

교통카드데이터를 활용한 대중교통 수단선택 행태분석: 버스의 도시철도에 대한 대체 및 경쟁을 중심으로

Analysis of Mode Choice Behavior on Mass Transit Using Smart Card Data: Focusing on Replacement and Competition of Bus for Metro

김경태 · 이인묵*

Kyoungtae Kim · Inmook Lee

Abstract The problem of metro crowding due to insufficient capacity of Seoul Metro Line 9 has been raised as a social problem. To reduce metro congestion, it is fundamentally necessary to increase the supply capacity of railways; however, in the short term, operating buses as a replacement mode can be an alternative. To apply this alternative, it is necessary to study the behavior of passengers choosing metro and bus. In this study, conditions under which metro had competitiveness compared to bus were verified by classifying metro station OD by distance. The TRIPS (Travel Record based Integrated Public transport operation System) program was used to extract data on sections where competition between metro and bus occurred. Using the extracted data, the travel behavior between the two means was analyzed. The conclusion of this study is that, to alleviate crowding and increase demand for metro, it is necessary to consider mode choice pattern (behavior) of mass transit passengers by travel distance and by distance between metro stations and bus stops.

Keywords : Smart card data, Mode choice model, Bus, Metro, Mass transit

초 록 9호선 연장운행, 김포 골드라인 등과 관련하여 차량 부족으로 인한 도시철도의 혼잡 문제가 사회적인 쟁점이 된 바 있다. 도시철도 혼잡을 줄이기 위해서는 근본적으로 철도의 공급 용량을 늘려야 하겠지만, 단기적으로는 대체 수단으로서 버스를 투입하는 것도 하나의 대안이 될 것이다. 이를 위해서는 이용자가 철도와 버스를 선택하는 행태에 대한 기초연구가 필요하다. 본 연구에서는 철도역 OD를 거리대별로 구분하여 도시철도가 버스 대비 경쟁 우위를 가지는 조건을 확인하였다. 이를 위해서 수도권 도시철도와 버스의 경쟁이 성립될 수 있는 구간을 TRIPS(통행실적 기반 복합대중교통 운영계획수립 지원시스템) 프로그램을 이용하여 기초자료를 추출하였으며, 데이터 정제 후 두 수단 간 이용 행태를 분석하는 과정을 거쳤다. 본 연구는 도시철도 혼잡 완화 및 수요 증대를 위해서는 철도역과 버스 정류장 사이의 거리 및 통행거리에 따른 대중교통 이용자의 수단선택 패턴(행태)을 고려하여야 한다는 결론을 도출하였다.

주요어 : 교통카드데이터, 수단분담모형, 버스, 도시철도, 대중교통

1. 서 론

서울 도시철도 9호선 개통 초기 차량 용량 부족(4량 1편성)으로 인한 급행 차량의 혼잡, 9호선에 연계되는 노선인 김포 골드라인 개통에 따른 출퇴근 시의 혼잡 문제가 사회적인 쟁점이 된 바 있다. 이렇듯 용량(운행 횟수 또는 공급 좌석수) 부족 또는 수요 과다로 인한 도시철도 혼잡구간의 발생은 9호선, 김포 골드라인 등의 신설 노선뿐만 아니라 기존의 2호선 강남구간도 문제시된 바 있다. 도시철도 혼잡을 줄이기 위해서는 근본적으로 철도의 공급 용량을 늘려야 하겠지만, 단기적으로는 통행 수요가 큰 구간에 신규 버스노선 신설 또는 대체 버스를 투입하는 대안을 제시할 수 있다.

본 연구는 도시철도 혼잡 문제의 해결 대안인 대체 버스 노선을 투입하는 데 필요한 수요전환 효과를 분석할 수 있도록 철도와 버스의 경쟁관계 분석 및 수단선택 모형을 구축하는 것을 목적으로 한다. 이를 위하여 교통카드데이터를 이용하여 수도권 도시철도 OD를 통행거리에 따라 세분한 후에 이와 대체 관계에 있는 버스 OD를 매칭하였다. 특히 수도권 통행의 80% 이상이

*Corresponding author. E-mail: mook79@krii.re.kr.

© 2022 The Korean Society for Railway. All rights reserved.

<https://doi.org/10.7782/JKSR.2022.25.11.770>

수단간 환승이 발생하지 않는 통행인 것을 고려하여 무환승 통행을 대상으로 철도와 버스의 수단선택 행태를 분석하였고, 이로부터 도시 대중교통 수단선택에 관한 시사점을 도출하였다. 실제로 경쟁노선 설정을 위해서는 그 효과를 분석할 수 있어야 하므로, 두 수단 간의 수단분담 모형을 제시하여 이를 활용할 수 있도록 하였다.

2. 본 론

2.1 선행연구 검토

교통카드데이터의 활용에 관해서는 서울시가 통합거리비례 요금제를 시행하면서 교통카드 시스템을 전면 개선한 2004년 이후부터 본격적으로 연구되고 있다. 초기 연구로서 The Korea Transport Institute [1]은 대중교통정책 수립의 참조 자료로서 교통카드데이터의 활용 가능성을 검증하였으며, 교통카드데이터의 기초 분석을 통한 대중교통 통행 현황 분석 결과를 제시하였다. The Seoul Institute [2]은 교통카드데이터로부터 정류장 단위 OD 정보를 추정하였으며, 이 과정에서 데이터의 오류 유형을 정리하고 통행이력 정보를 활용한 하차 정보 등의 보정 방안을 제시하였다. The Seoul Institute [3]은 교통카드데이터를 활용한 기초적인 대중교통 서비스 평가 모형을 제안하였으며 향후 지리정보와의 연계를 통한 분석의 확장 가능성을 언급하였다. 또한 교통카드데이터의 한계 중 하나인 도시철도 경로 정보의 추정을 위한 도시철도 통행행태 모형도 제안하였다. The Seoul Institute [4]은 교통카드데이터를 이용하여 운행횟수, 재차인원, 혼잡률 등의 서비스지표에 대한 값을 산출한 바 있다.

