

대규모 동적 O/D 생성을 위한 추정 방법론 연구: 침두 출근통행을 기준으로

한혁¹ · 홍기만¹ · 김태균¹ · 황준문² · 홍영석^{3*} · 조중래¹

¹명지대학교 교통공학과, ²경기연구원 휴먼교통연구실, ³명지대학교 산업기술연구소

A Methodology for Estimating Large Scale Dynamic O/D of Commuter Working Trip

HAN, He¹ · HONG, Kiman¹ · KIM, Taegyun¹ · WHANG, Junmun² · HONG, Young Suk^{3*} · CHO, Joong Rae¹

¹Department of Transportation Engineering, Myongji University, Gyeonggi 17058, Korea

²Office of Transportation Policy Research, Gyeonggi Research Institute, Gyeonggi 16207, Korea

³Industrial Technology, Myongji University, Gyeonggi 17058, Korea

*Corresponding author: hahnhyok@mju.ac.kr

Abstract

This study suggests a method to construct large scale dynamic O/D reflecting the characteristic that the passengers' travel patterns change according to the land use patterns of the destination. There are limitations in the existing research about dynamic O/D estimation method, such as the difficulty of collecting data, which can be applied only to a small area, or limiting to a specific transportation network such as highway networks or public transportation networks. In this paper, we propose a method to estimate dynamic O/D without limitation of analysis area based on transportation resources that can be easily collected and used according to the big data era. Clustering analysis was used to calculate the departure time trip distribution ratio based on arrival time and departure time trip distribution function was estimated by each cluster. As a result of the comparison test with the survey data, the estimated distribution function was statistically significant.

Keywords: commuter trip, departure time distribution function, dynamic O/D, land use patterns, working trip

초록

본 연구는 통행자의 통행패턴이 도착지의 토지이용패턴에 따라 변화하는 특징을 반영하여 대규모 동적 O/D를 구축하는 방법을 제안하였다. 기존 동적 O/D 추정 방법 관련 연구들을 살펴보면 자료수집의 어려움으로 소규모 지역에 국한하여 적용 가능하거나, 고속도로망 혹은 대중교통망 등 특정 교통망에 제한하여 O/D를 구축하는 등 한계가 존재한다. 이에 본 연구에서는 빅데이터 시대에 발맞추어 쉽게 수집, 이용이 가능한 교통 관련 자료들을 기반으로 분석 지역의 범위 제약 없이 동적 O/D를 추정하는 기법을 제시하였다. 군집 분석(Clustering Analysis) 기법을 이용하여 군집별 도착시간 기준 통행분포 비율로 출발시간 통행분포 비율을 산정하고 침두 출발시간 분포함수를 추정하였다. 추정된 분포함수를 조사자료에 적용하여 비교 검증을 진행해본 결과 통계적으로 유의하게 나타났다.

J. Korean Soc. Transp.
Vol.36, No.3, pp.203-215, June 2018
<https://doi.org/10.7470/jkst.2018.36.3.203>
pISSN : 1229-1366
eISSN : 2234-4217

ARTICLE HISTORY

Received: 9 March 2018

Revised: 1 May 2018

Accepted: 15 May 2018

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.

주요어: 침투 O/D, 출발시간 분포함수, 동적 O/D, 토지이용패턴, 출근 통행

서론

1. 연구의 배경 및 목적

최근 빅데이터 시대의 도래로 많은 분야에서 다양한 정보와 데이터를 제공하고 활용한다. 교통 분야에서도 통행자의 특성에 따른 교통시스템의 변화를 더 구체적인 수준까지 추정하고 실제 정책 결정에 활용하고자 하는 연구가 활발히 진행되고 있다.

그중 3세대 실시간 신호제어 분야에서는 교통과 관련된 정성적, 정량적 데이터를 통합한 효율적인 빅데이터뱅크 시스템을 이용하여 교통정보를 생성, 제공하고 빅데이터 분석을 통한 교차로 통행패턴 생성 및 실시간 교통정보를 이용한 예측 신호시스템 개발을 목적으로 한다. 예측신호 시스템은 짧은 미래(5-15분 후)에 대한 Real-time O/D를 이용해 교차로 회전 교통량을 예측하고 이를 통해 혼잡을 사전에 완화 할 수 있는 최적신호를 산정하여 중앙시스템을 통해 현장을 제어하는 시스템 개발이다. 본 연구에서는 이러한 신호제어 시스템에서 사용할 수 있는 Real-time O/D 산정에 관한 방법론을 제시할 예정이다.

현재 거시적인 교통수요분석에서도 물론, 미시교통류 시뮬레이션으로 동적 현상을 분석하기 위해 더욱 상세한 도로교통 가로망에 대한 정보, 교통운영 자료 등 교통 데이터가 필요하게 되기 마련이다. 특히 시간대별 기종점 통행자료, 즉 시간대별 O/D 구축을 위한 수학적이고 학술적 연구가 과거부터 이루어져 왔다. 하지만 아직까지 실무적인 응용이 불가하며 활용도가 높은 방법론이 구축되지 못한 상황이기 때문에 현실적 정책분석 적용에 한계가 있다. 이러한 원인에도 불구하고 미시적 정밀분석에 대한 정책분석의 필요성은 점차 커지고 있으며, 정책결정 과정에서 더욱 그 효용성이 증대할 것으로 기대되고 있는 추세이다. 거시적 시스템 분석결과와 연계된 미시적 상세분석에 대한 연구도 국제적으로 관심을 받고 있다.

2. 연구의 범위 및 방법론

현재까지 국가교통DB센터(KTDB)에서는 1998년 전국지역 간 조사를 시작으로 광역권 조사, 수도권 조사, 육상 및 해상 교통조사, 교통시설물 조사 등을 실시하여 O/D자료를 구축하고 있다. 이렇게 구축된 정적 O/D는 실시간이 아닌, 과거 출발지 및 도착지 정보에 대한 O/D자료로 거시적인 분석이 가능하지만 미시적 정밀분석으로 사용하기에는 적합하지 않다. 또한 실시간 동적 교통정보로 구축되는 자료는 통행정보, 경로 등 교통흐름에 국한되어 통행자의 출발지와 도착지 정보 및 통행목적에 대한 데이터 수집방법은 알 수 없다.

