

대중교통 정책지원 고도화를 위한
모바일 빅데이터 DB구축



2019년 「국가교통조사 · DB시스템 운영 및 유지보수」

대중교통 정책지원 고도화를 위한 모바일 빅데이터 DB구축

2019. 12



2019년 「국가교통조사·DB시스템 운영 및
유지보수」

대 중 교 통 정 책 지 원 고 도 화 를 위 한 모 바 일 빅 데 이 터 DB 구 축

13

제 출 문

국토교통부장관 귀하

본 보고서를 「2019년도 국가교통조사 및 DB시스템 운영 및 유지 보수」 최종보고서로 제출합니다.

2019년 12월

한국교통연구원

원장 오 재 학

**본 『2019년도 국가교통조사 및 DB시스템 운영 및
유지보수』는 다음 연구진에 의해 수행되었습니다.**

참 여 연 구 진

<한국교통연구원>	
연구책임자	◦ 김주영 연구위원
연 구 진	<ul style="list-style-type: none"> ◦ 박인기, 최정민, 조종석, 천승훈 연구위원 ◦ 박용일, 황순연, 장동익, 송태진, 성홍모, 원민수, 김병관, 우왕희 부연구위원 ◦ 신영권, 김동호, 김규진, 김정은 주임전문원, 이종우 전문연구원 ◦ 강국수, 강명제, 고두환, 광명신, 김관용, 김성민, 김은미, 박미란, 박준호, 오연선, 이선아, 이슬기, 이채영, 이해선, 정승환, 조용훈, 채정표, 홍성표 연구원 ◦ 김예은, 송수환 연구조원
<한국해양수산개발원>	
연 구 진	<ul style="list-style-type: none"> ◦ 이호춘, 전형진, 이종필 부연구위원 ◦ 류희영 연구원
<한국항공협회>	
연 구 진	<ul style="list-style-type: none"> ◦ 성인영 실장 ◦ 박수경 과장, 손병열 과장, 유인아 대리

『2019년도 국가교통조사 및 DB구축사업』

보고서 구성 및 담당연구진

번 호	과 제 명	연 구 진
제 1권	요약보고서	김주영, 신영권, 박준호
제 2권	전국여객O/D 보완갱신	조종석, 강국수, 박미란
제 3권	빅데이터 기반 여객 O/D 신뢰도 제고 연구	김병관, 채정표, 정승환
제 4권	항공여객 O/D 및 특성조사	한국항공협회
제 5권	물류거점 화물실태조사	박인기, 성홍모, 김정은, 조용훈 강명제
제 6권	전국연안화물O/D 조사	한국해양수산개발원
제 7권	빅데이터 기반 화물OD 신뢰도 제고 연구	박인기, 성홍모, 김정은, 조용훈 강명제
제 8권	교통분석용 네트워크 구축	최정민, 우왕희, 이선아, 이슬기
제 9권	KTDB 플랫폼 기반지도 구축	김동호, 김관용
제10권	국가교통통계조사	박용일, 곽명신
제11권	특별교통대책기간 통행실태 조사	최정민, 우왕희, 김은미
제12권	교통혼잡지도 DB구축	천승훈, 김성민, 김관용, 이채영
제13권	대중교통 정책지원 고도화를 위한 모바일 빅데이터 DB구축	김동호, 송태진, 원민수, 이해선, 이종우
제14권	교통유발원단위조사 예비조사	황순연, 오연선, 고두환
제15권	국가교통물류경쟁력지표 조사연구	장동익, 홍성표
제16권	DB시스템 운영 및 유지보수	신영권, 김규진, 박준호

『2019년도 국가교통조사 및 DB구축사업』

과제별 공동참여·위탁용역 사업자

【공동사업 참여기관】

- 전국 여객O/D 현행화 공동사업 (제주특별자치도 부문)
 - 홍익대학교산학협력단
- 전국 여객O/D 현행화 공동사업 (부산·울산권 부문)
 - 경성대학교산학협력단, (주)나우건설터트
- 전국 여객O/D 현행화 공동사업 (대전·세종·충청권 부문)
 - 코에스페셜주식회사, (주)신명이앤씨

【위탁용역 사업자】

- 전국 장래 시군 및 읍면동 인구예측에 관한 연구
 - 고려대학교산학협력단
- 물류거점화물실태조사
 - (주)코리아데이터네트워크
- 영업용 화물차 운행기록계 빅데이터를 이용한 화물 기종점통행량 및 운행특성분석연구
 - (주)노트스퀘어
- 도로 및 철도 교통분석용 네트워크 보완갱신
 - (주)신명이앤씨
- KTDB 교통빅데이터 플랫폼 (View-T) 분석맵 구축
 - (주)큐빅웨어

【위탁용역 사업자】

- 국가교통 DB Brief 발간 대행
 - ㈜우공이산
- 특별교통통행실태조사 및 이용자 만족도 조사
 - ㈜마크로밀엠브레인
- View-T 서비스 제공을 위한 차량모빌리티 데이터 구축 및 기능개선
 - ㈜큐빅웨어, (사) 한국ITS학회
- 모바일 빅데이터 기반 교통분석용 DB구축
 - ㈜KT
- View-T 2.0 서비스 제공을 위한 통신 빅데이터 구축 및 기능 개선
 - ㈜큐빅웨어
- 교통유발원단위 예비조사
 - ㈜아이로드테크, ㈜지알아이리서치

최종보고서 목차

- 제 1권 요약보고서**
- 제 2권 전국여객O/D보완갱신**
- 제 3권 빅데이터 기반 여객OD신뢰도 제고 연구**
- 제 4권 항공여객 O/D 및 특성조사**
- 제 5권 물류거점 화물실태조사**
- 제 6권 전국연안화물O/D 조사**
- 제 7권 빅데이터 기반 화물OD신뢰도 제고 연구**
- 제 8권 교통분석용 네트워크 구축**
- 제 9권 KTDB 플랫폼 기반지도 구축**
- 제 10권 국가교통통계조사**
- 제 11권 특별교통대책기간 통행실태조사**
- 제 12권 교통혼잡지도 DB구축**
- 제 13권 대중교통 정책지원 고도화를 위한 모바일 빅데이터 DB구축**
- 제 14권 교통유발원단위 예비조사**
- 제 15권 국가교통물류경쟁력지표 조사연구**
- 제 16권 DB시스템 운영 및 유지보수**

