok-eup	1	0	0	0	0
Total		6310	70	51	76	197

*Σ = N13+N37+N61

*The parts described in text are distinguished by Bold and Shading

<Table 5>는 <Table 4>를 노선합계를 기준으로 내림차순 하여 나타낸 것이다. 여기서 Bold처리한 상계6·7동(No.537)과, 신원동(No.559), 광남동(No.680)을 먼저 살펴보면, 세 개의 동은 모두 중심지역에서 발생한 심야통행이 1통행임에도 불구하고, 각각 이 지역을 지나는 노선이 2개, 1개, 0개이다.

또한 음영 처리를 한 반포4동(No.3), 오금동(No.8), 양재1동(No.9)은 강남역에서 발생한 심야통행량이 각각 184통행, 135통행, 117통행이 발생했음에도 불구하고 강남역을 지나는 심야버스가 한 대도 가지 않는 것을 알 수 있다. 이를 통행 실제 이용수요를 심야버스가 충족시키지 못하는 것을 알 수 있고, 이는 심야버스노선이 단지 통행량을 기준만으로 선정된 것이 아님을 알 수 있다.

4) 해당지역으로의 실제 서비스 제공률

현재 중심지역을 지나는 심야버스 노선은 N13, N37, N61, 총 세 개의 노선이 있다. 이들의 운행정보는 <Table 6>과 같다. 심야시간동안 세 개의 노선이 총 21대의 버스를 운행한다.

각 행정동으로의 실제 서비스 제공률을 식(3)과 식(4)를 이용하여 하여, A~F 서비스 수준에서 평균 서비스 수준인 C를 기준으로 실제 심야통행에 얼마나 서비스가 잘 제공되고 있는지 살펴보았다.

실제 서비스가 잘 제공되고 있기 위해서는 $RS_{i,j} \geq 1$ 이 나와야한다. <Table 7>은 <Table 4>를 서비스수준 C를 기준으로 하여 버스 1대당 수송인원을 40명으로 하여 $RSP_{i,j}$ 와 $RS_{i,j}$ 를 구한 표이다.

노선이 지나는 모든 행정동은 $RS_{i,j} \geq 1$ 이 나왔지만, 노선이 지나지 않는 동은 서비스가 제공되지 않아 $RS_{i,j} = 0$ 의 값이 나왔다. 따라서 서비스가 제공되기만 한다면, 심야시간대의 통행량에 적절한 서비스를 제공할 수 있다는 것을 알 수 있다.

하지만, <Table 7>에서 볼 수 있듯이 통행량이 많은 순서로 나열한 행정동 목록에서 Bold처리한 부분을 보면, 반포4동(No.3)이나 오금동(No.8), 양재 1동(No.9) 등은 강남역을 중심으로 한 지역에서 발생한 전체 683개의 동으로의 심야통행 중에서 높은 심야통행량을 가지고 있는 동이지만 심야버스 노선의 서비스를 제대로 받지 못하고 있음을 알 수 있다.

또한 이와는 반대로, 음영 처리한 상계5동(No.321), 불광1동(No.323), 응암1동(No.324), 응암2동(No.325)은 심야노선은 존재하나, 강남역에서 발생하는 심야통행 수요는 적은 것으로 나타났다.

이를 통해, 강남역에서의 심야통행량이 많으나 심야버스 노선 서비스가 잘 제공되지 않는 반포4동(No.3), 양재1동(No.9) 같은 경우는 강남역에서 비교적 가까운 지역이므로 심야버스 노선에서 제외된 것으로 보인다. 또한, 강남역에서 심야통행량은 적으나 불광1동(No.323), 응암1동(No.324), 응암2동(No.325) 같은 경우는 심야버스가 지나는 지역이므로 다른 지역에서의 이 지역으로의 심야통행량이 많을 수도 있다고 볼 수 있다. 이러한 결과는 강남역에서 발생한 심야통행량만 조사하여 심야버스와 비교한 것이라 과도하게 제공되는 것처럼 보인다.

