

POI 메타정보를 활용한 서울시 인구의 시공간적 이동패턴 이해

오석준*¹, 주승영*², 김수환**^{*}, 김민경***^{*}

Understanding Spatiotemporal Human Mobility Patterns in Seoul based on POI Metadata

Seokjoon Oh*¹, Seungyoung Joo*², Soohwan Kim**^{*}, and Minkyung Kim***^{*}

본 논문은 교육부 및 한국연구재단 4단계 두뇌한국21 사업(4단계 BK21 사업)으로부터 지원받은 연구임

요약

효율적이고 지속 가능한 도시 계획을 위한 유동인구의 이동패턴 연구는 활발히 진행되어 왔으나, 유동인구의 거시적 활동 이면의 시공간적 맥락을 반영하는 연구는 부족하다. 이에 본 연구는 서울시의 POI(Point of Interest) 메타 데이터를 활용하여 상황적 맥락과 결합한 시공간적 인간 이동패턴을 종합적인 관점에서 분석하고자 한다. 연구 결과, 서울시 유동인구는 주중, 주말에 따라 규칙적, 불규칙적인 이동패턴의 주기성을 보이며, 이는 주요 POI 메타 고유의 주기성에서 기인한다고 해석된다. 또한, 서울시 유동인구 이동패턴 유사도에 기반하여 자치구들을 클러스터링한 결과, 서울 외곽과 중심지역으로 양분되는 전반적인 양상을 확인하였으며 이는 인간 이동패턴에 대한 사회·경제적 요인의 영향력을 간과할 수 없음을 시사한다. 이 연구는 시간, 공간, POI 메타정보를 융합하여 상황적 맥락과 함께 생활 이동패턴을 다면적으로 이해하려는 연구의 초석이 될 수 있을 것이라 기대한다.

Abstract

Research on capturing human mobility patterns for efficient and sustainable urban planning has been widely conducted. However, studies that unveil spatial context beyond macro-level mobility patterns are relatively scarce. This study aims to analyze the spatiotemporal human mobility patterns with rich context, using POI metadata in Seoul from comprehensive perspectives. As a result, floating population of Seoul exhibits regular and irregular cyclical mobility patterns on weekdays and weekends respectively, stemming from the periodicity of the dominant POIs. Additionally, graphs based on mobility similarity and their regional clustering show clusters vary by POIs, but generally divided into peripheral and central regions of Seoul. This indicates that socioeconomic factors cannot be ignored for understanding human mobility patterns. We expect that this study to lay the groundwork for future research that aims to understand realistic human mobility patterns with rich context.

Keywords

human mobility, POI, metadata, spatiotemporal patterns, context

* 경희대학교 일반대학원 빅데이터융합학과(제1공동저자) · Received: Jun. 28, 2024, Revised: Jul. 22, 2024, Accepted: Jul. 25, 2024
- ORCID¹: <https://orcid.org/0009-0007-6764-4141> · Corresponding Author 1: Soohwan Kim
- ORCID²: <https://orcid.org/0009-0004-1895-8804> Kwangwoon University, Seoul 01897, Republic of Korea
- ORCID: <https://orcid.org/0000-0002-2891-7537> Tel.: +82-2-940-8476, Email: kimsoohwan@kw.ac.kr
** 광운대학교 정보융합학부 교수(교신저자) · Corresponding Author 2: Minkyung Kim
- ORCID: <https://orcid.org/0000-0002-2891-7537> Kyung Hee University, Seoul 02447, Republic of Korea
*** 경희대학교 빅데이터융합학과 교수(교신저자) · ORCID: <https://orcid.org/0000-0001-8292-7851> Tel.: +82-2-961-0650, Email: minkkim@khu.ac.kr

I. 서론

현대 도시는 인구와 사회적 자본의 밀집으로 인해 다양한 사회·경제적 현상이 발생하는 복합적인 구조를 가진다. 특히 서울시는 단위 면적당 거주 인구 수가 매우 높아 면적 대비 유동인구의 비중이 높은 특징을 보인다. 유동인구는 도시 활동을 통해 다양한 이동패턴을 형성하며, 유동인구의 이동 데이터를 다각도로 분석하는 것은 도시의 특성을 이해하는 데 필수적인 요소이다. 이를 도시 계획과 정책 입안에 활용하여 보다 효율적이고 지속 가능한 도시 계획 수립을 위한 근거를 마련할 수 있다. 특히, 서울과 같은 대도시의 유동인구 분석은 도시의 교통시스템 관리, 공공 서비스 설계 및 전염병 예방 등 다양한 측면에서 중요한 역할을 한다[1]-[3].

이러한 배경에서, 유동인구의 거시적인 이동패턴을 파악하려는 다양한 연구가 시도되었다. 가령, 서울에서 제공하는 서울 생활이동 데이터에서도 ‘생활이동’을 ‘특정 시점, 특정 지역 간에 이동하는 서울의 모든 인구’로 정의한다[4]. 그러나 실제 유동인구는 동일한 시간대와 지역에서도 다양한 목적으로 이동하기 때문에, 이를 확일적이고 거시적인 시공간으로 한정한다면 생활이동에 내재되어 있는 도시의 구조적, 기능적, 맥락적 특성을 이해하기 어렵다는 한계가 존재한다.

이에 본 연구는 상황적 맥락이 결합된 시공간적 인간이동 패턴을 거시적 및 미시적 관점에서 분석하여 이해하고자 한다. 이를 위해, 포스퀘어(Foursquare) 이용자들의 체크인(Check-in) 및 POI(Point of Interest) 정보가 기록된 이동 데이터를 이용하여 서울시 유동인구의 생활이동 패턴을 분석한다. 나아가 서울시 25개 자치구 간 POI 메타별 이동 패턴의 유사도에 기반하여 자치구 간의 유사도 네트워크를 생성하고, 그래프 클러스터링을 통해 군집분석을 수행한다.

분석결과로서, 서울시 유동인구의 주기적 이동패턴은 크게 주중과 주말로 이분화되어 있으며, 이는 각각 규칙적, 불규칙적 생활패턴이 반영된 결과로 해석할 수 있다. POI 메타정보를 활용하여 공간적 특성을 반영한 결과, 이용자의 방문확률이 높은 상위 POI 메타들을 중심으로 주중의 주기적 패턴이

한층 강화되는 현상이 나타났다. 서울시 자치구 간 생활 이동패턴의 유사도는 POI 메타별로 다르게 나타나며, 전체적으로 서울 외곽과 중심 지역으로 군집이 양분화되는 경향을 보였으며, 거주지와 경제생활 중심지 간 거리의 불균형적 분포를 반영한다. 즉, 도시 유동인구의 이동 패턴에 사회·경제적 요인의 영향을 배제하기 어렵다는 것을 시사한다.