목 차

요 약

제1장 과업의 개요 1

제1절 과업의 배경 및 목적 / 3

제2절 과업의 범위 및 내용 / 4

제2장 모바일 빅데이터 가공 알고리즘 개발 7

제1절 통행 왜곡데이터 보정 방법 / 9

제2절 채류 정보 추출 방법 / 21

제3절 3통행 미만 통행량 보정 방법 / 23

제3장 모바일 빅데이터 기반 기종점 통행량 DB 구축 29

제1절 기준년도 DB 구축 / 31

제2절 구축 DB 검증 / 40

제4장 모바일 빅데이터 기반 통행 분석 서비스 개발 51

제1절 개요 / 53

제2절 UI 설계 및 기능 개발 / 62

제3절 기반 DB 구축 / 90

제4절 운영 및 유지보수 / 122

제5장 결론 및 향후 과제 129

제1절 결론 / 131

제2절 향후 과제 / 134

표 목 차

<표 2- 1> 기록 간 속도 산출 방법	12
<표 2- 2> 시그널 점프 데이터 행 압축 방법	13
<표 2- 3> 시그널 점프 보정 전 후 비교	14
<표 2- 4> 거리별 핑퐁 핸드오버 발생빈도 및 비율 (강남구 vs 평창군)	16
<표 2- 5> 시간 차에 따른 핑퐁 핸드오버 발생빈도 및 비율 (강남구 vs 평창군) ·	17
<표 2- 6> 핑퐁 핸드오버 탐색 결과 (예시)	18
<표 2- 7> 핑퐁 핸드오버 데이터 행 압축 방법	19
<표 2- 8> 핑퐁 핸드오버 데이터 보정 전 후 비교	20
<표 2- 9> 기존 3통행 미만 통행 처리 방식	23
<표 2-10> 타 과제에서의 3통행 미만 처리 방법 - 비식별화 처리 방법	24
<표 2-11> 3통행 미만 통행량 보정계수 산출 기준 표	26
<표 2-12> 기준 달에서 누락된 보정계수 재산출 방법	26
<표 2-13> 3통행 미만 통행량 보정계수 적용 전후 비교	27
<표 3- 1> 모바일 기반 DB 형태	31
<표 3- 2> 모바일 DB 테이블 정의서	32
<표 3- 3> 체류지 식별 기준	35
<표 3- 4> KT 시장점유율	38
<표 3- 5> 인력식 조사 기준 전국 기·중점 통행량 (시·도)	41
<표 3- 6> 모바일 데이터 기준 전국 기·중점 통행량 (시·도) - (주) KT 데이터 기준	42
<표 3- 7> 모바일 데이터 기준 전국 기·중점 통행량 (시·도) - 전수화	42
<표 3- 8> 전국 기·중점 통행 비율 (비교 결과)	43
<표 3- 9> 시간대별 통행량 변화 - (주) KT 데이터 기준	45
<표 3-10> 요일별 통행량 변화 - (주) KT 데이터 기준	47
<표 3-11> 성별 통행량 변화 - (주) KT 데이터 기준	48
<표 3-12> 연령별 통행량 변화 - (주) KT 데이터 기준	50
<표 4- 1> 통신 빅데이터의 컬럼명 및 데이터 타입	54
<표 4- 2> 주요통행지역분석	56

<표 4- 3> 주요통행구간분석 개선 방향	57
<표 4- 4> Hot place 분석 개선 방향	57
<표 4- 5> 부가 서비스 개선 방향	58
<표 4- 6> 분석 맵 업데이트	59
<표 4- 7> 모바일 빅데이터 기종점 통행량 업데이트	59
<표 4- 8> 사용자 편의 개선	61
<표 4- 9> 사용자 편의 개선	62
<표 4-10> 메뉴의 구성	64
<표 4-11> 수집된 교통폴리곤 데이터의 속성정보	90
<표 4-12> MySQL 기반의 교통폴리곤 공간정보의 속성	91
<표 4-13> VectorTile, WMS의 지도 일반화	92
<표 4-14> GeoServer 교통폴리곤 레이어의 속성정보	93
<표 4-15> GeoServer 교통폴리곤 노드 레이어의 속성정보	94
<표 4-16> 수집된 행정구역 데이터의 속성정보	96
<표 4-17> MySQL 기반의 행정구역 폴리곤 공간정보의 속성	97
<표 4-18> GeoServer 행정구역 폴리곤 레이어의 속성정보	97
<표 4-19> MySQL 기반의 행정구역 노드 공간정보의 속성	98
<표 4-20> 통신빅데이터의 컬럼명 및 데이터 타입	100
<표 4-21> 모바일 빅데이터 기종점 통행량 DB의 연령별 코드	101
<표 4-22> 모바일 빅데이터 기종점 통행량 DB의 성별 코드	101
<표 4-23> 모바일 빅데이터 기종점 통행량 DB의 시간대 코드	102
<표 4-24> 모바일 빅데이터 기종점 통행량 DB의 체류특성별 코드	103
<표 4-25> 교통폴리곤 단위의 모바일 통행형태 통행량 컬렉션의 구성	107
<표 4-26> 통신빅데이터의 도큐먼트 수의 경량화 결과	108
<표 4-27> 읍면동 출발 기준의 통행형태 통행량의 컬렉션의 구성	108
<표 4-28> 읍면동 도착 기준의 통행형태 통행량의 컬렉션의 구성	109
<표 4-29> 시군구 출발 기준의 통행형태 통행량의 컬렉션의 구성	111
<표 4-30> 시군구 도착 기준의 통행형태 통행량의 컬렉션의 구성	111
<표 4-31> 시도 출발 기준의 통행형태 통행량의 컬렉션의 구성	111
<표 4-32> 시도 도착 기준의 통행형태 통행량의 컬렉션의 구성	113
<표 4-33> 읍면동 기준의 통행형태 통행량의 컬렉션의 구성	114
<표 4-34> 시군구 기준의 통행형태 통행량의 컬렉션의 구성	114

