ARTICLE

통행 수요를 고려한 대중교통 운행 방식 선정 방법에 관한 연구: 고정형과 수요응답형을 중심으로

서성주¹ · 김진희²*

¹현대자동차(주) 내비게이션개발팀 연구원, ²연세대학교 도시공학과 부교수

Selection of Operation Plan Considering Travel Demand: Focusing on the Fixed-Route Transit and the Demand-Responsive Transit

SEO, Sungiu¹ · KIM, Jinhee^{2*}

¹Research Engineer, Navigation Development Team, Hyundai Motor Group, Seoul 06179, Korea ²Associate Professor, Department of Urban Planning and Engineering, Yonsei University, Seoul 03722, Korea

Abstract

To maximize the socioeconomic benefits of public transportation, such as economic feasibility and eco-friendliness, it is necessary to establish an appropriate operational plan considering the travel demand. If not, it may lead to negative problems such as weakening service sustainability, worsening transportation equity, and increasing inconvenience in use. Furthermore, the fixed-route transit and the demand-responsive transit are different in aspects of the strengths; the former operates on pre-determined routes and timetables and the latter responds flexibly in responding to the demand. To exploit their strength, the appropriate operation plan should be determined with sufficient consideration of demand. Typically, the former is efficient where there is enough demand, while the latter is efficient where there is not enough demand. Therefore, this study aims to establish a methodology for deriving an appropriate type of public transportation service among fixed-route and demand-responsive transit based on demand data and additional socioeconomic characteristics. To this end, travel pattern was clustered by DBSCAN based on travel characteristics data divided into three main aspects: spatial, temporal, and socioeconomic characteristics. As a result of case analysis in Seodaemun-gu, Seoul, a total of 182 origin-destination pairs were divided into 7 clusters and 29 noises. Some specific clusters seem that the demand-responsive transit is appropriate in that they have low traffic volumes and irregular occurrence patterns, whereas the noises seem that the fixed-route transit is better in that they have enough traffic volumes and less irregular occurrence patterns.

Keywords: dbscan, demand-responsive transit, fixed-route transit, travel demand, travel pattern

초록

대중교통의 경제성, 친환경성 등 사회경제적 이점을 극대화하기 위해서는 통행 수요를 고려하여 적절한 운영 제반사항이 정립될 필요가 있다. 수요가 충분히 고려되지 않는다면, 손익 악화에 따른 서비스 지속가능성 약화. 교통 형평성 악화, 이용 불편 증가 등 부정적 문제로까지

J. Korean Soc. Transp. Vol.42, No.1, pp.29-46, February 2024 https://doi.org/10.7470/jkst.2024.42.1.029

pISSN: 1229-1366 eISSN: 2234-4217

ARTICLE HISTORY

Received: 31 October 2023 Revised: 7 November 2023 Accepted: 19 December 2023

Copyright © Korean Society of Transportation

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.

^{*}Corresponding author: kim.jinhee@yonsei.ac.kr

이어진다. 나아가 사전에 정해진 노선과 배차시간표대로 운행되는 고정형 방식과 시시각각 수요에 따라 유연하게 대응하는 수요응답형 방식은 서로 장단점이 다른데, 수요를 충분히 고려하여 지역별, 시간대별 적합한 운행 방식을 결정하여야 각각의 장점을 극대화할 수 있다. 일반적으로 전자는 수요가 시공간적으로 고정적이면서 양이 많은 곳에서 경제적으로 효율적인 반면, 후자는 수요가 적고 가변적인 곳에서 효율적인 것으로 알려져 있다. 이에 본 연구는 사전에 확보된 수요의 발생 패턴 정보와 그 외 추가적인 지역사회의 사회경제적 특성에 기반하여 고정형과 수요응답형 중 적절한 대중교통 서비스 유형을 도출해내는 방법론을 정립하고자 한다. 이를 위해 크게 공간적 특성과 시간적 특성, 사회경제적 특성 3가지로 구분한 통행 특성 데이터를 밀도 기반 클러스터링(DBSCAN)하여 통행 행태를 그룹화하였다. 서울특별시 서대문구 사례 분석 결과, 총 182개의 기종점쌍은 7개의 클러스터와 29개의 노이즈로 구분되었다. 특정 클러스터에 속하는 기종점쌍은 통행량이 적으면서도 발생 패턴이 불규칙하다는 점에서 수요응답형운행 방식이 적합하고, 노이즈에 속하는 기종점쌍은 통행량이 충분히 많고 그 정도가 시간의 흐름에 따라 큰 변화를보이지 않는다는 점에서 고정형 운행 방식이 적절한 것으로 판단되었다.

주요아: DBSCAN, 수요응답형 대중교통, 고정형 대중교통, 통행 수요, 통행 패턴

연구의 배경 및 목적

공통된 통행수단을 통하여 다수의 인원을 한 번에 효율적으로 수송하는 대중교통(Public Transportation)은 주로 4-5인승에 그치는 자가용 차량 대비 더 많은 인원을 한 번에 이동시킨다. 이를 통하여 이용자에게는 통행 비용을 절 감하고, 공급자에게는 운영 비용을 절감할 수 있는 이점이 있다. 나아가 상대적으로 저렴한 비용 덕에 경제적 취약계 층과 교통약자(장애인, 고령자 등)를 포함한 모든 대중에게 최소한의 이동권을 보장하는 역할을 한다. 또, 한정된 자원으로 높은 효율의 서비스 운영이 가능함에 따라 자가용 대비 경제적이고, 친환경적이다. 예를 들어, 대중교통은 도로상에서 차지하는 동일 면적 대비 더 많은 수송 인원을 처리할 수 있다. 또, 더 적은 대수의 차량만으로도 동일한 인원을 수송할 수 있다. 이에 운영사를 포함한 서비스 공급 주체 입장에서는 더 적은 비용으로 운영할 수 있다 (Forouzandeh et al., 2022; Kepaptsoglou and Karlaftis, 2009; Schofer, 2023).

대중교통의 사회경제적 이점과 공공적 성격을 고려하여 대중교통 이용 분담률을 높이고자 하는 노력이 이어지고 있다. 대중교통 특성상 이용객 수와 무관하게 최소 고정 비용이 발생하는데, 실제 수요가 이를 충당하지 못하는 경우, 민영제 방식만으로는 대중교통 운영이 경제성 측면에서 지속가능하지 않다. 이에 이동권 보장 수단이라는 점에서 결국 교통 형평성의 문제로까지 이어진다. 또한, 수요량 자체가 많더라도 해당 수요의 특성을 고려하지 않는다면 여전히 문제가 유발된다. 특정 노선을 이용하는 승객 수는 많으나 특정 시간대에만 몰리는 경우, 이러한 수요의 시간적 특성을 고려하지 않고 모든 시간대에 대해 동일하게 차량을 배차한다면, 이용 편의 악화는 물론, 이용객들의 시간비용 손실도 발생하고, 심지어는 안전 문제까지도 이어진다. 따라서 운영 노선과 스케줄, 특히, 이용 수요를 고려한 적절한 운영 제반사항이 정립될 필요가 있다.

대중교통 이용 분담률을 높이려는 노력의 일환으로써 대중교통 운영 방식 또한 다양화 및 정교화 되어 왔다. 일반적으로 대중교통은 고정된 노선(네트워크)과 배차시간표(운영스케줄)에 기반한 형태인데, 이는 다양한 통행패턴 특히, 변동이 큰 수요에 대해서는 빠른 대응이 어렵다는 한계를 지니고 있었다(Tuydes-Yaman et al., 2023). 그러나최근 정보통신기술(ICT)의 발달에 따라 이용객이 자신의 휴대전화로 손쉽게 수요를 표출할 수 있게 되었고, 이에이를 바탕으로 노선과 스케줄을 유연하게 변화시키는 형태가 대두되었다(Atasoy et al., 2015; König and Grippenkoven, 2020). 대표적으로 Figure 1과 같이 실시간 집계 수요에 기반하여 노선과 시간표를 유동적으로 변화시키는 수요응답형 대중교통(DRT; Demand Responsive Transit)이 주목받기 시작하였다.

