

KOTI

April 2026

KTODB BRIEF

국가교통데이터베이스

Korea Transport Database



- 1 2025년 전국 여객
기종점통행량 예비조사
- 2 전국 도로 교통량 조사, AI가 대신한다:
인공지능과 모빌리티 빅데이터를 활용한 교통량 추정의 진화
- 3 모바일 통신데이터 기반
경로형 데이터 구축
- 4 서울 대중교통
1시간 통근권 분석





Contents

국가교통데이터베이스

1	2025년 전국 여객 기종점통행량 예비조사	03
2	전국 도로 교통량 조사, AI가 대신한다	10
3	모바일 통신데이터 기반 경로형 데이터 구축	19
4	서울 대중교통 1시간 통근권 분석	29

2025년 전국 여객 기종점통행량 예비조사

김동호 한국교통연구원 책임전문원 · 정승환 한국교통연구원 부연구위원 · 조영찬 한국교통연구원 연구원

전국 여객기종점통행량(O/D) 조사는 국가교통조사 중 하나로, 사람 및 여객교통수단의 통행행태 파악뿐 아니라 국가교통계획 수립 및 교통사업 타당성 조사를 위한 필수자료인 여객 O/D자료를 구축하기 위해 5년 주기로 시행되고 있다. 본 과제에서는 2026년 전국 여객 O/D 조사의 효율성과 신뢰성을 확보하기 위해 예비조사를 수행함으로써 조사항목, 표본설계, 조사방법 등을 사전에 검토하고 개선방안을 마련하였다.

01 개요

1) 조사의 배경 및 목적

- 전국 여객 기종점통행량 조사는 정기 국가교통조사 중 하나로, 사람 및 여객통행 수단의 통행행태를 파악하고 전국 여객 O/D자료를 구축하기 위해 5년을 주기로 시행
- 2025년 전국 여객 O/D 예비조사는 2026년 본 조사에 앞서 변화된 환경에 따라 조사항목, 표본설계, 조사방법 등을 사전에 검토하고 개선방안을 마련하는 것을 목적으로 수행

2) 조사의 범위

- 시간적 범위: 2025년 5월 ~ 2025년 6월 평일(화~목) 하루
- 공간적 범위: 서울특별시, 대구광역시, 청주시

3) 조사내용 및 물량

- 조사규모: 개인통행실태조사 10,871명, 고속도로 이용실태조사 576명
- 조사대상:
 - 개인통행실태조사: 만 5세 이상 서울, 대구, 충북 청주 거주자
 - 고속도로이용실태조사: 서울, 대구, 청주 인근 고속도로 요금소 이용(진입) 승용차

표 1 조사내용

구분		조사내용
개인통행 실태조사	가구특성	· 가구원수, 주택종류, 가구월평균소득, 차종별 보유대수 등
	개인특성	· 출생년도, 성별, 운전면허 보유여부, 학생여부, 직업, 근무일수 등 · 직장 및 학교 주소(건물명) 등
	통행특성	· 통행일자, 조사당일 통행유무, 무통행 사유 및 체류장소 · 출발지 및 도착지, 출발시간 및 도착시간, 통행목적, 통행수단, 탑승인원 등
고속도로 이용실태 조사	개인특성	· 출생년도, 성별, 직장 및 학교 주소 등
	통행특성	· 통행일자, 출발지 및 도착지, 진입 및 진출 요금소, 출발시간 및 도착시간, 통행목적, 탑승인원 등

4) 조사방법

- 로직을 적용한 자기기입식 비대면 모바일 조사

5) 조사검증

- 온라인 설문 로직 적용, 통신사 위치 데이터 검증, 기초분석을 통한 검증

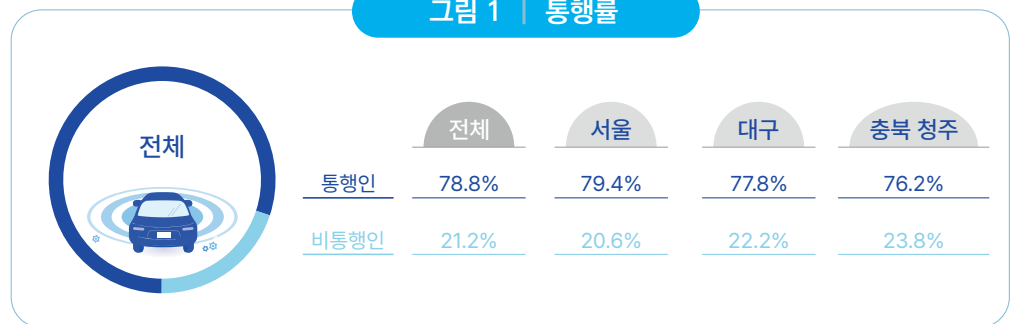
02

개인통행 실태조사 주요 결과

1) 통행률

- 전체 응답자의 78.8%는 조사기준일에 통행을 하였으며, 나머지 21.2%는 통행을 하지 않았다고 응답

그림 1 | 통행률



2) 통행목적 분포¹⁾

- 평일에 통행한 응답자의 통행목적은 '귀가'가 41.4%로 가장 많고, 이어 '출근(19.5%)', '등교(8.2%)' 순이며, 서울/대구/청주 지역별 통행목적 분포는 큰 차이 없었음

1) 기타는 귀사, 공공기관/금융/의료서비스 받으러, 누군가를 태우거나 내려주려고, 식사, 친지방문, 기타를 포함

그림 2 | 통행목적 분포 (개인통행실태조사)

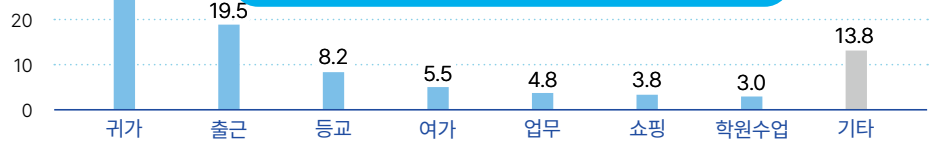
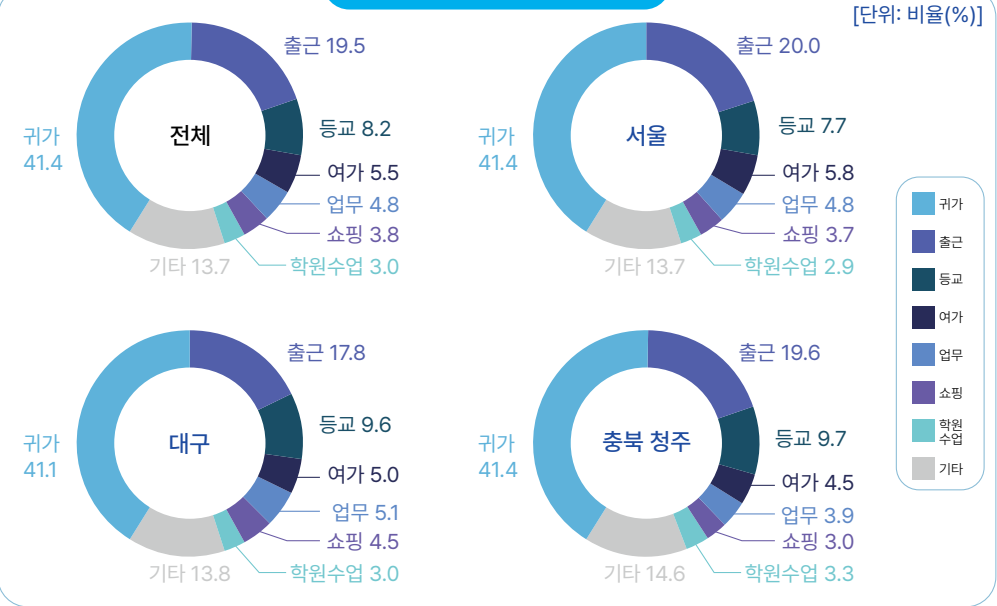


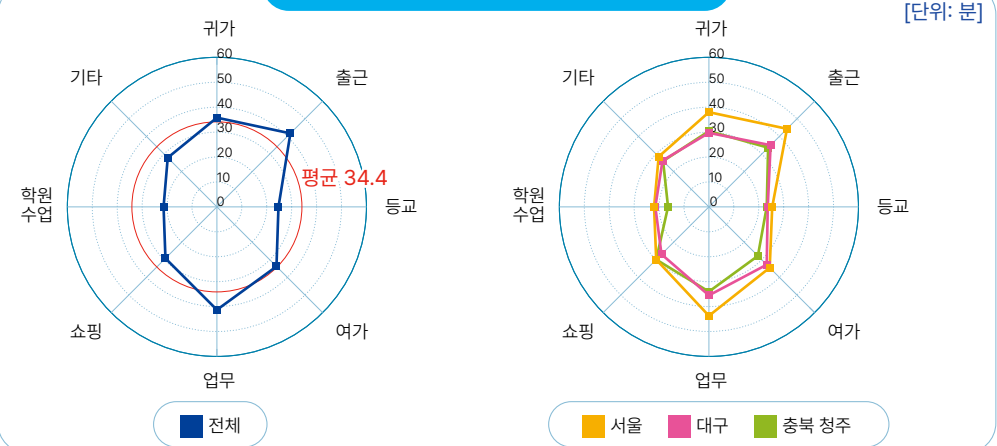
그림 3 | 통행목적 분포



3) 통행목적별 평균 통행시간²⁾

- 평균 통행시간은 34.4분이며, '출근(41.7분)', '업무(41.3분)' 등의 순으로 길게 나타남

그림 4 | 통행목적별 평균 통행시간



2) 모든 통행은 2시간 초과 아웃라이어를 제외한 평균값

표 2 통행목적별 평균 통행시간

(단위: 분)

