



디지털융복합연구 제13권 제2호

ISSN : 2713-6434(Print) 2713-6442(Online)

지하철 혼잡도 개선방안에 관한 빅데이터융합 기반의 탐색적 연구

김근원, 노규성, 김동우, 이주연

To cite this article : 김근원, 노규성, 김동우, 이주연 (2015) 지하철 혼잡도 개선방안에 관한 빅데이터융합 기반의 탐색적 연구, 디지털융복합연구, 13:2, 35-42

① earticle에서 제공하는 모든 저작물의 저작권은 원저작자에게 있으며, 학술교육원은 각 저작물의 내용을 보증하거나 책임을 지지 않습니다.

② earticle에서 제공하는 콘텐츠를 무단 복제, 전송, 배포, 기타 저작권법에 위반되는 방법으로 이용할 경우, 관련 법령에 따라 민, 형사상의 책임을 질 수 있습니다.

www.earticle.net

지하철 혼잡도 개선방안에 관한 빅데이터융합 기반의 탐색적 연구

김근원*, 김동우*, 노규성*, 이주연**
선문대학교 경영학과*, 아주대학교 산업공학과**

An Exploratory Study on Improvement Method of the Subway Congestion Based Big Data Convergence

KeunWon Kim*, DongWoo Kim*, Kyoo-Sung Noh*, Joo-Yeoun Lee**

Dept. of Business Administration, Sunmoon University*,

Dept of Industrial Engineering Ajou University**

요약 빅데이터의 가치가 중요하게 인식되면서 정부를 비롯한 공공기관, 민간기업 등이 빅데이터에 관심을 가지기 시작하였다. 과거와는 다르게 다양한 데이터의 원천이 있고, 이러한 데이터의 융합을 기반으로 한 다양한 기획 및 분석기법이 등장하게 되면서, 빅데이터는 새로운 고급 정보의 창출 및 의사결정 고도화 기반으로 자리매김할 것이 확실하다. 본 연구는 다양한 대책에도 불구하고 개선되지 않는 지하철 혼잡도 문제에 대한 대안을 모색하는 것이다. 이에 본 연구는 서울시의 공공데이터를 활용하여 서울시 지하철의 혼잡도를 개선할 수 있는 방안에 대해 탐색적 접근을 시도하였다. 연구는 서울시 공공데이터 분석을 오픈 소스 R을 통해 이루어졌다. 분석 결과 혼잡도가 높은 지하철 역을 중심으로 운행하는 버스 노선을 신설하는 정책 대안을 도출하게 되었다.

주제어 : 빅데이터, 정책융합, 지하철 혼잡도, 버스 노선

Abstract As the value of Bigdata has been recognized importantly, public agencies including the government, private sector, etc. began to have an interest in Big Data. As there are sources of various data, and a variety of planning and analysis methods based on these sources has emerged, It is true that Bigdata will become a tool for creation of the new high quality information and decision making based on new insights. The purpose of this study is to find an alternative to the subway congestion problem that is not improved even though the various measures. In this study, we tried to explore approaches for ways to improve the congestion of the Seoul Subway using Seoul Metropolitan public data. Lastly, this study derived a policy alternative to establish new bus route that runs around the metro station that have a high level of congestion.

Key Words : Bigdata, Policy Convergence, Survey Congestion, Bus Route

Received 5 December 2014, Revised 10 January 2015

Accepted 20 February 2015

Corresponding Author: Kyoo-Sung Noh (SunMoon University)

Email: ksnoh@sunmoon.ac.kr

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

최근 빅데이터가 최대 화두 중 하나로 떠오르고 있다. 비단 국내 뿐만 아니라 전세계적으로 빅데이터 열풍이 불고 있다[9]. 2014년 브라질 월드컵에서 독일이 우승을 한 이면에는 빅데이터 분석을 통한 게임 전략이 있었다는 것은 널리 알려진 사실이다. 또한 2013년 서울시의 심야버스 노선의 최적화 방안도 서울시의 교통데이터와 KT의 통신데이터 등 빅데이터를 분석했기에 가능한 일이다[7,8,11,12].