SEO, Sungju · KIM, Jinhee Article ·



Figure 1. Public transportation operation types: fixed-route transit and demand-responsive transit

수요응답형 대중교통은 기존 고정형 방식에 비해 차량 운용횟수당 수송인원은 적지만, 상대적으로 자원과 유지비용을 최소화할 수 있다는 이점이 있다. 또 기존 방식으론 대응이 어렵던 수요가 적거나 불규칙하던 지역까지도 수요응답형 대중교통은 커버할 수 있다. 그리고 이용자 입장에서는 택시에 비해서 여전히 다인승이라는 점에서 적은 통행비용으로도 이용이 가능하다(Quadrifoglio and Li, 2009; Tang et al., 2023). 다만, 일반적으로 운영 원리상 수요응답형 방식은 수요가 충분치 않은 곳에서는 효율적이지만, 수요가 충분히 많은 곳에서는 고정형 방식이 더 효율적인 것으로 알려져 있다. 수요가 밀집한 곳에서는 고정형 방식이 차량 운용횟수당 수송인원이 많다는 점에서 도로상 혼잡도 완화하고 운행 비용도 낮출 수 있다. 반면, 수요응답형 방식은 승하차하는 공간적 범위가 넓어지는 만큼차량주행거리(Vehicle Kilometers Traveled)가 길어지고 운행 비용도 늘어날 수 있다(Abdelwahed et al., 2023; Basu et al., 2018; Moon et al., 2021; Nourbakhsh and Ouyang, 2012). 실제 운영 사례를 살펴봐도 수요응답형 대중교통은 덜 복잡할수록 성공 확률이 높고, 대략적으로 시간당 10명 내외의 적은 인원을 수송하는 것으로 나타났다(Currie and Fournier, 2020; Potts et al., 2010; Wang et al., 2023).

이처럼 대중교통 운행 방식이 달성하고자 하는 목표가 다르듯이 각각의 장단점도 달라 운행 방식별로 도입하기 적합한 시공간적 범위도 상이하다. 상대적으로 수요가 가변적인 경우에는 수요응답형이, 반대로 수요가 고정적인 경우에는 고정형 방식이 유리하다. 따라서 지역별, 시간대별 수요 발생 패턴에 따라 어떠한 방식을 도입하는 것이 바람직한 지에 관해 연구가 필요하다. 이에 실제로 많은 연구가 이뤄졌으나, 수요와 공간적 구조, 그리고 지역적 특성을 종합적으로 고려하여 대중교통 운행 방식을 결정하는 것에 관한 연구는 적은 것으로 확인되었다. 이에 본 연구는 수요 정보와 그 외 사회경제적 특성에 기반하여 고정형과 수요응답형 중 적절한 대중교통 서비스 유형을 선정하는 방법론을 정립하고자 한다. 물론, 실제 대중교통 운행 방식을 결정하기 위해서는 차량의 크기, 운영대수, 운영 범위 등 다양한 상세 요건이 고려되어야 하나, 본 연구에서는 세부 운영 계획을 수립하기 전 단계로써 수요 정보와 그 외 특성에 기반하여 대략적으로 파악하는 과정에 한하여 제시한다.

선행연구

1. 대중교통 운영 제반사항 설계에 관한 연구

대중교통의 운영 효율성과 이용자 편의성을 높이기 위해 운영 제반사항의 개선 방안에 관한 많은 연구가 이뤄져왔다. 주로 한정된 자원 아래 수혜 인원을 늘리기 위하여 스케줄링과 노선망 선정에 관한 연구가 많다. 다수가 이용자 편익을 극대화하거나, 운영 비용을 최소화하는 등 여러 목적함수를 정립하고 이를 최적화하는 문제로 접근한다. 대표적으로, TRNDP(Transit Route Network Design Problem)가 있으며, 대중교통 노선의 경로와 수, 이용 가능한 차량 수 등 제약조건 아래에서 효과적인 대중교통 네트워크를 설계하는 방법을 강구한다(Cyril et al., 2020; Estrada et al., 2021; Kepaptsoglou and Karlaftis, 2009).

이외에도 교통카드 데이터와 GPS 데이터 등 통행에 관련된 다양한 데이터를 활용하는 방법론에 관한 연구도 활발하다. 각종 통행 데이터와 함께 설문조사 데이터 등 다양한 데이터를 종합하여 최적 노선을 설계하고자 하는 연구

도 있다. 특히, 각종 데이터를 바탕으로 주요 기종점 분포를 분석하여 수요응답형 버스의 노선을 선정하는 연구도 이 뤄졌다(Hamedmoghadam et al., 2021; Liu et al., 2023b; Ma et al., 2017; Yu et al., 2020).

2. 통행 패턴 파악에 관한 연구

정보통신기술 발달에 따라 통행 데이터 구득이 쉬워짐에 따라 이에 기반하여 통행 패턴을 도출해내기 위한 연구도 많이 이뤄지고 있다. 더 나아가 통행 패턴의 특성을 도출하거나 통행 패턴에 영향을 미치는 변수를 파악해내기 위한 연구도 이뤄졌다. 데이터로부터 통행 패턴의 시간적 주기성을 도출하는 방법론에 관한 연구도 있고, 차량 궤적의 시공간 위치, 이동 방향 특성, 사회경제 활동의 순서 및 활동 시간, 사회경제적 변수 등 다양한 특성에 기반하여 통행 패턴의 종류를 구분해내는 연구도 있다(Babu and Anjaneyulu, 2021; Jiang et al., 2023; Tang et al., 2021; Wilson, 2008). 데이터에 기반해서 공유 자전거 이용자 수요의 시공간적 패턴을 분석하여 관련된 전략을 수립하고 자한 연구도 있다(Lu et al., 2022; Wang et al., 2022; Zhao et al., 2022).

통행 패턴을 파악하기 위하여 데이터를 군집화하는 방법론도 상당수 있다. 그래프 기반 클러스터링 기법인 계층적 클러스터링과 커뮤니티 탐지 기법을 적용하기도 하고, 통행 패턴의 공간적 특성을 고려하여 밀도 기반 공간적 클러스터링 방법론을 활용한 사례도 있다(Chen et al., 2019; Fekih et al., 2022; Li et al., 2022; Tang et al., 2021; Zhao et al., 2017). 다수의 연구는 상대적으로 노이즈에 따른 왜곡이 적고 사전적으로 군집 개수를 정의할 필요가 없는 밀도 기반 클러스터링을 많이 활용하였다.

3. 수요응답형 대중교통에 관한 연구

기존 고정형 대중교통 방식과 더불어 수요응답형 대중교통 방식에 관한 연구도 늘고 있다. 예로, 고정형 버스와 수요응답형 버스 사이의 환승을 고려하여 수요응답형 버스의 노선을 도출하는 연구가 이뤄졌다(Tang et al., 2023). 이러한 수요응답형 대중교통은 다양한 기준에 따라 더 세분화 될 수 있다. 기종점 형태에 따라서는 일대일형, 다대일형, 일대다형, 다대다형으로 구분되며, 특히 노선 형태에 따라서는 탄력적 노선운영형, 지선노선 서비스형, 탄력적 정류장 정차형으로 구분된다. 탄력적 노선운영형은 기본적인 노선 아래 수요에 따라 일부 변경을 가하는 것이고, 지선노선 서비스형은 기존 대중교통 노선을 중심으로 해당 노선과 주변 지역을 연계하는 일종의 라스트 마일(last mile) 서비스 형태이다. 탄력적 정류장 정차형은 운영시간대에 따라 정차하는 정류장을 다르게 하는 것이다. 이와유사하게 첨두시간대 수요에 집중적으로 대응하는 첨두시간대 수요응답형 대중교통 서비스의 경우, 기존 도시 내대중교통을 보완하며 대중교통의 이용 편의성을 높일 수 있다는 점에서 많은 연구가 이뤄지고 있다(Choi et al., 2022; Korea Transportation Safety Authority, 2015; Moon et al., 2021).