<표 4-35> 시도 기준의 통행형태 통행량의 컬렉션의 구성	116
<표 4-36> 읍면동 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-37> 읍면동 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-38> 시군구 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-39> 시군구 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-40> 시도 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-41> 시도 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-42> 읍면동 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	118
<표 4-43> 시군구 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	119
<표 4-44> 시도 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	119
<표 4-45> Web Server 설치 정보	122
<표 4-46> Web Server 설치 정보(Apache Tomcat)	123
<표 4-47> 데이터베이스 백업 대상 파일	124
<표 4-48> 응용프로그램 백업 대상 파일	125
<표 4-49> 백업 및 복구 절차	125
<표 5- 1> 모바일 기반 DB 형태	132
<표 5- 2> View-T 2.0 사람 통행 분석 서비스 개발 주요 내용 (기반 DB 구축)	132
<표 5- 3> View-T 2.0 사람 통행 분석 서비스 개발 주요 내용 (분석 기능 개발)	133

그림목차

<그림 2- 1> 시그널 점프 데이터 유형	10
<그림 2- 2> 핑퐁 핸드오버 데이터	10
<그림 2- 3> 핑퐁 핸드오버 데이터 유형	11
<그림 2- 4> 시그널 점프 보정 전 후 비교	14
<그림 2- 5> 핑퐁 핸드오버 데이터 분석 범위	15
<그림 2- 6> 강남구와 평창군의 핑퐁 핸드오버 발생 특성 -거리	16
<그림 2- 7> 강남구와 평창군의 핑퐁 핸드오버 발생 특성 - 시간 차	17
<그림 2- 8> 핑퐁 핸드오버 데이터 보정 전 후 비교	20
<그림 2- 9> 2017년 기준 이동/체류 기준 설정 방법	21
<그림 2-10> 경로 체류시간 임계치 설정	22
<그림 2-11> 2017년 기준 DB의 AADT와 월별 일평균 통행량	25
<그림 3- 1> 선분 이력으로의 변환	33
<그림 3- 2> 체류순서 부여 (개인별 이동궤적 형성)	34
<그림 3- 3> 이동정보와 체류정보 구분 (예시)	35
<그림 3- 4> 체류지 유형 구분 (예시)	36
<그림 3- 5> 위치정보, 시간정보, 연령정보 변환 (예시)	37
<그림 3- 6> 출발/도착 구분 및 통행량 집계 (예시)	38
<그림 3- 7> 통행시간 산출 (예시)	39
<그림 3- 8> 시간대별 통행량 변화 (유입 기준)	44
<그림 3- 9> 요일별 통행량 변화 (유입 기준)	46
<그림 3-10> 연령별 통행량 변화 (유입 기준)	49
<그림 4- 1> 2018년 개발한 주요 분석 기능 4가지	53
<그림 4- 2> 모바일 분석 맵 4가지 (2018년 구축 공간정보)	55
<그림 4- 3> 사람 이동 통행 분석 메뉴	63
<그림 4- 4> 사용자 기능 유도 표출 예시	65
<그림 4- 5> 상세설정 UI	66
<그림 4- 6> 범례설정 UI	66

<그림 4- 7> 심볼 스키마 예시	67
<그림 4- 8> 심볼 색상 사용자정의 예시	67
<그림 4- 9> 스크린샷 UI	68
<그림 4-10> 라벨 폰트 표출 예시	68
<그림 4-11> Hotplace 분석 개선된 UI	69
<그림 4-12> Hot Place 분석 시군구 단위 표출 예시	70
<그림 4-13> Hot Place 분석 읍면동 단위 표출 예시	70
<그림 4-14> 주요통행지역분석 권한에 따른 선택 지역 UI	71
<그림 4-15> 주요통행지역분석 희망선 색상 구분	71
<그림 4-16> 주요통행지역분석 파이차트 정보표출	72
<그림 4-17> 주요통행지역분석 주제도의 행정구역 명칭 제어	73
<그림 4-18> 주요통행지역분석 분석 화면 예시(행정구역 명칭 ON / 정보표출 ON)	74
<그림 4-19> 주요통행구간분석 개선된 UI	75
<그림 4-20> 주요통행구간분석 읍면동 단위 표출 예시	76
<그림 4-21> 내부통행지역분석의 설정 UI	78
<그림 4-22> 내부통행분석 분석 화면 예시(파이차트ON / 정보표출 ON)	79
<그림 4-23> 주제도 클릭시 나타나는 그래프 UI	81
<그림 4-24> 그래프 분석하기 클릭시 나타나는 그래프 UI	81
<그림 4-25> 통행 시간/거리 분석의 설정 UI	82
<그림 4-26> 통행 시간 분석의 주제도·그래프 표출 화면 예시	83
<그림 4-27> 통행 거리 분석의 주제도·그래프 표출 화면 예시	84
<그림 4-28> 유형별 통신데이터 다운로드 설정 UI	85
<그림 4-29> 유형별 통신데이터 다운로드 링크 메일	86
<그림 4-30> 사용자가 설정한 조건에 해당하는 유형별 통신 엑셀 데이터	87
<그림 4-31> 지역 간 OD 통행량 데이터 다운로드 설정 UI	88
<그림 4-32> 사용자가 설정한 조건에 해당하는 지역 간 OD 엑셀 데이터	89
<그림 4-33> 교통폴리곤 구축 절차	90
<그림 4-34> GeoServer 기반의 교통폴리곤 레이어	92
<그림 4-35> GeoServer 기반의 교통폴리곤 노드 레이어	95
<그림 4-36> GeoServer 기반의 행정구역 폴리곤 레이어	98
<그림 4-37> GeoServer 기반의 행정구역 노드 레이어	99
<그림 4-38> 모바일 빅데이터 기종점 통행량 DB 형태	100

<그림 4-39> 모바일 빅데이터 기종점 통행량 DB의 구축 절차	103
<그림 4-40> 교통폴리곤과 읍면동 행정구역 코드의 매칭정보	104
<그림 4-41> 출·도착 교통폴리곤 간 거리 산출 예시	105
<그림 4-42> 3통행 보정계수 데이터 예시	105
<그림 4-43> 1차 가공데이터의 MongoDB 입력데이터 예시	106
<그림 4-44> MongoDB Sharding의 구성	121
<그림 4-45> 모바일 통행형태 기반 DB의 Sharding 구성 결과	121
<그림 4-46> 유지보수 범위 및 조직	126
<그림 4-47> 장애관리 절차	127
<그림 4-48> 장애처리 절차	128
 <그림 5- 1> 향후 모바일 DB 설계 방안	 134