현재 빅데이터는 다양한 산업/공공분야에서 적용을 시도하고 있으나, 아직까지 성공적인 사례라고 할 만한 분야는 교통 등 몇몇 분야를 제외하고는 많지 않은 실정이다. 교통부문의 빅데이터는 상당부분 정형데이터로 한국도로공사의 OASIS, 도로교통공단의 TAAS 등에서 수집되고 있다. 빅데이터가 등장하기 이전에는 정형데이터만을 가지고 분석을 수행하였으나, 최근에는 정형데이터와 비정형 데이터를 같이 분석하고 있다. 그러나 정형데이터에 비해 비정형데이터는 분석을 하기가 까다로운 편이어서 아직 활용도가 높지 않은 편이다.

교통부문에서 빅데이터를 활용한 사례로는 서울시 심야버스 노선 수립과 전국 교통혼잡 지도 등이 있다. 서울시는 심야버스 노선을 신설하면서 버스가 다니지 않는 새벽시간대인 00시~05시까지의 KT의 통신데이터 등을 활용하여 유동 인구의 밀집도를 분석하고 요일별/노선별 유동인구를 파악을 통해 일정한 패턴을 찾아내고 이를 바탕으로 심야버스 노선 최적화 방안을 도출하게 되었다. 또한 전국 교통혼잡 지도는 내비게이션을 사용하는 차량의 이동 궤적 데이터를 활용하여 도로/교차로/행정구역별로 혼잡 강도를 GIS(지리정보시스템)를 기반으로 하는 표출시스템을 구현하였다[2].

일반적으로 대중교통하면 버스를 생각할 수 있지만, 서울과 같은 대도시의 경우는 지하철도 중요한 대중교통 수단이다. 그러나 출퇴근 시간대에 일부지역 역사의 경우 혼잡도가 다른 역사에 비해 매우 높아, 불편한 단계를 넘어 결국에는 승객의 안전까지 우려되고 있다[4].

최근 광역버스 입석운행 금지제도가 시행되기 시작하

면서 버스보다는 지하철을 이용하는 승객들이 점차 증가하고 있으며, 이로 인해 지하철의 혼잡도가 더욱 증가하게 되었다. 한편 서울시 등에서는 지하철 9호선 운영방안 개선 등을 통해 혼잡도를 개선하고자 노력했지만 오히려 새로운 지하철 혼잡지역을 생성하는데 그쳤다는 평가가 많다[10].

이에 본 연구는 이러한 서울시 지하철 혼잡도를 개선할 수 있는 대안을 탐색적으로 모색하고자 한다. 이를 위해 먼저 서울시 지하철 역별 승·하차 인원과 혼잡도를 분석하고 이를 토대로 혼잡도가 높은 지역을 중심으로 혼잡도 개선방안을 제안하고자 한다. 이는 서울시 공공데이터의 융합을 기반으로 빅데이터의 분석도구인 R을 활용하여 서울시 지하철 혼잡도와 다른 요인들 간의 비계층적 군집분석과 연관분석을 통해 이루어지며, 그와 관련된 시각화 기법도 사용될 것이다.

2. 이론적 배경과 선행 연구

빅데이터가 급부상하면서 관련 연구도 급증할 것으로 보인다. 현재는 정치, 산업, 교육, 보안 분야 등의 다양한 응용 연구와 빅데이터 기획, 빅데이터 성공요인 등에 관한 연구가 이루어지고 있다[6,7,8,9,12,15]. 그런 가운데 데이터 분석을 통한 교통 관련 연구는 이미 많이 진행이 되고 있었으며, 최근에는 빅데이터를 이용한 교통정책 개발에 관한 연구들이 이루어지고 있다는 실정이다.

이석주 등(2014)은 교통 부문의 빅데이터 현황을 파악하고, 교통부문 정책 수립시 어떻게 빅데이터를 새로운 의사결정에 활용할 수 있는지 제시하기 위해 교통부문에서 빅데이터로 간주될 수 있는 교통카드 및 내비게이션 데이터를 분석하였으며, 실제 교통부문에서 활용을 할 수 있는 방안을 제시하고, 이에 따른 법적, 제도적 개선방안을 제시하였다[13].

한편 혼잡도에 대한 국내외 연구는 승객의 승하차시간과 정차시간과의 상관관계를 분석하는 실증적인 연구들이 주를 이루고 있다. 이두용 등(2008)은 지하철 혼잡도 분석을 위해 기존의 다중열차운용 시뮬레이터를 프로그램 구조 및 인프라, 열차제원, 시각표, 운행 루트 등을 수정하여 지하철 혼잡도 분석을 위해 적합화하고 이를 시뮬레이션을 하였다. 그리고 이를 통해 역사 혼잡도를

1) 전철이 보다 바른 용어이나, 지하철도 같은 의미로 쓰이고 있어 여기에서는 보다 보편적으로 쓰이는 지하철이란 용어를 사용한다.