본 연구에서는 우선 국가통계포털에서 제공하는 전국사업체조사 중 읍면동별 산업대분류별 현황 자료를 이용하여 읍면동 행정구역 존별 산업체 종사자수로 군집분석(Clustering Analysis)을 수행한 후 도출된 군집을 기준으로 2014년 가구통행실태조사 자료에서 군집별 출근통행 도착시간 분포비율을 산출한다. 다음 KTDB에서 배포하는 수도권 네트워크로 통행배정을 수행하여 산출한 기종점간 평균통행시간을 이용해 출근통행 출발시간 분포비율을 추정하고, 추정된 분포비율을 현재 KTDB에서 배포하고 있는 주수단 전일 O/D에 적용하여 출근 시간대별 O/D를 추정한다. 그다음 추정한 시간대별 O/D로 출발시간 변화를 반영한 침투 시간대 출근 통행 출발시간 분포함수를 제시하였다. 본 연구에서는 승용차 교통수단에 한하여 분석을 진행하였으며, 전체 수행과정의 흐름도는 Figure 1과 같다.

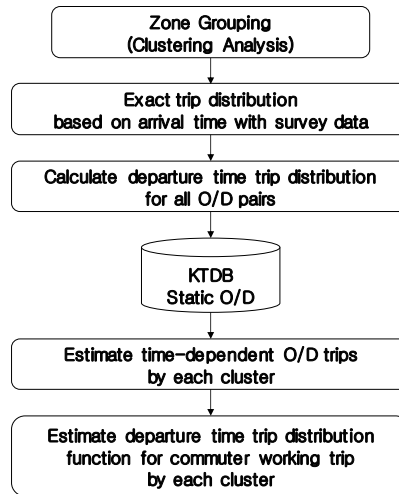


Figure 1. Flow chart of the methodology

기존 연구 고찰

Vickery(1969)는 출발시간 선택 모형을 효용함수로 변형시켜 연속적 로짓모형(Continuous Logit)을 사용하여 동적 O/D를 생성하였다. 이는 정적 O/D를 이용하여 동적 O/D를 추정하는 연구 중 하나이며, 또한 de Palma and Marchal(1999)는 통행비용을 자유속도 상태와 정체시를 구분하여 통행시간비용과 통행지체비용으로 구성하고 모형에 의해 결정된 출발시각으로 동적 O/D를 추정하였다. 비록 총량의 정적 O/D를 활용하였으나 출근 통행목적과 같이 통행자의 통행목적지 도착시간이 정해졌을 때 적용가능하다는 전제조건이 있다. 이러한 이론을 Marchal and de Palma(2001)는 실제 프랑스 파리에 응용한 사례도 보여주었다. 하지만 이 방법을 적용하기 위해 기초자료로 입력해야 하는 기대 도착시간 등 데이터가 필요한 현실적 한계점으로 실제 정책분석 사용에는 어려움이 있다. Nambisan and Pulugurtha(2001), Kim(2002)의 동적 O/D 추정방법도 출발시각과 도착시각에 대한 자료를 요구하거나 de Palma and Marchal(1999)와 마찬가지로 도착시각을 알고 있는 상태에서 출발시각 선택모형을 적용하여 추정하는 방법론으로 역시 실무상 자료의 한계를 가지고 있다.

Choi et al.(2009)은 엔트로피를 포함시킨 목적함수로 경로기반 노선배정기법을 이용하여 광역 범위 중 부분지역에 대한 정적 O/D를 추정하고 이러한 과정의 반복수행으로 시간대별 관측 지점 교통량과 유사하게 할 수 있는 동적 O/D 추정방법을 제안하였다.

Kim et al.(2005)은 휴대폰 기지국 기반 O/D를 행정동 기반 O/D화하는 방법으로 시간대별 목적별 수단별 O/D 산출이 가능한 방법론을 제시하였다. 하지만 실제 정책분석에 적용할 수 있는 O/D를 구축하려면 전국 범위 내 모든 기지국에 대한 자료수집 및 특정 산출방법을 통해 행정동 기반 O/D로 전수화를 진행하여야 하는 등 현실적으로 많은 제약을 가지고 있다.

Lee et al.(2014)의 연구에서는 거시적 상용프로그램을 이용하여 먼저 거시적 분석을 통해 전체와 부분 지역 간의 교통패턴 관계를 분석한 다음 가구통행실태조사와 지점 교통량 조사자료를 통행배정결과와 연계하여 부분 국소 지역 동적 O/D를 추정하였다. 하지만 미시적 분석을 위한 상세 네트워크를 우선 구축해야 하고, 부분상세지구(Sub-area) 경계에 해당하는 특정 링크에 대해 Selected Link Analysis 분석을 진행하여 통행패턴을 파악하고 분석해야 하는 등 과정을 먼저 수행해야 하는 한계가 있다.

본 연구에서는 쉽게 수집, 이용이 가능한 교통 관련 데이터로 전체 분석 지역에 대한 거시적인 분석을 진행함과 동시에 모든 지역 간 통행패턴 분석 혹은 부분 상세 지역 정밀분석이 가능하도록 첨두 시간대 출발시간 분포함수를 추정하는 방법론을 제시하고자 한다.

존 특성 분석 및 그룹화

1. 산업별 출근통행 특성 및 종사자 규모를 반영한 교통존의 군집분석

기종점간 시간대별 O/D 추정을 진행하기 위해 먼저 각 존의 토지이용 및 출근 통행 분포 특성에 대해 그룹화를 수행해야 한다. 우선 국가통계포털에서 제공한 읍면동별 산업대분류별 현황 자료에서 2014년에 배포한 데이터를 이용해 읍면동 행정구역별 종사자수를 추출하였다. 다음 산업체분류별 특성 상 오전에 출근하는 산업, 고정 출근 시간대가 없는 산업, 2-3교대 산업으로 재분류하여 존별 출근 통행 특성별 종사자수를 재집계한다. 재집계한 결과로 군집분석을 수행하기 위해 오전에 출근하는 산업, 고정 출근 시간대가 없는 산업, 2-3교대 산업을 각각 변수 $\times 1$, $\times 2$, $\times 3$ 로 정의하고 Table 1과 같이 산업체를 분류한다.