	귀가	출근	등교	여가	업무	쇼핑	학원	기타	평균
전체	35.8	41.7	24.7	33.8	41.3	29.2	21.4	27.9	34.4
서울	38.0	44.3	25.4	34.5	43.7	29.9	22.0	28.5	36.3
대구	29.7	35.0	23.6	32.8	35.5	26.9	21.5	26.3	29.6
충북 청주	30.4	33.6	23.2	27.8	34.0	29.9	16.4	26.0	29.2

4) 통행수단 분포³⁾

- 응답자가 이동 시 주로 이용하는 교통수단은 '도보(32.3%)', '승용/승합(29.2%)', '지하철/전철/경전철(17.7%)' 등의 순으로 나타남
- 특히 서울은 '도보(33.4%)' 통행이 많은 반면, 대구와 충북 청주는 '승용/승합'의 이용 비율이 각각 47.6%, 52.1%로 가장 높게 나타남

그림 5 | 통행수단 분포 (개인통행실태조사)

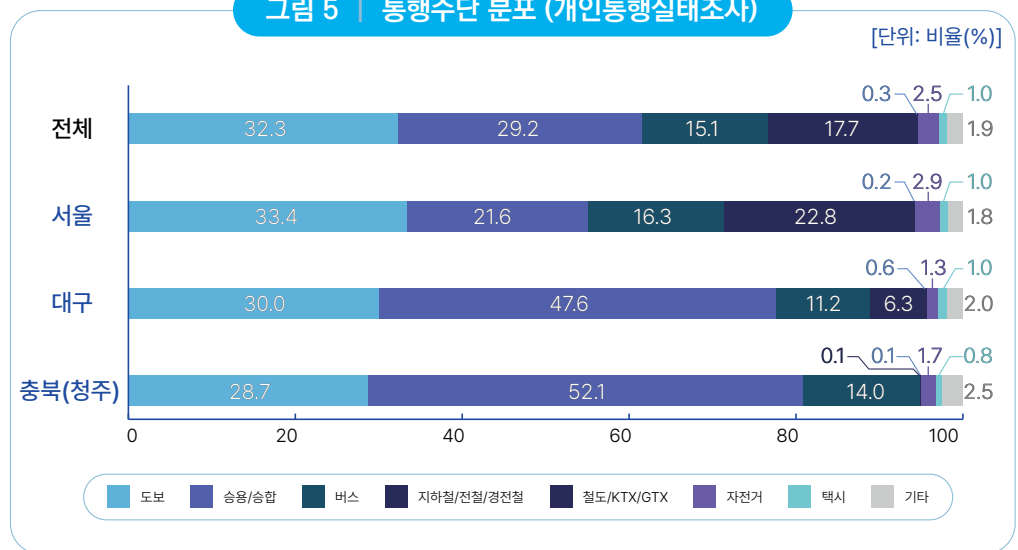


표 3 통행수단 분포

[단위: 비율(%)]

지역	도보	승용/승합	버스					지하철/전철/경전철	철도/KTX/GTX	자전거	택시	기타
			소계	시내/광역	시외/고속	수요응답형	기타					
전체	32.3	29.2	15.1	12.1	0.6	0.3	2.1	17.7	0.3	2.5	1.0	1.9
서울	33.4	21.6	16.3	13.7	0.6	0.3	1.8	22.8	0.2	2.9	1.0	1.8
대구	30.0	47.6	11.2	8.5	0.4	0.4	1.9	6.3	0.6	1.3	1.0	2.0
충북 청주	28.7	52.1	14.0	7.3	0.8	0.3	5.5	0.1	0.1	1.7	0.8	2.5

3) 수요응답형 버스는 2025년 추가된 항목이며, 기타는 화물차, 오토바이, 전동킥보드, 전동휠체어, 항공, 기타를 포함하여 산출

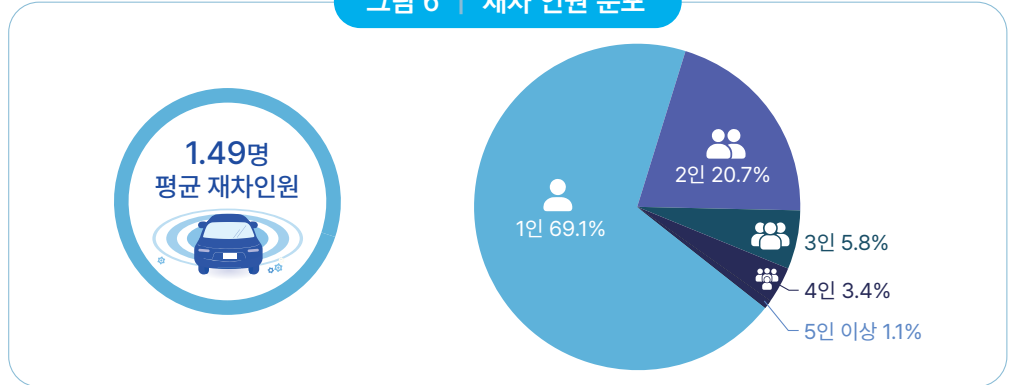
03

고속도로 이용실태조사 주요 결과

1) 재차인원 분포

- 고속도로 이용 승용차의 평균재차인원은 1.49명이며, 이 중 '1인'이 탑승한 승용차 비율이 69.1%로 가장 높았음

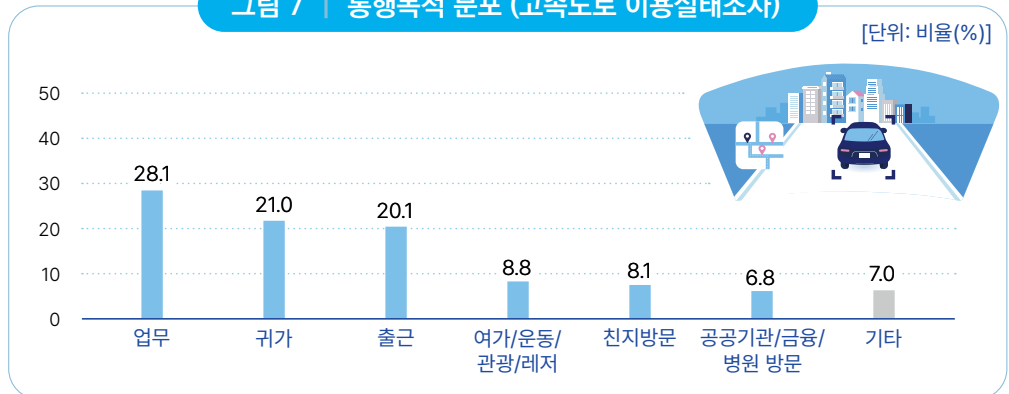
그림 6 | 재차 인원 분포



2) 통행목적 분포

- 통행목적은 '업무(28.1%)', '귀가(21.0%)', '출근(20.1%)' 등의 순으로 많았음

그림 7 | 통행목적 분포 (고속도로 이용실태조사)

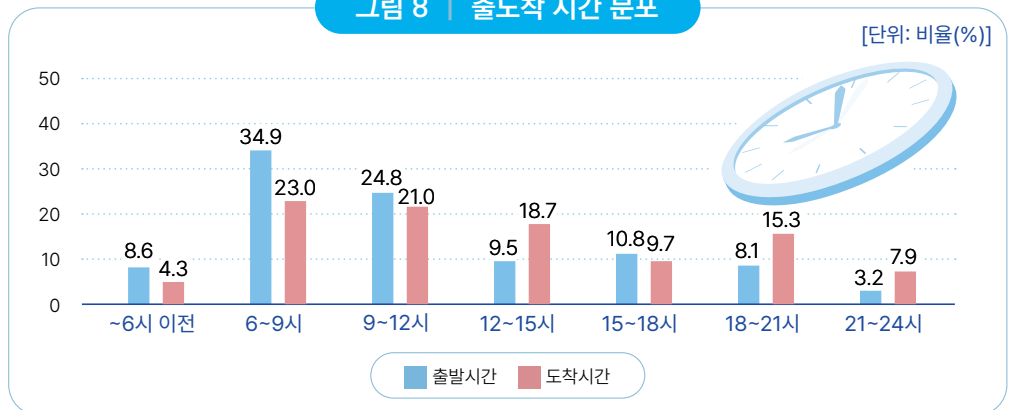


3) 출도착 시간분포

- 통행의 출발 시간대는 '오전(68.3%)'이 많았으며, 이 중 '6시~9시' 출발 비율이 34.9%로 가장 높음

- 도착 시간대는 '6시~9시(23.0%)', '9시~12시(21.0%)', '12~15시(18.7%)' 순으로 많았음

그림 8 | 출도착 시간 분포



04

주요 성과 및
향후 추진계획

1) 예비조사의 성과

- 신규 교통수단 등 최신 통행행태를 반영하여 조사표 개선
- 통신데이터와 통행실태조사자료를 연계 활용할 수 있는 조사방법론 마련
- 내비게이션 DB를 활용한 고속도로 이용실태조사 추진을 통해 표본의 대표성을 확보하고, 기존 고속도로 휴게소 조사의 한계점을 개선
- 전국 여객기종점 통행량 조사의 가이드라인과 논리검증 매뉴얼 마련