개선하기 위한 방안으로 환승통로/계단/에스컬레이터 등 특정지점에 혼잡이 집중되는 것을 최소화하기 위해 RFID를 이용하여 승객을 분산시키는 방안을 제시하였다[1]. 그리고 Takagi et al.(2006)은 환승역에 도착하는 열차 시간이 늦어질 때 앞 열차를 언제 출발시킬 것인지에 대한 연구를 수행하였다[14].

이상의 이전 연구들은 지하철 혼잡도에 관한 해결방안으로 주로 자체적으로 증편을 계획한다거나 지하철 역사내 승객 이동에 대한 분석 등이 주를 이루고 있으며, 이직 빅데이터를 활용한 예를 없다. 이에 본 연구에서는 빅데이터를 활용하여 혼잡도 개선에 관한 새로운 접근 방법을 탐색적으로 제시하고자 한다.

3. 연구 방법 및 절차

본 연구는 도출된 문제 현상을 정의한 다음, 이를 해결할 대안 모색을 위해 필요한 정보, 데이터 및 활용 분석 기법 등을 기획하는 과정을 거쳤다. 그런 다음, 오픈소스 분석 도구인 R을 활용하여 분석 및 시각화를 시도하고 이를 해석하는 절차로 진행하였다[3].

3.1 문제 정의

지하철을 이용하는 승객은 일반적으로 역사내에서는 최단거리로 이동하려는 욕구를 가지고 있다. 따라서 승객들은 이미 역사내 이동경로의 최단거리를 인지하고 있으며, 하차 후 신속하게 역을 빠져나가기 위해 계단, 환승통로, 출구를 이용하고자 한다. 이러한 이유 때문에 이들 구역에서 혼잡도가 가중되고 있는 상황이다.

특히 출·퇴근 시간대에는 지하철 혼잡도가 증대되면서 시민들의 불편이 증가하고 있는 상황이다. 따라서 본 연구에서의 문제는 ‘혼잡도를 증가시키는 원인이 무엇이며, 이를 해결할 방안은 무엇인가’로 정의하고자 한다.

3.2 문제에 필요한 필요정보

광원 수도권에서 승객이 가장 많은 지하철 노선은 1~4호선으로 파악되고 있다. 따라서 분석 대상은 1~4호선이지만, 그 중 어느 노선 또는 어떠한 시간대 가장 많은 혼잡도가 이뤄지는지 알아볼 필요가 있다. 각 호선별 혼잡도와 혼잡 시간대 등의 정보를 파악해 버블차트형태로

나타내보고 혼잡도가 높은 이유를 찾아내는 것이 정책대안 마련에 필요할 것이다.

3.3 정보도출을 위한 필요 데이터

문제 파악에 필요한 정보를 도출하기 위해서는 관련 데이터가 필요하다. 앞에서 정의한 바와같이 각 노선별 혼잡도와 혼잡 시간대 및 혼잡 원인 파악을 위해서는 ① 지하철 1~4호선의 승·하차인원과 혼잡도가 1차적인 분석 대상 데이터일 것이다. 또한 지하철 혼잡도에 간접적으로 영향을 미치는 ② 서울시 각 구별 건축허가 및 사업체 수, 주간 및 ③ 야간 상주인구 관련 데이터 분석도 혼잡도 원인 파악에 필요할 것이다.

3.4 정보도출을 위한 분석 기법

먼저 1~4호선의 혼잡 정도를 파악하기 위해 게이지 함수와 파이차트를 이용한 시각화와 승·하차 인원과 혼잡도가 어떤 집단으로 묶이는 지를 보기 위해 군집분석을 실시하였다. 이는 시각화의 버블차트에 의해 보완적으로 분석되었다. 아울러 지하철 역 근처가 상업지구이고 유동인구가 많은 곳이 혼잡도가 높을 것이라는 가설을 검증하기 위해 연관분석을 실시하였다. 연관분석은 둘 이상의 변수간의 연관성을 파악하는데 도움을 주기 때문이다.