본 연구는 공간적 특성과 상황적 맥락을 결합하여 시공간적 생활 이동패턴의 다면적 이해에 기여하였으며, 다양한 응용분야에 적용할 수 있는 기초 연구가 될 수 있을 것으로 기대한다.

II. 관련 연구

교통 시스템의 발달에 따른 지역 간 경계가 점차 모호해지고, 도시의 생활권역이 확대됨에 따라 도시 간 물리적, 기능적 경계의 구분이 동적으로 진화하고 있다. 따라서 이를 이해하고 도시계획 수립에 적용하는 것은 매우 도전적인 과제이다. 하지만 최근 다양한 모바일 디바이스를 통한 개별 사용자들의 라이프로그(Lifelogging)이 활발해지고, 이를 통해 생활 이동 데이터의 대량 수집이 가능해짐에 따라, 인간 이동패턴을 분석하여 도시의 지역적 특성을 이해하고 도시계획을 수립하기 위한 다양한 연구가 존재한다[5][6].

전 세계의 도심지역에 걸쳐 유동인구의 이동패턴을 시간대별로 비교·분석한 연구에서는 도시마다 보이는 이동패턴의 차이를 보고하였다[7]. 이는 이동패턴을 거시적인 차원에서 분석한 것으로, 특정 시간대에 얼마나 많은 유동인구가 이동하는가에 초점이 맞춰져 있다. 또한, 도심지역의 평균 이동반경을 도시의 지역적, 사회·경제적 특성들과 연계하여 심도 있는 도시 이동패턴을 이해하고자 하였다[8].

한편, POI 데이터는 미시적인 관점의 인간 이동패턴 분석에 중요한 자료이다. POI 정보를 고려한 분석은 다음 이동 장소 예측 모델 개선[9] 및 이동성 기반 맞춤형 장소 추천 시스템에 활용될 수 있다[10]. 더 나아가, 이동 경로를 연결하는 에지(Edge)에 메타정보를 추가하여 이동 네트워크를 형성함으로써, 빈부격차와 같은 사회·경제적 요인을 발견한 사례도 존재한다[11].

이는 POI 데이터 분석이 인간 이동패턴의 다면적 이해를 위한 중요한 자원임을 보여준다.

하지만 POI 메타와 시공간 정보를 결합하여 상황적 맥락이 반영된 인간 이동패턴을 종합적인 관점에서 연구한 사례는 아직 부족하다. 서울시의 POI 데이터를 기반으로 이에 본 논문에서는 POI 데이터에 내재된 유동인구의 시공간적 이동패턴을 단계적·다면적으로 분석하고, 이를 사회·경제적 요인과 결합하여 서울 자치구의 구조적, 기능적 특성을 이해하고자 한다.

III. 데이터 수집 및 전처리

3.1 데이터 수집

도심 지역에서 발생하는 시공간적 생활이동 패턴을 분석하기 위해, 대한민국 유동인구의 90% 이상을 차지하는 서울시를 대상으로 POI 방문 데이터를 이용하고자 하였다. 포스퀘어는 개발자 계정의 유료 API를 제공하여 호출하는 방식으로 사용자가 데이터셋을 수집할 수 있도록 한다. 본 연구에서는 포스퀘어가 가장 활발히 사용되어 데이터가 풍부한 시점을 기준으로, 전 세계 체크인(Check-in) 데이터에 기반한 선행연구[12][13]에서 공개한 데이터를 우선 수집하였다[14].

전 세계 사용자 대상의 위치 기반 소셜 네트워크 서비스인 포스퀘어 시스템의 핵심 기술은 두 가지로, 지오태깅(Geotagging)과 자동완성(Autocomplete)이다. 지오태깅은 특정 시점에 포스퀘어를 사용하는 사용자의 장치가 어디에 있는지, 주변에 무엇이 있는지의 지리적·상황적 정보를 감지하는 POI 태깅 기술로, 이를 통해 사용자가 포스퀘어를 사용하며 제출한 키워드를 기준으로 장소 정보를 태그한다. 이후, 제공된 키워드 및 위치 입력과 일치하는 최상위 주소, POI 정보 등을 자동으로 반환해 사용자의 체크인 기록을 완성하고, 최종적으로 시간 정보와 함께 시스템 로그에 기록한다[15].

해당 데이터는 2012년 4월부터 2014년 1월까지 약 22개월간 114,324명의 서비스 사용 로그이며, 총 3,820,891개의 장소에서 발생한 22,809,624건의 체크인 기록을 포함한다(POI별 평균 6건의 체크인). 또

한, 체크인이 발생한 시점(세계표준시 기준), 체크인 지점의 위도·경도 좌표와 함께 이용자들이 방문한 POI 카테고리 정보가 포함되어 있다. 각 사용자들에게는 익명처리된 고유 ID가 할당되어 있어, 개별 사용자의 이동 경로를 구분할 수 있다. 즉, 이용자들이 이동하며 방문한 정확한 위치의 좌표 정보뿐만 아니라 방문 장소의 구체적인 POI 정보까지 제공하는 특징점이 있다. 이는 건물의 시설이용 현황을 통해 간접적으로 도시 인구가 이용하는 POI를 측정하는 방식[16]을 보완할 수 있다.

데이터 수집기간이 본 논문의 연구시점과 차이가 존재하나, 장기간에 걸친 대규모 POI 데이터 구득의 현실적 어려움과 포스퀘어의 위치기반 서비스 플랫폼으로서의 대표성(데이터 수집기간 기준 전 세계 많은 이용자 보유)을 감안할 때 수집 데이터는 연구적 가치가 있다. 더 나아가, 본 연구는 포스퀘어 플랫폼과는 독립적으로, 일반적인 POI 이동 데이터로부터 유동인구 이동패턴에 대한 맥락적 인사이트를 도출할 수 있는 분석 방법론에 초점을 두고 있다.

3.2 데이터 전처리

서울시 자치구 경계선의 위도·경도 좌표 정보를 기반으로 Python shapely 모듈과 GeoPandas 모듈을 이용하여 전 세계 이용자들의 체크인 정보가 담긴 포스퀘어 데이터셋 중에서 서울시에 해당하는 데이터를 추출하였다. 이후 다시 모듈을 적용하여 체크인 좌표에 해당하는 서울시 행정 자치구 정보를 새롭게 추가하였다. 그 결과 2012년 4월 3일부터 2014년 1월 29일까지 서울시 25개 행정 자치구의 유동인구 POI 체크인 데이터셋을 구축하였다. 또한, 정확한 시간 정보 구분을 위해 원 데이터의 체크인 시간을 한국표준시(KST)로 변환하였다. 전처리 과정을 통해, 포스퀘어 이용자 1,992명이 기록한 124,186건의 서울시 체크인 기록으로 정제하였다. 보다 구체적으로는, 이용자들이 방문한 장소는 총 29,970곳이며 이들은 386개의 POI로 구분되어 있다. 상황적 맥락을 위해, 이들 POI 정보를 산업 분류 기준[17]을 참조하여 다시 16개의 POI 메타 범주로 재분류하였다.