Qiu et al.(2019)은 첨두시간대 수요응답형 방식의 일환인 대도시 내 맞춤형 통근버스의 도입을 위한 준비 방안으로써 교통카드 데이터에 기반하여 유사 통행을 클러스터링하였다. Moon et al.(2021)도 교통카드 데이터에 기반하여 유사 통행을 군집화했는데, 수요응답형 서비스의 효과를 평가하고 서비스 도입 기준을 수립하기 위해서였다. 이같이 기존 교통카드 데이터에 기반하여 잠재적인 수요응답형 서비스의 후보 노선을 도출하기 위한 연구도 많은 한편, 기존 고정형 대중교통 방식과 여러 수요응답형 대중교통 방식 중에서 적합한 방식을 택하는 방법론에 관한 연구도 많다. Choi et al.(2022)는 교통카드 데이터와 택지개발지구 면적, 도시철도 개통 유무 등을 바탕으로 군집화를 수행하여 첨두시간대 수요응답형 방식을 포함한 4가지 대안에 대하여 도농복합도시 내 세부 지역별로 적합한 것을 선정하였다. Li and Quadrifoglio(2010)과 Quadrifoglio and Li(2009)는 기존 대중교통과 지선노선 서비스형 중에서 도보시간 및 대기시간, 차내통행시간, 수요의 양을 고려하여 적절한 운영 방식을 선택하는 방법론에 관하여 연구하였다.

앞선 연구들처럼 기존 대중교통 이용 행태에 기반하여 통행을 군집하고, 대중교통 운영 제반사항을 도출하기 위한 많은 노력이 이뤄져 오고 있다. 하지만 시간의 흐름과 이동 방향에 따른 수요의 발생 패턴 특히, 잠재수요를 최대

SEO, Sungju · KIM, Jinhee Article

한 포함한 이동 행태를 고려한 연구는 아는 한에서 부족하다. 이에 더해 노선 경로의 유사성과 지역적 특성도 함께 연관 지어 종합적으로 고려하는 연구는 추가로 이뤄질 필요가 있다. 이에 본 연구는 수요의 시공간적 특성과 사회경 제적 특성을 종합적으로 고려하여 데이터 군집화를 통해 최적의 대중교통 운행 방식을 선택하는 방법론을 정립하고 자 하였다. 다만, 본 연구는 세부 운영 제반사항 결정 전 대략적으로 적합한 우선순위를 도출하는 단계라는 점에서 유동적인 노선과 시간표를 통하여 불규칙하거나 적은 수요에 대응하는 방식 모두를 수요응답형 대중교통 방식으로 정의하고, 고정형 대중교통 방식은 대형 차량으로 많은 수요에 대응하는 방식으로 정의한다.

연구 방법 및 데이터

본 연구는 궁극적으로 지역별, 시간대별 수요 특성에 맞는 대중교통 운행 방식을 도출하고자 한다. 본 연구에서 다루고 있는 대중교통 운행 방식은 크게 고정형 방식과 수요응답형 방식으로 구분된다. 전자는 배차 스케줄 및 노선이 사전에 정의되어 실제 운영 당시에는 변화 없이 고정되어 있는 방식을 의미한다. 일반적으로, 한 번에 많은 인원을 수송하고, 고정된 스케줄 아래 높은 정시성과 신뢰성을 확보하는 데 초점이 맞춰져 있다. 반면, 후자는 배차 스케줄 및 노선이 실시간으로 수집되는 수요에 발맞추어 상대적으로 유연하게 변화하는 방식을 의미한다.

1. DBSCAN

지역별, 시간대별 적합한 대중교통 운행 방식 선정을 위해서는 우선 기존의 지역별 및 시간대별 통행 수요가 파악될 필요가 있다. 이를 위해 통행 데이터를 군집화하여 기존의 통행 행태를 그룹화하고자 하였다. DBSCAN(Density-Based Spatial Clustering of Applications with Noise)은 데이터 사이의 밀도에 기반하여 군집화를 수행하는 대표적인 방법으로, 클러스터의 개수를 사전적으로 정의해 줄 필요도 없고, 노이즈에 따른 영향력도 미미하다는 장점이 있어 많은 연구에서 활용된다(Pedregosa et al., 2011; Qiu et al., 2019). 이에 본 연구에서도 DBSCAN을 활용하여통행 데이터를 군집화함으로써 통행 패턴을 도출하고자 한다.

DBSCAN은 충분히 높은 밀도의 데이터를 하나의 클러스터로 구분하고, 이에 속하기 어려운 데이터는 노이즈로 구분해낸다. 예를 들어, 각각의 데이터를 하나의 점 단위로 표현하였을 때, 해당 점 데이터 사이의 거리를 바탕으로 최대 인접 거리 내에 속하는 점들끼리는 동일한 클러스터에 속한다고 추정한다. 주변의 인접 데이터로 영역을 점진 적으로 확장해나며 클러스터를 구성하는 셈이다. 그리고 그렇게 묶인 점들이 사전에 정의된 최소 점 개수보다 많은 경우, 하나의 클러스터로 구분해내는 방식이다. 이런 점에서 사전에 클러스터의 개수를 정의할 필요가 없고, 노이즈 데이터로부터의 영향도 적다는 이점이 있다. 그러나 여전히 클러스터 내 포함되어야 하는 최소 데이터 개수와 동일한 클러스터 내 인접 데이터 사이의 최대 거리를 연구자가 정의해주어야 한다.

본 연구는 각 기종점쌍에 해당하는 데이터 사이의 유사도를 거리 단위로 처리하여 클러스터링을 수행하였다. 다시 말해, Figure 2와 같이 기종점쌍별 다차원 데이터에 대하여 서로 간의 거리를 계측하고, 이를 바탕으로 군집화를 수행하였다.

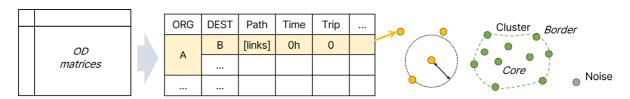


Figure 2. DBSCAN based on origin-destination pair data

2. 데이터별 거리 지표

클러스터링의 입력 데이터로 활용할 특성(Feature)은 통행 패턴과 관련성이 높은 변수로 선정하고자 하였다. 크게 공간적 특성과 시간적 특성, 사회경제적 특성 3가지로 구분하였다. 우선 공간적 특성의 일환으로, 본 연구의 대상이 대중교통이라는 점에서 여러 인원을 합승시킬 수 있도록 기종점쌍별로 경로가 유사한지 여부를 고려하고자 하였다. 또, 시간대별로 이동방향의 경향성이 다르다는 점에서 기종점쌍별 이동 방향성을 구분하여 데이터를 대입하고자 하였다. 다음으로 시간적 특성으로, 통행 수요가 정기적으로 발생하는지에 따라 적합한 대중교통 운행 방식이 상이하기에 통행의 발생 빈도와 주기성을 반영하고자 하였고, 특히, 일반적으로 주중인지 주말인지 여부와 시간대별 발생하는 통행 패턴이 상이하다는 점도 고려하고자 하였다(Ishii et al., 2022; Liu et al., 2023a; Liu and Cheng, 2020; Nakanishi et al., 2018; Wan et al., 2021; Wilson, 2008). 마지막으로, 통행이 사회경제 활동에 따른 파생수요라는 점을 고려하여 사회경제 활동과 관련된 변수도 반영하고자 하였고, 이외에도 인구구성, 고용여부, 교육수준, 자동차 보유 여부 등이 통행패턴 및 통행수단 선택에 영향을 미친다는 점을 고려하고자 하였다(Babu and Anjaneyulu, 2021; Dean and Kockelman, 2021; Fekih et al., 2022; Jiang et al., 2023; Rafiq and McNally, 2021)

DBSCAN을 수행하기 위해서는 각 데이터 사이의 거리(유사도) 지표가 정립될 필요가 있다. 일반적으로 유클리드 거리를 많이 활용하지만, 본 연구에서는 수요의 공간적, 시간적, 사회경제적 등 3가지 특성을 종합적으로 고려해야 한다는 점에서 다양한 변수가 함께 고려할 수 있는 Gower distance(Gower's dissimilarity coefficient)를 활용하였다. Gower distance는 양적 변수와 질적 변수 모두를 고려하여 유사도를 계산할 수 있는 지표로, 각각의 변수마다거리 지표를 정립하여 0부터 1 사이의 값을 계산하고, 해당 값을 가중평균하는 형태이다. 예를 들어, 질적 변수 (x_i, x_j) 의 유사도는 Equation 2와 같을 때, Gower distance는 Equation 3과 같이 계산될 수 있다(Gower, 1971; Tuerhong and Kim, 2014).