분석은 먼저 데이터를 수집한 다음, 수집한 데이터를 전처리 과정을 통해 분석이 가능하도록 정리하고 각 데이터에서 분석에 필요한 데이터만을 모아 하나의 새로운 분석 대상 데이터를 만든다. 그리고 마지막으로 해당 분석 도구 R²⁾을 통해 원하는 정보를 산출한다.

4. 데이터 수집 및 분석

4.1 데이터 수집 및 전처리

분석에 필요한 데이터는 서울시의 열린 데이터 광장(<http://data.seoul.go.kr>)과 서울시 통계 사이트(<http://stat.seoul.go.kr>)에서 구할 수 있다.

2) 통계프로그램인 R은 오픈소스 기반의 무료소프트웨어로서 다양한 형태의 빅데이터의 처리와 통계 및 시각화 구현이 가능하기 때문에 본 연구에서는 R을 사용하여 데이터 분석을 실시했다.

본 연구에서는 서울시의 열린 데이터광장에서 지하철 1~4호선의 승하차 인원 및 혼잡도 데이터를 구할 수 있었다. 그리고 서울시 통계사이트에서 구별 건축허가 및 사업체 수, 주간 및 장주(야간) 인구 등 통계 자료를 수집할 수 있었다.

수집된 데이터는 전처리함으로써 분석이 가능한 형태로 전환하였고, 분석에 사용된 데이터는 CSV로 변환하여 분석에 사용하였다[5]. 데이터는 분석 상황에 맞게 전처리되었다. 시간대별 혼잡도를 파악하는 데 있어서는 임의의 구간을 설정하여 평균 값을 내어 사용하였다. 연관 분석에서는 분석할 데이터의 값이 숫자가 아닌 문자열로 되어있어야 했기 때문에 각 항목별로 임의의 구간을 설정하여 문자열로 전처리를 하였다.

<Table 1>과 <Table 2>는 전처리 과정을 거친 분석 대상 데이터의 예시이다.

<Table 1> Congestion Data/per Route · Section · Time

13년_역간별 혼잡도 [화요일] - Microsoft Excel

구간별 시간대별 혼잡도(평일) - 상선(내선)												
역명	방향	5시대	6시대	7시대	8시대	9시대	10시대	11시대	12시대	13시대	14시대	15시대
서울역	상선	10	20	43	72	54	35	42	38	40	34	3
신길	상선	10	20	40	66	55	39	47	43	45	41	4
종각	상선	10	17	29	50	46	35	44	42	45	42	4
종로3가	상선	10	17	31	57	50	42	47	49	51	51	5
종로5가	상선	10	16	26	45	42	37	43	45	48	49	5
동대문	상선	11	16	24	43	38	36	41	44	47	49	5
신설촌	상선	10	15	22	39	33	37	39	41	43	46	4
신설동	상선	10	13	18	31	36	38	39	41	39	43	4
계기동	상선	10	13	17	29	30	31	33	34	32	37	4
강남(강남)	상선	9	11	14	24	24	25	26	27	29	31	3
신길	내선	22	21	50	81	44	43	34	37	39	38	4
종로3가	내선	18	15	19	32	33	36	38	37	38	39	4

<Table 2> Get on and off People Data/per Station

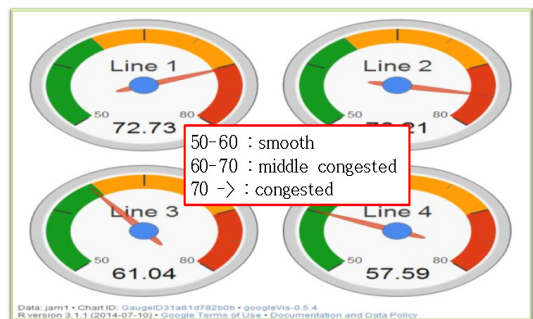
역명	방향	00-01	01-02	02-03	03-04
서울역	상선	185	0	0	0
신길	상선	14	0	0	0
종각	상선	67	0	0	0
종로3가	상선	26	0	0	0
종로5가	상선	35	0	0	0
동대문	상선	106	0	0	0
신설촌	상선	36	0	0	0
신설동	상선	3	0	0	0
계기동	상선	111	0	0	0
강남(강남)	상선	34	0	0	0
신길	내선	0	0	0	0
종로3가	내선	0	0	0	0

4.2 데이터 분석

먼저 지하철 데이터를 R프로그램의 게이지(gauge)함수를 사용하여 시각화를 하였다. 지하철 1~4호선의 혼잡도의 평균을 내고 임의의 범위를 설정하여 시각화를 시도하였다. 시각화하여 분석한 결과, [Fig 1]에서 볼 수

있듯이, 1, 2호선의 혼잡도가 3, 4호선에 비해 높다는 것을 볼 수 있다.