따라서 본 연구에서 제시한 심야버스 이용수요 추정 방법을 토대로 이용수요의 특성을 파악하여, 이를 토대로 강남역뿐만 아니라 서울시 전체에서의 심야통행량을 추정하여 비교해보아야 한다.

<Table 6> Night Bus Route Information

Route	Operation Hour	Interval	Total number of operations
N13	12:00-03:25	40~45min	6
N37	12:00-03:10	40min	6
N61	11:50-03:50	30~45min	9

〈Table 7〉 Actual Service Offered Based on Service Level C from 〈Table 4〉

No.	Dong	Traffic	Σ	RSP	RS
1	Banpo1	241	2	480	2.0
2	Seocho3	216	1	360	1.7
3	Banpo4	184	0	0	0
4	Yeoksam1	176	3	840	4.8
5	Nonhyun1	170	2	480	2.8
6	Yeoksam2	158	3	840	5.3
7	Samseong2	150	2	600	4.0
8	Ogeum	135	0	0	0
9	Yangjae1	117	0	0	0
10	Seocho1	97	1	360	3.7
⋮	⋮	⋮	⋮	⋮	⋮
320	Dobong1	3	0	0	0
321	Sanggye5	3	2	600	200
322	Sanggye3·4	3	0	0	0
323	Bulgwang1	3	1	240	80
324	Eungam1	3	1	240	80
325	Eungam2	3	1	240	80
⋮	⋮	⋮	⋮	⋮	⋮
678	Byeongjeom2	1	0	0	0
679	Silchon-eup	1	0	0	0
680	Gwangnam	1	0	0	0
681	Jangheung-myeon	1	0	0	0
682	Gasan-myeon	1	0	0	0
683	Jeongok-eup	1	0	0	0
Total		6310	197		

*RSP based on Level of Service C in <Table 4>

RSP(Real Service People)=(N13×6+N31×6+N61×9)×40

*RS(Real Service)=RSP/Traffic

*The parts described in text are distinguished by Bold and Shading

VI. 결론 및 향후과제

본 연구에서는 스마트카드 자료를 이용하여 실제 심야시간에 발생한 통행량을 계산하였다. 또한 이를 심야시간대 통화량을 기준으로 설정한 기존 심야버스 노선과 비교하기 위해 정류장을 행정동을 기준으로 분류하였고, 또한 심야버스노선과 비교하여 제공되는 서비스율 계산식을 제시하였다. 이 지표를 이용하여 중심지역에서 발생하는 통행량과 기존노선이 운반하는 인원을 계산하여 노선의 서비스 제공률을 확인하였다.

본 연구는 빅 데이터인 스마트카드 자료에서 분석에 필요한 데이터만을 뽑아내어 새로운 데이터베이스를 구축하여 데이터를 효과적으로 다루고, 지하철역을 기준으로 공간적 범위를 재설정하여 중심지역인 강남역

일대에서 발생한 심야통행 데이터를 추출하였다. 위에서 제시한 방법론은 사례지역인 강남역일대가 아닌 다른 지역에서의 적용도 가능성을 보여준다. 또한, 기존 노선의 영향권과 실제 통행분포를 비교하는 방법을 제시하여, 노선의 서비스 제공률을 평가하였다.

본 연구의 한계점은 대중교통을 이용하지 않고 귀가한 모든 통행자를 심야버스를 이용할 가능성이 있다고 본 점이다. 야외활동이 많아짐에 따라 반드시 귀가통행이 일어나지 않는 경우도 존재하지만, 이를 드문 경우로 보고 포함하지 않았다. 또한 이용객이 심야시간대에 1km를 걸어갈 의향이 있다고 본 점에서 접근성보다 심야시간대의 안전을 더 중요시 여기는 이용객의 선호는 반영되지 않았다. 마지막으로, 지하철보다 버스가 접근성이 높은 점을 감안하지 않고 지하철역으로 공간적 범위를 설정한 것을 한계점으로 볼 수 있다.

본 연구에서는 노선의 일치도와 중심지역으로 선정된 강남역에서만 통행량을 가지고, 노선의 서비스 제공률을 계산하였다. 그렇기 때문에 강남역에서 제공되는 서비스만을 평가해서 얻은 결론은 서울시에서 현재 운행되고 있는 심야버스노선 전구간의 서비스 제공률을 대표할 수는 없다.