$$s_{ijk} = \begin{cases} 0, & x_i \text{ and } x_j \text{ are same in the } k^{th} \text{ character} \\ 1, & \text{otherwise} \end{cases}$$
 (1)

$$s_{ijk} \in [0, 1]$$
where
$$\begin{cases} 0, & x_i \text{ and } x_j \text{ are same in the } k^{th} \text{ character} \\ \text{maximum, } x_i \text{ and } x_j \text{ are at opposite ends of their range in the } k^{th} \text{ character} \end{cases}$$

$$s_{ij} = \sum_{k=1}^{v} s_{ijk} w_k(x_{ik}, x_{jk}) / \sum_{k=1}^{v} \delta_{ijk} w_k(x_{ik}, x_{jk})$$
where
$$\begin{cases} \delta_{ijk} \text{ is 1 when character } k \text{ can be compared for } i \text{ and } j, \text{ and 0 otherwise} \\ w_k(x_{ik}, x_{jk}) \text{ is the weight function for character } k \text{ that differs by character} \end{cases}$$

Gower distance를 계산하기 위해서 특성별 거리 지표도 정의하였다. 우선 공간적 특성인 기종점 방향별 이동 경로는 Dis-Commonality Factor(DCF) 지표를 채택하였다. 통행 패턴을 클러스터링하는데, 이동 방향성과 이동 경로 사이의 유사성을 반영하고자 하였기에 두 경로 사이의 유사도를 계산할 수 있는 지표를 활용하였다. DCF는 두 경로를 링크 단위로 구분하고, 두 경로의 전체 길이 대비 서로 겹치는 링크 부분의 비중을 산정하는 형태로, Equation 4와 같이 계산할 수 있다(Advani et al., 2022).

$$s_{ijk} = DCF = 1 - \frac{\delta l_{ij}}{\sqrt{l_i l_j}} \in [0, 1]$$
 (4)

where $\{\partial_{ij}$ is the sum of the length of the common links between path i and j

시간적 특성으로써 요일별, 시간대별 수요의 발생패턴 사이의 유사도를 계산하기 위하여 Dynamic Time Warping(DTW)을 활용하였다. DTW는 두 시계열 데이터의 패턴 유사성을 산정해내는 지표이다. 유클리드 거리처럼 그 절대적인 양의 차이를 비교하지만, 유클리드 거리가 동일 시점에 대해서만 비교하는 것과 달리 DTW는 Equation 5와 같이 비교 대상 시점 기준 인접한 주변 시점까지도 비교한다(Giorgino, 2009; Meert et al., 2020; Tormene et al., 2009). 여러 인원을 한 번에 수송한다는 점을 고려하고자 비슷한 시간대에 비슷한 인원 규모를 지닌 그룹으로 나누기 위하여 DTW를 활용하였다.

$$s_{ijk} = DTW(X_i, X_j) = \min d_\phi(X_i, X_j) \in [0, \infty]$$

$$\begin{cases} d_\phi(X_i, X_j) = \sum_{k=1}^T d(\phi_{x_i}(k), \phi_{x_j}(k)) m_\phi(k) / M_\phi \\ X_i, X_j \text{ are time series} \\ d(p, q) \text{ is local dissimilarity between any pair of elements } x_{ip} \text{ and } x_{jq} \\ \text{ for time index } p \text{ and } q \\ \phi(k) \text{ is warping curve for time index } k \\ m_\phi(k) \text{ and } M_\phi \text{ is a per-step weighting coefficient} \\ \text{ and the corresponding normalization constant} \end{cases}$$

사회경제적 특성은 가장 대표적인 거리 함수인 유클리드 거리(ED; Euclidean Distance)를 활용하였으며, 계산식은 Equation 6과 같다.

$$s_{ijk} = ED = \sum_{b=1}^{n} \sqrt{x_{ib} - x_{jb}} \in [0, \infty]$$

$$(6)$$

요약하면, 본 연구는 Figure 3과 같이 공간적 특성에 해당하는 경로 사이의 유사도는 DCF로, 시간적 특성에 해당하는 통행 수요의 시계열적 패턴 사이의 유사도는 DTW로, 사회경제적 특성에 해당하는 여러 데이터의 유사도는 ED로 계측하였다. 각 데이터별로 계산한 유사도 값을 0과 1 사이의 범위로 스케일링하기 위하여 일부 데이터는 Min-Max Scaling을 수행하였다. 값의 범위가 통일된 거리 지표들을 바탕으로 여러 기종점쌍 사이의 Gower distance를 계산하고, 이를 바탕으로 DBSCAN을 수행하여 통행 패턴을 군집화하였다.

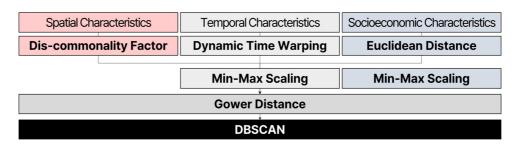


Figure 3. Research methodology flow-chart

3. 사용 데이터

본 연구에서 사용한 데이터는 크게 네트워크 데이터, 통행 데이터, 사회경제 데이터이다. 그리고 본 연구의 시간적 범위는 구득한 데이터별로 상이하다. 본 연구의 공간적 범위는 서울특별시 서대문구로, 내부 통행만을 연구 대상으로 하였으며, 공간적 최소 분석 단위는 행정동 단위이다. 서울특별시 서대문구는 2021년 기준 연평균 40,521건의 내부 통행이 시내버스로 이뤄졌으나, 여전히 수요응답형 방식과 가까운 택시를 통한 통행이 연평균 5.121건이나 이

뤄져 고정형과 수요응답형 2가지 방식 모두가 필요한 지역으로 추정되었다(Seoul, 2022). 또, 행정동별 거주인구 분포와 사업체 종사인구 분포가 다양하며, 65세 이상 인구가 해당 행정동 인구 대비 15%에 지나지 않는 지역과 함께 과반이 넘는 지역까지도 있다. 뿐만 아니라, 지역 내 대학가도 자리하고 있어 다양한 내부 통행 패턴이 발생할 것으로 추정되었다(Seodaemun-gu, 2022). 이러한 점에서 서울특별시 서대문구를 본 연구의 공간적 범위로 선정하였다.

네트워크 데이터는 국가교통데이터베이스(Korea Transport Database, 2020)에서 제공하는 2019년 수도권 네트워크 데이터를 활용했으며, 이를 바탕으로 행정동쌍별 최단 경로를 도출하여 수요의 공간적 특징인 두 기종점쌍경로 사이의 유사도를 계산하였다. 공간적 분석 범위인 서울특별시 서대문구의 네트워크 현황은 Figure 4와 같다. 통행 데이터로는 2022년 11월 기준 서울 생활이동 데이터를 활용하였다. 서울 생활이동 데이터는 특정 시점, 특정 지역 간에 이동하는 서울의 모든 인구를 추정한 것으로, 특정 통신사의 휴대전화 신호에 기반하여 수집 및 가공된 데이터이다. 통신 데이터에 기반하고 있어 통행수단 구분 없이 모든 통행이 집계된 내역이고, 특정 통신사만을 대상으로 수집한 것이다 보니 해당 통신사의 점유율을 바탕으로 모집단에 대한 추정이 이뤄진 정보이다(Seoul, 2023). 요일별 생활이동 인구수의 분포를 살펴본 결과는 Figure 5와 같으며, 선 굵기가 굵을수록 생활이동 인구수가 많다는 의미이다.