2) 2026년 전국 여객기종점통행량(O/D) 조사 추진계획

- 2026년 전국 여객조사는 예비조사 성과를 기반으로 조사설계 등의 과정을 거쳐 10월에 실시될 예정이며, 세부 추진일정과 조사내용은 아래와 같음

그림 9 | 조사추진 일정

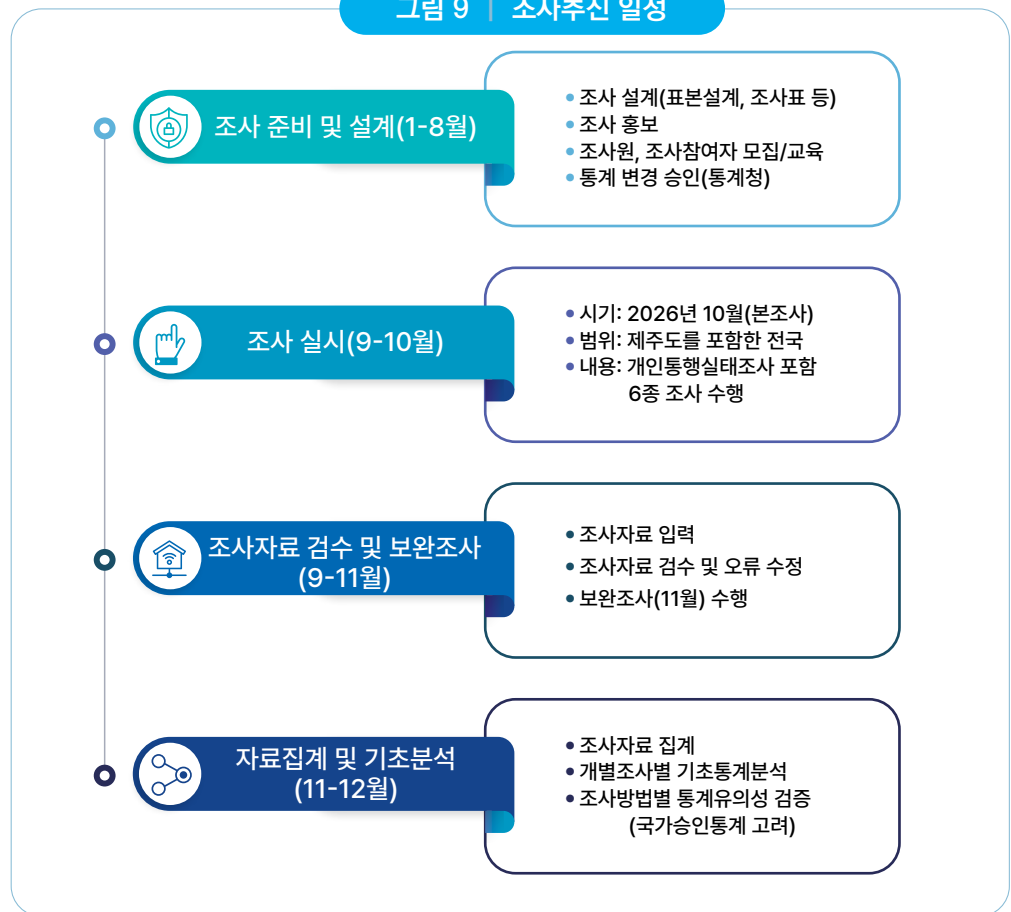
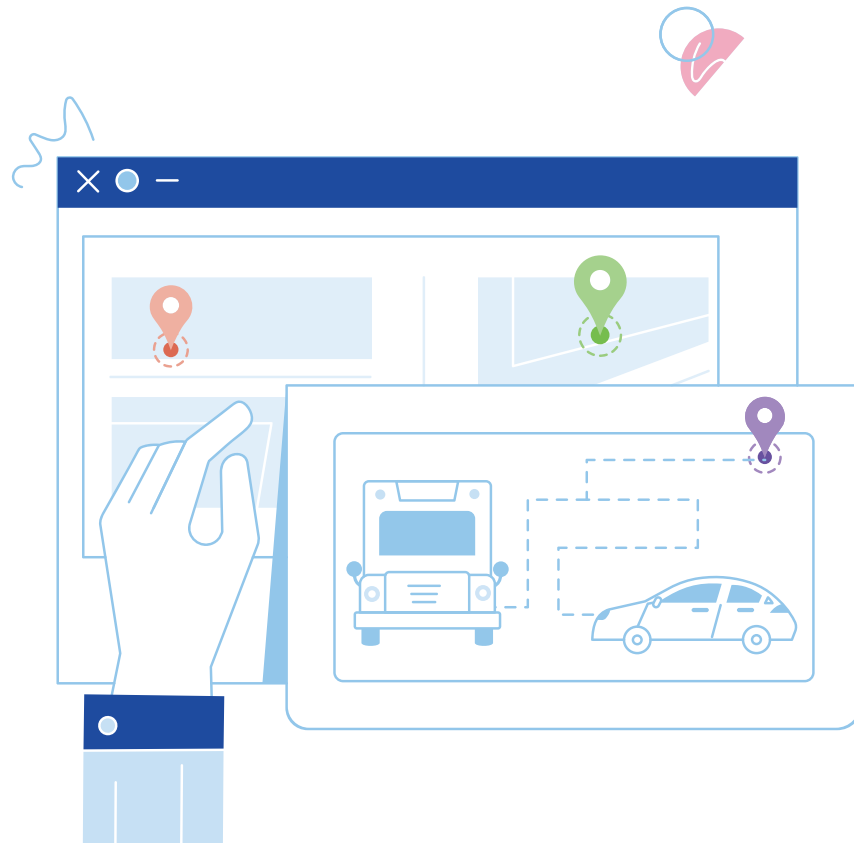


표 4 조사내용 및 방법

조사구분	조사대상	조사물량	조사내용	조사방법
개인통행 실태조사 (평일/주말)	· 만 5세 이상 국내거주 국민 · 평일 및 주말 하루통행	· (평일) 전국 총조사 인구의 0.2% · (주말) 평일조사 표본의 10%	· 가구특성 · 개인특성 · 통행특성	· 비대면 온라인 조사 · 전화조사(고령자 대상)
여객교통 시설물 이용실태 조사	· 버스터미널, 철도역, 공항, 여객선터미널 출발 승객	· 전국 여객시설물별 이용인원 2.7%	· 개인특성 · 통행특성	· 대면 면접조사
고속도로 이용실태 조사	· 고속도로 요금소 이용 승용/승합/택시운전자	· 전국 고속도로 요금소 이용인원 0.17%	· 통행특성	· 고속도로 요금소 통과차량 비대면 온라인조사
교통량 조사	· 전국 주요 코든/ 스크린 지점 통과차량	· 전국 주요 코든/ 스크린라인 지점	· 시간대별/차종별/ 방향별 교통량	· 24시간 영상촬영조사
전세버스 조사	· 전세버스 운전자	· 전국 전세버스 조합 등록 전세버스 수의 3%	· 차량/운전자 정보 · 통행특성	· 대면 면접조사



전국 도로 교통량 조사, AI가 대신한다:

인공지능과 모빌리티 빅데이터를 활용한 교통량 추정의 진화

이승봉 교통빅데이터연구본부 부연구위원

이체영 교통빅데이터연구본부 연구원 · 전준수 교통빅데이터연구본부 전문연구원

공공 교통량 관측 지점은 전체 도로망의 약 2%에 불과하여 전국 교통현상 설명에 구조적 한계가 있으나, 차량 GPS 기반 모빌리티 빅데이터와의 융합을 통해 전국 미관측 도로구간의 교통량 추정이 가능해지고 있다. 본 연구는 인공지능 기반 교통량 추정 모형을 단계적으로 고도화하여 전국 단위 교통량 DB의 신뢰도를 실증적으로 개선하고, 교통정책·SOC 투자·탄소배출 분석 등 정책 활용 기반을 확장하였다.