본 연구에 사용한 서울시 포스퀘어 데이터셋의 기초통계는 표 1과 같다.

표 1. 분석을 위해 정제한 포스퀘어 데이터셋 기초통계
Table 1. Basic statistics of preprocessed foursquare datasets used in this study

Category	Number of cases
Users	1,992
Check-ins	124,186
Venues	29,970
POI	386
POI meta classification	16
Latitude range	37.413294 ~ 37.75133
Longitude range	126.734086 ~ 127.269311

IV. 실험방법 및 결과

4.1 이동패턴 기반 서울시 자치구 클러스터링

앞에서 전처리한 데이터셋을 기반으로 서울시 유동인구의 거시적인 이동패턴을 이해하기 위해, 먼저 익명화된 개별 이용자 i 의 이동 경로를 그래프 $G_i = (V_i, E_i)$ 로 표현하였다. 이때 그래프를 구성하는 Vertices $V_i = \{u, v | (u, v) \in E_i\}$ 는 이용자 i 가 방문한 POI가 속해 있는 자치구들의 집합이며, Edges E_i 는 순서대로 방문한 POI들 간의 이동 경로를 의미한다. 즉, POI가 위치한 자치구 u 에서 v 로 이동한다면 해당 경로인 에지 $(u, v) \in E_i$ 가 형성된다.

이후 개별 이용자 i 의 POI 방문 네트워크 G_i 를 전체 이용자 n 명에 대해 결합하여 유향 가중 그래프 $G = \{G_i\}_{i=1}^n = \{V, E | V_i \in V, E_i \in E\}$ 를 구축하였다. 이렇게 생성한 그래프는 포스퀘어 이용자들의 서울시 자치구 간 이동의 흐름을 반영하므로, 유동인구의 이동이 활발하게 일어나는 지역들 간의 강한 연결성(Connectivity)을 포착하기 위해 Louvain 알고리즘을 사용하여 결합 그래프 G 를 클러스터링하였다.

그림 1(a)-(b)는 주중(a)과 주말(b)을 구분한 그래프 클러스터링 결과를 나타낸다. 해당 그림에서 색상은 클러스터를 구분하며, 그중 특히 주중의 클러스터들이 주말에 하나의 클러스터로 확장되는 효과를 시각적으로 표현하기 위해 주중의 해당 군집들을 같은 색 계통의 그라데이션으로 표현하였다. 그림 1(c)는 서울특별시에서 수립한 ‘2030 서울도시기본계획’에서의 권역생활권 구분[18]을 참조로 첨부한 그림이다. 먼저, 그림 1(a)-(b)의 전반적인 군집 유형을 통해, 인접 지역 간의 강한 연결성을 확인할 수 있으며, 특히 한강을 가로지르지 않고 한강 이북과 이남으로 명확히 구분되어 있음을 알 수 있다. 즉, 지리학적 근접성은 지역 간 이동에 주요 요인 중 하나로 해석될 수 있다. 하지만 모든 인접 지역들이 하나의 클러스터로 연결되지 않는 것은 지리적 특성 이외에 인구 및 교통시스템 등의 지역적 특성이 생활 이동의 경계로 작용하고 있음을 암시한다.

이동패턴의 구조적 특징은 주중과 주말로 구분하여 군집분석을 수행을 하였을 때 더욱 선명해진다.

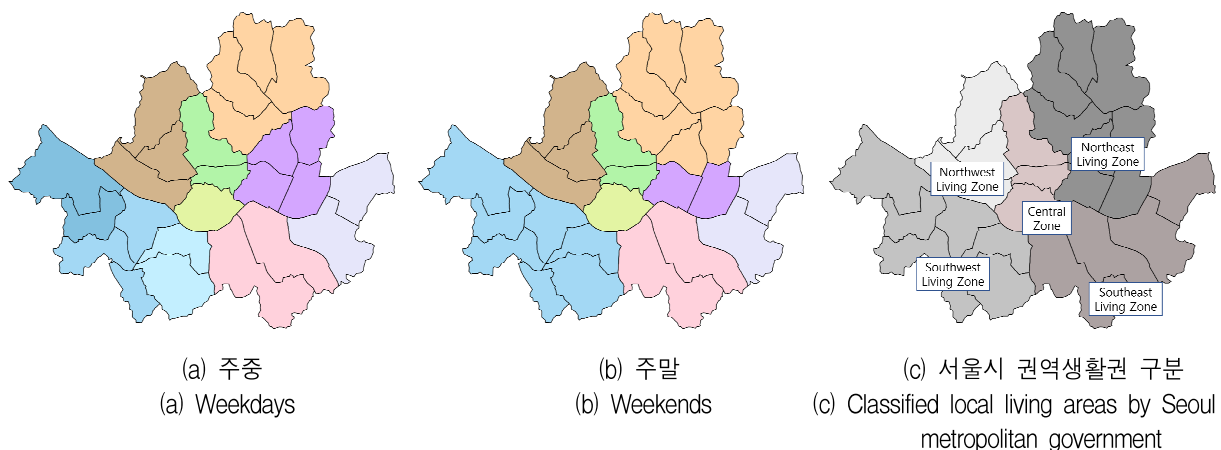


그림 1. 서울시 유동인구의 POI 이동 기반 자치구 간 클러스터링 (색상: 클러스터 구분)

Fig. 1. Location-based population movement clustering across 25 districts in Seoul (color-coded by clusters)

주중(그림 1(a))의 군집 크기는 대체적으로 균일한 반면, 주말(그림 1(b))은 서남생활권과 동북생활권을 중심으로 군집의 크기가 확대됨을 알 수 있다. 이는 주중의 규칙적인 출퇴근 및 등하교가 2-3개 자치구 이내로 이동반경을 형성하지만, 시간의 제약이 상대적으로 적은 주말의 경우 유동인구의 이동반경이 확대되는 것으로 이해할 수 있다.

그림 1(c)에서 물리적 환경과 자치구 기능을 중심으로 설정된 서울시의 생활권역 구분은 비교적 이동반경이 확장된 주말 군집(그림 1(b))과 유사한 양상을 보인다. 하지만, 이동반경이 축소되는 주간 군집(그림 1(a))과는 군집 크기에서 두 배 정도의 차이를 보인다. 이는 서울시 유동인구의 실제적인 생활이동은 각 지역을 정적으로 구분한 서울시 생활권역과는 차이가 있으나, 인구 및 교통시스템 등 지역의 동질적 특성의 경계를 투영하는 것으로 해석할 수 있다.