요약

요 약

1. 과업의 개요

가. 과업의 배경 및 목적

- 현재 교통 DB는 도로 중심의 데이터가 주를 이루고 있어 대중교통 정책 지원이 어려운 실정
 - 전체 교통 DB 중에서 가구통행실태조사를 제외하고 가장 많은 양을 차지하는 데이터는 차량에 탑재된 기기(내비게이션, DTG 등)나 도로에 설치된 기기(검지기, DSRC 등)를 통해 수집된 데이터임
- 차량 외 타 교통수단의 이동을 모니터링하기 위해서는 도보, 버스, 철도 등 모든 통행자의 통행 정보가 상세하게 기록된 통신 빅데이터를 적극적으로 활용할 필요
 - 전체 사람의 이동수요 대비 대중교통 이용 변화량을 파악 가능해야 데이터 기반의 대중교통 정책 수립이 가능
- 이에 교통모니터링·분석·데이터 제공 플랫폼(View-T) 내 통신 빅데이터 DB를 구축하고, 이를 기반으로 분석할 수 있는 다양한 서비스 기능을 개발하고자 함

나. 과업의 범위 및 내용

1) 과업의 범위

- 시간적 범위
 - 모바일 빅데이터 가공 알고리즘 개발: 2018년 4월에 생성된 원천데이터
 - 2018년도 기준 기반 DB 구축: 2018년 1월 1일 ~ 2018년 12월 31일 (1년)
 - 웹 서비스를 위한 각종 분석도구 개발: 2017년 1월 1일 ~ 2017년 12월 31일
 - 2018년도 연구 성과물인 2017년도 기준 기반 DB 활용
- 공간적 범위: 전국

2) 과업의 내용

○ 모바일 빅데이터 가공 알고리즘 개발

- 통행 왜곡데이터 보정 방법 개발: 통행 정보를 왜곡할 수 있는 Case를 제시하고, 이를 보정할 수 있는 방법론을 제시
- 체류 정보 추출 방법 개발: 새로 개발한 전처리 방법론을 적용한 데이터를 기준으로 체류/이동 정보를 구분하는 방법론을 개발
- 3통행 미만 통행량 보정 방법 개발: 개인정보보호법에 위배되지 않는 선에서 3미만 통행량을 보정할 수 방법을 개발

○ 모바일 빅데이터 기반 기종점 통행량 구축

- 모바일 빅데이터 기종점 통행량 구축: 2018년 1~12월에 생성된 모바일 빅데이터를 대상으로 본 과업을 통해 개발한 가공 알고리즘 적용하여 2018년 기준 DB를 구축
- 구축 DB 검증: 2018년 기준 DB에서 이상치가 존재하지 않는 지 확인, 통상적인 통행 패턴을 나타내는지 확인

○ View-T 모바일 빅데이터 기반 서비스 기능 개발

- 기존 서비스 개선 및 유지보수: 2018년에 개발한 분석 서비스(주요통행지역 분석, 주요통행구간 분석, Hot Place 분석, 통행량 검색 기능) 개선
- View-T 서비스 유지를 위한 기반 DB 업데이트: 2018년 과업 성과물인 2017년 기준 통신 빅데이터 기반 기종점 통행량 DB, 통신 분석맵 DB을 시스템에 탑재
- 서비스 기능 추가 개발: 세분화된 체류지 구분 코드, 통행 목적 정보를 통해 구현할 수 있는 서비스 기능 개발

2. 모바일 빅데이터 가공 알고리즘 개발

가. 통행 왜곡데이터 보정 방법

1) 통행 왜곡데이터

- 통행 왜곡 데이터란, 로그 기록에서 통행자의 움직임에 의해 기록된 데이터가 아닌 신호 이상에 의해 기록된 데이터를 의미하며, 이러한 데이터는 실제 통행 중에 기록된 데이터 사이사이에 기록되어 있어 실제 통행정보를 왜곡시킴
- 본 과업에서 보정하고자 하는 통행 왜곡 데이터는 시그널 점프 데이터와 핑퐁 핸드오버 데이터임
 - 시그널 점프 데이터 (Signal jump): 비현실적인 속도로 기지국의 위치가 변환되어 기록된 데이터
 - 핑퐁 핸드오버 데이터 (Ping-pong Handover): 짧은 시간동안 복수개의 기지국이 연달아 번갈아서 기록된 데이터로, 크게 두 가지 유형으로 구분됨
 - 유형 1: 두 개의 기지국이 연달아 번갈아 기록된 데이터 (예: 1→2→1→2)
⇒ 신호 세기가 센 주 기지국이 2개인 경우
 - 유형 2: 둘 이상의 기지국이 순환하면서 기록된 데이터 (예: 1→2→3→1)
⇒ 신호 세기가 센 주 기지국이 1개인 경우

2) 시그널 점프 데이터 보정 방법

- Iovan et al. (2013)이 제시한 속도기반 방법(Speed-based correction method)에 의거, 허용 임계치보다 신호 기록 속도가 크게 나타날 경우 시그널 점프 데이터로 간주하고, 시그널 점프가 발생한 데이터는 하나의 데이터로 행 압축
- (1단계) 원천데이터 로그 기록을 고객ID별로 분할한 후 기록 순서(체류일자, 체류시작시간, 체류종료시간)에 따라 나열하여 순번을 부여
- (2단계) 기지국간의 기록 속도가 매우 빠르게 나타나는 데이터를 탐색
 - 현재 기록의 체류종료시간, 기지국 좌표 정보와 다음 순번 기록의 체류시작시간, 기지국 좌표 정보를 기준으로 기록 시간 차(time-gap)과 직선거리를 산출하여 기록 속도를 추정하고, 시간 차가 0분으로 나타나 속도 산출 불능이거나, 산출한 속도가 300km/h이상인 경우 시그널 점프 데이터로 분류

- (3단계) 시그널 점프가 발생한 데이터를 하나의 로그 기록으로 압축
 - (시간 정보 보정) 시그널 점프가 발생한 데이터에서 가장 먼저 기록된 데이터의 체류시작시간 정보와, 가장 마지막에 기록된 데이터의 체류종료시간 정보를 이용하여 다시 체류시작시간과 체류종료시간을 업데이트
 - (위치 정보 보정) 고개별 일자별 기지국별 기록 빈도와 기지국별 총 체류시간 (해당 기지국이 기록된 체류시간을 모두 더한 값)을 산출한 후, 시그널 점프가 발생한 기지국간 결과를 비교하여 빈도가 높고, 체류시간이 길게 나타난 기지국 정보로 위치 정보를 업데이트
- ㈜KT에서 제공한 일주일간의 원천데이터(분 단위의 로그 기록)를 활용하여 전술한 보정 알고리즘을 적용한 결과, 통행자가 체류했을 가능성이 낮은 10분 이하의 기록은 줄어들고, 10분을 초과하는 기록은 증가하는 것으로 나타남
 - 시간적 범위: 2018. 04. 16. (월)
 - 공간적 범위: 전국에 기록된 1분 단위의 로그 기록