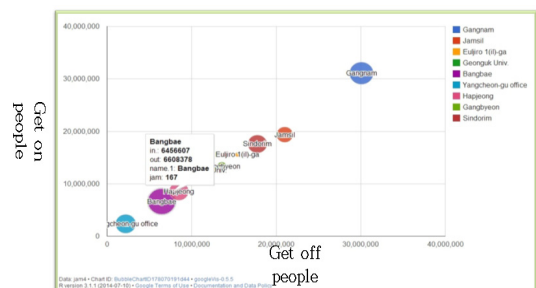
보다 자세히 확인하기 위하여 출·퇴근시간별 각 운행 노선별 혼잡도를 파이차트를 통해 시각화를 해 보았다. 시각화하여 분석한 결과, 출근시간에는 지하철 1호선 하행, 2호선 내·외선의 혼잡도가 높다는 것을 알 수 있고, 퇴근시간에는 지하철 1호선 상행, 2호선 내·외선의 혼잡도가 높다는 것을 알 수 있다. 이상의 분석을 근거로 본 연구에서는 주요 분석 노선을 지하철 2호선으로 정하고 이의 혼잡도를 완화하는 방안을 찾고자 하였다.



[Fig. 1] Congestion Chart of 1~4 Line Using Gauge Function

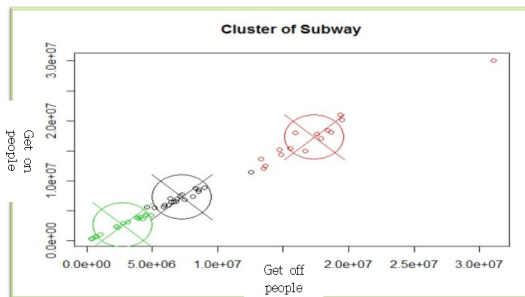
4.2.1 2호선 역 별 승·하차 인원과 혼잡도 관계

혼잡도가 높아지는 원인을 찾기 위해 승·하차 인원과 혼잡도를 버블차트를 통해 시각화하여 분석을 했다. 시각화하여 분석한 결과 강남역의 경우 승·하차 인원이 많고 혼잡도도 높다고 나왔지만, 방배역이나 합정역과 같이 승·하차 인원이 많지 않음에도 혼잡도가 높게 나타나는 것을 확인할 수 있다. 즉, 승·하차 인원과 혼잡도의 관계가 높지 않은 것을 확인할 수 있다.



[Fig. 2] Relation of Get on and off People and Congestion Using Buble Chart

또한 승·하차 인원과 혼잡도의 관계가 어떤지를 확인하기 위해 군집분석을 추가적으로 실시했다. 군집의 개수를 3개로 정하고 비계층적 군집분석을 실시한 결과를 시각화하였다. 그 결과, 3개의 군집 모두 중심으로 많이 모여 있는 것을 볼 수 있다. [Fig 3]에서 볼 수 있듯이, 비계층적 군집분석 결과, 1번 군집의 경우 승·하차 인원이 가장 많고 혼잡도가 높은 군집으로 군집화 되었다. 3번 군집의 경우 승·하차 인원은 1번 군집과 많은 차이가 있었으나 혼잡도에 있어서는 큰 차이가 없는 것을 알 수 있다. 결국 본 연구의 가정과는 다르게 승·하차 인원이 많을수록 혼잡도가 높은 것이 아니라는 결과를 확인할 수 있었다. 즉, 승·하차 인원과 혼잡도의 관계가 높지 않은 것으로 분석되었다.