4.2 유동인구의 시간대별 이동패턴

위에서 이동패턴 기반 자치구 클러스터링을 통해, 서울시 유동인구의 포괄적인 주중, 주말 이동반경에 대해 살펴보았다. 다음으로 이동패턴의 주기적 특성을 살펴보기 위해 먼저 시계열 데이터를 일주일 단위로 구분하고, 시간대별 움직임을 확대·포착하기 위해 일주일을 168(24시간/일×7일)시간으로 세분화하여 그림 2와 같이 유동인구의 시간대별 POI 방문 확률 $P(\tau)$ 를 계산하였다.

그림 2에서, 이동패턴의 주기성이 주중($0 \leq h \leq 120$)과 주말($120 < h \leq 168$)로 이분화되어 있음을 확인할 수 있다.

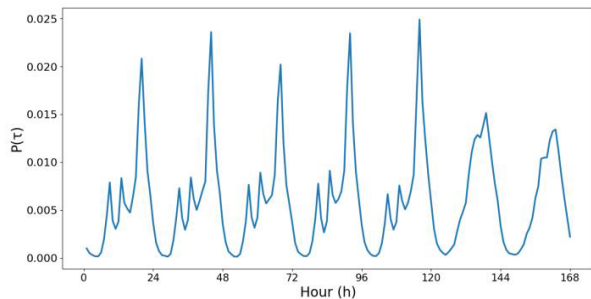


그림 2. 시간대별 POI 방문 확률분포(일주일 기준)
Fig. 2. Probability distribution of POI visits on a weekly basis

이는 (1)주중의 출퇴근 및 등하교와 같은 규칙적 생활패턴과 (2)주말의 보다 자유로운 시간대에 다양한 POI 방문으로 인한 자율적·불규칙적 생활패턴이 반영된 결과로 해석할 수 있다. 구체적으로, 주중 그래프의 경우 하루 세 번의 가파른 상승점이 존재하며, 이때의 시간대는 각각 09시, 13시, 20시이다. 이들은 각각 출근, 점심, 퇴근 이후의 시간대와 관련된 유동인구의 급격한 증가를 드러낸다. 한편, 주말 그래프의 경우 주중 그래프에 비해 특정 시간대에 집중되지 않고 방문 확률이 가장 높은 19시까지 완만한 증가폭을 보인다. 이는 주중의 출퇴근과 같은 다수의 동시다발적 이벤트가 주말에는 상쇄되어, 유동인구가 보다 여유있는 자율적인 시간대에 다양한 활동을 한 결과로 해석할 수 있다.

4.3 유동인구의 POI 메타별 이동패턴

서울시 유동인구의 시간대별 움직임에 공간적 특성을 고려하기 위해 그림 2의 방문 확률을 POI 메타별로 다시 산출하였다. 그림 3은 전체 POI를 16개 메타로 범주화하여 공간적 특성인 POI 메타와 주기적 시간 정보를 동시에 반영한 방문 확률분포 $P(\tau)$ 를 나타낸다. 그림 3이 보여주는 바와 같이, 전반적으로는 앞서 분석한 주중, 주말 이동패턴의 주기성을 유지하고 있으나, POI 메타별로 이동패턴의 고유한 주기성이 내재되어 있음을 확인할 수 있다.

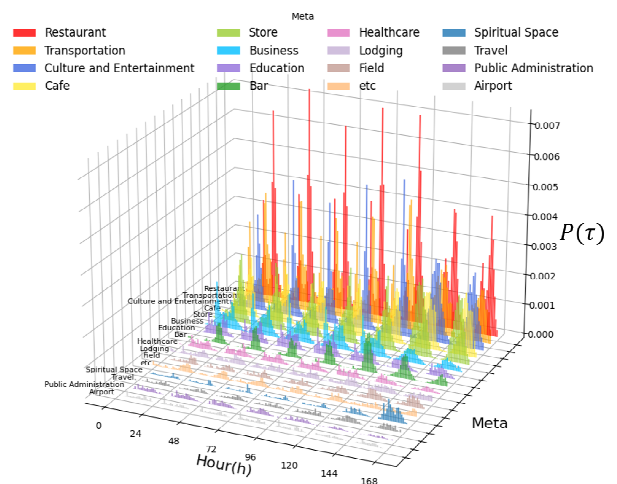


그림 3. POI 메타별, 시간대별 방문 확률분포(일주일 기준)
Fig. 3. Probability distributions of POI meta visits on a weekly basis

이러한 경향은 특히 이용 확률이 높은 상위권 POI 메타일수록 하위권 메타에 비해 주기적 경향이 강하게 드러나며, 구체적으로는 레스토랑, 교통시스템, 문화·엔터테인먼트, 카페, 상점, 비즈니스, 교육 메타가 해당된다. 특히, 레스토랑과 교통시스템의 이용확률이 다른 메타보다 높은 것은 그림 2에서 확인한 이동 확률의 가파른 상승 시간대(09시, 13시, 19-20시)와 강한 연관이 있음을 알 수 있다. 즉 점심·저녁식사 시간과 출퇴근에 관련된 POI 메타 이용에서 기인하는 것으로 해석할 수 있다.

그림 3에서 확인한 해당 7개의 상위 메타는 전체 이용자 수의 약 50%(약 1천명) 이상의 POI 체크인 기록을 기반으로 하며, 이는 대다수의 포스퀘어 이용자의 이동 패턴을 반영한다. 이에 따라, 서울시 포스퀘어 이용자의 대표성을 확보하기 위해 상위 7개 POI 메타를 중심으로 이후의 연구를 진행하였다.

4.4 POI 메타 간 이동패턴의 유사도

앞서 분류한 POI 메타 중에서도 방문 확률이 높은 상위 7개 메타 간 주기적 이동패턴의 유사성을 분석하기 위해, KL 발산(Kullback-Leibler divergence)을 이용하여 유사도를 계산하였다. 아래 식 (1)은 POI 메타 r 과 s 의 방문 확률분포인 $P^r(\tau)$ 와 $P^s(\tau)$ 의 유사도를 계산한다.

$$D_{KL}(r,s) = \sum_{\tau} P^r(\tau) \log \frac{P^r(\tau)}{P^s(\tau)} \quad (1)$$

이때, KL 발산은 방향성이 존재하므로, 아래 수식 (2)를 통해 POI 메타 간 유사도 행렬이 대칭이 되도록 하였다.

$$D_{KLS}(r,s) = \frac{1}{2} (D_{KL}(r,s) + D_{KL}(s,r)) \quad (2)$$

그림 4는 식 (2)를 통해 계산한 POI 메타 간 유사도 행렬을 Louvain 알고리즘을 이용하여 클러스터링한 결과를 나타내며, 크게 세 종류의 군집을 확인할 수 있다. 그림에서 노드의 크기와 에지의 굵기는 각각 POI 메타 방문자 수와 메타 간 유사도에 비례하며, 강한 유사도를 가진 POI 메타 군집을 추

출하기 위해 유사도가 낮은 에지는 필터링하였다. 그림 4의 군집 결과와 연계하여, 주중과 주말의 POI 메타 방문 확률분포를 군집 별로 구분하여 각각 그림 5와 그림 6에 나타내었다.