Table 1. Define variables by industry for clustering analysis

All industries	Variables
Agriculture, forestry and fishing (01-03)	$\times 2$
Mining and quarrying (05-08)	$\times 2$
Manufacturing (10-33)	$\times 1$
Electricity, gas, steam and water supply (35-36)	$\times 1$
Sewage & waste treatment, material recovery and restoration activities of environment (37-39)	$\times 1$
Construction (41-42)	$\times 1$
Wholesale and retail sale (45-47)	$\times 2$
Transportation (49-52)	$\times 3$
Accommodation and food service activities (55-56)	$\times 2$
Publishing, video, broadcast communications and information services (58-63).	$\times 1$
Financial service and insurance activities (64-66)	$\times 1$
Real estate activities and renting and leasing (68-69)	$\times 1$
Professional, scientific and technical activities (70-73)	$\times 1$
Business facilities management and business support services (74-75)	$\times 1$
Public administration and defence, compulsory social security (84)	$\times 1$
Education (85)	$\times 2$
Human health and social work activities (86-87)	$\times 1$
Arts, sports and recreation related services (90-91)	$\times 2$
Membership organizations, repair and other personal services (94-96)	$\times 1$

Table 1에서 알 수 있듯이 농업, 임업 및 어업(01-03), 광업(05-08), 도매 및 소매업(45-47), 숙박 및 음식점업(55-56), 교육 서비스업(85), 예술, 스포츠 및 여가관련 서비스업(90-91) 등 6개 산업체는 고정 출근 시간대가 없는 산업($\times 2$)으로, 운수업(49-52)은 2-3교대 산업($\times 3$)으로, 기타 12개 산업체를 오전에 출근하는 산업($\times 1$)으로 분류하여 총 3개 변수를 설정하였다. 그리고 국가통계포털에서 제공한 2014년 읍면동 행정구역 존 코드와 KTDB에서 배포하는 행정구역 존 코드를 수도권을 기준으로 매칭한 총 1120개 존으로 재분류한 종사자수 비율을 산출하여 군집분석을 수행하였다.

본 연구의 군집분석에서 수행할 샘플 수가 비교적 많은 데이터 특성 상, 군집분석 방법 중 k-means 분할방법을 채택하여 진행하였다. k-means 분할은 군집의 평균이나 중심 값으로 객체들에서 k (초기 Cluster 설정 값)를 임의로 추출하고 남겨진 객체들을 객체와 군집 평균에 기초해서 가장 가까운, 즉 유사한 군집에 할당하는 군집분석 방법 중 하나이다. 분석 방법은 SAS 9.4 (Statistical Analysis System) 프로그램의 Fastclus을 적용하였다. 해당 모듈은 상호배반적(Disjoint) 군집분석으로 Euclidean distances를 이용한 최소 자승 추정법을 기반으로 최적 초기 값을 결정하여 Clustering을 진행하는 방법이다. 해당 k-means 분할방법을 이용한 군집분석 수행 결과는 Table 2와 같다.

Table 2. Result of clustering analysis (1)

Cluster	Frequency	RMS std deviation	Maximum distance from seed to observation	Nearest cluster	Distance between cluster centroids
1	398	0.0489	0.1821	2	0.2339
2	472	0.0536	0.2020	1	0.2339
3	56	0.0432	0.1542	1	0.2351
4	69	0.0708	0.3067	1	0.2542
5	1	-	0	4	0.6911
6	124	0.0480	0.1771	2	0.2675

Table 2에서 알 수 있듯이 5번째 군집에 하나의 존만 포함되어 가장 가까운 4번째 군집으로 병합하여 최종 5개의 군집으로 분류하였다. 분류 결과 Cubic Clustering Criterion은 15.876으로 나타났고, 해당 값이 2보다 클 경우 분류가 잘 되었다는 것을 설명한다. 각 군집에 포함되는 존의 개수와 각 존의 산업분류별 종사자수 비율 범위는 Table 3과 같다.

Table 3. Result of clustering analysis (2)

Cluster	No. of zones	×1	×2	×3
1	398	36.4-46.6%	46.8-57.6%	2.2-10.4%
2	472	52.3-64.5%	30.1-42.1%	1.8- 9.2%
3	56	22.0-31.5%	65.1-75.4%	0.3- 5.8%
4	70	33.0-49.0%	28.5-40.5%	17.4-31.6%
5	124	72.2-84.0%	13.0-23.1%	0.8- 6.9%

2. 도착시간 기준 출근 통행 분포 산출

출근 통행의 특성 상 목적지 기준으로 보면 출근 출발시간은 기종점간 통행시간의 함수로 표현되지만, 출근 도착 시간은 큰 변화가 없는 특징이 있다. 때문에 본 연구에서는 행정구역 존에 대해 도착시간 기준으로 출근 통행 분포를 구축한 후 기종점별로 출발시간 기준 출근 통행 분포를 각각 산출하여 시간대별 O/D를 산정하기로 하였다.

행정구역 존의 토지이용패턴을 이용한 군집분석 결과로 산출된 5개 군집으로 출근 통행 도착시간 기준 통행 분포를 산출하기 위해, 먼저 2014년 가구통행실태조사 자료에서 행정구역 기종점간 출근목적 전일 통행량을 집계하였다. 다음 도착시간대별로 해당 시간대 통행량 비율을 각각 산출하여 군집별로 저장한다. 도착시간 기준 m시간대 k군집의 출근 통행 분포 비율 A_k^m 을 수식으로 표현하면 Equation 1과 같다.

$$A_k^m = \frac{\sum_i^N V_{ik}^m}{\sum_t^T \sum_i^N V_{ik}^t} \quad (1)$$

여기서, N : 존 개수

V_{ik}^m : 도착시간 기준 m시간대 i존에서 k군집에 포함된 존까지 통행량

$\sum_t^T V_{ik}^t$: i존에서 k군집에 포함된 존까지 전일 통행량의 합

가구통행실태조사 자료에서 출발/도착시간을 5분 단위로 승용차 수단 통행량을 집계하여 수식과 같은 방법으로 도착시간 기준 군집별 전일 출근 통행 분포를 그래프로 표출하면 Figure 2와 같다.

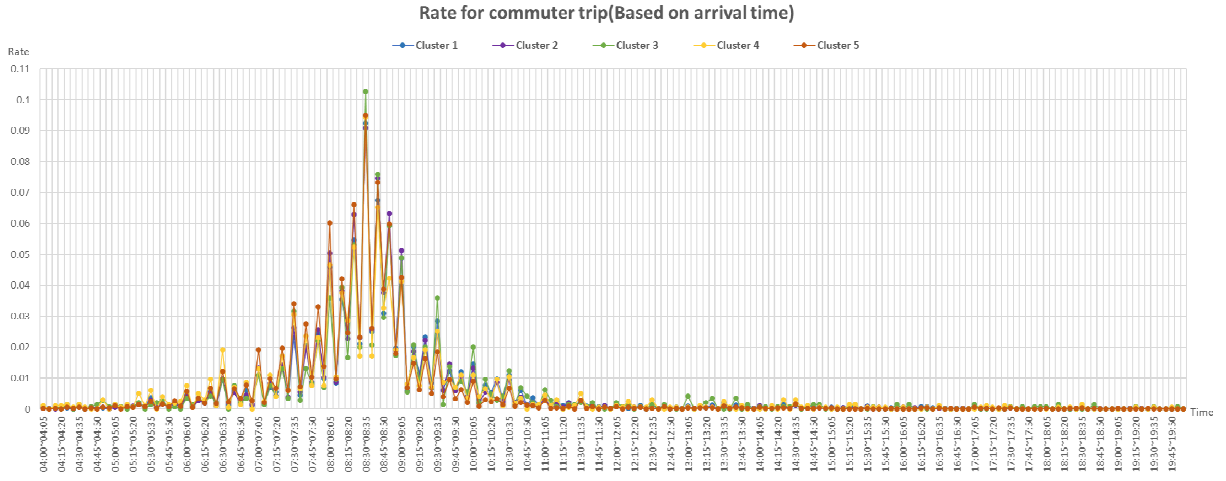


Figure 2. Rate for commuter trip by arrival time (04:00-20:00)