[Fig. 3] Clustering Chart of Get on and off People and Congestion

4.2.2 2호선 역 별 연관 분석 관계

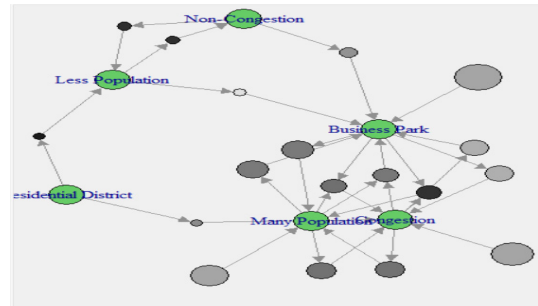
군집분석 등의 결과에 의해 승·하차 인원과 혼잡도의 관계가 높지 않은 것을 확인하였기 때문에, 혼잡도에 영향을 주는 다른 원인을 파악하고자 연관분석을 실시하였다.

연관분석을 위한 데이터는 ① 상업지구 및 주거지구 (구별 건축허가 및 사업체 수 비교), ② 혼잡 및 비혼잡 (지하철 혼잡도 비교), ③ 유동인구 비교(주간 및 상주(야간) 인구 비교) 등이다.

이상의 데이터 중 분석을 위해 각 자치구별 주간 및 야간(상주)인구 및 주택 및 상업지구 현황 통계 데이터를 토대로 전처리를 실시하였다. 주간 및 야간(상주)인구 통계 자료를 토대로 임의의 기준치를 정하여 유동인구 많은 곳과 적은 곳으로 나누었다. 주택 및 상업지구 현황 통계 자료를 토대로 임의의 기준치를 정하여 주택지구와

상업지구로 나누었고, 혼잡도 또한 임의의 기준치를 정하여 혼잡과 비혼잡으로 나누어 데이터 전처리를 했다.

이상의 데이터 전처리를 시행한 다음, 자치구 형태 및 유동인구가 혼잡도에 영향을 주는지 파악하기 위해 연관분석을 실시한 결과 상업지구이고, 유동인구가 많은 곳이 혼잡하다는 데이터간의 연관성을 확인할 수 있었다.



[Fig. 4] Relation Chart of Area · Population and Congestion

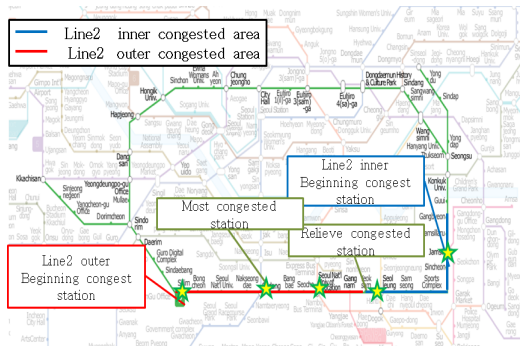
5. 지하철 혼잡도 개선방안

이상과 같이 서울시 공공 데이터의 융합적 분석을 통해 지하철 2호선의 출·퇴근 시간대 혼잡도가 가장 높다는 것을 확인할 수 있었다. 특히 강남구에 소재된 지하철 역들의 혼잡도가 다른 구에 비해 높다는 것을 알 수가 있었다. 연관분석에서는 상업지구이고 유동인구가 많은 곳이 혼잡하다는 연관성을 확인하였다. 강남구는 기업체가 많은 상업지구이며, 유동인구가 많은 곳이기 때문에 다른 자치구에 비해 혼잡도가 높게 나타났다.

선행 연구 등을 종합해 볼 때 그간 지하철 혼잡도 완화를 위해 운행대수의 증편, 신설 노선의 검토, 출근시간대의 조정 등 다양한 대안이 시도 내지 논의된 바 있다. 그러나 일부 대안은 광역수도권 차원에서 해결할 수 없는 대안이며 일부 대안은 오히려 혼잡도를 증가시켰다는 비판의 대상이 되기도 하였다.

이에 본 연구에서는 혼잡도를 완화하기 위해 혼잡도가 높은 몇몇의 역에만 정차하는 급행 버스를 신설하는 방안을 제시한다. 급행버스 운행에서 혼잡도가 높은 역이 아니라 높아지는 역부터 운행하는 방안을 제시하는 이유는 혼잡도는 갑자기 높아지는 것이 아니라, 앞선 역

에서부터 점차적으로 증가하는 것이기 때문에 높은 역이 아닌 높아지는 역부터 운행하는 급행버스를 신설하는 방안을 제시한다. 급행버스를 신설하여 운행함으로써, 강남권 역들의 혼잡도를 완화시키고자 한다. 출·퇴근 시간 도로 혼잡에 영향을 받지 않기 위해 버스전용차로로 운행하여 출·퇴근 시간에도 효과적인 운행이 가능하도록 하는 것이 좋은 방안이다.