01 개요

- 도로 교통량은 교통혼잡 관리, 도로투자 타당성, 교통안전 등 국가 교통정책의 핵심 근거이자 도로망의 소통·지체·혼잡 수준을 판단하고 정책 효과를 검증하는 기본 지표임
 - 정책 집행과 투자 의사결정의 신뢰도는 교통량 데이터의 정확도에 크게 좌우됨
- 그러나, 국가에서 조사하고 있는 교통량 조사 지점은 전체 도로망 대비 매우 제한적(약 2% 수준¹⁾)으로, 전국 단위 교통망 성능을 파악하기 어려움
 - 루프·영상검지기 등 기존 센서 기반 수집 방식은 특정 지점에 고정되어 있어 공간적 대표성 부족, 설치·유지관리 비용 부담, 포장 변형 시 성능저하 등의 문제가 제기됨
 - 또한, 기관별 조사방식·장비가 상이하여 시·공간적 일관성이 낮고, 미관측구간이 광범위하게 발생하여 정책평가 과정에서 대표성·정확성·일관성 제약이 지속적으로 지적됨

1) 2023년 교통량 조사지점 11,378 지점(양방향)/2023년 12월 기준 국가표준노드링크수 548,614 개 = 2.07%

표 1 통행수단 분포

(단위:개)

구분		조사 지점수	방향별 관측교통량 DB 구축 지점 수		
			상행	하행	양방향
건설기술연구원	소계	3,893	3,893	3,893	7,786
	상시 조사지점	601	601	601	1,202
	수시 조사지점	3,292	3,292	3,292	6,584
한국도로공사		427	213	213	426
서울특별시		139	139	139	278
7대 특·광역시	소계	713	1,066	1,074	2,140
	인천광역시	195	246	249	495
	대전광역시	40	37	37	74
	대구광역시	124	244	244	488
	광주광역시	33	61	61	122
	부산광역시	135	237	237	474
	울산광역시	38	49	49	98
	세종특별자치시	48	192	197	389
합계		5,262	6,519	4,870	11,389

- 데이터 수집 환경은 고정 센서에서 이동 센서 기반으로 빠르게 전환되고 있음
 - 차량 GPS, 스마트폰, DSRC 등 이동 프로브 데이터의 급격한 확산으로, 개별차량의 실제 주행 경로와 전국 커버리지 확보가 가능해졌으며, 데이터 양·질이 기존 대비 비약적으로 개선되었으며, 실시간성, 전국 단위 연속성, 대규모 표본 확보가 가능한 환경으로 변화
- 국가통합교통체계효율화법(제12조·제14조) 및 시행령 제8조에 따라, 국가교통데이터베이스(KTDB) 구축·운영에는 전국 도로망을 대표할 신뢰성 높은 교통량 DB 구축이 필수임
 - 교통망 성능평가 지표는 교통혼잡, 온실가스, 물류 효율성, 안전 등 주요 정책의 근거이므로, 전국 단위에서 대표성과 정확성을 갖춘 교통량 DB 구축이 요구됨

02

교통량 추정
모형의 진화

1) 교통량 추정 모형 개요

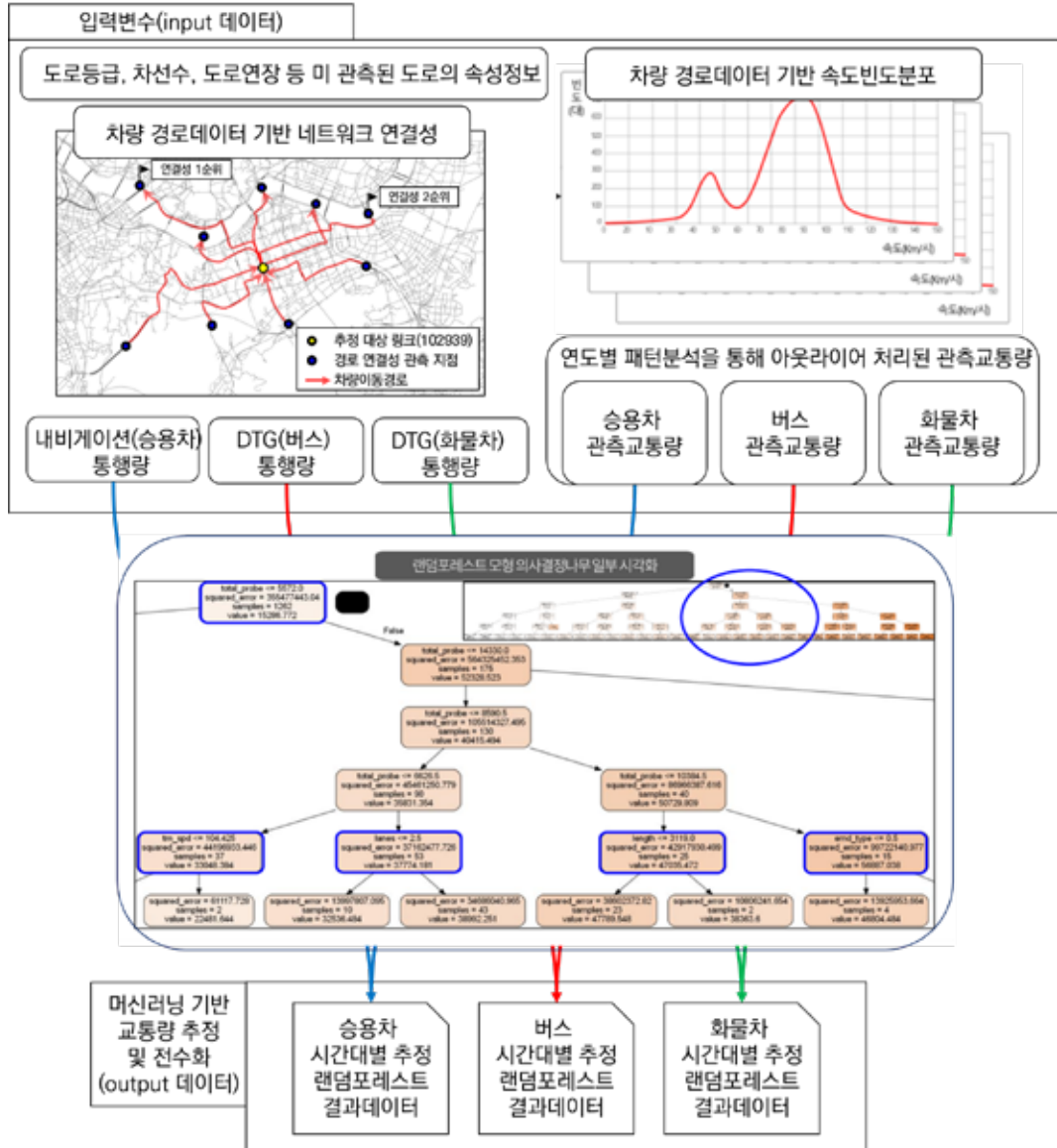
- 전국 도로 교통량 DB 구축의 핵심은 2% 수준의 관측교통량과 전국 단위 커버리지를 가진 차량 GPS 데이터를 융합하여, 나머지 98% 미관측 지점을 신뢰성 있게 추정하는 데 있음
 - 국가 교통량 조사는 전체 도로망의 약 2%만 관측되나, 차량 GPS · 스마트폰 · 내비게이션 등 이동 프로브 데이터는 전국적 연속성을 갖춰 이러한 한계를 보완하는 핵심 자원임
 - 이에 따라 KTDB 사업은 관측교통량과 차량GPS 기반 주행 패턴을 결합해 미관측 구간을 추정하는 체계를 구축하고, 교통량 추정 모형을 단계적으로 고도화하고 있음
 - **Model①** KNN(K-nearest neighbors)과 지역가중선형회귀(Locally weighted linear regression) 알고리즘이 결합된 방법론²⁾으로 미관측 링크 구간에 대한 교통량을 추정함
 - **Model②** 랜덤포레스트(Random Forest) 기반 AI 회귀모형을 도입해 입력변수를 다변화하고 비선형 관계를 반영하여 일반화 성능을 강화
 - **Model③** 기존 Model②에 다년도 누적 학습데이터 + 네트워크 연결성* 변수를 추가하여 시공간적 연속성과 추정 정밀도를 높인 고도화 모델
- * 개별차량 경로 데이터를 이용하여 특정 도로구간을 통과하는 차량의 연결성이 높은 관측지점 정보와 프로브 통행량을 생성하여 입력데이터로 활용

2) 모델별 교통량 추정 모형의 특징

- **Model① (KNN+LWLR 기반 초기 모형)**
 - KNN은 미관측 구간과 통행 특성이 유사한 관측지점을 선택하는 군집화 역할 수행하고, LWLR은 선택된 관측지점 정보를 기반으로 국부 가중 회귀식을 적용해 교통량 추정
- **Model② (AI 모형: 랜덤포레스트)**
 - 다수의 의사결정나무를 앙상블해 예측하는 랜덤포레스트를 적용하여 비선형 관계를 효과적으로 반영하고, 무작위 데이터 샘플링과 변수 선택을 통해 과대적합을 줄여, 변동성이 큰 도로에서도 일별·시간대별 교통량을 안정적으로 추정
- **Model③ (누적 학습 + 네트워크 연결성 변수 기반 고도화 모형)**
 - 기존 Model②에 다년도 누적 데이터와 네트워크 연결성 변수를 함께 활용하여 도로망의 구조적·시공간적 특성을 정밀하게 반영하고 모형의 일반화 성능과 추정 정확성 개선