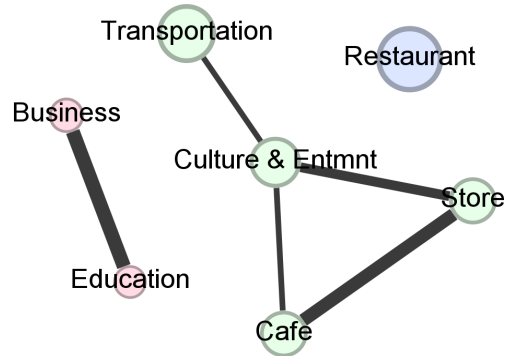


그림 4. POI 메타 방문 확률분포 유사도 기반 상위 7개 메타 군집 결과
Fig. 4. Clustering of top 7 POI meta based on $P(\tau)$ in Fig. 3

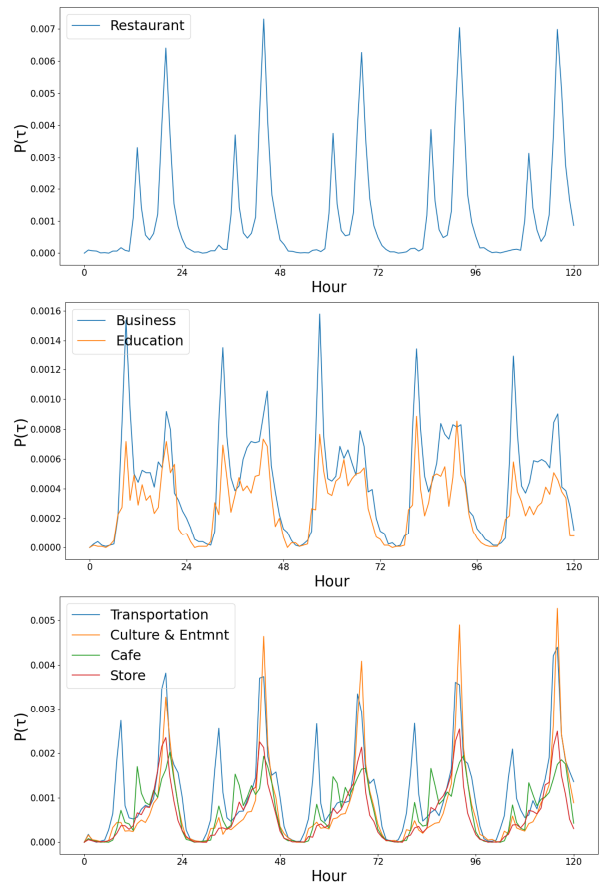


그림 5. 그림 4의 군집 별 POI 메타의 주중 방문확률
Fig. 5. Weekday visit probability distributions of POI meta by each cluster in Fig. 4

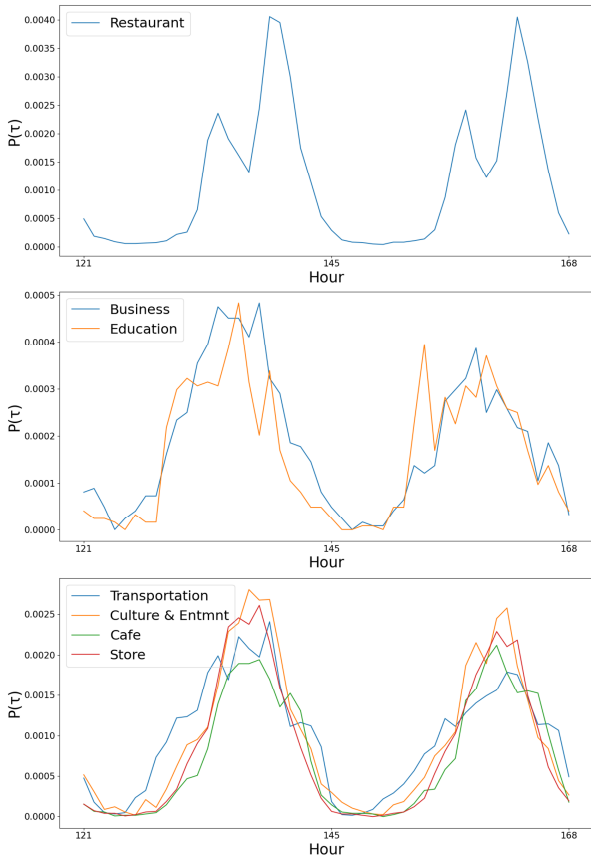


그림 6. 그림 4의 군집 별 POI 메타의 주말 방문확률
 Fig. 6. Weekend visit probability distributions of POI meta by each cluster in Fig. 4

이를 종합하여 살펴보면, 첫 번째 군집인 레스토랑 메타는 이용 확률이 가장 높으며, 주중과 주말 모두 상승 시간대가 규칙적인 특성을 보인다. 이는 식사 시간에는 주중과 주말에 큰 변화가 없으므로 주기성을 유지하는 것으로 보인다. 두 번째 군집인 비즈니스, 교육 메타는 직장인의 출퇴근, 학생의 등하교와 같은 주기적 이동패턴이 주중에 나타나는 반면, 주말에는 특정 시간대에 집중되지 않는 불규칙적 이동패턴을 보인다. 마지막으로, 세 번째 군집은 문화·엔터테인먼트를 중심으로 주기적 이동패턴의 유사도가 높으나, 주중과 주말의 패턴 차이를 알 수 있다. 이는 주중에는 일과시간 이후 제한된 시간 안에 이용이 가능하나, 주말에는 상대적으로 여유로운 시간대에 이용이 가능한 것으로 해석할 수 있다.

4.5 서울시 자치구 간 시공간적 이동패턴의 유사도

앞에서 상위 7개 POI 메타 간 시간대별 이동패

턴의 유사도를 살펴보았다. 이번에는 각 POI 메타에 대하여, 서울시 자치구 별로 주중 POI 방문 데이터를 추출하여 자치구 간 시공간적 이동패턴의 유사도를 분석하고자 한다. 이를 위해, 자치구 i 에 위치한 POI 메타 m 으로의 시간대별 방문 확률 $P_i^m(\tau)$ 를 계산하였다.