2010년까지는 교통카드데이터를 실적자료로서 정책수립을 보조할 수 있는 현황 분석, 통계량 산출 등 기초적 활용방안에 관한 연구가 중심이 되었다면 그 이후로는 교통카드데이터를 대중교통 통행행태 또는 패턴 분석에 활용하는 방안 또는 그 결과 중심의 연구가 수행되었다. Gyeonggi Research Institute [5]은 단순 통행(trip) 단위가 아닌 통행사슬(trip chain) 개념을 적용한 최초 출발지에서 최종 목적지 기준의 통행패턴을 산출하였으며, 산출된 통행패턴 정보를 활용하여 환승 정류소 유형 구분 방안을 제시하였다. The Seoul Institute[6]은 교통카드데이터를 도시철도 열차 스케줄 정보와 연계한 동적 최적경로탐색 알고리즘에 의하여 도시철도 내부경로 정보의 추정을 개선하였으며, 추정된 경로 정보를 활용하여 열차 및 환승역의 혼잡도를 추정하였다. The Seoul Institute [7]은 사회경제지표로서 인구수, 종사자 수 자료와 연계 분석을 통하여 대중교통 공급 수준을 지역(구단위)적으로 비교하였다. Lee et al. [8]은 k-means 클러스터링을 활용하여 대중교통 이용 시각에 따른 이용자 군집을 생성하였으며, 이를 통해 장기간의 교통카드데이터를 활용하여 대중교통 이용자 패턴의 생성이 가능함을 확인하였다. Oh and Lee [9]는 대중교통계획 시 사용하는 최적전략 통행배정기법의 결과와 교통카드데이터에서 추출한 실제 대중교통 이용자의 통행패턴을 비교 분석하여, 최적전략 통행배정에서 추정된 통행 배정량과 실제 통행량은 통행패턴 측면에서 유의미한 차이가 있음을 확인하였다.

교통카드데이터를 이용한 해외 연구는 주로 대중교통 통행행태의 분석 및 모형화를 다루고 있다. Agard et al. [10] 및 Bagchi and White [11]는 교통카드데이터 분석이 사용자 행동을 더 잘 이해하는 데 도움이 될 수 있다고 밝혔다. Agard et al. [10]은 대중교통 이용자 행동을 분석하기 위해 k-means 클러스터 기반의 방법론을 개발했다. 이 연구는 실현할 수 있는 데이터 마이닝 기법을 입증했고 대중교통 운영기관이 승객의 행동을 더 잘 이해하는 데 도움을 줄 수 있다고 하였다. Devillaine et al. [12]은 교통카드데이터를 활용해 대중교통 이용자가 수행한 활동 장소, 시간, 기간, 목적 등을 검출 및 추정 후, 목적별로 구분하는 방법론을 개발하였다. Jánošíková et al. [13]은 경로의 로짓(logit) 모델에 교통카드데이터 활용 방법을 기술했다. 차내 이동 시간, 환승 도보 시간, 환승 횟수, 노선의 배차 간격 등 4가지 속성을 고려했다. El Mahrsi et al. [14]는 교통카드데이터에 기록된 승차 정보를 활용하여 승객의 시간 프로필을 구축하였으며, unigram 모델인 토픽 모델(topic model) 기반의 군집화를 적용하여 이용자의 통행패턴을 분류하였다. Briand et al. [15]은 승객의 시간적 통행행태 및 통행패턴을 분석하기 위하여, 승객을 통행 시각에 따라 군집화 하는 연구를 수행하였다. 대부분의 기존 방법에서 사용되는 시간대(time bin) 기준의 모형 대신 시간의 연속적 표현을 고려하며 승객의 시간 프로필에서 가우스 혼합 모델(Gaussian mixture model)을 추정하는 새로운 방법을 제시하였다. He et al. [16]은 교통카드 사용자들의 일일 통행 프로필을 분류(classification)하는 연구를 수행하였다. 특히, 통행패턴의 군집화에 있어서 기존의 벡터 요소의 군집화만을 고려한 방법론이 아닌 벡터 내에서의 각 통행의 시간적 위상(position)을 고려하고자 하였다. Chen et al. [17]은 교통카드데이터와 온라인 POI(Point of interest) 데이터를 활용하여 도시철도 이용과 건조환경(built environment)의 상관성을 밝혔다. Nassir et al. [18]은 다빈도 버스 및 철도 서비스에서 교통카드데이터를 사용하여 재귀 링크기반 경로선택 모형을 제안하고, 경로 선택에서 무작위성을 반영하기 위해 매력도(attractiveness) 개념을 도입하였다. Yap et al. [19]은 교통카드데이터를 기반으로 도시대중교통(트램, 버스) 혼잡도를 평가하였다. 또한 통행 빈도가 낮은 이용자는 경로 선택에 있어서 혼잡도를 고려하지 않음을 발견하였다. Arriagada et al. [20]은 교통카드데이터를 활용하여 대규모 대중교

통망에서 이용자들의 경로 선택 전략의 이질적 행태를 규명하였다. 경쟁경로 패턴(예: 지하철+버스 | 버스 | 지하철)에서는 공통적인 성향을 보이거나 패턴 내의 개별 경로 선택은 매우 개인적인 행태를 확인하였다.