2) 천승훈 외(2016), 『첨단교통자료를 활용한 교통지표DB구축(1) - 전국도로망 기초교통DB 구축 및 활용방안』, 한국교통연구원, pp.106-111

그림 1 | Model③의 교통량 추정 프로세스



자료 : 한국교통연구원(2024), "2024년 국가교통조사 및 분석 - 차량 GPS 빅데이터 구축" 한국교통연구원, pp. 61

03

교통량 추정
모형별 성능
비교 결과

1) 분석 개요

- 교통량 추정 모형의 단계별 고도화에 따른 성능 개선 효과를 신뢰도 분석을 통해 검증함
 - 분석에 활용된 핵심 데이터는 차량 GPS 기반 개별 차량 경로 데이터와 관측교통량 자료로 모든 모형에서 동일함
 - 다만, 모형별 적용 시점 차이에 따라 기준연도와 샘플 구조에는 차이가 존재하나, 동일한 검증 기준을 적용하여 성능을 비교함
 - 또한, Model①에서는 내비게이션 데이터만을 활용하고, Model ②, ③에서는 내비게이션 데이터는 승용차, DTG 데이터는 버스·화물 교통량 추정에 활용함
- 모델별로 적용된 시점의 차이로 수집 데이터 및 통행량의 차이는 존재할 수 있음

방법론	추정 모형	수집 데이터	기준연도	학습데이터	입력변수
Model①	(비모수 기법) KNN+LWLR	내비게이션, 관측교통량	2019년	단일연도 (2019년)	링크별 프로브 통행량, 인접구간 관측교통량 등
Model②	(인공지능) 랜덤포레스트	내비게이션, DTG, 관측교통량	2021년	단일연도 (2021년)	링크별 차종별 프로브 통행량, 도로 정보(차선수, 도로등급 등), 평균속도, 관측교통량 등
Model③	(인공지능) 랜덤포레스트	내비게이션, DTG, 관측교통량	2023년	누적연도 (2020-2023년)	링크별 차종별 프로브 통행량, 도로 정보(차선수, 도로등급 등), 평균속도, 관측교통량 등 네트워크 연결성 변수

2) 신뢰도 분석결과

- **(분석조건)** 신뢰도 검증은 관측교통량이 24시간 모두 확보된 지점을 대상으로 수행함
 - Model①은 학습과 검증 지점 별도 구분없이, 전체 자료를 이용하여 성능을 평가, Model②, ③은 학습데이터와 검증데이터를 8:2 비율로 분리하여 모형구축 및 검증 수행
 - 모형별 신뢰도 검증 외에도 교통량 수준별 검증을 통해 모형의 활용 가능성 검토
- **(평가기준)** 교통량 추정 결과의 신뢰도 평가는 평균절대비율오차(MAPE, %) 지표를 활용함
 - MAPE는 관측교통량 대비 추정교통량의 상대적 오차 수준을 정량적으로 평가하는 지표로, 모형 간 성능 비교에 주로 활용됨

$$MAPE(\%) = \frac{\sum_{i=1}^n \left(\frac{|q_i - \hat{q}_i|}{q_i} \right) \times 100}{n}$$

여기서, q_i = 지점 i 의 관측교통량(대/일)

\hat{q} = 지점 i 의 추정교통량(대/일)

n = 전체 지점 개수(개)

① Model별 분석 결과

- 관측교통량과 추정교통량 간 신뢰도 분석 결과, 교통량 추정모형의 고도화에 따라 평균 오차율(MAPE)이 단계적으로 크게 개선된 것으로 나타남
 - Model①의 평균 오차율은 33.4% 수준으로, 고속도로 등급을 제외하고는 미관측 구간 추정에 한계가 있었으나, 인공Model②, ③의 신뢰도는 크게 개선됨
 - Model③은 다년도 누적 학습과 네트워크 연결성 변수 반영을 통해 평균 오차율 8.3%까지 개선됨에 따라 교통량 추정의 정밀도와 전국 단위 활용 가능성이 실증적으로 확보됨

② 도로등급별 분석 결과

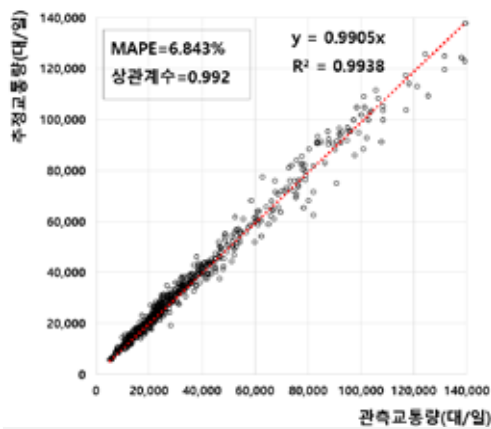
- Model③은 기존에 신뢰도가 낮았던 교통량 낮은 구간에서 뚜렷한 개선 효과를 보임
 - 특별·광역시도는 오차율 7.1%, 지방도는 8.4%, 시군도는 9.2%로 기존 Model① 대비 각각 약 41%p, 53%p 이상 오차가 감소하여 안정적인 추정 수준을 확보함
 - 도시고속국도 역시 약 6% 수준의 오차율을 보이며, 대도시 고속 소통축에서도 실무 적용이 가능한 수준으로 평가됨
 - 국가지원지방도는 57.1% → 11.3%로 -45.9%, 일반국도는 27.1% → 9.9%로 -17.1% 개선되어 간선도로 전반에서 일관된 성능 향상이 확인됨
- 이러한 결과는 차량 GPS 기반 교통량 추정모형의 고도화를 통해 전 도로등급을 포괄하는 전국 단위 교통량 DB로서의 활용 가능한 신뢰도 수준을 확보하였음을 의미함

표 3 도로등급별 교통량 추정 신뢰도 분석 결과

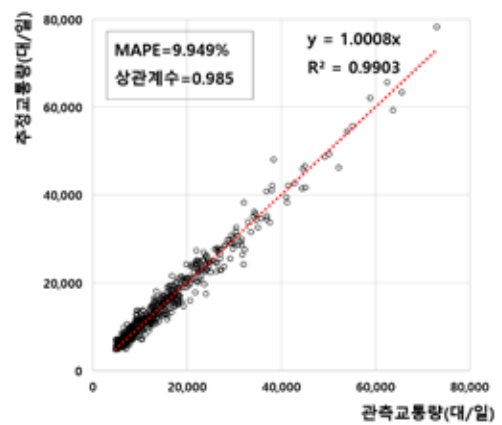
(단위: %, %p)

도로등급	오차율(MAPE)			초기 모델과 비교 (Model③-Model①)
	Model①	Model②	Model③	
고속국도	14.0	5.9	6.9	-7.1
도시고속국도	26.3	3.2	6.3	-20.0
일반국도	27.1	11.4	9.9	-17.1
특별광역시도	21.4	9.	7.1	-14.3
국가지원지방도	57.1	13.5	11.3	-45.9
지방도	49.2	15.3	8.4	-40.7
시군도	62.6	23.2	9.2	-53.3
평균	33.4	11.8	8.3	-25.2

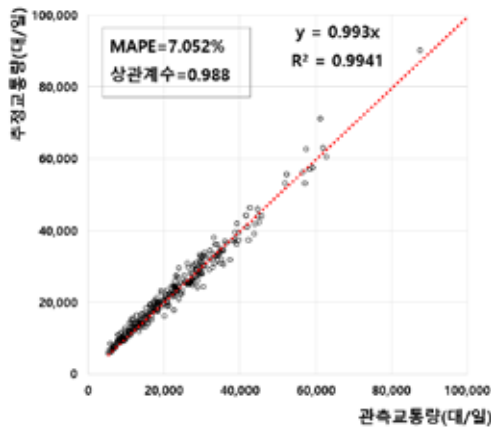
그림 2 | Model③의 도로등급별 교통량 추정 검증 결과



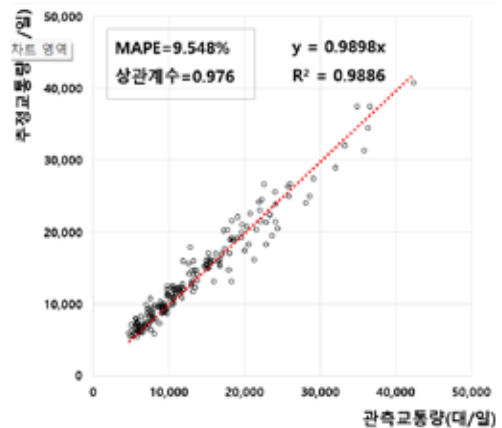
고속국도+도시고속국도



일반국도



특별광역시도



국가지원지방도+지방도

04 결론

1) 주요결과 요약

- 공공 교통량 조사자료의 공간적 한계와 데이터 수집 환경 변화에 대응하여, 모빌리티 빅데이터 기반 교통량 추정 모형을 단계적으로 고도화하여 구축함
 - 기존 고정식 관측 중심 체계에서 차량 GPS 기반 이동형 데이터 중심 체계로 전환
- 차량 GPS 기반 개별차량 경로데이터와 관측교통량 자료를 활용하여 전국 미관측 도로구간의 교통량을 추정하는 체계를 구축함
- 다양한 방법론 및 입력변수 적용을 통해 교통량 추정모형의 신뢰도를 실증적으로 개선함
 - 네트워크 연결성 변수를 적용한 인공지능 기반 추정모형을 통해 평균 오차율이 Model① 약 34% → Model③ 약 8% 수준으로 약 26%p 이상 개선
- 이를 통해 고속국도부터 시군도까지 전 도로등급을 포괄하는 전국 단위 교통량 DB로서의 활용 가능한 신뢰도 수준을 달성함
 - 특히, 기존 관측 사각지대였던 국가지원지방도·시군도 등 교통량 수준이 낮은 도로구간의 활용성이 크게 향상됨