도착시간 기준으로 출근 통행을 살펴보면 8시 30분에서 8시 35분 시간대에 최대 통행 비율이 산출되며 3번째 군집이 10.27%로 제일 많은 비율을 나타내며, 2번째 군집이 9.08%로 제일 낮은 비율로 나타난다. 그 외 1번째, 4번째, 5번째 군집은 각각 9.23%, 9.45%, 9.49%로 차이가 1% 내외로 도출되며 통행 분포 패턴은 대체적으로 유사한 모양을 가지고 있다.

현재 그래프에 fluctuation이 다소 심한 것으로 나타나는데 이는 가구통행실태조사 자료 특성 상 일반적으로 출발/도착시간을 5분, 10분 등 단위로 대략적인 시간을 기입하기 때문에 발생하는 것임을 알 수 있다. 이러한 fluctuation을 완화하기 위해 이동평균(Moving Average)법을 적용하였다. 계산너비를 5로 지정하여 적용한 결과, 군집별 도착시간 기준 통행 분포 비율 그래프는 Figure 3과 같다.

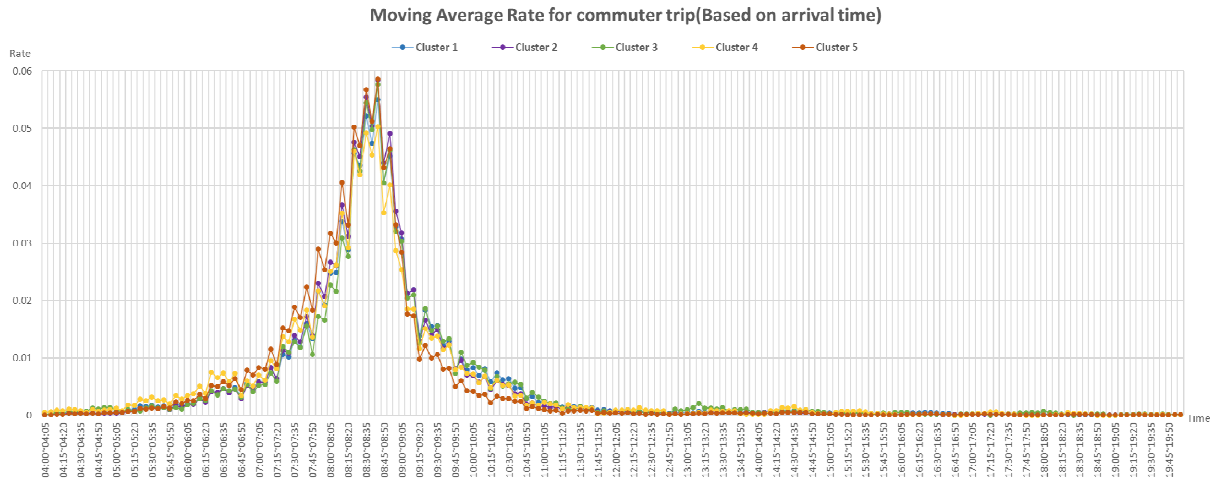


Figure 3. Moving average rate for commuter trip by arrival time (04:00-20:00)

이동평균법 적용 후 그래프가 비교적 원활하게 나타났고 8시 40분에서 45분 시간대에 군집별 최대 통행 비율이 산출되었으며 그 비율은 각각 5.50%, 5.84%, 5.76%, 5.04%, 5.86%로 계산되었다.

첨두 출근 통행 출발시간 분포 함수 추정

1. 출발시간 분포 비율 및 시간대별 O/D 추정방법

출근 통행 목적 기점 존에 대한 출발시간 분포함수를 추정하기 위해 먼저 출발시간 통행 분포 비율을 산정하였다. 각 기종점간 출발시간 기준 출근 통행 분포는 각 군집별 도착지 존에 대한 도착시간 통행 분포와 기종점간 통행시간의 함수로 표현할 수 있으며, 그 수식은 Equation 2와 같다.

$$D_{ij} = A_j - t_{ij} (\forall i, j \in N) \quad (2)$$

여기서, D_{ij} : 존 i에서 j까지의 출근 통행 출발시간 분포

A_j : 도착지 존 j의 출근 통행 도착시간 분포

t_{ij} : 존 i에서 j까지의 통행시간

즉 기종점간 출발시간 분포 비율 그래프는 목적지 기준 도착시간 분포 비율 그래프를 통행시간 만큼 음수 방향으로 이동한 것임을 알 수 있다. 그림으로 표현하면 Figure 4와 같다.

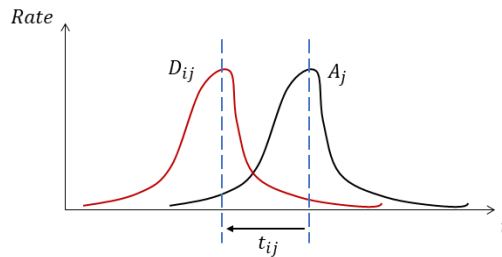


Figure 4. Graph for calculating D_{ij}

Equation 2 중 t_{ij} 는 기종점간 평균통행시간으로 가정하고 KTDB에서 배포하는 2015년 수도권 네트워크와 전일 O/D 데이터로 통행배정을 수행하여 적용하였다. 오전 첨두 시간대 출근 목적 평균통행시간이 120분 이상인 데이터는 이상치로 판단하고 아래 산출 과정에서 모두 제외하였다.