2015년 12월 관련 법률(대중교통의 육성 및 이용촉진에 관한 법률) 개정에 따라 기존 여러 연구에서 필요성으로 제기하였던 국가 차원에서 교통카드데이터를 수집·제공하는 체계가 마침내 갖추어 졌다. 교통카드데이터는 더 이상 특수한 자료가 아닌 일반적인 대중교통 부문 빅데이터가 되었으며, 기존의 연구를 통해 활용의 가능성도 충분히 입증되었다. 수집·제공에 이어 활용 체계가 정착되기 위해서는 기존 대중교통 분석체계의 보조 자료가 아닌 교통카드데이터의 특성 및 장점을 이용한 교통카드데이터 기반의 대중교통 운영방안, 즉 어떻게 주요 대중교통 현안에 적용할 수 있을지에 관한 연구가 필요하다. 이러한 관점에서 본 연구에서는 도시철도 혼잡 문제의 해결 대안인 대체 버스노선의 개발에 교통카드데이터를 활용하는 방안 및 모델을 제시하고자 한다.

2.2 분석의 전제

본 연구에서는 2019년 11월 11일(월)부터 15일(금)까지의 5일간의 서울시 교통카드데이터를 이용하였다. 그리고 철도와 버스 간 경쟁 관계에 따른 수단선택 행태를 분석하기 위해서는 두 수단이 경쟁 관계가 성립하는 조건을 먼저 설정할 필요가 있으며, 다음의 사항을 충족할 때 경쟁 관계가 성립하는 것으로 가정하였다.

첫째, 버스 통행의 경우 개별 정류장 단위의 분석이 아닌 인접한 정류장을 묶어서 통합한 정류장(통합정류장)을 구성한다. 개별 정류장으로 분석하면 자료가 세분되어 지엽적인 분석만 가능하기 때문이다. 예를 들어 강남역 주변 500m 이내에 20개 이상의 버스 정류장이 존재하는데, 각 정류장 단위의 분석으로는 이용자의 통행행태를 파악하는 데에 한계가 있다. 일반적으로 버스 이용 시 승차와 하차 정류장이 물리적으로 다른 경우가 대부분인 것도 고려해야 한다. 철도의 경우 환승역은 노선의 수에 상관없이 하나의 역으로 설정(예를 들어 시청역의 경우 1호선, 2호선 역이 존재하지만 하나로 묶어 줌)한다. 버스는 서울시 모든 운행노선의 버스 정류장 간 평균 거리가 노선별로 125~1,722m 수준이며 평균값은 약 385m인 것(2019년 11월 15일 노선데이터로부터 산출한 결과임)을 고려하여, 200m 이내의 인접한 정류장끼리 묶어서 하나의 통합정류장을 구성한다. 그리고 각 통합정류장을 대표하는 지점을 중심점이라 하고 중심점의 좌표는 통합정류장 내의 개별 정류장들의 좌표 평균으로 정의한다.

둘째, 통행의 출발지와 도착지의 조합인 OD의 설정은 철도역 OD를 기준으로 설정한다. 따라서 철도의 경우는 선택된 출발역과 도착역의 OD 값이 분석의 기준이 된다. 이와 경쟁할 수 있는 버스의 출발지 및 도착지를 설정함에 있어서는 철도역과의 직선거리(기준거리)를 기준으로 하여 설정한다. 일반적으로 철도역의 영향권이 500m인 점을 고려하여 철도역과 버스 통합정류장의 중심점 간의 거리가 500m 이내인 경우로 한정하였으며, 100m 단위로 구분하여 분석한다. 기준거리 이내에 버스 통합정류장이 존재하지 않는 경우는 분석 대상 OD에서 제외한다.

이상의 조건을 충족시키는 OD 쌍을 도출하기 위해서 한국철도기술연구원이 개발한 TRIPS(Travel Record based Integrated Public transport operation System, 통행실적 기반 복합대중교통 운영계획수립 지원시스템) 프로그램을 활용하였다. TRIPS 프로그램을 활용하여 5일 치의 교통카드데이터를 입력하였으며 입력된 자료의 요약은 Table 1과 같다.

Table 1. Summary of smart card data.

Date	Number of stations		Number of bus lines	Number of trips by number of transfers				
	Metro	Bus		Total	No transfer	1	2	3+
20191111	739	14,721	617	10,694,010	8,627,747	1,828,200	208,616	29,447
20191112				10,935,560	8,822,944	1,872,835	211,000	28,781
20191113				10,867,316	8,731,163	1,890,622	215,348	30,183
20191114				10,791,766	8,704,141	1,850,144	208,247	29,234
20191115		14,710	616	11,026,869	8,946,765	1,849,626	202,846	27,632

먼저 분석을 위한 철도역 구성 방식은 Fig. 1과 같다. 그림에서 청색 점들은 호선 별 각각의 철도역을 의미하며, 원으로 표시된 영역의 중심이 묶인 환승역들의 지리적인 대푯값이 된다.