3) 핑퐁 핸드오버 데이터 보정 방법

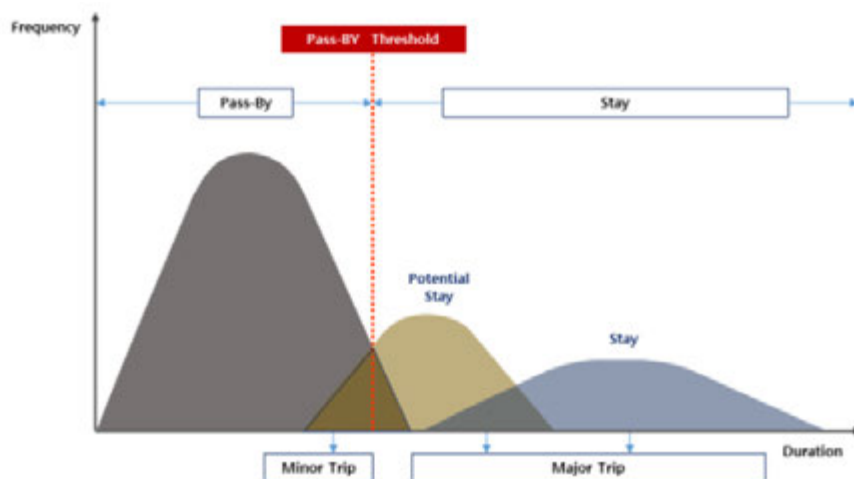
- Lee and Hou(2006)이 제안한 패턴기반 방법(Pattern-based correction method)에 의거, 두 개의 기지국간 연속적인 변동이 나타나는 패턴을 보이는 경우 핑퐁 핸드오버 데이터로 간주하고, 가장 긴 체류시간을 가지는 기지국으로 보정
 - (1단계) 시그널 점프 데이터를 보정
 - (2단계) 일자별, 고객별 두 개의 기지국이 번갈아 나타나는 데이터(A-B-A'-B') 중에서 기록된 데이터의 시간 차이가 1분 이내인 데이터를 탐색
 - (3단계) 핑퐁 핸드오버 데이터로 탐색된 데이터는 하나의 로그 기록으로 압축
 - (시간 정보 보정) 가장 먼저 기록된 A 기지국의 데이터에서 체류시작시간 정보와, 가장 마지막에 기록된 B'기지국의 데이터에서 체류종료시간 정보를 가져와 새로운 체류시작시간과 체류종료시간을 생성
 - (위치 정보 보정) 고개별 일자별 기지국별 총 체류시간(해당 기지국이 기록된 체류시간을 모두 더한 값)을 산출한 후, 핑퐁 핸드오버가 발생한 기지국(A와 B)간의 결과를 비교하여 체류시간이 길게 나타난 기지국 정보로 위치 정보를 업데이트
- ㈜KT에서 제공한 일주일간의 원천데이터(분 단위의 로그 기록)를 활용하여 전술한 보정 알고리즘을 적용한 결과, 180분을 초과한 체류 기록은 증가하고, 그 외의 기록

은 모두 줄어든 것으로 나타났으며, 특히 10분 이하의 체류일 때 보정 효과가 높은 것으로 나타남

- 시간적 범위: 2018. 04. 16. (월)
- 공간적 범위: 전국에 기록된 1분 단위의 로그 기록

나. 체류 정보 추출 방법

- ‘체류시간 25분’을 이동(Pass-by)과 체류(Stay)를 구분하는 기준으로 삼았던 기존 방식에서 벗어나 이동과 체류가 중첩되는 1차 변곡점을 체류와 이동을 구분하는 임계치로 설정함
- 체류시간 25분이란 기준은 일부 기지국에 체류했던 이용자들의 체류시간을 기준으로 누적분포함수 값을 산출하여 결정한 것이지만, 누적분포함수값은 체류시간 분포를 설명하는 값으로 대체로 통행자가 체류 중이라고 판단할 수 있는 근거가 될 수는 있지만, 체류와 이동을 구분하는 임계치 역할로는 부적합
- 로그 기록 빈도를 체류시간별로 출력해보면 세 개의 마루를 갖는 굴곡을 띤 그래프가 그려지고, 이러한 형태는 ‘이동, 잠재체류, 주 체류’라는 통행자의 세 가지 체류 형태 특성이 반영된 것으로 볼 수 있음
- 본 과업에서는 첫 번째 굴곡과 두 번째 굴곡이 만나는 지점을 이동과 체류 특성이 만나는 지점으로 보고, 이동 행태와 체류 행태 특성이 변화하기 시작하는 변곡점을 이동과 체류를 구분하는 임계치로 둬
- 원천데이터에 전처리 알고리즘을 적용한 후 체류시간별 빈도를 추출한 결과, 체류시간 15분에서 1차 변곡점이 나타나는 것으로 확인