[Fig. 5] High Congestion Station and New Bus Routes

우리가 제시하는 급행버스에 대해 ‘혼잡도 개선을 위해 지하철 운행 횟수를 증가시키면 되지 않을까?’라는 의문점이 들게 된다. 그러나 운행 횟수만을 증가시킨다고 해서 혼잡도가 완화되는 것은 아니다. 서울시에서 2010년 운행 횟수 증가로 ‘신도림-강남’ 혼잡도 10% 완화시키려고 하였다. 그러나 많은 증차는 오히려 혼잡도를 증가시켰다[10]. 따라서 본 연구에서는 신림역과 잠실역 사이의 급행버스 신설 방안을 제시하고자 한 것이다.

본 연구에서는 지하철의 혼잡도를 완화시키고자 ‘신림-사당-교대-선릉-잠실’ 5곳의 역에만 정차하는 급행버스 신설하는 것을 제안하고자 한다. 각 역별 운행거리 약 4~5km, 운행시간 약 15~20분정도 소요가 된다. 운행시간은 혼잡도가 높은 시간을 중심으로 하여 운행시간 설정(출근시간 : 6시30분 ~ 10시, 퇴근시간 : 17시 30분 ~ 20시)하였다. 출·퇴근 시간에 급행버스를 신설함으로써 지하철 혼잡도를 완화시킬 것으로 기대한다.

또한, 보다 효과적인 결과를 얻기 위해 급행버스와 일반시내버스를 연계하도록 시스템을 구축하여, 급행버스가 정차하지 않는 곳에도 불편함 없이 버스를 이용하여 시민들이 출·퇴근 할 수 있도록 한다면 혼잡도 완화에 더

욱 효과적일 것이라고 기대한다.

6. 결론

본 연구에서는 서울시의 공공데이터 분석을 통해 서울시 지하철의 혼잡도를 개선할 수 있는 방안을 제안하고자 하였다. 연구 결과, 시사점은 다음과 같다. 첫째, 서울시의 공공데이터를 기반으로 혼잡도를 개선할 수 있는 방안을 기획하고 R로 분석을 하였다는 점이다. 둘째, 지하철의 혼잡도를 개선하는 방안으로 급행버스 노선을 제안하였다. 기존의 선행연구들은 대부분 지하철의 혼잡도 개선을 위한 방안으로 증차, 게이트 수의 증가 등이 대부분이었다. 그러나 본 연구에서는 혼잡도가 높은 지하철역을 운행하는 급행버스 노선을 제시하였다는 점은 기존에 수많은 대안에서 없었으며, 현재의 교통상황을 감안할 때 나름대로 추진가능한 대안으로 신선하다 할 수 있다.

그러나 본 연구에서 제시한 정책 대안이 실현되는 데에는 여러 가지 제약이 따를 것이다. 우선 분석 대상이 단순히 지하철 노선과 관련되는 변수 몇 가지에 한정되었다는 점이다. 버스 노선의 신설은 도로사정 등 다양한 다른 변수들이 복합적으로 고려되어야 하는 복잡한 정책 사안이다. 이에 따라 본 연구는 공공데이터 외의 다양한 데이터를 분석하지 못했다는 한계를 가지고 있다. 향후 연구에서는 보다 폭넓은 빅데이터 기획을 기반으로 다양한 데이터를 확보하고 그에 걸맞는 분석 기법에 의해 분석 및 대안 마련이 이루어져야 할 것으로 보인다.

REFERENCES

- [1] Doo-Yong Lee, Zhong-Shi Li, Dong-Hee Kim, Soon-Heum Hong and Chang-Cho Lee, A Study on the Analysis of Subway Crowdedness using Simulator, Journal of Korea Safety Management & Science, 10(4), pp. 267-273, 2008.
- [2] IISoo Yoon, Yul Han, Bigdata and Transportation, Transportation Technology and Policy, 11(1), pp. 64-67, 2014.