2) 정책적 시사점

- 현재 공공에서 조사·운영 중인 교통량 관측 지점은 전체 도로망의 2% 수준으로 제한적이어서, 기존 관측자료만으로는 전국 단위 교통현상을 충분히 설명하는 데 한계가 있음
 - 이에 따라 차량 GPS 기반 교통량 추정기술과의 결합을 통해 미관측 구간을 포함한 전국 단위 교통량 DB 구축으로 공간적 공백을 해소할 필요가 있음
- GPS 데이터는 표본 자료인 반면, 관측교통량은 전수자료로 추정모형 학습·검증의 핵심 기준이므로, 교통량 조사 자료의 표준화와 통합 관리체계 구축이 교통량 DB 신뢰도의 출발점임
 - 조사 지점별 장비·관측 방식 차이에 따른 자료 불일치를 해소하기 위해 품질 기준 정립과 정합성 관리가 필수적임
 - 특히 시군도 등 저교통량 도로의 관측지점 비중이 낮으므로, 교통량 수준을 반영한 전략적 표본 설계를 통해 저·고 교통량 구간의 추정 정확도를 균형 있게 확보할 필요가 있음

- 현재 교통량 조사는 특정 1주일 중심으로 이루어지고 있어, 연평균 평일(AADT) 기준 통계에 한정되는 한계가 존재함
 - 향후 정책 활용도를 높이기 위해서는 주말·계절 변동성을 반영할 수 있도록 상시 교통량 조사 비중을 단계적으로 확대할 필요가 있음
 - 지자체 스마트교차로 교통량 데이터는 보조 관측수단으로 활용 가치가 높으나, 정식 통계 활용을 위해서는 표준화된 품질관리 기준 정립이 병행되어야 함
- 교통량은 도로 SOC 투자평가, 혼잡 개선 효과, 탄소배출량 산정 등 국가 교통정책의 핵심 정량지표로 활용되므로, 교통량 데이터의 품질 수준이 곧 정책 신뢰도를 좌우함
 - 모빌리티 빅데이터 수집 확대 및 인공지능 기반 교통량 추정모형의 정확도 고도화를 통해 승인통계 수준의 교통량 통계 생산 체계로 발전시킬 필요가 있음
- 차량 GPS 기반 교통량 DB는 교통혼잡비용, 차량주행거리, 수송부문 탄소배출량, SOC 투자효과 분석 등 다양한 정책 분야에서 활용 가능한 국가 핵심 인프라 데이터로 확장 가능
 - e-나라지표, 도로망 성능평가, 탄소공간지도 등과의 연계를 통해 혼잡 완화, 탄소중립, 투자 효율성 제고를 뒷받침하는 범정부 정책 활용 기반으로 지속 확장될 것으로 기대됨

모바일 통신 데이터 기반 경로형 데이터 구축

장동익 교통빅데이터연구본부 모빌리티데이터융합연구팀 팀장 · 이종우 교통빅데이터연구본부 부연구위원
곽명신 교통빅데이터연구본부 연구원 · 심지윤 교통빅데이터연구본부 연구원

모바일 통신데이터는 개인의 실제 이동 경험을 시·공간적으로 정밀하게 기록하는 새로운 교통 빅데이터 원천으로 주목받고 있다. 교통빅데이터연구본부는 통신데이터 기반 개인별 이동패턴과 전국 교통 네트워크를 결합하여, 도로·철도·항공 등 수단별 상세 경로 정보를 포함한 전국 단위 경로형 통행 데이터베이스 구축하고 있다. 이를 통해 KTDB는 기존 교통조사의 공간적 한계를 극복하고 통행경로, 교통시설 접근통행, 교통수단 구분 등 정교한 모빌리티 분석 기반을 마련하고자 한다.

01 개요

- 교통빅데이터연구본부에서는 2024년부터 모바일 통신데이터 기반 객체별 이동패턴 데이터를 구축 및 분석하고 있음
- 이동패턴은 개인식별이 불가능하도록 성·연령대 등 개인정보를 제거한 상태에서, 50m 격자 기준 1분 단위로 시·공간 이동정보를 구축함
- 구축된 이동패턴을 기반으로 실제 이동 경로를 정밀화하기 위해 도로, 철도, 지하철, 공항 등 주요 교통 인프라의 노드, 링크 정보를 구축하여 활용함
- 교통수단 간 연계성을 확보하기 위해 수단별 네트워크를 통합 네트워크 형태로 결합하여, 상세 경로 기반의 통행정보를 구축함



02

통합 네트워크 및 격자 단위 공간정보 구축

- 도로, 철도/지하철, 항공 네트워크를 하나의 통합 네트워크로 정비
- 수단 간 연계를 위한 환승 링크를 구축하여 수단 간 연속적인 경로 탐색이 가능하도록 설계 및 구축

그림 1 | 통합 네트워크 구성



통합 네트워크 구축 프로세스

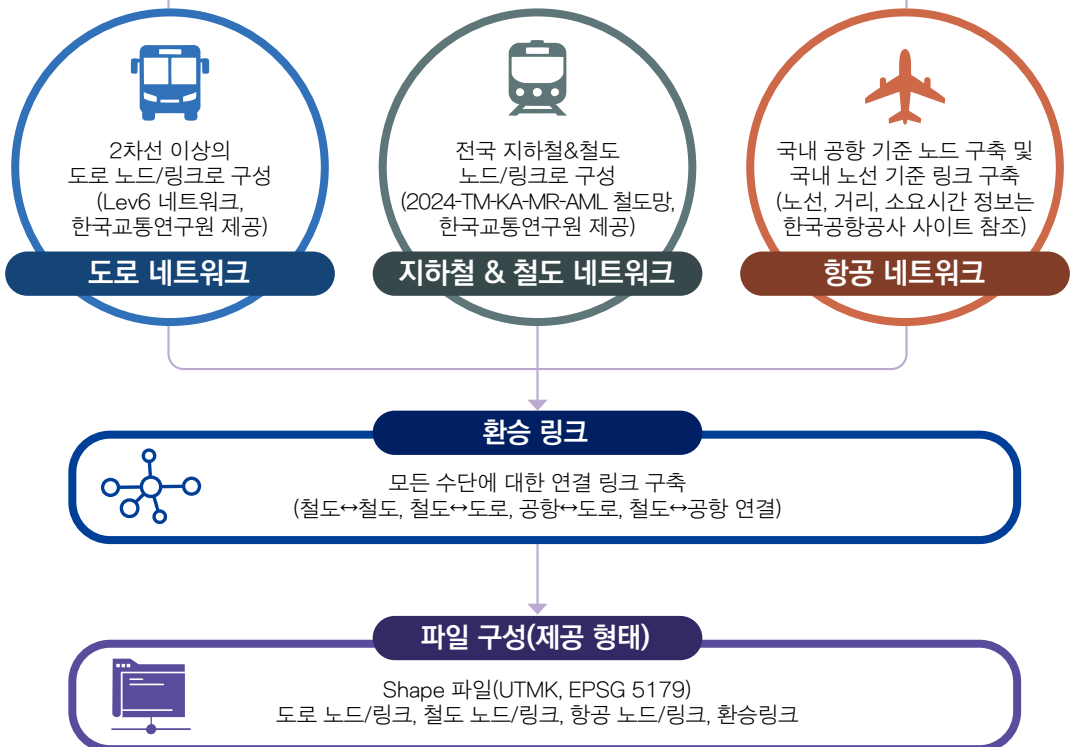


그림 2 | 통합 네트워크 데이터 구축 결과



03

경로형
데이터
구축과정

- 경로형 데이터 구축은 1) 통합 네트워크 및 격자 단위 공간정보 구축, 2) 모바일 궤적 데이터 전처리, 3) 이동수단 판별, 4) 이상치 판단, 5) 경로 생성으로 총 5단계로 구성됨

① 통합 네트워크 및 격자 단위 공간정보 DB 구축

- 통합 네트워크는 모바일 경로 데이터를 구축하기 위하여 도로, 철도, 지하철, 공항 간의 노드 및 링크를 구축하고 모든 이동 수단 별 연결될 수 있도록 연결 링크를 구축
- 격자는 1차 이동 수단을 판단하기 위하여 1km 및 50m 격자 단위로 통합 네트워크를 매칭 및 정보를 생성. 통합 네트워크를 사용하여 수단 구분을 하지 않는 이유는 모바일 이동 궤적의 오차범위가 수단 별로 차이가 크게 상이하여 이에 맞추어 격자단위 1km와 50m 격자 단위로 정보를 생성