위와 같은 방법으로 산출된 출발시간 기준 출근 통행 분포 비율을 KTDB 배포 자료 중 출근 통행 목적 전일 O/D에 적용하여 시간대별 출근 통행 O/D를 산정한다. 그 과정을 수식으로 표현하면 Equation 3과 같다.

$$V'_{ij}^t = V_{ij} \times P_{ij}^t \quad (3)$$

여기서, V'_{ij}^t : 출발시간 기준 t시간대 존 i에서 j까지의 추정 통행량

V_{ij} : 존 i에서 j까지의 전일 통행량

P_{ij}^t : 출발시간 기준 t시간대 존 i에서 j까지의 출근 통행 분포 비율

Equation 3을 이용하여 각 군집별 시간대별 O/D를 모두 산출한 후, 우선 첫 번째 군집을 기준으로 산출한 시간대별 O/D를 이용해 기종점간 평균통행시간을 30분 간격으로 최대 120분까지 총 4개 범위로 구분하여 집계해 보았다. 그 결과 출발시간 통행 분포 비율은 Figure 5와 같은 그래프로 나타난다.

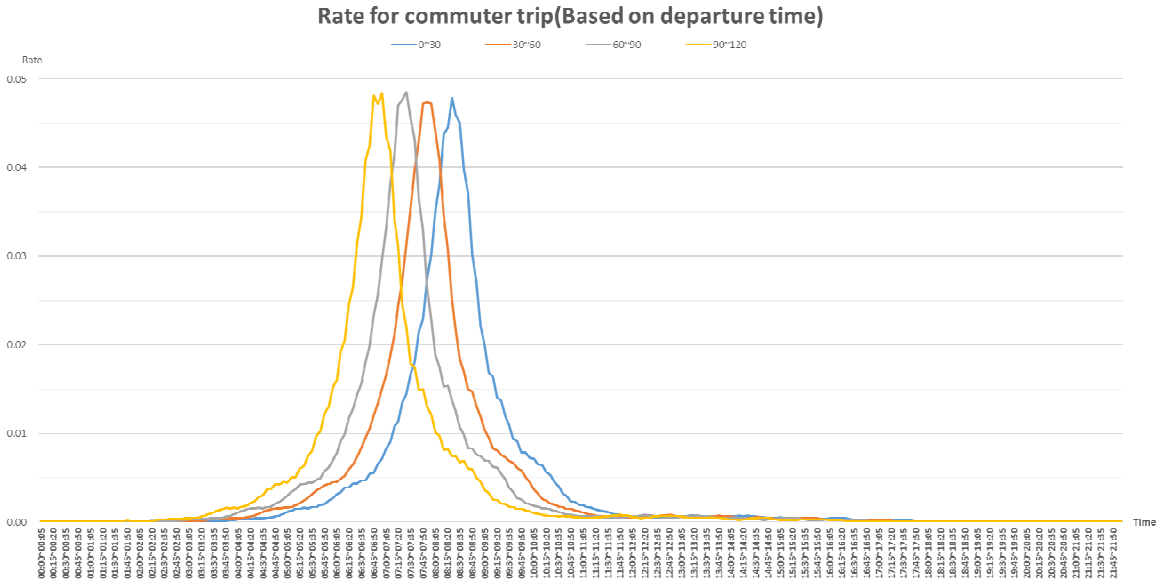


Figure 5. Rate for commuter trip by departure time and travel time (00:00-22:00)

Figure 5에서 알 수 있다시피 Equation 2의 적용으로 인해 기종점간 통행시간이 클수록 출발시간 통행분포 비율이 음수 방향으로 더 많이 이동한 것을 알 수 있다. 또한 산출한 시간대별 O/D 중 오전 첨두 시간대를 기준으로 각 군집별 전일 총 O/D 통행량 대비 각 5분 단위 시간대 통행 분포 비율을 계산해 보면 Table 4와 같다.

Table 4. Estimated time-dependent O/D for commuter trip by each cluster

Time	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5	
	O/D trips	Percentage	O/D trips	Percentage	O/D trips	Percentage	O/D trips	Percentage	O/D trips	Percentage
07:00-07:05	14,960.2	1.07%	24,772.8	1.09%	1,262.1	1.02%	2,400.9	1.18%	13,615.0	1.43%
07:05-07:10	16,655.6	1.19%	27,784.1	1.22%	1,406.6	1.14%	2,684.6	1.32%	14,945.4	1.57%
07:10-07:15	18,850.6	1.35%	31,702.5	1.40%	1,595.6	1.29%	3,029.5	1.49%	16,959.4	1.78%
07:15-07:20	20,739.4	1.48%	35,000.9	1.54%	1,766.6	1.43%	3,357.5	1.65%	18,367.5	1.92%
07:20-07:25	23,668.3	1.69%	40,300.0	1.78%	2,028.3	1.64%	3,744.6	1.84%	20,961.8	2.20%
07:25-07:30	25,888.2	1.85%	44,134.1	1.95%	2,189.7	1.77%	4,033.0	1.98%	22,448.9	2.35%
07:30-07:35	28,790.3	2.06%	49,173.0	2.17%	2,423.5	1.96%	4,356.5	2.14%	24,712.3	2.59%
07:35-07:40	31,317.8	2.24%	53,603.0	2.36%	2,638.2	2.13%	4,645.3	2.29%	26,173.4	2.74%
07:40-07:45	34,729.3	2.49%	59,691.7	2.63%	2,936.9	2.37%	5,023.2	2.47%	28,823.2	3.02%
07:45-07:50	37,524.9	2.69%	64,442.7	2.84%	3,174.3	2.57%	5,400.3	2.66%	30,307.1	3.18%
07:50-07:55	41,673.7	2.98%	72,135.3	3.18%	3,586.8	2.90%	5,963.3	2.93%	33,567.1	3.52%
07:55-08:00	44,038.5	3.15%	76,576.8	3.38%	3,895.3	3.15%	6,364.3	3.13%	34,636.0	3.63%
08:00-08:05	47,647.5	3.41%	82,910.7	3.65%	4,248.0	3.43%	6,859.7	3.38%	37,127.8	3.89%
08:05-08:10	50,303.2	3.60%	87,410.1	3.85%	4,499.6	3.64%	7,266.1	3.58%	38,135.6	4.00%
08:10-08:15	53,484.7	3.83%	93,348.5	4.11%	4,791.0	3.87%	7,613.8	3.75%	40,460.6	4.24%
08:15-08:20	54,665.6	3.91%	94,534.8	4.17%	4,937.4	3.99%	7,733.5	3.81%	40,065.4	4.20%
08:20-08:25	57,567.1	4.12%	99,356.8	4.38%	5,143.5	4.16%	7,977.4	3.93%	42,048.6	4.41%
08:25-08:30	55,962.1	4.01%	96,511.1	4.25%	4,986.0	4.03%	7,655.0	3.77%	40,048.4	4.20%
08:30-08:35	54,624.5	3.91%	94,608.7	4.17%	4,850.5	3.92%	7,313.7	3.60%	38,971.2	4.08%
08:35-08:40	50,413.0	3.61%	86,783.2	3.83%	4,455.7	3.60%	6,680.4	3.29%	34,895.1	3.66%
08:40-08:45	46,635.8	3.34%	80,494.0	3.55%	4,050.1	3.27%	5,999.4	2.95%	31,952.6	3.35%
08:45-08:50	39,812.8	2.85%	67,994.6	3.00%	3,430.1	2.77%	5,077.3	2.50%	25,960.3	2.72%
08:50-08:55	35,870.2	2.57%	60,869.6	2.68%	3,074.3	2.49%	4,513.0	2.22%	22,643.0	2.37%
08:55-09:00	29,679.7	2.12%	49,131.2	2.17%	2,543.5	2.06%	3,723.1	1.83%	17,408.9	1.82%
Total	915,503.0	65.54%	1,573,269.9	69.35%	79,913.9	64.60%	129,415.8	63.69%	695,234.7	72.84%