따라서 향후 서울시 전체를 이와 같은 방법으로 공간적 단위를 나누고, 현재 운행 중인 전체 심야버스 구간의 주요 정류소간 기종점을 파악하여 심야버스 이용수요를 계산하여 서비스율을 평가할 수 있으며, 나아가 노선 선정 시 필요한 다양한 요소와 각 행정동이 갖는 지역적 특성을 고려할 수 있다. 이러한 종합적인 분석을 통해 기존 심야시간 통화량을 기준으로 선정한 심야버스 노선과 스마트카드 자료에서 추출한 실제 심야통행량을 비교하여 이용수요에 따른 심야버스 서비스 제공률을 살펴볼 수 있을 것이다. 또한 재차인원 및 통행발생량 등으로 서비스수준을 고려하여 버스 내 혼잡도나 배차간격 등 앞으로 다양한 방향으로의 연구가 가능할 것이다.

REFERENCES

- Ahn H. J. and Lee Y. I.(2007), "A Study on Construction and Utilization of Bus/Subway Station Based OD and Network Using Smart Card Data in Seoul," *Transportation Technology and Policy*, vol. 4, no. 4, pp.31-59.
- Ahn K. J.(2015), "Seoul City's Operation and Usage Status and Its Improvement Plan," *The Seoul Institute Policy Report*, vol. 186, pp.1-19.
- Bae G. M.(2003), "A Study on the Feeder Transit Route Choice Technique," *Journal of Korean Navigation and Port Research*, vol. 27, no. 4, pp.479-484.
- Chang K. U., Kim H. B., Park H. S. and Park S. B.(2011), "A Study of Origin and Destination Decision for a Direct Bus Line in a City with Transit Mobility and Potential Demand," *Journal of the Korean Society of Civil Engineers*, vol. 31, no. 4D, pp.547-553.
- Chen C., Zhang D., Zhou ZH., Li N., Atmaca T. and Li S.(2013), "B-Planner : Night Bus Route Planning using Large-scale Taxi GPS Traces," *IEEE International Conference on Pervasive Computing and Communications*.
- Demissie MG., Phithakkitnukoon S., Sukhvibul T., Antunes F., Gomes R. and Bento C.(2016), "Inferring Passenger Travel Demand to Improve Urban Mobility in Developing Countries Using Cell Phone Data : A Case Study of Senegal," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, pp.2466-2478.
- Ha S. J. and Cheon S. H.(2011), "A study on Bus Route Appropriateness According to Bus Route

- Change,” *Journal of Korean Society of Transportation*, vol. 64, pp.680-685.
- Kim N. J.(2012), “Estimating the Subway Station Influence Area by the Distribution of Walking Distance and the Changes of Housing Sale Prices - Focused on the Subway Stations of Jungang Line in Gury and Namyangju City,” *Journal of Korea Planning Association*, vol. 47, pp.29-38.
- Lee S. Y. and Park K. A.(2003), “Quantitative Evaluation Indicators for the City Bus Route Network,” *Journal of Korean Society of Transportation*, vol. 21, no. 4, pp.29-44.
- Ministry of Land, Infrastructure and Transport.(2013), Korea Highway Capacity Manual(KHCM), Korea, p.594.
- Munizaga MA. and Palma C.(2012), “Estimation of a disaggregate multimodal public transport Origin-Destination matrix from passive smartcard data from Santiago, Chile,” *Transportation Research Part C : Emerging Technologies*, vol. 24, pp.9-18.
- News Jelly, http://contents.newsjel.ly/issue/seoul_nbus/, 2016.4.12.
- Yonhap News, <http://media.daum.net/m/channel/view/media/20160320070109079>, 2016.4.8.
- Yoon K. M., Park H. K. and Kim Y. H.(2011), “A Study on Improvement and Operation of Midnight Extending a Late Night Gyeongin Line Express Train,” *Journal of the Korean Society for Railway*, pp.483-495.