<그림 1> 경로 체류시간 임계치 설정

다. 3통행 미만 통행량 보정 방법

- 모바일 빅데이터를 활용하는 타 과제에서는 개인정보보호법에 저촉되지 않도록 최대한 3통행 미만이 나오지 않도록 DB를 설계하거나, 3통행 미만인 경우 실제 통행량을 정확하게 알 수 없도록 비식별화 처리하고 있음
- 하지만 위와 같은 방식은 활용성이 떨어질 우려가 있고, 무엇보다 비식별화 처리 방안은 모두 3통행으로 처리하는 기존 방식과 반대로 오히려 통행량이 과소하게 집계될 수 있기 때문에 본 과제에서는 후(後) 보정 방법으로 3통행 미만 통행을 보정함
 - 후 보정 방법: 이미 설계한 DB에 맞춰 모두 3통행으로 처리된 3미만의 통행량을 원래 값과 유사하도록 다시 되돌리는 방법
- 본 과제에서 정의한 3통행 미만 보정 방법은 다음 세 가지 단계로 정의할 수 있음
 - (1단계) 보정계수 기준 달 설정: 2018년도 과업을 통해 구축한 2017년 1월부터 12월까지의 모바일 빅데이터 기반 기종점 통행량 DB에서 연평균 일통행량(AADT)을 산출한 후 이와 가장 유사한 일평균 통행량 값을 갖는 달을 기준 달로 설정
 - 연평균 일통행량: 120,402,573 통행 / 기준달(6월) 일통행량: 119,626,502 통행
 - (2단계) 보정계수 기준 달 설정: 출발 정보와 성, 연령을 기준 필드로 두고 원천데이터에서 집계한 통행량을 기준 DB에서 같은 조건으로 집계한 통행량으로 나누어 보정계수를 산출
 - (3단계) 기준 달의 보정계수로 보정할 수 없는 경우에는 연령을 제외하고 출발 시도, 출발 요일, 출발 시간대, 출발지 체류 특성, 통행자 성별 정보를 기준으로 기산출한 보정계수의 평균값을 보정계수로 활용
- 다음 표는 2017년 기준 DB에 위에서 정의한 3통행 미만 보정 방법을 적용한 결과임

<표 1> 3통행 미만 통행량 보정계수 적용 전후 비교

구분	보정계수 적용 전	보정계수 적용 후
연평균 일통행량 (AADT)	120,402,573	41,558,998
KT 시장점유율 적용 결과 ¹⁾	427.353.477	147.508.329
인당 일통행량 ²⁾	8.26	2.85

주 1: 2017.12 기준 KT 시장점유율 28% 적용

주 2: 시장점유율을 적용한 연평균 일통행량을 대한민국 인구(51,709,098명)로 나눈 값

3. 모바일 빅데이터 기반 기종점 통행량 DB 구축

가. 기준년도 DB 구축

- 본 과업에서는 플랫폼(View-T)에 모바일 DB가 기반 DB로 활용될 것을 고려하여 법률에 저촉되지 않도록 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)를 기준으로 통행 정보가 동일한 인구를 집계하여 <표 2>와 같은 형태로 기준년도 DB를 구축함

<표 2> 모바일 기반 DB 형태

출발				도착				성별 코드	연령 코드	통 행 량	통행 시간 (분)
일자	시 간	교통 폴리곤 ID	체류지 유형	일자	시 간	교통 폴리곤 ID	체류지 유형				
20181121	13	48270310	집	20181123	14	12390981	기타	M	20	10	34
20180501	06	48270310	집	20180501	09	32789014	회사	F	50	25	39
20180311	17	48270320	집	20180312	18	45608912	학교	M	10	15	25

- 원천데이터에서 위와 같은 형태의 DB로 가공하는 단계를 정리하면 다음과 같음
 - (1단계) 원천데이터 수집 및 이상치 제거
 - 2018년 1월 1일부터 2018년 12월 31일까지 기록된 로그를 수집
 - 기지국 좌표가 Null값이거나 행정동 정보가 매칭되지 않는 데이터를 제거
 - (2단계) 선분 이력으로 변환
 - 포인트 단위의(밀리세컨 단위) 로그를 선분 단위의(분 단위) 로그로 변환
 - (3단계) 통행 정보 왜곡 데이터 보완
 - 통행 정보를 왜곡할 수 있는 ‘시그널 점프 데이터’, ‘핑퐁 핸드오버 데이터’를 본 과업에서 개발한 전처리 기술을 통해 보정함
 - (4단계) 체류순서 부여
 - 로그 기록 일자, 로그 기록 시작 시간(이하 ‘체류시작시간’), 로그 기록 종료 시간(이하 ‘체류종료시간’)을 기준으로 로그 기록 순서에 따라 개인별 데이터를 정렬하여 체류 순서를 구분
 - (5단계) 체류 정보 추출

- 이동 중에 기록된 로그 기록을 제외하고 체류 중에 기록된 로그 기록만을 추출 (체류시간이 15분을 초과하는 로그 기록을 추출)
- (6단계) 체류지 유형 구분
 - 통행자의 체류 형태에 따라 체류지를 5가지(집, 회사, 학교, 정기적 잠재체류지, 비정기적 잠재체류지)로 구분
- (7단계) 위치정보, 시간정보, 연령정보 변환
 - 개인의 이동궤적을 추적할 수 없도록 원천데이터에서 출발지와 도착지의 위치 정보는 분석 맵의 교통폴리곤 ID로 변경한 후, 시간 정보에서 분 정보를 생략한 시간대 필드를 생성하고, 1세 단위의 연령 정보는 10세 단위로 변경
- (8단계) 통행량 집계
 - 체류순서에 따라 출발과 도착을 구분한 다음, 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)가 동일한 로그 기록을 집계
- (9단계) 통행량 보정
 - 통행량이 3 이하인 경우 모두 3통행으로 변환한 후, 3통행 미만 보정계수를 적용하여 통행량을 보정
 - 실제 통행 분석에 활용할 때에는 필요에 따라 통신업체의 시장점유율을(KT 시장점유율) 기준으로 전수화 하여 사용
- (10단계) 평균통행시간 산출
 - 각 개인의 통행정보에서 통행시간(도착시간-출발시간, 분 단위)을 산출한 다음, 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)가 동일한 로그 기록의 통행시간을 집계한 후, 집계한 통행시간을 산술평균하여 산출

나. 구축 DB 검증

1) 기존 방식의 기·종점 통행량과 단순 비교

- 2016년도 기준 가구통행실태조사를 통해 추정된 기종점 통행량과 비교하여 지역간 통행량, 지역내 통행량 통행 패턴이 유사한지 확인
- 비교 분석 결과, 모바일 빅데이터 기반 인당 일평균 통행횟수는 3.9회로 기존 방식보다 약 2~3배 높게 나타나는 것으로 확인

- 시도별 통행 분포는 거의 일치하고, 서울과 경기지역에서만 통행 비율이 약 2~5% 차이는 것으로 분석됨

<표 3> 전국 기·종점 통행 비율 (비교 결과)

(단위: %)