Table 4를 살펴보면 오전 침두 시간대 출근 통행 중, 2-3교대 산업 종사자수 비율이 가장 큰 4번째 군집의 오전 침두 시간대 O/D 통행량 비율이 63.69%로 제일 작게 나타났고 그 다음으로 고정 출근 시간대가 없는 산업 종사자수 비율이 가장 큰 3번째 군집이 64.60%로 4번째 군집과 0.91% 차이가 났다. 또한 오전에 출근하는 종사자수 비율이 가장 큰 5번째 군집은 72.84%로 제일 크게 나타났다.

2. 군집(Cluster)별 출발시간 분포함수 추정

군집별 오전 침두 시간대 출근 통행 O/D를 산출한 후, Matlab 프로그램을 이용하여 각 군집별 오전 침두 시간대 O/D에 대해 Curve Fitting 기법으로 출근시간 분포비율 함수를 도출하였다. Matlab에서 2개의 변수를 Fitting하는 모수적 방법(Parametric Method)으로 다항식(Polynomial) 모형을 제공하며 그 일반화 공식은 Equation 4와 같다.

$$y = \sum_{i=0}^n \sum_{j=0}^m p_{ij} x_1^i x_2^j \quad (p_{ij} \in R) \tag{4}$$

여기서, $0 \leq n \leq 5, 0 \leq m \leq 5$, 즉 최고차수는 5차이며 $i + j \leq \max\{n, m\}$, 즉 각 항에 차수의 합은 최고차수보다 작아야 한다. 이와 같은 다항식 모형으로 군집별 Curve Fitting한 결과는 Table 5와 같다.

Table 5. Estimated departure time trip distribution for commuter trip by each cluster (1)

Cluster	Function $\hat{P}_{ij}^t = f(S_i, T_{ij})$
1	$-0.01483S_i + 0.04592T_{ij} + 0.00247S_i^2 + 0.0565S_iT_{ij} - 0.2055T_{ij}^2 - 0.0001102S_i^3 - 0.005724S_i^2T_{ij} + 0.3276T_{ij}^3 + (1.458e-06)S_i^4 + 0.0001661S_i^3T_{ij} + 0.003044S_i^2T_{ij}^2 + 0.004905S_iT_{ij}^3 - 0.1775T_{ij}^4 - (1.26e-06)S_i^4T_{ij} - (4.83e-05)S_i^3T_{ij}^2 - 0.0003876S_i^2T_{ij}^3 + 0.001344S_iT_{ij}^4 + 0.03072T_{ij}^5$
2	$-0.009908S_i - 0.05487T_{ij} + 0.001622S_i^2 + 0.05576S_iT_{ij} + 0.1165T_{ij}^2 - (3.596e-05)S_i^3 - 0.005554S_i^2T_{ij} - 0.04603S_iT_{ij}^2 - 0.03953T_{ij}^3 - (1.594e-06)S_i^4 + 0.0001582S_i^3T_{ij} + 0.003052S_i^2T_{ij}^2 + 0.009526S_iT_{ij}^3 + (4.534e-08)S_i^5 - (1.032e-06)S_i^4T_{ij} - (5.042e-05)S_i^3T_{ij}^2 - 0.0003689S_i^2T_{ij}^3$
3	$-0.01717S_i + 0.05047T_{ij} + 0.002766S_i^2 + 0.06327S_iT_{ij} - 0.2508T_{ij}^2 - 0.0001227S_i^3 - 0.006335S_i^2T_{ij} - 0.04563S_iT_{ij}^2 + 0.3956T_{ij}^3 + (1.634e-06)S_i^4 + 0.0001835S_i^3T_{ij} + 0.003389S_i^2T_{ij}^2 + 0.005123S_iT_{ij}^3 - 0.2129T_{ij}^4 - (1.412e-06)S_i^4T_{ij} - (5.334e-05)S_i^3T_{ij}^2 - 0.0004393S_i^2T_{ij}^3 + 0.001634S_iT_{ij}^4 + 0.03675T_{ij}^5$
4	$-0.007339S_i - 0.03272T_{ij} + 0.001363S_i^2 + 0.04397S_iT_{ij} + 0.08511T_{ij}^2 - (3.697e-05)S_i^3 - 0.004637S_i^2T_{ij} - 0.03643S_iT_{ij}^2 - 0.03037T_{ij}^3 - (1.093e-06)S_i^4 + 0.0001408S_i^3T_{ij} + 0.002493S_i^2T_{ij}^2 + 0.007583S_iT_{ij}^3 + (3.647e-08)S_i^5 - (1.059e-06)S_i^4T_{ij} - (4.265e-05)S_i^3T_{ij}^2 - 0.0002973S_i^2T_{ij}^3$
5	$-0.01523S_i + 0.4558T_{ij} + 0.002678S_i^2 + 0.06348S_iT_{ij} - 0.1849T_{ij}^2 - 0.0001247S_i^3 - 0.006507S_i^2T_{ij} - 0.05041S_iT_{ij}^2 + 0.3087T_{ij}^3 + (1.728e-06)S_i^4 + 0.000198S_i^3T_{ij} + 0.003609S_i^2T_{ij}^2 + 0.008437S_iT_{ij}^3 - 0.172T_{ij}^4 - (1.65e-06)S_i^4T_{ij} - (5.672e-05)S_i^3T_{ij}^2 - 0.0004916S_i^2T_{ij}^3 + 0.001009S_iT_{ij}^4 + 0.03004T_{ij}^5$

수식에서 \hat{P}_{ij}^t 는 통행분포비율이고 S_i 는 출발시간대(5분 단위로 부여된 순서번호, 예를 들어 07:00-07:05 시간대 S_i 값은 1이고 07:05-07:10 시간대 S_i 값은 2임), T_{ij} 는 기점 i에서 종점 j까지의 평균통행시간(단위는 시간)이다. Curve fitting으로 추정된 함수들의 통계적 검증 결과는 Table 6과 같다.