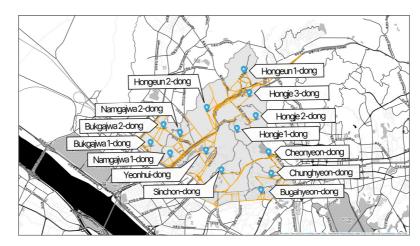


Figure 4. Administrative dong and network of Seodaemun-gu, Seoul

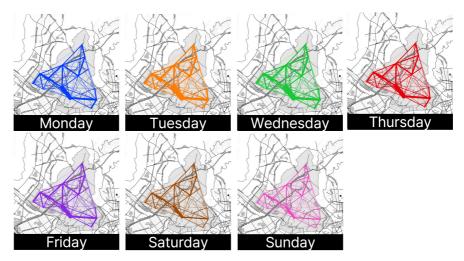


Figure 5. Seoul moving population by day of the week

사회경제 변수 관련 데이터로는 통행 패턴에 영향을 끼칠 수 있는 변수로써 지역의 사회경제적 특성을 대표할 수 있는 데이터를 주로 활용하고자 하였다. 첫 번째로, 행정동별 인구 정보로써 2022년 12월 31일을 기준 주민등록 전산망 기반 주민등록 인구수 데이터를 구득하였다(Figure 6). 두 번째로, 행정동별 사업체 종사자수로써 서울특별시산업체 조사 보고서 기반 2021년 12월 31일 기준 농림어업, 광업 등 19개 모든 산업분야에서 종사하고 있는 인구수에 관한 데이터를 구득하였다(Figure 7). 세 번째로, 행정동별 전체 사업체수 대비 도매업 및 소매업 사업체수의 비중과 교육서비스업 사업체수 비중도 서울특별시 산업체 조사 보고서를 기반으로 2021년 12월 31일 기준 데이터를 구득하였다(Figures 8, 9). 마지막으로, 통행 패턴에 영향을 미칠 수 있는 요소인 자가용 보유 여부에 대한 정보를 반영하기 위하여 행정동별 자동차 등록대수와 주차장 면수 정보를 수집하였다. 전자는 2021년 12월 기준 자동차 관리정보 시스템상 자동차 등록현황 데이터를 구득하였고, 후자는 2020년 기준 서대문 통계연보 데이터를 구득하였다 (Figures 10, 11).

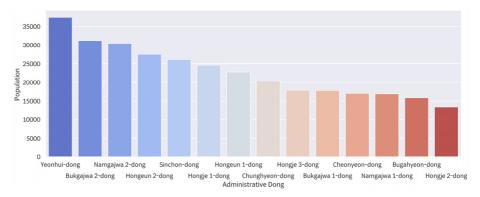


Figure 6. Population by administrative dong in 4Q, 2022

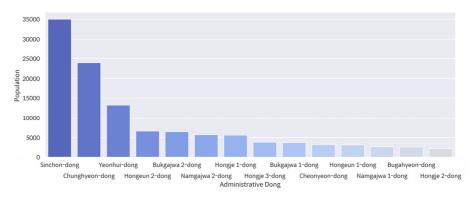


Figure 7. Number of workers by administrative dong in 2021

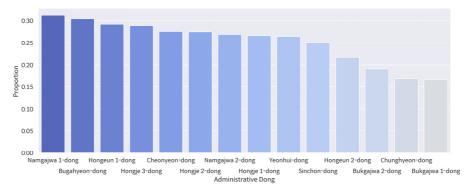


Figure 8. Percentage of wholesale and retail businesses by administrative dong in 2021

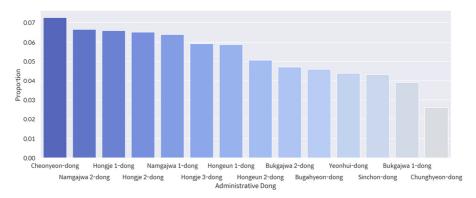


Figure 9. Percentage of educational service businesses by administrative dong in 2021

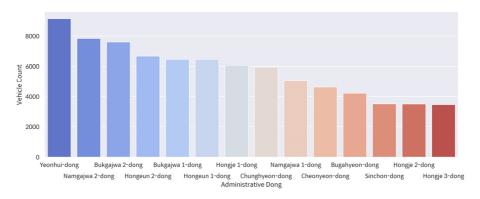


Figure 10. Number of cars by administrative dong in December, 2021

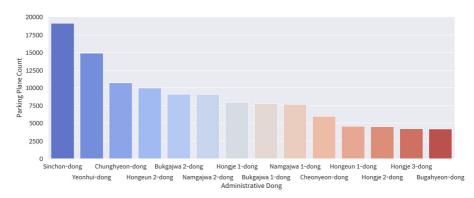


Figure 11. Number of parking planes by administrative dong in 2020

분석 결과

본 연구에서 Gower distance 가중치는 가중치 변화에 따른 클러스터링 결과를 바탕으로 선정하였다. 공간적 특성에 0.64, 시간적 특성에 0.3, 사회경제적 특성 각각에 0.01을 부여하여 실루엣 지수도 적당히 높으면서 클러스터도 적당한 개수로 나눠지도록 하였다(Equation 7). 이때, 실루엣 지수란, 각 개체가 자신이 속한 군집과 다른 군집에비해 얼마나 유사한지를 나타내는 지표이다. 시간적 특성 가중치를 높이면, 실루엣 지수는 커지지만, 클러스터가 단한 개만으로 수렴하는 것으로 나타났다. 이에 반해, 공간적 특성 가중치를 높이면, 실루엣 지수도 커지지만, 클러스터 수도 늘어나는 것으로 나타났다.

Gower distance =
$$DCF \times 0.64 + DTW \times 0.3 + \sum_{k=1}^{6} (ED_k \times 0.01)$$
 (7)

DBSCAN을 통하여 통행 패턴을 군집화하기 위해 클러스터 내 최소 표본 개수를 4개, 거리 임계치를 0.25로 설정하였다. 일반적으로 거리 임계치는 k-distance plot에서 팔꿈치 지점 값을 채택하는데, 즉, k번째로 인접한 데이터 까지의 거리가 급격하게 증가하는 지점의 값을 선택한다(Sander et al., 1998). 본 연구에서는 가장 인접한 4번째 데이터까지의 거리를 짧은 순으로 나열한 Figure 12를 바탕으로 0.25로 설정하였다.

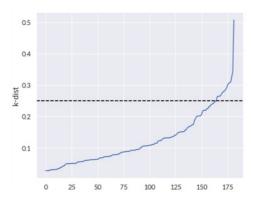


Figure 12. k-distance plot

서울특별시 서대문구 내 14개 행정동에 대하여 총 182개 기종점쌍을 클러스터링한 결과, 7개의 클러스터와 나머지 29개의 노이즈 데이터로 구분되었다. 클러스터별로 속하는 데이터의 개수는 Figure 13과 같고, 클러스터별로 추정 출발시간대별 평균 이동 인구수는 Figure 14와 같다. 클러스터 6 및 7은 다른 클러스터에 비해 기종점쌍의 개수도 적고, 절대적인 통행량도 적은 것으로 나타났다. 이와 반대로, 노이즈는 다른 클러스터에 비해 기종점쌍의 개수도 많고, 절대적인 통행량도 월등히 많은 편인 것으로 드러났다. 출발시간대별 통행량의 분포 관점에서는 클러스터 1 및 4는 오전에 상대적으로 통행이 많았다가 오후에 줄어드는 패턴을 보이는 반면, 클러스터 2 및 3은 반대로 오전에 통행이 적었다가 오후에 늘어나는 패턴을 보인다. 이와 달리, 클러스터 5와 노이즈는 오전과 오후 둘 다 통행이 많은 시점이 존재하나, 그 시간이 오전은 짧은 반면, 오후는 길게 퍼져 있다. 이동 방향 측면에서는 클러스터 2는 신촌동과 충현동 등을 기점으로 하는 통행이 다수인 반면, 클러스터 4는 신촌동과 충현동 등으로 향하는 통행이 다수로 정반대의 특성을 보인다(Figure 15). 한편, 노이즈에 해당하는 기종점쌍의 경우, 서로 군집화되지 못할 정도로 절대적인 통행량이 다른 기종점쌍에 비해 훨씬 많거나 공간적으로 특정 도로축을 공유하지 않는 것으로 추정된다.