② 데이터 전처리: 원시 상태의 모바일 이동 궤적 데이터는 자료구조 최적화 및 기초 필터링 과정을 통해 전처리

③ 이동유형 및 수단 구분: 전처리된 궤적 데이터를 대상으로 슬라이딩 윈도우 기법 및 Rule-based 알고리즘을 통해 각 이벤트의 이동유형과 수단 판단

④ 이상치 판단: 이동 경로의 연속성 및 현실성을 판단하기 위해 Reachability (도달 가능성), Forwardness (이동 방향성) 등의 검증 지표를 기반으로 이상치 이벤트를 탐지하고 라벨링

⑤ 경로 생성: 마지막 단계에서는 이상치가 제거된 궤적 이벤트와 통합 네트워크 데이터를 기반으로 실제 경로 생성

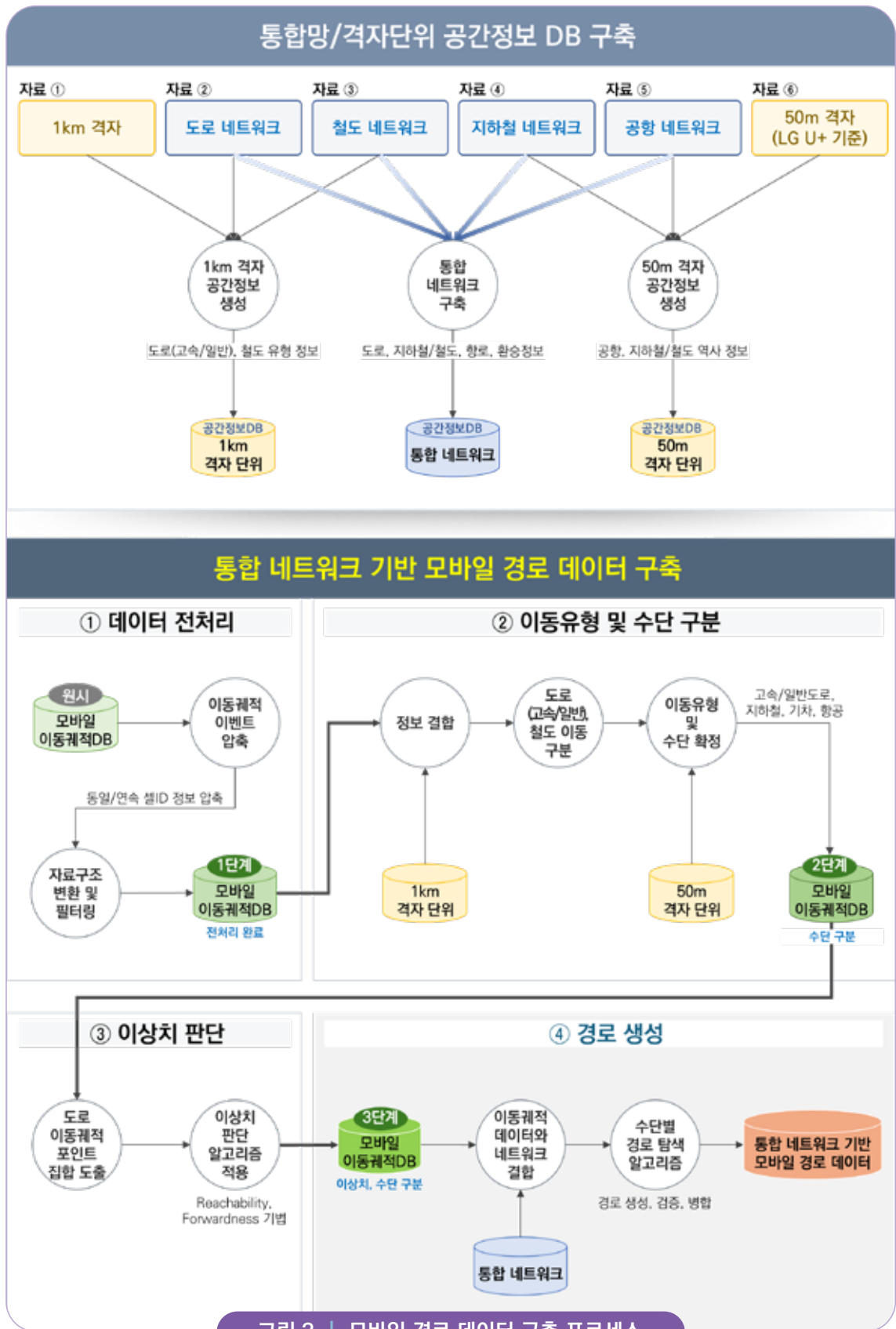


그림 3 | 모바일 경로 데이터 구축 프로세스

04 경로형 데이터 구축 결과

- 통합 네트워크를 이용한 경로형 데이터 구축 결과, 통신데이터 기반 이동체적이 실제 도로, 철도 등 네트워크와 연결되면서 현실적 이동경로로 재구성
- 통합 네트워크 기반 모바일 경로 데이터 포맷은 <표 1>과 같음

표 1 도로등급별 교통량 추정 신뢰도 분석 결과

순번	필드명	필드설명	자료형	코드	비고
1	TRIP_NO	통행 ID	INTEGER		
2	TRANSFER_TYPE	링크 수단 정보	INTEGER	1: 도로 2: 지하철 3: 철도 4: 공항 5: 환승 6: 기타	
3	SEQ	통행 ID별 순번	INTEGER		
4	LINK_ID	링크 ID	INTEGER		통합 네트워크 기준 0: 매칭정보없음
5	F_NODE_ID	시점 노드 ID	INTEGER		통합 네트워크 기준
6	T_NODE_ID	종점 노드 ID	INTEGER		통합 네트워크 기준
7	IN_TIME	진입시간	DATETIME		추정값
8	OUT_TIME	진출시간	DATETIME		추정값
9	SPEED	속도	REAL		km/h
10	LENGTH	링크 길이	REAL		m (미터)
11	SIDO_NAME	시도 명칭	TEXT		
12	TRAV_TIME	누적 통행시간	INTEGER		분
13	TRAV_DIST	누적 통행거리	INTEGER		m (미터)
14	OD_FLAG	시종점 링크 여부	INTEGER		0: 일반 1: 삭제대상후보

그림 5 | 통합 네트워크 기반 경로 구축 결과(도로-항공-지하철)



그림 6 | 통합 네트워크 기반 경로 구축 결과(도로-철도)



그림 7 | 통합 네트워크 기반 경로 구축 결과(지하철-철도-지하철)



05

경로형 데이터 분석 결과 (세종특별자치시)

- 2024년 11월 11일 경로형 데이터 구축 결과에 대해 세종특별자치시 출발지, 경유지, 도착지 통행목적별 통행 개수를 추출하여 시각화

그림 8 | 통행목적별(출발지/도착지/경유지) 경로 데이터 시각화 결과



- 동일 일자의 세종특별자치시 경로 데이터 중 교통수단별(공항, 고속철도) 이용 통행에 대하여 추출한 결과를 시각화

그림 9 | 항공·고속철도 이용 통행 경로 데이터 시각화 결과



05

모바일 경로형 데이터- 내비게이션 데이터 간 비교 분석 결과

- 모바일 경로형 데이터는 차량 이동뿐 아니라 버스, 철도 등 대중교통 정보까지 포함하여 실제 통행 흐름을 광범위하게 반영
- 다양한 교통수단 간 환승 패턴이 나타나며, 내비게이션의 차량 중심 이동보다 통행 범위가 더 넓게 확산되는 특징이 있음

그림 10 | 세종시 햇무리교 통과 경로 시각화



그림 11 | 세종시 햇무리교 통과 경로 시각화 (확대)



06

경로형 데이터 활용방안

- 모바일 통신데이터를 기반으로 2024년 기준 전국 단위의 상세 통행DB가 구축됨에 따라, 보다 세분화된 시·공간 단위의 모빌리티 패턴 분석이 가능해짐
- 이를 통해 지역 간 통행 구조, 시간대별 이동행태를 정교하게 파악할 수 있는 기반이 마련됨
- 경로형 데이터의 고도화는 교통정책 수립, 대중교통 서비스 개선, 신교통수단 도입 등 다양한 분야에서 높은 활용가치를 제공할 것으로 기대됨

서울 대중교통 1시간 통근권 분석

홍성표 한국교통연구원 주임연구원

한국교통연구원에서는 대중교통 부문 교통접근성지표 산정을 위하여 2018년부터 전국 GTFS(General Transit Feeds Specification) 기반 대중교통 네트워크를 구축하여 분석에 활용하고 있다. 가장 최근 구축자료인 2024년 3월 기준 대중교통 DB를 기반으로 서울 외 지역의 서울 대중교통 1시간 통근권을 분석하였다.