Table 6. Estimated departure time trip distribution for commuter trip by each cluster (2)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
SSE	338.393	488.237	53.550	59.311	109.670
R ²	0.8493	0.8480	0.8411	0.8313	0.8840
DF	10,371,149	12,182,647	1,415,787	1,851,127	2,930,645
Adj. R ²	0.8493	0.8480	0.8411	0.8313	0.8840
RMSE	0.0057	0.0063	0.0062	0.0057	0.0061

결과를 살펴보면 R²값이 모두 0.83 이상으로 비교적 높은 값으로 도출되었고 통계적으로 유의한 수준으로 판단 된다. 추정된 군집별 오전 첨두 출발시간 분포함수 중, R²값이 0.88로 가장 높은 5번째 군집에 대해 3차원 그래프를 그려보면 Figure 6과 같다.

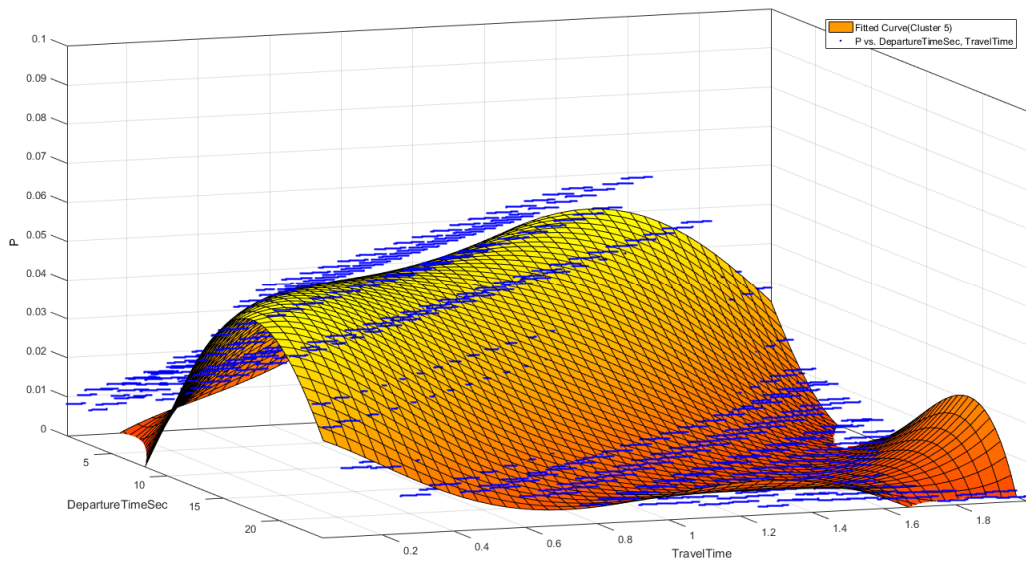


Figure 6. 3D surface graph of departure time trip distribution for commuter trip by cluster 5

Curve Fitting 기법으로 추정한 출발시간 분포함수를 적용하여 오전 첨두 5분 단위 시간대별 O/D를 산정하는 방법은 Equation 5와 같다.

$$\hat{V}_{ij}^t = V_{ij} \times \hat{P}_{ij}^{t,c} \tag{5}$$

여기서, \hat{V}_{ij}^t : t시간대 존 i에서 j까지의 추정 통행량

V_{ij} : 존 i에서 j까지의 전일 통행량

$\hat{P}_{ij}^{t,c}$: t시간대 존 i에서 c군집 도착지 존 j로 통행하는 출발시간 통행분포 비율

추정한 출발시간 통행분포함수를 검증하기 위해 우선 가구통행실태조사자료에서 출근 통행 목적 전일 O/D를 집계하였다. 다음 집계한 전일 O/D를 각각 Equation 5와 도착시간 기준 통행분포 비율(Figure 3)에 적용하여 오전 첨두 5분 단위 시간대별 O/D를 산정하고 비교해 보았다. 우선 시간대별(5분 단위로 7시부터 9시, 총 24개 시간대)로 상관관계 분석을 수행한 결과 Table 7과 같다.

Table 7. Survey O/D vs. Estimated O/D for commuter trip (1)

Time	Survey O/D		Estimated O/D		Pearson correlation coefficients
	O/D trips	Percentage	O/D trips	Percentage	
07:00-07:05	557.7	1.08%	174.9	0.34%	0.24188
07:05-07:10	623.5	1.21%	246.0	0.48%	0.17877
07:10-07:15	685.7	1.33%	358.2	0.70%	0.27553
07:15-07:20	773.8	1.50%	503.2	0.98%	0.33746
07:20-07:25	850.1	1.65%	672.8	1.31%	0.48923
07:25-07:30	932.0	1.81%	858.6	1.67%	0.66680
07:30-07:35	1,011.2	1.97%	1,052.5	2.05%	0.82171
07:35-07:40	1,112.6	2.16%	1,246.6	2.42%	0.90599
07:40-07:45	1,211.4	2.36%	1,433.4	2.79%	0.94770
07:45-07:50	1,351.3	2.63%	1,606.1	3.12%	0.95902
07:50-07:55	1,464.3	2.85%	1,758.3	3.42%	0.95970
07:55-08:00	1,593.9	3.10%	1,884.3	3.66%	0.96073
08:00-08:05	1,718.5	3.34%	1,979.7	3.85%	0.95588
08:05-08:10	1,858.2	3.61%	2,040.5	3.97%	0.95978
08:10-08:15	1,949.6	3.79%	2,064.4	4.01%	0.95795
08:15-08:20	2,075.7	4.04%	2,049.9	3.99%	0.97569
08:20-08:25	2,108.0	4.10%	1,996.9	3.88%	0.98699
08:25-08:30	2,093.6	4.07%	1,907.1	3.71%	0.98994
08:30-08:35	1,993.0	3.88%	1,783.4	3.47%	0.99180
08:35-08:40	1,861.1	3.62%	1,630.7	3.17%	0.99081
08:40-08:45	1,638.9	3.19%	1,455.5	2.83%	0.98983
08:45-08:50	1,448.2	2.82%	1,266.4	2.46%	0.99006
08:50-08:55	1,223.2	2.38%	1,073.9	2.09%	0.98739
08:55-09:00	1,029.4	2.00%	891.0	1.73%	0.98871
Total	33,164.9	64.50%	31,934.5	62.11%	-

분석 결과, 7시부터 7시 25분 사이 시간대에서 산정한 조사자료 O/D와 추정 O/D 상관계수는 0.5 이하로 약한 상관관계가 있는 것으로 나타났고, 7시 25분부터 7시 30분 시간대에 상관계수는 0.667로 보통의 상관관계가 있는 것으로 나타났다. 그리고 7시 30분부터 9시까지 기타 시간대는 상관계수가 0.8 이상으로 도출되었으며 이는 높은 상관관계가 있는 것으로 판단된다. 그 중 상관계수가 0.992로 상관관계가 가장 높은 8시 30분부터 8시 35분 시간대에 대하여 조사자료 O/D와 추정 O/D의 각 O/D 쌍에 대해 산포도를 그려보면 Figure 7과 같다.