○ \ D	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	계
서울	3.62	-0.01	-0.01	-0.44	-0.01	-0.02	0.00	-0.01	-2.50	-0.06	-0.04	-0.05	-0.01	-0.01	-0.02	-0.01	-0.01	0.42
부산	-0.01	0.44	-0.01	0.00	0.00	0.00	-0.04	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.31	0.00	0.02
대구	-0.01	-0.01	0.22	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.34	-0.02	0.00	-0.19
인천	-0.39	0.00	0.00	0.80	0.00	0.00	0.00	0.00	-0.36	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01
광주	-0.01	0.00	0.00	0.00	-0.35	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.02	-0.31	0.00	0.00	0.00	-0.71
대전	-0.02	0.00	0.00	0.00	0.00	-0.05	0.00	-0.08	-0.02	0.00	-0.07	-0.09	-0.01	0.00	-0.01	0.00	0.00	-0.38
울산	0.00	-0.05	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.05	-0.04	0.00	0.03
세종	-0.01	0.00	0.00	0.00	0.00	-0.08	0.00	0.13	-0.01	0.00	-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	-0.02
경기	-2.49	-0.01	-0.01	-0.37	-0.01	-0.02	0.00	-0.01	4.57	-0.05	-0.04	-0.12	-0.01	-0.01	-0.01	-0.01	-0.01	1.41
강원	-0.05	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	-0.05	0.35	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.22
충북	-0.04	0.00	0.00	-0.01	0.00	-0.08	0.00	-0.03	-0.04	-0.01	0.09	-0.02	0.00	0.00	0.00	0.00	0.00	-0.16
충남	-0.05	0.00	0.00	-0.02	0.00	-0.10	0.00	-0.03	-0.12	0.00	-0.03	-0.01	0.00	0.00	0.00	0.00	0.00	-0.35
전북	-0.01	0.00	0.00	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	0.00	0.00	0.02	-0.02	0.00	0.00	0.00	-0.07
전남	-0.01	-0.01	0.00	0.00	-0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	0.19	0.00	-0.01	0.00	0.04
경북	-0.01	-0.01	-0.32	0.00	0.00	-0.01	-0.05	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.24	-0.02	0.00	-0.20
경남	-0.01	-0.30	-0.02	0.00	0.00	0.00	-0.04	0.00	-0.01	0.00	0.00	0.00	-0.01	-0.01	-0.02	0.56	0.00	0.14
제주	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.16	-0.19
계	0.47	0.03	-0.18	-0.08	-0.50	-0.38	0.03	-0.02	1.42	0.22	-0.15	-0.34	-0.07	-0.17	-0.23	0.13	-0.19	0.00

주 : 모바일 자료 기반 통행 비율에서 KTDB 통행 비율을 뺀 값임.

2) 파라미터 변화에 따른 통행 수요 변화

- 시간대별, 요일별, 성별, 연령별 통행량 변화를 확인하여 통상적인 통행 패턴을 벗어나는지 확인
 - 시간대별 통행량 변화
 - (예상) 심야시간대에는 주간시간대보다 통행량이 낮게 나타남 → (결과) 심야시간대 통행량은 주간시간대 통행량의 1/3 수준인 것으로 분석됨
 - 요일별 통행량 변화
 - (예상) 평일에는 주기적으로 발생하는 통근·통학 통행으로 인해 주말에 비해 평일의 통행량이 높게 나타남 → (결과) 월요일부터 금요일까지 평일동안 일정 수준의 통행량이 유지되다가 토요일부터 통행이 줄어들어 일요일에는 통행량이 가장 낮게 나타남
 - 성별 통행량 변화

- (예상) 남/여 통행량은 차이가 없음 → (결과) 전체 통행량에서 남성과 여성의 통행량은 평균적으로 53.8%, 46.2%를 차지하는 것으로 나타났으며, 남성과 여성의 인구 성비가 50.1% 49.9%인 것을 고려할 때 통행량에는 거의 차이가 없다고 볼 수 있음
- 연령별 통행량 변화
 - (예상) 경제활동인구(20세 이상 60세 미만)의 통행량이 고령인구(60세 이상)에 비해 통행량이 높게 나타남 → (결과) 경제활동인구의 통행량이 가장 많고, 60세 이상부터 점차 줄어드는 것으로 나타남

4. 모바일 빅데이터 기반 통행 분석 서비스 개발

가. 기반 DB 구축

1) 분석 맵 구축

- 2018년에 개발한 교통폴리곤 영역과 2016년 기준 행정구역 정보를 시스템에 구축

<표 4> 분석 맵 업데이트

구분	기존 구축 DB	신규 구축 DB
교통폴리곤	- 2015년 집계구와 2016년 기지국 정보를 활용하여 구축한 영역 (2017년 개발)	- 2016년 집계구와 2017년 기지국 정보를 활용하여 구축한 영역 (2018년 개발)
행정구역	- 2015년 기준 읍면동, 시군구, 시도 영역	- 2016년 기준 읍면동, 시군구, 시도 영역

2) 모바일 빅데이터 기반 기종점 통행량 DB 구축

- 2017년에 생성된 로그 기록을 활용하여 구축한 기종점 통행량 DB를 시스템에 구축
 - 3통행 미만 보정계수를 적용하여 과다하게 집계되던 통행량을 보정

<표 5> 모바일 빅데이터 기종점 통행량 업데이트

구분	기존 구축 DB	신규 구축 DB
시간적 범위	2016년 3월 ~ 2016년 12월	2017년 1월 ~ 2017년 12월
공간적 범위	전국	좌동
기타	-	3통행 미만 보정계수 적용