01 분석 개요

1) 분석의 범위

- 공간 범위 : 전국 행정동, 법정리 기준
- 시간 범위 : 2023년 12월(대중교통 정보는 2024년 3월 자료 사용)

2) 사용 자료

공간정보 및 사회경제지표

- 행정동 경계 : 통계지리정보서비스(<https://sgis.kostat.go.kr/>) 내 센서스용 행정구역 경계(2023년 12월)
- 법정리 경계 : 주소기반산업지원서비스(<https://business.juso.go.kr/>) 내 법정구역 경계(2023년 12월)
- 건물 형상정보 : 주소기반산업지원서비스(<https://business.juso.go.kr/>) 내 도로명주소 건물 정보(2023년 12월)
- 인구 : 행정안전부 주민등록인구통계(<https://jumin.mois.go.kr/>)(2023년 12월)

도로 및 대중교통 정보

- 도로망 : 오픈스트리트맵(<https://www.openstreetmap.org/>) 도로망 정보(2024년 1월)
- 대중교통¹⁾ : 국가교통DB센터(<https://www.ktdb.go.kr/>) 버스·철도 GTFS 대중교통 네트워크 정보(2024년 3월)

1) 대중교통 DB 구축과정은 국토교통부·한국교통연구원, 2024, 『2024년 국가교통조사 및 분석: 교통접근성지표 구축』, 107~126쪽 참고.

02 분석 방법

1) 행정동·법정리 중심점 DB 구축²⁾

건물 형상정보와 행정동·법정리 경계정보 간 공간중첩분석 후 구역 내 건물별 연면적 가중 중앙 중심점*(Median Center) 산출

표 1 구역 중심점 산출방법

구분	구역 내 건물 존재여부		구역 중심점 산출방법
	전체	주거용도	
1차	O	O	- 주거용도 건물이 존재하는 구역 · 주거용도 대상 거리×연면적 합이 최소인 건물의 좌표
2차	O	X	- 주거용도 외 건물만 존재하는 구역 · 전체용도 대상 거리×연면적 합이 최소인 건물의 좌표
3차	X	-	- 집계구 내 건물이 존재하지 않는 경우 · 구역 경계의 면적중심 좌표

자료: 국토교통부·한국교통연구원, 2024, 『2024년 국가교통조사 및 분석: 교통접근성지표 구축』, 98쪽에서 재구성

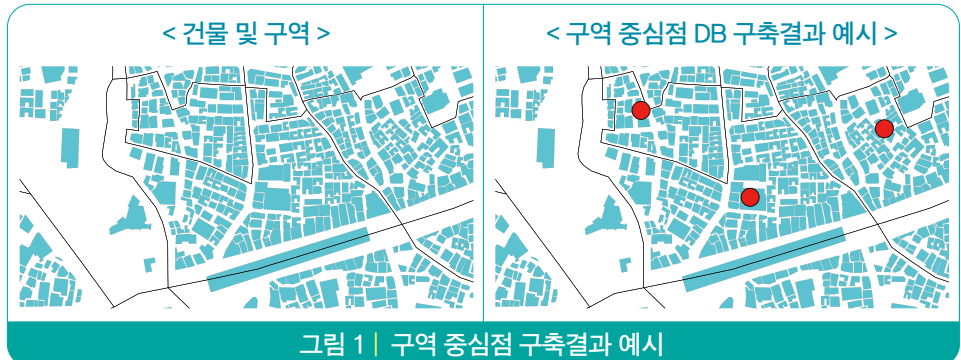


그림 1 | 구역 중심점 구축결과 예시

자료: 국토교통부·한국교통연구원, 2024, 『2024년 국가교통조사 및 분석: 교통접근성지표 구축』, 98쪽에서 재구성

* 각 구역 소속 건물 간 거리 × 연면적의 합이 최소가 되는 건물의 좌표

구역 내 건물 존재여부 등에 따라 아래와 같이 중심점 DB 구축

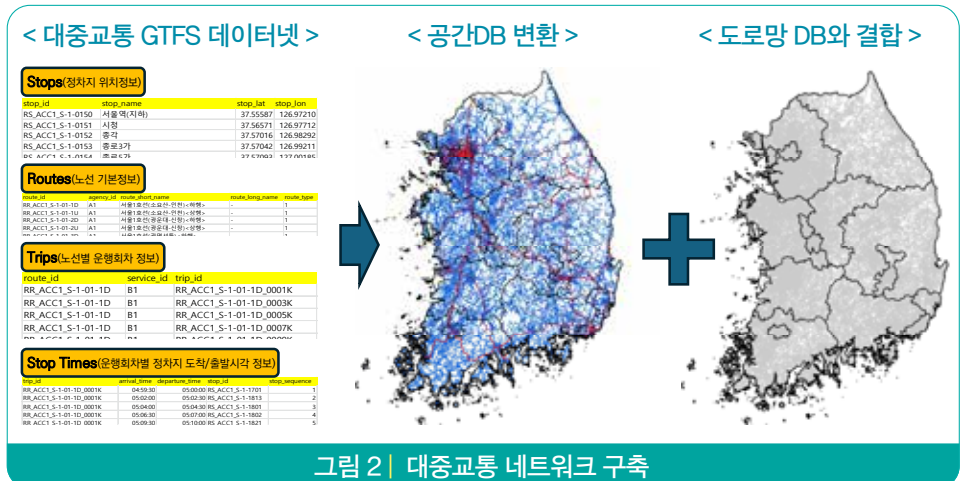


그림 2 | 대중교통 네트워크 구축

2) 국토교통부·한국교통연구원, 2024, 『2024년 국가교통조사 및 분석: 교통접근성지표 구축』, 98쪽.

2) 대중교통 네트워크 구축

버스·철도 GTFS DB를 공간 DB로 변환 후 도로망 정보와 결합하여 분석용 네트워크 구축

3) 대중교통 네트워크 분석

출 발 지 : 서울 외 각 행정동·법정리 중심

도 착 지 : 서울 내 대중교통 정차지

출발시각 : 07:00~09:00,
10분 간격(총 13개 출발시각)

한계시간 : 60분

분석 모듈 : ArcGIS 내 Network Analysis Tools



4) 1시간 통근권 지표 산정방법

$$CR_i^{1hour} = \frac{I_i^{CR}}{N} \times 100 (\%)$$

행정동·법정리별 서울 대중교통 1시간 통근 가능 비율(1시간 통근권)

- 각 행정동·법정리별 전체 출발시각별 서울 진입 통행시간 중 서울 1시간 내 접근 가능한 시각대 비율

여기서, i =각 행정동·법정리, CR_i^{1hour} = 1시간 내 통근 가능 비율,

N =분석 대상 시각 수(13개 시각),

I^{CR} =서울 통근 가능성(1시간 이하=1, 1시간 초과=0)

$$PCR_j^{1hour} = \frac{\sum_{i \in (j)} (Pop_i \times CR_i^{1hour})}{\sum_{i \in (j)} (Pop_i)} \times 100 (\%)$$

시·군별 서울 대중교통 1시간 통근 가능 인구 비율

- 각 시·군별 총 인구 중 서울 대중교통 1시간 내 통근 가능 인구 비율

여기서, j =각 시·군, PCR_j^{1hour} = 1시간 내 통근 가능 인구 비율

2) 시·군별 서울 대중교통 1시간 통근 가능 인구 비율

표 2 시·군별 서울 대중교통 1시간 통근 가능 인구 비율

권역	시·군	인구수(명)	대중교통 1시간 내 서울 통근 가능 인구(명)	대중교통 1시간 내 서울 통근 가능 인구 비율%(명)
수도권	성남	919,747	919,747.00	100.00
	의정부	464,213	464,213.00	100.00
	안양	544,660	544,660.00	100.00
	부천	779,968	779,968.00	100.00
	광명	280,197	280,197.00	100.00
	과천	81,000	81,000.00	100.00
	구리	187,093	187,093.00	100.00
	군포	261,229	261,229.00	100.00
	의왕	158,870	158,870.00	100.00
	하남	329,861	329,861.00	100.00
	고양	1,074,907	1,069,826.85	99.53
	남양주	732,264	708,164.69	96.71
	동두천	88,626	84,407.08	95.24
	김포	486,167	451,945.00	92.96
	수원	1,197,257	1,078,703.38	90.10
	인천	2,908,028	2,539,318.23	87.32
	양주	268,026	218,321.54	81.46
	광주	391,377	317,350.15	81.09
	용인	1,075,566	845,909.46	78.65
	안산	629,308	470,843.77	74.82
	파주	497,749	336,426.38	67.59
	시흥	519,715	344,827.23	66.35
	화성	944,327	562,951.62	59.61
	오산	229,656	83,189.38	36.22
	포천	143,322	40,015.46	27.92
	평택	591,022	141,866.62	24.00
	이천	222,098	50,424.46	22.70
	양평	125,236	24,259.15	19.37
	가평	62,300	7,722.15	12.40
	연천	41,584	2,210.77	5.32
	안성	189,033	4,628.38	2.45
	강화	69,002	0.00	0.00
	옹진	20,372	0.00	0.00
여주	114,412	0.00	0.00	
수도권 외	천안	655,959	44,865.54	6.84
	아산	345,795	7,488.54	2.17
	음성	91,169	946.15	1.04
	청주	852,184	1,013.62	0.12
	당진	170,302	197.23	0.12
원주	361,483	153.54	0.04	