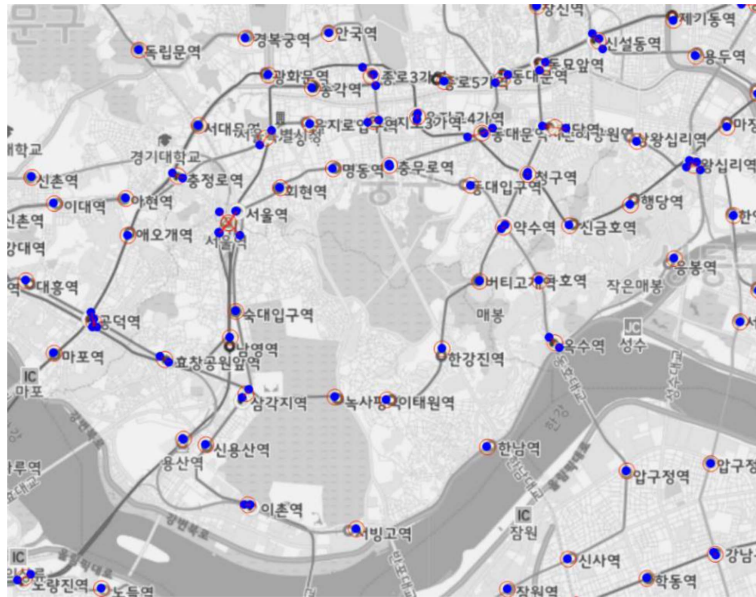


Fig. 1. Metro stations.

버스정류장도 같은 방식으로 통합정류장을 구성한다. 철도역 및 버스 통합정류장이 모두 구축된 예시를 Fig. 2에 제시하였다. 왼쪽은 강남지역 일부를 표시한 것이고, 오른쪽 그림은 왼쪽 그림에서의 사각형 부분(선릉역 인근)을 확대한 것이다. 그림에서 보는 것처럼 붉은 원이 철도역, 녹색 원이 버스의 통합정류장이다. 이렇게 구축된 통합정류장 간 거리는 중심점의 좌표를 이용하여 계산할 수 있다. 철도역 및 버스 통합정류장 간 OD가 서로 경쟁 관계라고 설정하는 과정은 다음의 두단계로 구성된다.

- 단계 1: 철도역을 기준으로 하여 실제 통행이 발생한 철도의 출발지와 도착지 쌍(A)을 모두 구축.
- 단계 2: 출발지 및 도착지의 철도역 모두에서 기준거리 이내에 있는 버스 통합정류장이 있는 경우만을 경쟁이 가능한 것으로 설정하여 버스 및 철도의 출발지-도착지 쌍을 구축(즉, A에서 경쟁이 성립하지 않는 출발지-도착지 쌍은 제외).

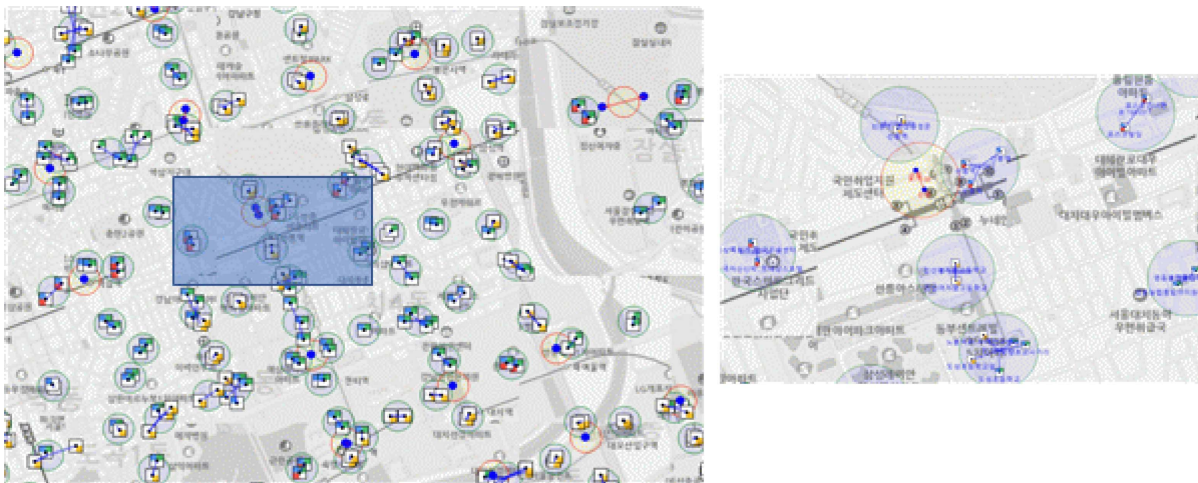


Fig. 2. Aggregated bus stops and metro stations.

2.3 철도와 버스의 경쟁관계 분석

앞서 철도와 버스의 경쟁이 성립하는 기준거리는 최소 100m에서 최대 500m까지 설정하였다. 일반적인 도시철도역의 도보

이용권역이 500m인 점을 고려하여 이를 최대치로 하였고, 100m 단위로 대안을 설정하였다. 기준거리가 작아질수록 철도역에 매칭되는 버스 통합정류장의 수가 작아지게 된다.

수도권과 달리 지방의 경우 도시철도 노선의 수가 많지 않기 때문에 무환승 통행이 대부분을 차지한다. 본 연구에서 사용한 교통카드데이터에서도 무환승 통행이 전체의 80%를 상회하고 있다. 따라서 본 연구에서는 가장 일반적인 통행패턴인 무환승 통행을 기준으로 행태를 분석한다. 무환승 통행에 대해서 이동거리별로 철도 및 버스 이용자 추이를 분석한 것이 Fig. 3이다. 여기서 D_{br} 은 기준거리로서 버스정류장(b)과 철도역(r)의 중심 간 거리를 의미한다. 통행거리는 1km 단위로 구분하였으며, 도시철도를 이용한 실제 통행거리를 반영한 것이다. 기준거리가 커질수록 경쟁이 되는 버스 통합정류장 간 OD가 커지기 때문에 버스의 분담률은 점차 올라가는 형태를 보인다.