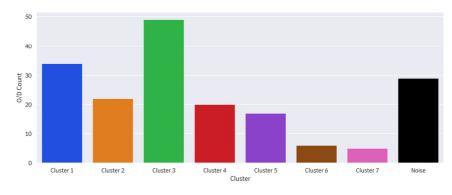


Figure 13. Number of origin-destination pairs per cluster

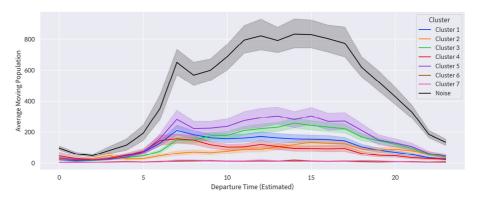


Figure 14. Number of people moving by departure time and cluster

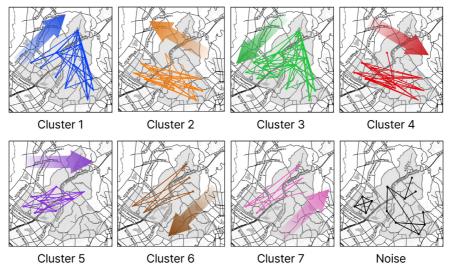


Figure 15. Spatial distribution of origin-destination pairs by cluster

클러스터별 기종점 지역의 사회경제적 특성을 총 주민등록 인구와 사업체수, 자동차 등록대수 관점에서 분석할 결과는 Figure 16과 같다. 클러스터 1과 4는 주민등록인구가 많은 지역에서 사업체가 많은 지역으로의 통행 즉, 주 거 지역에서 상업업무 지역으로의 통행으로 추정된다. 반대로 클러스터 2와 3은 사업체가 많은 지역에서 주민등록 인구가 많은 지역으로의 통행이다. 클러스터 5는 주민등록인구가 많은 지역 사이의 통행이고, 클러스터 6과 7은 주민등록인구가 중간인 지역 사이의 통행이다. 마지막으로 노이즈에 해당하는 기종점쌍은 주민등록인구도 많고, 사업체도 많은 지역 간의 통행이다.

앞선 클러스터별 특성을 종합해보면, 클러스터 1과 4는 출퇴근 및 등하교 통행과 같이 주거 지역으로부터의 유출 통행이라고 볼 수 있고, 클러스터 2와 3은 귀가 통행과 같이 주거 지역으로의 유입 통행이라 볼 수 있다.

클러스터끼리 상호 비교를 수행한 결과, 클러스터 6과 7 모두 주중과 주말 구분 없이 유사한 패턴을 보이며, 오전과 오후 구분 없이 매우 불규칙적으로 통행이 발생하면서도 절대적인 양 측면에서는 많지 않은 것으로 나타났다 (Figures 17, 18). 또, 두 클러스터는 서로 정반대의 이동 방향성을 띄지만, 각 클러스터 내 속하는 기종점쌍끼리는 공간적으로 유사한 경로를 공유하는 것으로 나타났다(Figure 15). 이러한 점을 모두 고려하였을 때, 대중교통 운행 방식 중 수요응답형 대중교통 서비스가 바람직한 것으로 판단된다. 수요응답형 대중교통 서비스는 불규칙하면서도 양이 적은 수요까지도 대응할 수 있다는 이점을 지니고 있기 때문이다.

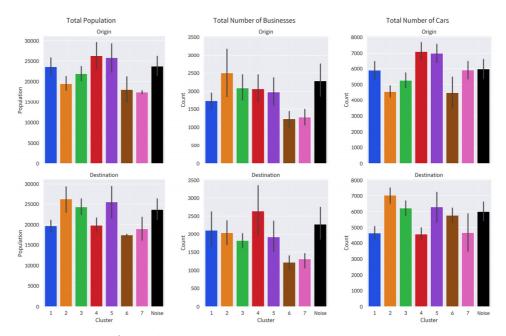


Figure 16. Average socioeconomic characteristics by cluster

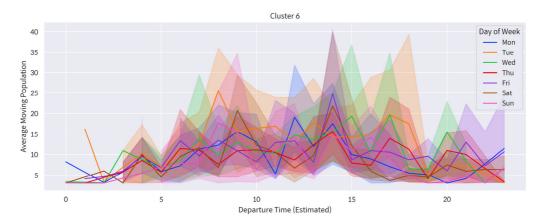


Figure 17. Number of people moving by departure time and day of week (Cluster 6)

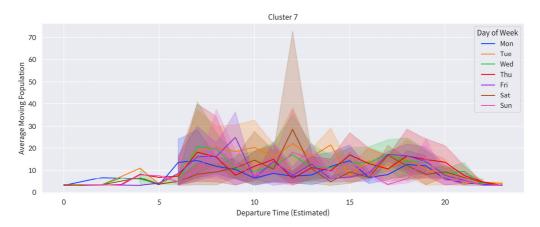


Figure 18. Number of people moving by departure time and day of week (Cluster 7)

노이즈에 해당하는 기종점쌍의 경우에는, 공간적으로 다양한 지역에 넓게 퍼져 있다(Figure 15). 통행량은 주중 대비 주말에 절대적인 통행량이 극명하게 적지만, 주중과 주말을 구분하여 살펴본다면, 요일 구분 없이 반복적인 패턴을 보인다는 것을 확인할 수 있다(Figure 19). 주중에만 유독 오전에 잠깐 많은 통행량을 보이다가 오후에 다시 긴시간동안 많은 통행량을 보이는 것으로 나타났다. 절대적인 통행량은 클러스터 6과 7등 다른 클러스터에 비해 월등히 많은 편인 것으로 드러났다(Figure 14). 이에 더해 노이즈에 해당하는 기종점쌍 지역의 사회경제적인 특성이 주민등록인구와 사업체수 모두 많은 지역이었다는 점까지 모두 종합하면, 해당 통행은 고정형 노선 및 스케줄로 대응하는 것이 바람직하다. 그 이유는 주중과 주말을 구분하면, 오전 및 오후 모든 시간대에 대하여 계속해서 통행 수요가 풍부하게 발생하고, 요일별로 매우 반복적인 패턴을 보이기 때문이다. 다만, 해당 기종점쌍들이 노이즈인만큼 하나로 묶어 고정형 노선을 투입하는 형태보다는 각 기종점쌍마다 개별 고정형 노선을 투입하는 것이 바람직하다.

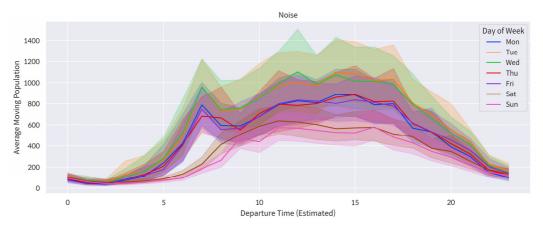


Figure 19. Number of people moving by departure time and day of week (Noise)

결론

본 연구는 각기 다른 장단점을 지닌 대중교통 서비스 운행 방식인 고정형과 수요응답형 방식 중 지역별, 시간대별 적합한 방식을 파악하고자 하였다. 고정형 방식은 배차시간표 및 노선이 고정되어 높은 정시성과 신뢰성이 확보되고, 주로 대형 버스나 철도 차량을 활용하여 운영됨에 따라 한 번에 많은 인원을 수송할 수 있다는 장점이 있다. 반면, 수요응답형 방식은 배차시간표 및 노선이 실시간으로 변화하는 수요에 발맞춰 상대적으로 유연하게 대응하는 방식으로, 주로 소형 버스나 승합차로 운영됨에 따라 효율적으로 인원을 수송한다. 더욱이, 수요가 일정치 않거나 그 양이 적은 지역까지도 대응할 수 있어 국민의 이동권을 보장하는 데 이바지할 수 있다. 심지어 Door-to-Door 형태로 운영되어 버스 정류장이나 철도 역사까지의 접근시간을 단축시킬 수 있어 이용자의 편의성도 증대할 수 있다는 장점이 있다. 이처럼 두 운행 방식이 각기 다른 형태 및 강점을 지니고 있어 본 연구는 통행 데이터를 포함한 각종 데이터를 활용하여 통행 패턴을 군집화하고, 이에 대한 해석을 바탕으로 바람직한 대중교통 운행 방식을 도출하는 방법론을 제안한다. 그리고 실제 서울특별시 서대문구 지역을 대상으로 해당 방법론을 적용하여 분석을 수행하였다. 입력 데이터는 통행 수요의 특성을 시간적 특성, 통행이 실제로 행해지는 공간의 특성, 그리고 통행이 발생하는 지역의 사회경제적 특성으로 나누어 대입하였다. 데이터별 특성을 종합적으로 고려하여 데이터 간 유사도를 계측할 수 있는 Gower distance 지표에 기반하여 데이터 밀도에 따라 클러스터를 구성하고 노이즈를 구분해내는 DBSCAN을 수행하였다.