각 POI 메타 별 자치구 상호 간의 이동패턴의 유사도를 계산하기 위해, 식 (1)을 아래 식 (3)과 같이 POI 메타 m 에 대한 지역 i 와 j 의 확률분포 간의 거리인 KL 발산으로 세분화하였다.

$$D_{KL}^m(i, j) = \sum_{\tau} P_i^m(\tau) \log \frac{P_i^m(\tau)}{P_j^m(\tau)} \quad (3)$$

이때, KL 발산은 역시 방향성이 존재하므로, 아래 식 (4)를 통해 자치구 간 유사도 행렬이 대칭이 되도록 하였다.

$$D_{KLS}^m(i, j) = \frac{1}{2} (D_{KL}^m(i, j) + D_{KL}^m(j, i)) \quad (4)$$

식 (3)과 (4)를 이용하여 서울시 25개 자치구로 구성된, 유사도 기반 무향 가중 그래프(Undirected weighted graph)를 생성하고, 그래프 클러스터링을 통해 군집분석을 수행하였다. 그림 7은 군집 정도가 높은(modularity: 0.3 이상) POI 메타 별((a)교육, (b)교통시스템, (c)문화·엔터테인먼트)로 군집 결과를 색상으로 구분하여 시각화한 결과이다. 또한, 그림 8은 그림 7을 설명하기 위한 근거 자료로서 해당 POI 메타의 각 군집별 $P_i^m(\tau)$ 의 평균 확률분포를 나타낸 것이다.

그림 7(a)의 교육 POI 메타의 경우, 서울 외곽(파란색)은 나머지 군집들에 비해 중·고등학교 이용 비율이 높다. 또한 초록색과 보라색 군집은 대학들이 위치한 자치구들에 해당되며, 특히 초록색 군집에는 주요 4년제 대학들이 밀집되어 있다. 그림 8(a)가 보여주는 바와 같이 시간대별 이용 확률분포에 차이가 존재하는 것은, 대학들의 교육운영 특성이 반영된 것으로 해석된다.

그림 7(b)의 교통시스템 POI 메타의 경우, 세 군집 모두 출퇴근 시간대에 이용확률이 급증하는 양상을 보인다(그림 8(b)).



그림 7. POI 메타별 서울시 자치구 간 이동패턴 유사도 기반 클러스터링
Fig. 7. Clustering of Seoul districts based on mobility pattern similarities for each POI meta

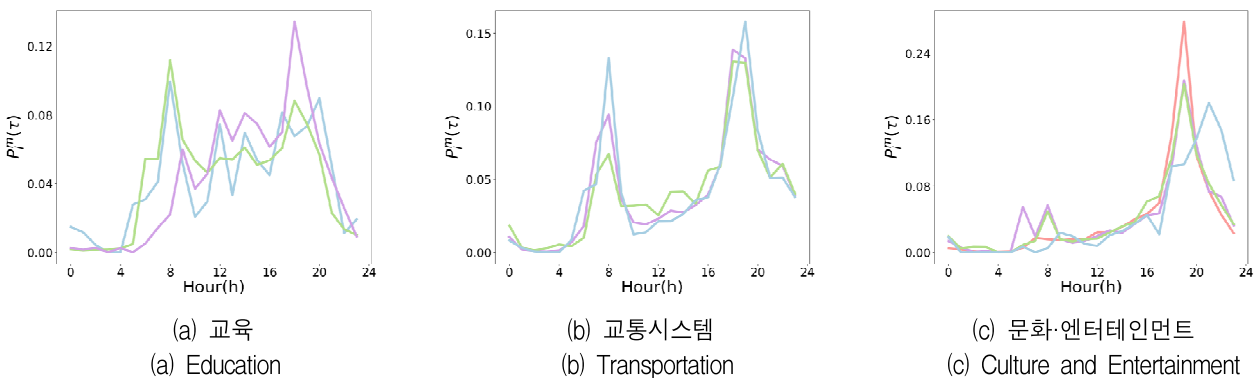


그림 8. 그림 7의 POI 메타별 군집(색상참조)에 대한 $P_i^m(\tau)$

Fig. 8. $P_i^m(\tau)$ of the clusters in Fig. 7 (each plot is color-coded by the cluster colors)

그러나 서울 외곽 지역(파란색)은 중심부에 위치한 다른 두 군집에 비해 교통시스템을 이용하는 시간대가 빠르고 출근 시간대의 이용률이 압도적으로 높으며, 군집 공통으로 $P_i^m(\tau)$ 가 최고치인 퇴근 시간대에는 이용 시간대가 늦춰짐을 알 수 있다. 하지만 그림 7(b)에서 출퇴근 사이의 오후 시간대에는 서울 외곽지역(파란색)의 교통시스템 이용률이 중심 지역보다 상대적으로 낮음을 알 수 있다. 이는 서울 외곽지역의 경우 거주지와 근무지가 상이한 서울시의 구조적 특징을 반영하는 것으로 해석할 수 있다.

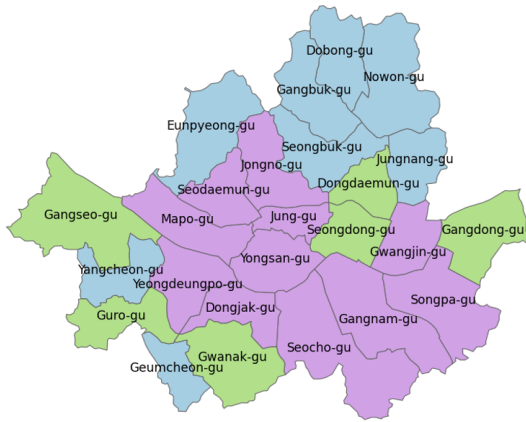
그림 7(c)의 문화·엔터테인먼트 POI 메타의 경우, 다른 메타에 비해 비교적 고르게 분할된 군집 분포를 보여준다. 그림 8(c)에서 모든 군집 공통으로 $P_i^m(\tau)$ 가 18시 이후 증가하는데, 이는 퇴근 시간 이후에 다양한 자치구로 분산된 유동인구가 문화·엔터테인먼트를 영위하기 위한 POI 방문패턴이 나타난 것이라 할 수 있다. 더 나아가, 외곽지역(파란색)의

이용 시간대가 늦어지는 것은 앞서 교통시스템 메타의 이용 시간대와 연계하여 중심부에서 퇴근 후 귀가하기까지의 시간적 격차가 반영된 것으로 보인다.

4.6 서울시 유동인구의 종합적 생활이동 패턴

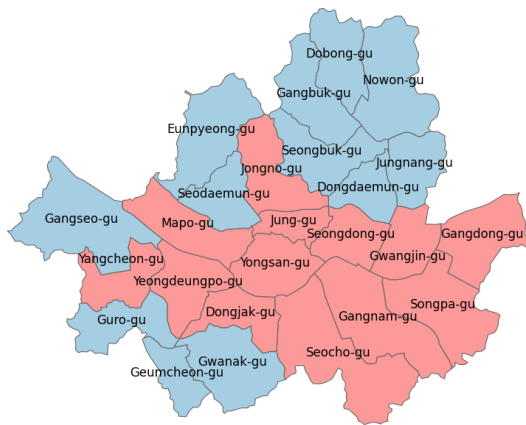
그림 9에서 (a)는 POI 방문확률 상위 7개 메타 전체에 대해서 서울시 자치구 간 유사도 군집을, (b)는 데이터셋 수집 기간과 근접한 서울시 아파트의 평당 매매가격[19]을 클러스터링한 결과를 나타낸다. 서울시 유동인구의 시공간적 이동패턴의 유사도를 종합하면 중심과 외곽지역으로 자치구들이 크게 2개의 군집으로 구분되며, 이러한 현상은 서울시 아파트 평당 매매가격의 양분화된 분포와 유사함을 알 수 있다.