Fig. 3에서 통행거리가 증가할수록 철도의 분담률이 상승하는 경향을 모든 경우에서 확인할 수 있다. 철도의 환승 횟수별로 같은 형태의 그림을 그릴 수 있으며, 분석한 그림을 별도로 제시하지는 않았으나, 같은 현상을 확인할 수 있다. Table 2는 무환승 통행에 대해서 기준거리별로 철도의 분담률이 각각 50, 75, 85, 90, 95%에 도달하게 되는 통행거리를 표시한 것이다. 기준 거리에 따라 차이는 있으나, 기준거리가 500m일 때도 통행거리가 7km를 넘어간다면 철도의 분담률이 95% 이상으로서 압도적으로 철도의 분담률이 높아지게 된다. 7km 이하의 통행에 대해서는 버스 정류장과 철도역 사이의 거리가 철도의 경쟁력에 영향을 크게 미치나, 7km를 초과하는 통행에 대해서는 버스 정류장과 철도역 사이의 거리가 철도의 경쟁력에 큰 영향을 미치지 않음을 의미한다.

Table 2. Distances at which modal share of metro is achieved (non-transfer trips).

D_{br}	Modal share of metro				
	50%	75%	85%	90%	95%
~ 100m	0~1km	0~1km	0~1km	1~2km	1~2km
~ 200m	0~1km	0~1km	1~2km	1~2km	2~3km
~ 300m	0~1km	1~2km	2~3km	3~4km	4~5km
~ 400m	0~1km	2~3km	3~4km	4~5km	5~6km
~ 500m	1~2km	3~4km	4~5km	5~6km	7~8km

결국, 도시철도의 수요 분산을 위한 대체 노선의 설계방안으로 다음의 두 가지를 도출할 수 있다. 첫째, 철도역과 버스 정류장의 거리가 가까울수록 버스로 흡수되는 수요가 커진다. 결국 대체 노선의 정차 정류장은 해당 도시철도 노선의 역과 지리적으로 가깝게 배치하는 경우 효과가 극대화된다. 둘째, 도시철도 수요를 분산해야 하는 구간이 상대적으로 중장거리인 경우, 철도역과 다소 거리를 두고 버스 정류장을 배치해도 충분한 분산 효과를 얻을 수 있다.

즉, 철도역과 다소 거리가 있는 지점에서 통행하는 승객의 경우에 대체 버스노선이 존재하지 않을 때는 어쩔 수 없이 철도역을 이용하려는 경향이 높지만, 대체 노선이 존재한다면 철도역의 이용 필요성이 줄어든다는 것을 의미한다. 대부분의 철도역 주변이 교통 및 보행이 혼잡한 지역임을 감안할 때, 해당 철도역에서의 OD 패턴에 따라서는 철도역에서 다소 거리가 있는 지점의 수요가 철도역으로 흡수되지 않도록 대체 버스노선의 정류장을 배치해도 효과를 거둘 수 있음을 의미한다. 따라서 수요 예측의 정확성 향상을 위해서는 경쟁 수단의 정류장(역)과의 거리에 따른 각각의 수단분담 모형개발이 필요하다.

본 연구의 실험 결과를 다른 관점에서 해석한다면, OD 기준으로 철도와 버스가 인접해서 배치된다면 상호 경합 효과가 커진다는 것이다. 도시철도 노선 설계 때 대중교통망의 효율성 관점에서는 불필요한 경합이 발생하므로 버스노선의 재배치가 필요하다. 또한 도시철도 기본계획이나 타당성 조사 단계에서 주변의 버스 정류장의 위치 및 버스노선까지 고려한 경합 효과를 분석하는 것이 도시철도의 정확한 수요 예측에 필수적이다.

2.4 철도와 버스의 수단분담 모형

철도와 버스의 수단선택 행태를 분석한 자료를 토대로 하여 버스노선과 도시철도의 대체 또는 경합 효과를 정량적으로 분석하기 위해서는 두 수단 간의 선택행태를 설명할 수 있는 모형이 필요하다. 앞서 도출한 결과에 따라, 수단분담 모형을 거리대별로 구분하여 각각 개발한다. 본 연구에서는 이항로짓 모형을 적용하였으며, 효용함수 구축에 사용된 변수는 다음의 효용함수 식에서 보는 바와 같이 통행거리와 운행횟수이며, 운행횟수는 버스에만 적용하는 것으로 모형을 설정하였다.