이 과정에서 Gower distance 가중치를 선정하기 위하여 가중치 변화에 따른 클러스터링 결과를 살펴보았다. 그 결과, 시간적 특성 가중치를 높이면, 군집화 정도를 나타내는 지표인 실루엣 지수는 커지지만, 클러스터가 단 하나로 수렴하는 형태를 보였다. 반대로, 공간적 특성 가중치를 높이면, 실루엣 지수와 클러스터 수 모두 늘어났다. 이러

SEO, Sungju · KIM, Jinhee Article ·

한 점에서 시간적 특성뿐만 아니라, 공간적 특성 즉, 기반 환경에 관한 정보가 모두 고려되어야 적절히 클러스터링되는 것으로 나타났다.

DBSCAN 결과, 총 14곳의 행정동으로 이루어진 서울특별시 서대문구의 총 182개의 기종점쌍은 7개의 클러스터와 29개의 노이즈로 구분되었다. 클러스터별 사회경제적 지표와 통행 패턴을 바탕으로 라벨링을 수행한 결과, 특정 클러스터에 해당하는 기종점쌍은 절대적인 통행량도 적고, 통행이 발생하는 패턴 또한 불규칙적인 통행 패턴인 반면, 노이즈에 해당하는 기종점쌍은 전반적인 시간대에 대하여 꾸준히 통행량이 많고, 절대적인 통행량도 많으며, 기종점 지역 자체의 거주인구 및 사업체 종사인구도 많은 통행 패턴으로 드러났다. 전자는 통행량이 적으면서도 발생패턴이 불규칙하다는 점에서 수요응답형 운행 방식이 적합하다고 판단하였고, 후자는 절대적인 통행량이 충분히 많고 그 정도가 시간의 흐름에 따라 큰 변화를 보이지 않는다는 점에서 고정형 운행 방식이 적절하다고 결론지었다.

다만, 본 연구는 일부 한계점을 지니고 있다. 첫째, 이미 드러난 데이터에 기반한 분석 결과이기에 잠재수요에 관해서는 파악이 어렵다. 둘째, 본 연구에서 활용한 데이터의 기준 시기가 서로 상이하다. 셋째, 데이터별 거리 지표를 Gower distance로 종합하는 과정에서 각 지표 특성이 충분히 고려되지 않았을 수 있다. 넷째, 본 연구 분석 사례의 대상지였던 서울특별시 서대문구는 상대적으로 행정동 사이의 이질성이 낮은 편이라 다소 부적합했을 수 있다. 더욱이 수요응답형 방식이 중소 도시 및 농어촌 지역을 중심으로 전개되고 있는 만큼 그러한 지역을 대상으로 한 추가연구가 필요하다. 특히, 농어촌 지역의 경우, 단순 주중 및 주말 구분이 아닌, 장날에 따른 통행 패턴도 상이하다는점을 고려한 연구도 필요하다. 다섯째, 구체적인 운영 제반사항 및 관련 기술에 따라 수요응답형 방식도 수요가 많은지역에서 높은 효율성을 보일 수 있는 등 세부 조건에 따라 우선순위가 바뀔 수 있다. 그럼에도 불구하고 본 연구는데이터 기반 통행 패턴 군집을 바탕으로 세부 조건 결정 전에 대략적으로 적절한 대중교통 운행 방식을 도출할 수 있는 방법론을 제시함으로써 더 나은 의사결정 방향을 제시했다는 점에서 의의가 있다.

Funding

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport(Grant 23AMDP-C161756-03).

알림

본 논문은 제89회 대한교통학회 학술발표회(2023. 10. 12.)에서 발표된 내용을 수정·보완하여 작성된 것입니다.

ORCID

SEO, Sungju http://orcid.org/0000-0002-9005-789X KIM, Jinhee http://orcid.org/0000-0003-3436-6652

REFERENCES

Abdelwahed A., Van Den Berg P. L., Brandt T., Ketter W. (2023), Balancing Convenience and Sustainability in Public Transport Through Dynamic Transit Bus Networks, Transportation Research Part C: Emerging Technologies, 151, 104100.

Advani C., Bhaskar A., Haque Md. M. (2022), Bi-level Clustering of Vehicle Trajectories for Path Choice set and its Nested Structure Identification, Transportation Research Part C: Emerging Technologies, 144, 103895.

Atasoy B., Ikeda T., Song X., Ben-Akiva M. E. (2015), The Concept and Impact Analysis of a Flexible Mobility on Demand System, Transportation Research Part C: Emerging Technologies, 56, 373-392.

- Babu D., Anjaneyulu M. V. L. R. (2021), Exploratory Analysis on Worker's Independent and Joint Travel Patterns During Weekdays and Weekends, Transportation Engineering, 5, 100073.
- Basu R., Araldo A., Akkinepally A. P., Nahmias Biran B. H., Basak K., Seshadri R., Deshmukh N., Kumar N., Azevedo C. L., Ben-Akiva M. (2018), Automated Mobility-on-Demand vs. Mass Transit: A Multi-Modal Activity-Driven Agent-Based Simulation Approach, Transportation Research Record: Journal of the Transportation Research Board, 2672(8), 608-618.
- Chen Y., Zhao Y., Tsui K. L. (2019), Clustering-based Travel Pattern Recognition in Rail Transportation System Using Automated Fare Collection Data, 2019 Prognostics and System Health Management Conference (PHM-Qingdao), 1-7.
- Choi J. H., Kang M. H., Song J. I., Hwang K. Y. (2022), A Study on Introduction of Demand Responsive Transport(DRT) Rural-Urban Mixed Area, J. Korean Soc. Transp., 40(3), Korean Society of Transportation, 289-304.
- Currie G., Fournier N. (2020), Why most DRT/Micro-Transits Fail: What the Survivors Tell Us about Progress, Research in Transportation Economics, 83, 100895.
- Cyril A., Mulangi R. H., George V. (2020), Demand-Based Model for Line Planning in Public Transport, Transportation Research Procedia, 48, 2589-2596.
- Dean M. D., Kockelman K. M. (2021), Spatial Variation in Shared Ride-Hail Trip Demand and Factors Contributing to Sharing: Lessons from Chicago, Journal of Transport Geography, 91, 102944.
- Estrada M., Mension J., Salicrú M. (2021), Operation of Transit Corridors Served by Two Routes: Physical Design, Synchronization, and Control Strategies, Transportation Research Part C: Emerging Technologies, 130, 103283.
- Fekih M., Bonnetain L., Furno A., Bonnel P., Smoreda Z., Galland S. et al. (2022), Potential of Cellular Signaling Data for Time-of-Day Estimation and Spatial Classification of Travel Demand: A Large-Scale Comparative Study with Travel Survey and Land use Data, Transportation Letters, 14(7), 787-805.
- Forouzandeh F., Arman H., Hadi-Vencheh A., Masoud Rahimi A. (2022), A combination of DEA and AIMSUN to Manage Big Data When Evaluating the Performance of Bus Lines, Information Sciences, 618, 72-86.
- Giorgino T. (2009), Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package, Journal of Statistical Software, 31(7), 1-24.
- Gower J. C. (1971), A General Coefficient of Similarity and Some of Its Properties, Biometrics, 27(4), 857-871.
- Hamedmoghadam H., Vu H. L., Jalili M., Saberi M., Stone L., Hoogendoorn S. (2021), Automated Extraction of Origin-Destination Demand for Public Transportation from Smartcard Data with Pattern Recognition, Transportation Research Part C: Emerging Technologies, 129, 103210.
- Ishii Y., Hayakawa K., Koide S., Chikaraishi M. (2022), Entropy Tucker Model: Mining Latent Mobility Patterns with Simultaneous Estimation of Travel Impedance Parameters, Transportation Research Part C: Emerging Technologies, 137, 103559.
- Jiang Z., Tang Y., Gu J., Zhang Z., Liu W. (2023), Discovering Periodic Frequent Travel Patterns of Individual Metro Passengers Considering Different Time Granularities and Station Attributes, International Journal of Transportation Science and Technology.
- Kepaptsoglou K., Karlaftis M. (2009), Transit Route Network Design Problem: Review, Journal of Transportation Engineering, 135(8), 491-505.