이는 자치구별로 거주지역과 출퇴근 지역 간의 거리적 분리 정도가 반영된 것이라 할 수 있다.



(a) 상위 7개 메타 종합 이동패턴 유사도 군집결과

(a) Clustering result based on the similarity of top 7 POI Meta mobility patterns



(b) 서울시의 아파트 평당 매매가격 군집결과

(b) Clustering result of apartment sale prices per square meter in Seoul's districts

그림 9. 서울시 자치구의 상위 7개 POI 메타 이동패턴 유사도 기반 군집결과 및 서울시 아파트 평당 매매가격 기준 군집결과 비교

Fig. 9. Comparison of clustering results based on the similarity of Top 7 POI meta mobility patterns and apartment sale prices per square meter in Seoul's districts

즉, 외곽지역의 경우 거주지역과 경제활동 지역 간의 차이가 크다고 할 수 있다. 다시 말해, POI 주요 메타에 기반한 유동인구의 이동패턴은 사회·경제적 요인의 영향을 받음을 시사한다. 특히 이러한 군집의 양분화는 서울 외곽에서부터 다양한 경제활동의 인프라가 밀집된 중심지역으로의 유동인구의 이동을 견인하는 공간적인 도시 특성이 반영된 결과라고 할 수 있다.

본 연구는 서울시에 위치한 POI에 대한 선행연구

들과 비교할 때, 유동인구의 시공간적 이동을 메타 이동으로 추상화함으로써 도시의 구조적, 기능적 특성의 재해석 방법론과 가능성을 제시한 점에서 차별성을 갖는다. 서울 시민들의 집단적 POI 사용을 활용한 이동 군집 및 인기 장소 분석[20][21], 도시활동 중심지 도출 및 산업 입지 결정요인 분석[22][23] 등의 사례들은 POI의 메타 정보 없이 시간적, 혹은 공간적 차원 중 한 가지만을 연구하였기 때문에 종합적·상황적 맥락을 분석하는 데 한계가 있다. 또한 서울에 제한하지 않고 전 세계 도시 간의 이동패턴 비교 연구[7]에서도 시간과 공간적 차원을 개별적으로 사용하였기에 다차원적인 분석에는 한계가 있다. 더 나아가 본 연구는 복잡계 시스템 내에서 메타인구(Meta-population) 상호 작용으로 인한 확산현상[24]으로 확장할 수 있으며, 전염병 예측 모델에도 주요 변인으로 적용[25]할 수 있다.

V. 결론 및 향후 과제

본 연구에서는 POI 체크인 시계열 데이터를 메타화함으로써 서울시 유동인구의 시공간적 이동패턴을 거시적·미시적 관점에서 상황적 맥락을 고려하여 종합적으로 분석하였다.

위치 이동패턴을 기반으로 서울시 자치구를 클러스터링한 결과, 주중과 주말의 군집 분포에 차이가 있음을 확인하였다. 즉, 주중의 균등한 군집 크기는 등하교 및 출퇴근과 같은 규칙적 생활 이동을 위한 효율적 반경을 반영하는 반면, 주말의 불균등한 군집 크기의 분포는 교통 인프라를 중심으로 이동 반경의 자율적 확대를 의미한다.

또한, 시계열 데이터를 일주일 단위로 구분하여 유동인구의 시간대별 POI 메타 이용 확률분포를 통해 주중, 주말의 규칙적·불규칙적인 패턴을 확인하였다. 또한 POI 메타의 시간대별 이용 확률분포를 KL 발산에 적용하여 POI 메타간 유사도를 계산하였으며, 이를 다시 서울시 자치구별 POI 메타의 시간대별 이용확률 분포로 세분화하여 KL 발산에 적용함으로써 자치구 간 이동패턴의 유사도를 함께 도출하였다. 종합적으로 살펴보면, 서울시 자치구는 외곽과 중심 지역으로 양분화된 현상이 두드러지며, 이는 거주지와 경제적 활동 중심지 간의 거리적 분

리 정도를 반영하는 것으로서 사회·경제적 영향을 배제하기 어렵다는 것을 시사한다. 더 나아가 이러한 상황적 맥락을 해석 가능하게 하는 시공간적 생활 이동 패턴 분석의 중요성을 시사한다.

향후 연구계획으로는 POI 메타의 정적 분류에서 벗어나 보다 유연하고 동적인 정의가 가능한 방법론을 연구하고, 이를 인구이동을 예측하는 시계열 모형에 적용하고자 한다. 또한, 특정 플랫폼에 국한되지 않고 현재 활발히 사용되는 소셜네트워크 위치 기반 데이터에 기반하여 이동패턴의 시간적 변화 양상을 연구하고자 한다.