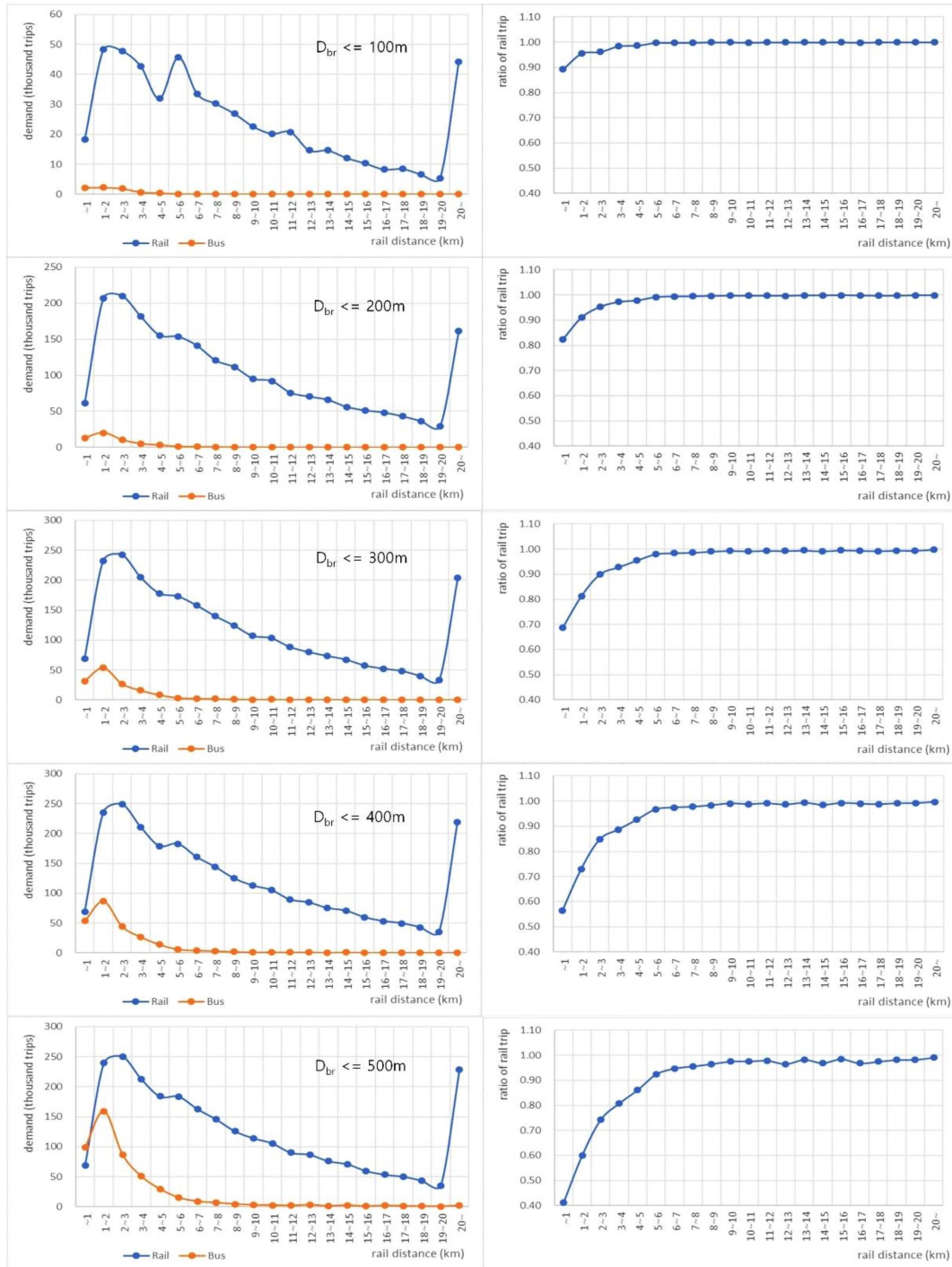


Fig. 3. Trends of modal share of metro for metro travel distance.

$$U_m = \alpha_1 \cdot DIST_m + \alpha_2 \cdot FREQ_m + c_m$$

U_m : 수단 m의 효용함수

α_1, α_2 : 통행거리 및 운행횟수에 대한 계수

c_m : 수단 m의 상수항

$DIST_m$: 수단 m을 이용한 이용자의 통행거리 평균 (단위: m)

$FREQ_m$: 수단 m의 운행횟수 합으로, 버스 경우에만 적용(철도의 경우 0) (단위: 회)

효용함수 구축에 활용된 자료는 다음의 조건을 만족하는 것으로 하였다. 첫째, 철도역과 버스정류장 간의 물리적인 직선거리가 300m 이내인 경우로 한정한다. 100m 이내의 경우 샘플의 수가 229개 OD로서 매우 제한적으로 작아지므로 적정 수의 샘플을 확보할 수 있는 200m 및 300m 이내인 경우에 대하여 모형을 구축하였다. 둘째, 통행거리가 철도 기준 1~8km인 통행을 대상으로 한다. 앞에서 분석한 바와 같이 철도의 통행거리가 8km를 초과하면 버스의 경쟁력이 없어 추가노선 투입에 따른 효과를 달성할 수 없다. 또한 1km 이하의 통행은 도심에서 1~2개 정류장 정도의 이동거리로서 기준거리와의 차이를 고려할 때 오류 발생의 가능성이 있다. 셋째, 환승 통행은 고려하지 않는다. 서울의 환승 통행은 전체 통행의 20% 미만인 수준이며, 철도의 혼잡을 줄이기 위해서 버스를 투입하는 것을 고려할 때 환승의 경우는 경쟁력이 매우 떨어지기 때문이다. 마지막으로, 서울시에서 운행하고 있는 버스노선 중에서 일반적인 버스의 성격을 갖는 간선(파란색) 및 지선(초록색) 버스를 대상으로 한다. 광역노선은 장거리 이용객이 주 대상이므로 통행거리가 8km 이하에서는 큰 의미를 갖기 어렵다. 마을버스의 경우 도시철도와 경쟁하기보다는 보완(연계교통) 수단으로 주로 이용되는 점을 고려하여 제외한다.

이상의 조건을 만족하는 데이터에 대한 기본정보는 Table 3과 같다. 기준거리를 200m로 하는 OD 쌍은 1,320개이며, 도시철도 수요는 465,155 통행, 버스 수요는 33,093 통행 수준이다. 도시철도 이용객의 평균 통행거리는 3.3km, 버스는 2.5km 수준이다. 버스 이용객들이 선택할 수 있는 노선들의 운행 횟수 합은 평균은 563회 수준이다. 마찬가지로 기준거리 300m인 경우는 해당 값들을 표에서 확인할 수 있다.