- [3] In-Kyoo Park, Clustering Algorithm for Data Mining using Posterior Probability-based Information Entropy, *Journal of Digital Convergence*, 12(12). pp. 293-301, 2014.
- [4] Jin-Hyo Kim, Suk-Yun Han, Woo-Ding Lee, Review of the Congestion Status of the Subway Station, *Korean Society for Railway*, Proceedings of 2008 Spring Conference, pp. 1810-1816, 2008.
- [5] Jun-Seok Lee, A Study on the Data Mining Preprocessing Tool For Efficient Database Marketing 12(11), pp.257-264, 2014.
- [6] Ki-hyoung Han, Hyung-Jong Jeong, Doog-Sik Lee, Myung-hui Chae, Cheol-hee Yoon, Kyoo-sung Noh, A Study on implementation model for security log analysis system using Big Data platform, *Journal of Digital Convergence* 12(8), 2014.
- [7] Kyoo-Sung Noh, A Study on Utilization Strategy of Big Data for Local Administration by Analyzing Cases, *Journal of Digital Convergence* 12(1) pp. 89-97, 2014.
- [8] Kyoo-Sung Noh, Smart Learning Strategies utilizing Convergence of e-Learning and Bigdata, *Journal of Digital Convergence*, 13(1), pp. 487-493, 2015.
- [9] Kyoo-Sung Noh, Sanghwi Park, An Exploratory Study on Application Plan of Big Data to Manufacturing Execution System, *Journal of Digital Convergence* 12(1), pp. 305-311, 2014.
- [10] Press Report, Pilot Program for Relieving Congestion Was Effective, *Seoul Metro*, 2011.
- [11] Seong-Hwan Ju, Kyoo-Sung Noh, A Study on Policy for Data Convergence infrastructure of e-Learning Industry, *Journal of Digital Convergence*, 13(1), pp. 77-83, 2015.
- [12] Su Hyeon Namn, Kyoo-Sung Noh, A Study on the Effective Approaches to Big Data Planning, *Journal of Digital Convergence*, 13(1), pp. 227-235, 2015.
- [13] Suk-Joo Lee, Ji-Yoon Yeon, Seung-Hun Chun, Big Data for Transportation Policies and Their Applications, *Korea Transport Institute Research Series* 192, pp.50-57, 2013.
- [14] Takagi R., Goodman C. J. and Roberts C, "Optimization of train departure times at an interchange considering passenger flows", *Proc. IMechE Vol. 220 Part F. J. of Rail and Rapid Transit* 2006, 2006.
- [15] Young-ki Jung, Myung-gun Suk, Chang-Jae Kim, A study on the success factors of Big Data through an analysis of introduction effect of Big Data, *Journal of Digital Convergence*, 12(11), pp.241-248, 2014.

김 근 원(Kim, Keun Won)



- 2015년 2월 : 선문대학교 경영학 (경영학사)
- 관심분야 : 빅데이터, ERP
- E-Mail : dark-kgw@hanmail.net

김 동 우(Kim, Dong Woo)



- 2015년 2월 : 선문대학교 경영학부 (경영학사)
- 관심분야 : 빅데이터, ERP
- E-Mail : dongwoo1108@naver.com

노 규 성(Noh, Kyoo Sung)



- 1984년 2월 : 한국외대 경영학과(경영학사)
- 1995년 8월 : 한국외대 대학원 경영정보학과(경영정보학 박사)
- 2003년 ~ 2010년 : 中國 延邊科學技術大學 兼職教授
- 1997년 ~ 현재 : 선문대학교 경영학과 비즈니스데이터과학 전공 교수
- 2004년 ~ 현재 : 한국디지털정책학회 회장
- 2012년 ~ 현재 : 스마트융합학술전국연합 의장
- 관심분야 : 디지털정책&스마트융합, 디지털경제민주화, 창의기반 경영혁신, 빅데이터
- E-Mail : ksnoh@sunmoon.ac.kr

이 주 연(Lee, Joo Yeoun)



- 1993년 2월 : 아주대학교 대학원 경영정보전공(경영학석사)
- 2004년 2월 : 인하대학교 대학원 경영학과(경영학박사)
- 2007년 ~ 2011년 : 한국산업정보학회 회장(Chairman)
- 2005년 ~ 2011년 : SK C&C 전략마케팅본부장(상무)
- 2011년 ~ 2014년 : 포스코ICT 그린사업부문장(전무)
- 2014년 ~ 현재 : 아주대학교 공과대학 산업공학과 교수
- 2015년 ~ 현재 : 산업통상부 산업융합촉진 옴부즈만
- 관심분야 : Business Intelligence, Smart Convergence (Smart Grid, Factory, Fin Tech, Cognitive)
- E-Mail : jooyeoun325@ajou.ac.kr