References

- [1] Y. Yao, H. Zhang, J. Chen, W. Li, R. Shibasaki, and X. Song, "Mobility Tableau: Human Mobility Similarity Measurement for City Dynamics", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 24, No. 7, pp. 7108-7121, Jul. 2023. <https://doi.org/10.1109/TITS.2023.3258551>.
- [2] A. C. Perofsky, et al., "Impacts of human mobility on the citywide transmission dynamics of 18 respiratory viruses in pre- and post-COVID-19 pandemic years", *Nature Communications*, Vol. 15, pp. 4164, May 2024. <https://doi.org/10.1038/s41467-024-48528-2>.
- [3] H. Oh and M. Kim, "Dynamics of COVID-19 Spread based on Human Mobility Patterns by Epidemic Stages in Seoul", *JKIIT*, Vol. 23, No. 4, pp. 153-160, Apr. 2023. <https://doi.org/10.14801/jkiit.2023.21.4.153>.
- [4] Migration data in Seoul, <https://data.seoul.go.kr/dataVisual/seoul/seoulLivingMigration.do> [accessed: Apr. 05, 2024]
- [5] E. Dorostkar, M. Najarsadeghi, M. Molavi, and N. Zali, "Human Mobility and Energy: How Do Human Mobility and Energy Affect Urban Policy and Planning?", *Journal of Urban Management*, Vol. 4, No. 12, pp. 413-419, Dec. 2024. <https://doi.org/10.1016/j.jum.2023.08.001>.
- [6] C. Carbrera-Arnau, C. Zhong, M. Batty, R. Silva, and S. Kang, "Inferring Urban Polycentricity from the Variability in Human Mobility Patterns", *Scientific Reports*, Vol. 13, pp. 5157, Apr. 2023. <https://doi.org/10.1038/s41598-023-33003-7>.
- [7] F. Betancourt, A. P. Riascos, and J. L. Mateos, "Temporal Visitation Patterns of Points of Interest in Cities on a Planetary Scale", *Scientific Reports* Vol. 13, pp. 4890, Mar. 2023. <https://doi.org/10.1038/s41598-023-32074-w>.
- [8] S. Wang, Y. Zheng, G. Wang, T. Yabe, E. Moro, and A. Pentland, "Infrequent activities predict economic outcomes in major American cities", *Nature Cities*, Vol. 1, No. 4, pp. 305-314, Mar. 2024. <https://doi.org/10.1038/s44284-024-00051-7>.
- [9] Z. Huang, S. Xu, M. Wang, H. Wu, Y. Xu, and Y. Jin, "Human Mobility Prediction with Causal and Spatial-constrained Multi-task Network", *EPJ Data Science*, Vol. 13, No. 22, Mar. 2024. <https://doi.org/10.1140/epjds/s13688-024-00460-7>.
- [10] S. R. Bashir, S. Raza, and V. B. Mistic, "BERT4Loc: BERT for Location—POI Recommender System", *Future Internet*, Vol. 15, No. 6, pp. 213, Jun. 2023. <https://doi.org/10.3390/fi15060213>.
- [11] T. Abbiasov, C. Heine, S. Sabouri, A. Salazar-Miranda, P. Santi, E. Glaeser, and C. Ratti, "The 15-minute City Quantified Using Human Mobility Data", *Nature Human Behaviour*, Vol. 8, No. 3, pp. 445-455, Feb. 2024. <https://doi.org/10.1038/s41562-023-01770-y>.
- [12] D. Yang, B. Qu, J. Yang, and P. Cudré-Mauroux, "Revisiting User Mobility and Social Relationships in LBSNs: A Hypergraph Embedding Approach", *WWW '19: The World Wide Web Conference*, San Francisco, USA, pp. 2147-2157, May 2019. <https://doi.org/10.1145/3308558.3313635>.
- [13] D. Yang, B. Qu, J. Yang, and P. Cudré-Mauroux, "Lbsn2vec++: Heterogeneous

- hypergraph embedding for location-based social networks," IEEE Trans. Knowl. Data Eng, Vol. 216, No. 6, Jul. 2016. <https://doi.org/10.1109/TKDE.2020.2997869>.
- [14] D. Yang "Global-scale Check-in Dataset with User Social Networks", <https://sites.google.com/site/yangdingqi/home/foursquare-dataset> [accessed: Jul. 13, 2024]
- [15] Foursquare systems overview, <https://docs.foursquare.com/developer/reference/places-api-overview> [accessed: Jul. 13, 2024]
- [16] Y. Cao and S. Lee, "Analysis of Urban Vitality and Its Determinant Factors Using POI Bigdata in Seoul, Korea", Journal of Korea Planning Association, Vol. 56, No. 7, pp. 87-102, Dec. 2021. <https://doi.org/10.17208/jkpa.2021.12.56.7.87>.
- [17] Korean Standard Industrial Classification (KSIC), Statistics Korea, http://kssc.kostat.go.kr/ksscNew_web/kssc/common/selectIntroduce.do?gubun=2&bbid=kssc_s&categoryNameCode=001&categoryMenu=002 [accessed: Jul. 13, 2024]
- [18] 2030 Seoul Community Plan, <https://urban.seoul.go.kr/view/html/PMNU2030000000> [accessed: Jul. 13, 2024]
- [19] Apartment Sale Prices per Square Meter. Hyundai Development Company, 2021. https://realty.chosun.com/site/data/html_dir/2021/04/16/2021041600394.html [accessed: Jul. 13, 2024]
- [20] I. Seo and J. Cho, "Clustering Foursquare Users' Collective Activities: A Case of Seoul", Korea Big Data Society, Vol. 5, No. 1, pp. 55-63, May 2020. <https://doi.org/10.36498/kbigdt.2020.5.1.55>.
- [21] E. Gil, Y. Ahn, and Y. Kwon, "Tourist Attraction and Points of Interest (POIs) Using Search Engine Data: Case of Seoul", Sustainability, Vol. 12, No. 17, pp. 7060, Aug. 2020. <https://doi.org/10.3390/su12177060>.
- [22] S. Kim and S. Lee, "Identifying Urban Economic Centers and Their Functions Using POI Big Data: The Case of Seoul Metropolitan Area", Journal of Korea Planning Association, Vol. 56, No. 6, pp. 36-52, Nov. 2021. <https://doi.org/10.17208/jkpa.2021.11.56.6.36>.
- [23] A. Son, J. Park, and S. Lee, "Analysis of Location Pattern and Determining Factors of New Industry in Seoul, Korea", Journal of the Korean Geographical Society, Vol. 57, No. 6, pp. 549-565, Dec. 2022. <https://doi.org/10.22776/kgs.2022.57.6.549>.
- [24] M. Kim, D. Paini, and R. Jurdak, "Real-world diffusion dynamics based on point process approaches: A review", Artificial Intelligence Review, Vol. 53, pp. 321-359, Jan. 2020. <https://doi.org/10.1007/s10462-018-9656-9>.
- [25] M. Kim, D. Paini, and R. Jurdak, "Modeling stochastic processes in disease spread across a heterogeneous social system," Proceedings of the National Academy of Sciences, Vol. 116, No. 2, pp. 401-406, Jan. 2019. <https://doi.org/110.1073/pnas.1801429116>.

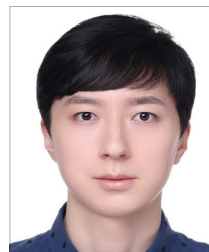
저자소개

오 석 준 (Seokjoon Oh)



2024년 2월 : 경희대학교
응용영어통번역학과(문학사)
2024년 3월 ~ 현재 : 경희대학교
일반대학원 빅데이터응용학과
석사과정
관심분야 : 빅데이터, 딥러닝,
그래프 마이닝

주 승 영 (Seungyoung Joo)



2024년 2월 : 경희대학교
국제학과(국제학사)
2024년 3월 ~ 현재 : 경희대학교
일반대학원 빅데이터응용학과
석사과정
관심분야 : 인공지능, 빅데이터,
그래프 마이닝

김 수 환 (Soohwan Kim)



2015년 11월 : The Australian National University(공학박사)
2023년 3월 ~ 현재 : 광운대학교
인공지능융합대학 정보융합학부
교수
관심분야 : 인공지능, 지능로봇,
모빌리티, 자율주행자동차

김 민 경 (Minkyung Kim)



2015년 6월 : The Australian National University(공학박사)
2022년 9월 ~ 현재 : 경희대학교
경영대학 빅데이터융용학과
교수
관심분야 : 인공지능, 그래프
머신러닝, 정보이론, 소셜 컴퓨팅