Table 3. Summary of input data.

	Number of OD pairs	Number of trips		Average number of operating bus	Average travel distance	
		Rail	Bus		Rail	Bus
~ 200m	1,320	465,155	33,093	563	3,332	2,476
~ 300m	2,234	784,718	89,344	552	3,525	2,564

이 데이터들을 사용하여 추정된 효용함수 계수는 Table 4와 같다. 모든 계수의 *t* 값이 유의함을 알 수 있고, ρ^2 값도 모형을 사용하기에 적절한 수준이다.

Table 4. Summary of utility functions.

	Coefficient (standard error, t-value)		Constant (standard error, t-value)		Rho-squared (Rho-bar-squared)	Log-likelihood at zero (Log-likelihood at convergence)
	α_1	α_2	rail	bus		
~ 200m	-0.00072043 (0.000028, -25.2909)	0.00106168 (0.000035, 30.6972)	2.9483 (0.0258, 114.4260)	0	0.6714 (0.6713)	-34,545.1 (-11,353.0)
~ 300m	-0.00068128 (0.000020, -34.6536)	0.00091698 (0.000023, 39.4633)	2.4049 (0.0167, 143.6913)	0	0.5507 (0.5507)	-60,626.1 (-27,236.9)

모형의 검증은 실제 통행량과 각 모형을 기준으로 하여 예측된 통행량을 상호 비교하여 RMSE, MAPE, MAE, Theil의 부등계수(Theil의 부등계수는 0에 가까울수록 예측 통행량과 관측 통행량이 유사한 것이고, 1에 가까울수록 예측 통행량과 관측 통행량은 관련이 없는 것으로 판단할 수 있음)를 이용하였다. 철도 통행량을 기준으로 계산한 결과 기준거리 200m인 경우 RMSE, MAPE, MAE, 부등계수가 기준거리 300m인 경우보다 상당히 개선되는 것을 확인할 수 있다(Table 5). 기준거리를 200m로 설정하면 300m인 경우에 비해서 모든 경우에서 모형의 정확도가 훨씬 더 개선되는 것으로 보아 실제 적용 시에 신규노선의 정류장과 철도역 간의 거리를 기준으로 자료를 구분하여 따로 적용하는 것이 분석의 정확도를 개선할 수 있을 것으로 판단된다.

Table 5. Validation of models.

D_{br}	Number of OD pairs	RMSE	MAPE	MAE	Theil's U*
~ 200m	1,320	32.5924	40.1981	17.9710	0.0307
~ 300m	2,234	49.2391	49.4152	27.9632	0.0477

$$* \text{Theil's } U = \frac{\sqrt{\sum_{a=1}^N (\hat{V}_a - V_a)^2 / N}}{\sqrt{\sum_{a=1}^N (V_a)^2 / N} + \sqrt{\sum_{a=1}^N (\hat{V}_a)^2 / N}}$$

V_a : the actual values, \hat{V}_a : the predicted values,
 N : the number of observations

3. 결 론

최근 신규 경전철, 도시철도 노선 계획과 관련하여, 수요 예측의 정확성 향상을 위하여 철도와 버스 간의 경쟁 및 보완 관계에 관한 연구의 필요성이 증가하고 있다. 본 연구에서는 서울시의 교통카드데이터를 이용하여 도시철도와 버스의 수단선택 행태를 분석하였다. 이를 위해서 정류장을 그룹화하고, 경쟁 관계가 성립되는 조건의 정립 및 5일 치의 교통카드데이터를 이용하여 행태를 분석하였다. 실험 결과, 다음의 결론을 도출하였다. 첫째, 경합 관계에 있는 버스노선에 따라 도시철도 수요가 달라질 수 있음을 확인하였다. 둘째, 도시철도 노선과 버스노선이 경합 관계에 있는 경우, 철도역과 버스 정류장의 거리가 가까울수록 버스로 흡수되는 수요가 커진다. 셋째, 도시철도의 용량 한계로 수요의 분산이 필요한 경우, 분산해야 하는 구간이 상대적으로 중장거리라면 철도역과 다소 거리를 두고 버스 정류장을 배치해도 충분한 분산 효과를 얻을 수 있다. 결국, 도시철도 노선 계획 단계에서 정확한 수요 예측을 위해서는 버스노선과 정류장 위치가 정확히 반영된 공간적으로 세밀한 분석이 필요하다.

또한 본 연구에서는 버스노선과 도시철도의 대체 또는 경합 효과를 정량적으로 분석하기 위하여 두 수단 간의 선택행태를 설명할 수 있는 수단분담 모형을 개발하였다. 특히, 버스 정류장과 철도역 사이의 거리에 따라 수단 선택 행태가 달라짐을 확인함에 따라, 버스 정류장과 철도역 사이의 거리에 따른 수단분담 모형을 각각 개발하였으며 실측 데이터를 활용하여 모형의 정확성을 검증하였다. 실제 도시철도의 대체 버스노선을 설계하는 경우, 신규 버스노선의 OD와 기존 도시철도의 OD를 매칭하여 거리대별로 본 모형을 적용할 수 있다. 수요전환이 어느 정도 발생하는지 정량적으로 분석할 수 있으므로, 모형 적용에 따른 수요전환 효과 분석과 노선 재설계의 과정을 반복하면서 최적의 노선을 설정할 수 있을 것이다. 또한 외부적인 요인으로 도시철도가 정상적으로 운행하지 못할 경우의 대체 노선 개발에도 제한적이거나 활용될 수 있을 것이다.