König A., Grippenkoven J. (2020), Travellers' Willingness to Share Rides in Autonomous Mobility on Demand Systems Depending on Travel Distance and Detour, Travel Behaviour and Society, 21, 188-202.

- Korea Transport Database (2020), 2019 Traffic Analysis Data in Seoul Metropolitan Area: O/D and Network Data Manual, KTDB.
- Korea Transportation Safety Authority (2015), A Study on the Establishment and Operation of Demand Response Transportation (DRT) System.
- Li X., Quadrifoglio L. (2010), Feeder Transit Services: Choosing Between Fixed and Demand responsive Policy, Transportation Research Part C: Emerging Technologies, 18(5), 770-780.
- Li Z., Yan H., Zhang C., Tsung F. (2022), Individualized Passenger Travel Pattern Multi-Clustering Based on Graph Regularized Tensor Latent Dirichlet Allocation, Data Mining and Knowledge Discovery, 36(4), 1247-1278.
- Liu S., Zhang F., Ji Y., Ma X., Liu Y., Li S. et al. (2023a), Understanding Spatial-Temporal Travel Demand of Private and Shared E-Bikes as a Feeder Mode of Metro Stations, Journal of Cleaner Production, 398, 136602.
- Liu Y., Zuo X., Ai G., Zhao X. (2023b), A Construction-and-Repair Based Method for Vehicle Scheduling of Bus Line with Branch lines, Computers & Industrial Engineering, 178, 109103.
- Liu Y., Cheng T. (2020), Understanding Public Transit Patterns with Open Geodemographics to Facilitate Public Transport Planning, Transportmetrica A: Transport Science, 16(1), 76-103.
- Lu C., Gao L., Huang Y. (2022), Exploring Travel Patterns and Static Rebalancing Strategies for Dockless Bike-Sharing Systems From Multi-Source Data: A Framework and Case Study, Transportation Letters, 1-14.
- Ma J., Yang Y., Guan W., Wang F., Liu T., Tu W. et al. (2017), Large-Scale Demand Driven Design of a Customized Bus Network: A Methodological Framework and Beijing Case Study. Journal of Advanced Transportation, 2017, 3865701.
- Meert Wannes, Hendrickx Kilian, Van Craenendonck Toon, Robberechts Pieter, Blockeel Hendrik, Davis Jesse (2020), DTAIDistance, v2.3.10.
- Moon S.D., Kim D.K., Kho S.Y., Cho S.H. (2021), Performance Measurement and Determination of Introduction Criteria for Peak Demand Responsive Transit Service, J. Korean Soc. Transp., 39(1), Korean Society of Transportation, 100-114.
- Nakanishi W., Yamaguchi H., Fukuda D. (2018), Feature Extraction of Inter-Region Travel Pattern Using Random Matrix Theory and Mobile Phone Location Data, Transportation Research Procedia, 34, 115-122.
- Nourbakhsh S. M., Ouyang Y. (2012), A Structured Flexible Transit System for Low Demand Areas, Transportation Research Part B: Methodological, 46(1), 204-216.
- Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O. et al. (2011), Scikit-learn: Machine learning in Python, Journal of Machine Learning Research, 12, 2825-2830.
- Potts J., Marshall M., Crockett E., Washington J. (2010), A Guide for Planning and Operating Flexible Public Transportation Services.
- Qiu G., Song R., He S., Xu W., Jiang M. (2019), Clustering Passenger Trip Data for the Potential Passenger Investigation and Line Design of Customized Commuter Bus, IEEE Transactions on Intelligent Transportation Systems, 20(9), 3351-3360.

- Quadrifoglio L., Li X. (2009), A Methodology to Derive the Critical Demand Density for Designing and Operating Feeder Transit Services, Transportation Research Part B: Methodological, 43(10), 922-935.
- Rafiq R., McNally M. G. (2021), Heterogeneity in Activity-Travel Patterns of Public Transit Users: An Application of Latent Class Analysis, Transportation Research Part A: Policy and Practice, 152, 1-18.
- Sander J., Ester M., Kriegel H.-P., Xu X. (1998), Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications, Data Mining and Knowledge Discovery, 2(2), 169-194.
- Schofer Joseph L. (2023), Mass Transit. Encyclopedia Britannica.
- Seodaemun-gu (2022), 2022 Seodaemun Statistical Yearbook.
- Seoul (2022), Seoul Transportation in 2021.
- Seoul (2023), Seoul Moving Population Data.
- Tang J., Bi W., Liu F., Zhang W. (2021), Exploring Urban Travel Patterns Using Density-Based Clustering with Multi-Attributes from Large-Scaled Vehicle Trajectories, Physica A: Statistical Mechanics and Its Applications, 561, 3125301.
- Tang X., Yang J., Lin X., He F., Si J. (2023), Dynamic Operations of an Integrated Mobility Service System of Fixed-Route Transits and Flexible Electric Buses, Transportation Research Part E: Logistics and Transportation Review, 173, 103081.
- Tormene P., Giorgino T., Quaglini S., Stefanelli M. (2009), Matching Incomplete Time Series with Dynamic Time Warping: An Algorithm and an Application to Post-Stroke Rehabilitation, Artificial Intelligence in Medicine, 45(1), 11-34.
- Tuerhong G., Kim S. B. (2014), Gower Distance-Based Multivariate Control Charts for A Mixture of Continuous and Categorical Variables, Expert Systems with Applications, 41(4, Part 2), 1701-1707.
- Tuydes-Yaman H., Kocak B., Karagumus E., Dalkic-Melek G. (2023), Boarding Stop Assignment for Public Bus Trips Using Smart Card Data: Comparison of Two Algorithms, Transportation Research Procedia, 69, 472-479.
- Wan L., Tang J., Wang L., Schooling J. (2021), Understanding Non-Commuting Travel Demand of Car Commuters Insights from ANPR Trip Chain Data in Cambridge, Transport Policy, 106, 76-87.
- Wang H., Li J., Wang P., Teng J., Loo B. P. Y. (2023), Adaptability Analysis Methods of Demand Responsive Transit: A Review and Future Directions, Transport Reviews, 43(4), 676-697.
- Wang Y.-J., Kuo Y.-H., Huang G. Q., Gu W., Hu Y. (2022), Dynamic Demand-Driven Bike Station Clustering, Transportation Research Part E: Logistics and Transportation Review, 160, 102656.
- Wilson C. (2008), Activity Patterns in space and Time: Calculating Representative Hagerstrand Trajectories, Transportation, 35(4), 485-499.
- Yu Q., Zhang H., Li W., Song X., Yang D., Shibasaki R. (2020), Mobile Phone GPS Data in Urban Customized Bus: Dynamic Line Design and Emission Reduction Potentials Analysis, Journal of Cleaner Production, 272, 122471.
- Zhao J., Qu Q., Zhang F., Xu C., Liu S. (2017), Spatio-Temporal Analysis of Passenger Travel Patterns in Massive Smart Card Data, IEEE Transactions on Intelligent Transportation Systems, 18(11), 3135-3146.
- Zhao S., Zhao K., Xia Y., Jia W. (2022), Hyper-Clustering Enhanced Spatio-Temporal Deep Learning for Traffic and Demand Prediction in Bike-Sharing Systems, Information Sciences, 612, 626-637.