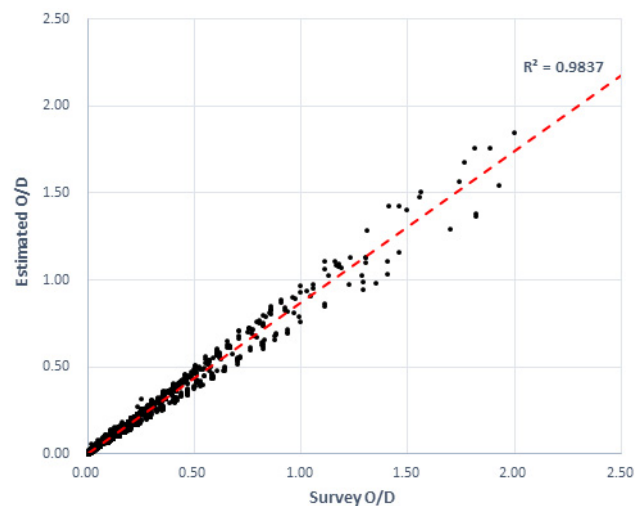
**Figure 7.** Survey O/D vs. Estimated O/D for commuter trip (08:00-08:30) (2)

Figure 7에서 알 수 있듯이 해당 시간대 조사자료 O/D와 추정 O/D의 산포도는 선형관계를 갖고 있으며 선형 예측결과에 해당하는 R^2 값은 0.984로 높게 산출되었다.

결론 및 향후과제

본 연구에서는 통행자의 기종점간 통행이 목적지 토지이용특성에 따라 통행패턴이 변화하는 특징을 반영하여 대 규모 동적 O/D를 구축하는 방법을 제안하는데 의의를 두고 있다. 또한 현재 수집 가능한 교통 관련 데이터를 최대한 활용하는 것에 목표를 두고 산업체분류 현황자료, 가구통행실태조사자료, 국가교통DB센터(KTDB)에서 배포하는 네트워크 및 O/D 자료 등을 이용하였다. 분석 방법으로는 군집분석(Cluster Analysis)방법을 이용한 그룹화 과정을 통해 우선 군집별 도착시간 기준 통행 분포 비율을 산출한 다음, 기종점간 출발시간 통행 분포 비율을 산출해내고 군집별 침투 출발시간 분포함수를 추정하는 방법론을 제시하였다. 결과 군집별 침투 출근 통행 출발시간 분포함수들의 R^2 값이 모두 0.84 이상으로 비교적 높은 값으로 산출되었으며 통계적으로 유의하다고 판단하였다. 그리고 이러한 방법으로 추정된 출발시간 분포함수를 가구통행실태조사자료에서 집계한 O/D로 침투 출근 통행 동적 O/D를 산출하여 실제 조사자료에서 추출한 O/D와 비교분석을 해보았다. 결과 대다수 시간대별로 비교적 높은 상관관계가 있는 것으로 나타났다.


향후 출근 목적뿐만 아니라 등교, 업무, 쇼핑, 여가운동 등 다양한 통행목적과 승용차 외 기타 교통수단으로 분석을 실행하여 목적별 수단별 시간대별 O/D 추정에 관한 연구진행이 필요하며, 따라서 빅데이터 기술을 이용해 더 많은 교통 관련 데이터 수집이 가능해지게 되면 더 정확한 토지이용 특성 분석이 가능해지고 더 세분화된 통행목적별 수단별 시간대별 O/D 자료 생성이 가능할 것으로 판단된다.


Funding


This research was supported by Projects for Research and Development of Police science and Technology under Center for Research and Development of Police science and Technology and Korean National Police Agency funded by the Ministry of Science, ICT and Future Planning (PA-A000001).


ORCID


HAN, He  <http://orcid.org/0000-0003-1192-579X>

HONG, Kiman  <http://orcid.org/0000-0003-1187-3918>

KIM, Taegyun  <http://orcid.org/0000-0002-6755-6340>

WHANG, Junmun  <http://orcid.org/0000-0001-6953-8982>

HONG, Young Suk  <http://orcid.org/0000-0001-6145-0863>

CHO, Joong Rae  <http://orcid.org/0000-0002-0314-5606>

References

- Choi C., Jayakrishnan R., Kim H., Yang I., Lee J. (2009), Dynamic Origin & Destination Estimation Using Dynamic Traffic Simulation Model in an Urban Arterial Corridor, 2133, Transportation Research Record, 133-141.
- de Palma A., Marchal F. (1999), Analysis of Travel Cost Components Using Large-scale Dynamic Traffic Models, 1676, Transportation Research Record, 177-183.

- Kim S. G., Yoo B. S., Kang S. P. (2005), Origin-Destination Estimation Based on Cellular Phone's Base Station, J. Korean Soc. Transp., 23(1), Korean Society of Transportation, 93-102.
- Kim S. (2002), Generating Dynamic O-D Table for Sub-area Analysis by Using System-wide Static O-D Table, Hanyang University, Master's Thesis.
- Lee J. H., Kim I. K., Kim D. H. (2014), Macroscopic-Microscopic Sequential Traffic Simulation Analysis and Dynamic O/D Estimation for Sub-area, J. Korean Soc. Transp., 32(6), Korean Society of Transportation, 567-578.
- Marchal F., de Palma A. (2001), Dynamic Traffic Analysis with Static Data: Guidelines with Application to Paris, 1752, Transportation Research Record, 76-83.
- Nambisan S. S., Pulugurtha S. S. (2001), Estimating Time-Dependent Origin-Destination Trip Tables with Trips Originating in Multiple Time Slices, 1752, Transportation Research Record, 133-139.
- Vickrey W. S. (1969), Congestion Theory and Transportation Investment, 59, The American Economic Review, 251-260.