다만, 본 연구에서 개발한 모형은 분석 대상 지역에서만 검증되었으므로, 일반화된 형태의 모형개발을 위하여 다양한 지역에서의 분석이 필요하다. 통행행태의 관점에서 평일뿐만 아니라 주말 및 공휴일에 대해서 별도의 분석도 필요하다. 또한 본 연구에서는 환승이 없는 통행만을 연구의 범위로 설정하여 수단분담 모형을 개발하였으나, 일반적인 통행행태로 확장하기 위해서 환승통행 및 환승 유형까지 포함하는 형태의 연구가 필요할 것이다.

후 기

본 연구는 한국철도기술연구원 주요사업(데이터 기반 대중교통 수요예측 및 서비스 개선기술 개발, PK2201C3)의 연구비 지원으로 수행되었습니다.

References

- [1] The Korea Transport Institute (2006) Study of potential usage of contactless smart card data for establishing public transport policy, KOTI Basic Research Reports, pp. 108-134.
- [2] The Seoul Institute (2007) The estimation and application of origin-destination tables by using smart card data, The Seoul Institute Research Reports, pp. 41-96.
- [3] The Seoul Institute (2007) Integrated transit service evaluation methodologies using transportation card data, The Seoul Institute Research Reports, pp. 63-93.
- [4] The Seoul Institute (2009) A study on the public transportation service index development (2nd phase), The Seoul Institute Research Reports, pp. 61-91.
- [5] Gyeonggi Research Institute (2011) A study on analysis travel patterns and applications using Gyeonggi-do traffic card data, GRI Research Reports, pp. 39-76.
- [6] The Seoul Institute (2011) Congestion index of urban rail transit using public transportation card data, The Seoul Institute Research Reports, pp. 33-50.
- [7] The Seoul Institute (2012) Travel characteristics and behavior of transit card users, The Seoul Institute Research Reports, pp. 73-87.
- [8] I. Lee, J. Min, K. Kim, S.-Y. Kho (2020) Generating Travel Patterns of Public Transportation Users Using a k-means Clustering Based on Smart Card Data, *Journal of the Korean Society for Railway*, 23(3), pp. 204-215.
- [9] D. Oh, I. Lee (2021) Comparative Analysis of the Optimal Strategy Transit Assignment and the Actual Travel Behavior Using Smart Card Data, *Journal of Korean Society for Urban Railway*, 9(4), pp. 1049-1057.
- [10] B. Agard, C. Morency, M. Trépanier (2006) Mining public transport user behaviour from smart card data, *IFAC Proceedings Volumes*, 39(3), pp. 399-404.

- [11] M. Bagchi, P.R. White (2005) The potential of public transport smart card data, *Transport Policy*, 12, pp. 464-474.
- [12] F. Devillaine, M. Munizaga, M. Trépanier (2012) Detection of Activities of Public Transport Users by Analyzing Smart Card Data, *Transportation Research Record*, 2276, pp. 48-55.
- [13] L. Jánošíková, J. Slavík, M. Koháni (2014) Estimation of a route choice model for urban public transport using smart card data, *Transportation Planning and Technology*, 37(7), pp. 638-648.
- [14] M.K. El Mahrsi, E. Côme, J. Baro, L. Oukhellou (2014) Understanding Passenger Patterns in Public Transit Through Smart Card and Socioeconomic Data, *3rd International Workshop on Urban Computing (SigKDD)*, New York, USA.
- [15] A.-S. Briand, E. Côme, M.K. El Mahrsi, L. Oukhellou (2016) A mixture model clustering approach for temporal passenger pattern characterization in public transport, *International Journal of Data Science and Analytics*, 1, pp. 37-50.
- [16] L. He, B. Agard, M. Trépanier (2018) A classification of public transit users with smart card data based on time series distance metrics and a hierarchical clustering method, *Transportmetrica A: Transport Science*, 16, pp. 1-20.
- [17] E. Chen, Z. Ye, C. Wang, W. Zhang (2019) Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data, *Cities*, 95, Article 102359.
- [18] N. Nassir, M. Hickman, Z.-L. Ma (2019) A strategy-based recursive path choice model for public transit smart card data, *Transportation Research Part B: Methodological*, 126, pp. 528-548.
- [19] M. Yap, O. Cats, B. van Arem (2020) Crowding valuation in urban tram and bus transportation based on smart card data, *Transportmetrica A: Transport Science*, 16, pp. 23-42.
- [20] J. Arriagada, M. A. Munizaga, C. A. Guevara, C. Prato (2022) Unveiling Route Choice Strategy Heterogeneity from Smart Card Data in a Large-Scale Public Transport Network, *Transportation Research Part C: Emerging Technologies*, 134: 103467.

(Received 5 October 2022; Revised 17 October 2022; Accepted 31 October 2022)

Kyoungtae Kim : ktkim@krri.re.kr

Principal Researcher, Innovative Transportation and Logistics Research Center, Incheon University, Korea Railroad Research Institute, 176 Cheoldobangmulgwan-ro, Uiwang-si, Gyeonggi-do 16105, Republic of Korea

Inmook Lee : mook79@krri.re.kr

Senior Researcher, Innovative Transportation and Logistics Research Center, Incheon University, Korea Railroad Research Institute, 176 Cheoldobangmulgwan-ro, Uiwang-si, Gyeonggi-do 16105, Republic of Korea