나. 분석 기능 개발

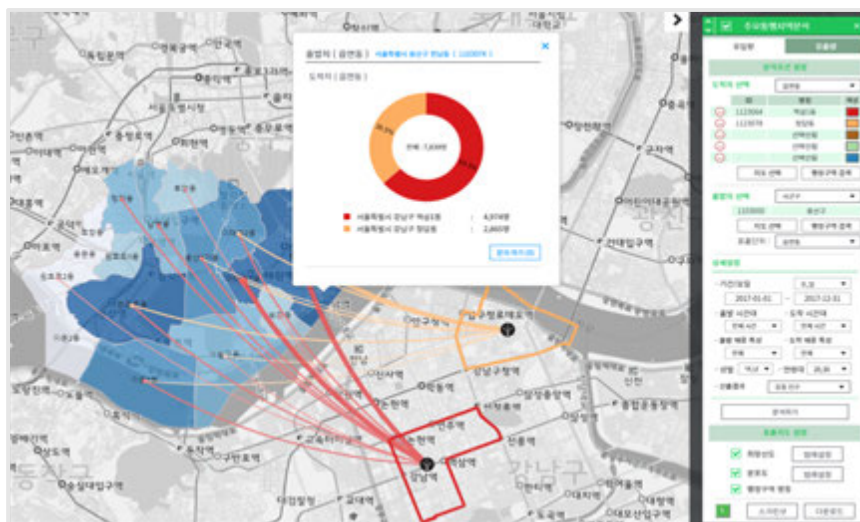
1) 기존 기능 개선

① 주요통행지역분석

- 기존보다 더 많은 지역을 비교 분석할 수 있도록 개선
- 사용자가 선택한 분석 대상 지역의 결과를 더 쉽게 구분할 수 있도록 표출 방법을 개선
- 사용자가 분석 결과를 더 빨리 파악할 수 있도록 행정구역 명칭 ON/OFF 기능을 추가
- 선택지역에 대한 분석 결과를 더 쉽게 비교할 수 있도록 상세 정보도 시각화하여 표현

<표 6> 주요통행지역분석 개선 내역

구분	기존	개선
분석 대상 지역	- 2개 지역까지 선택 가능	- 5개 지역까지 선택 가능 - 권한에 따라 선택 개수 제한
표출 방법	- 선택지역의 분석결과가 동일한 색상으로 표현	- 선택지역별 분석결과를 각기 다른 색상으로 표현
기타	- 주제도 클릭시 상세결과를 텍스트 형식으로 확인	- 행정구역 ON/OFF 기능 추가 - 주제도 클릭시 상세결과를 파이차트 형식으로 확인 가능



<그림 2> 주요통행지역분석 분석 기능 개발 결과

② 주요통행구간분석 (전 O/D 분석)

- 지역간 통행정보도 확인할 수 있도록 개선
 - 기존에는 사용자가 선택한 지역을 기준으로 지역 내 주요 통행 구간만 확인 가능
- 기존보다 더 많은 지역을 분석할 수 있도록 분석 대상 지역 선택 폭 확대
- 분석 결과에 대한 표출 단위를 다양하게 설정할 수 있도록 개선

<표 7> 주요통행구간분석 개선 내역

구분	기존	개선
분석 대상 지역	- 1개 지역만 선택 가능	- 5개 지역까지 선택 가능 - 권한에 따라 선택 개수 제한
분석 결과	- 지역 내 통행 정보 확인 가능	- 지역 간 통행 정보 확인 가능
표출 단위	- 교통폴리곤 단위	- 시군구 단위 - 읍면동 단위 - 교통폴리곤 단위



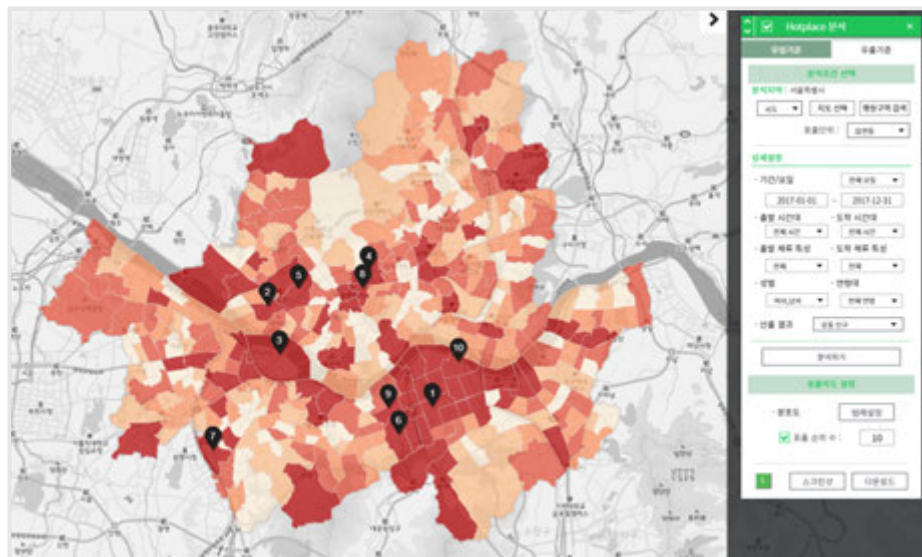
<그림 3> 주요통행구간분석 기능 개발 결과

③ Hot Place 분석

- 분석 결과에 대한 표출 단위를 다양하게 설정할 수 있도록 개선
- 산출 결과를 기존보다 다양하게 확인할 수 있도록 개선

<표 8> Hot place 분석 개선 내역

구분	기존	개선
표출 단위	- 교통폴리곤 단위	- 시군구 단위 - 읍면동 단위 - 교통폴리곤 단위
표출 결과	- 통행량 (총 합계)	- 통행량 (총 합계) - 일평균 통행량 (통행량/분석일) - 면적대비 통행량 (통행량/분석단위 면적)



<그림 4> Hot Place 분석 기능 개발 결과

④ 부가 서비스 (데이터 검색 및 다운로드)

- 다양한 형태로 통신데이터를 집계해서 다운로드 할 수 있도록 유형별 통신데이터 다운로드 기능과 지역간 OD 통행량 데이터 다운로드 기능 추가
 - 기존에 사용자가 원천데이터를 다운받은 후 필요에 따라 집계해서 사용해야 했던 불

필요한 단계를 단축

- 사용자 대기시간을 단축시키기 위해 표출화면에 표기되지 않는 데이터인 경우 분석결과를 메일로 받아볼 수 있도록 함

<표 9> 부가 서비스 개선 내역

구분	기존	개선
출력(통행량 집계) 조건 설정	- 설정 불가	- 사용자가 설정한 행 조건(가로축), 열 조건(세로축)에 따라 다양하게 출력(집계) 가능 - 지역간 통행량 출력(집계) 가능
기타	-	- 분석 결과를 메일로 송부

① 데이터 설정

- 기준년도, 기간, 요일 설정
- 분석지역 공간 범위 설정
- 시간대, 체류특성, 성별, 연령대 설정

② 테이블 설정

- 가로축/세로축 별 다양한 집계유형 조건 설정

③ 다운로드

- 데이터 설정과 테이블 설정의 조건으로 분석한 결과를 사용자의 E-Mail로 발송

④ 테이블 설정에 따른 예시

- 사용자가 설정한 조건을 테이블 형태로 예시를 나타냄

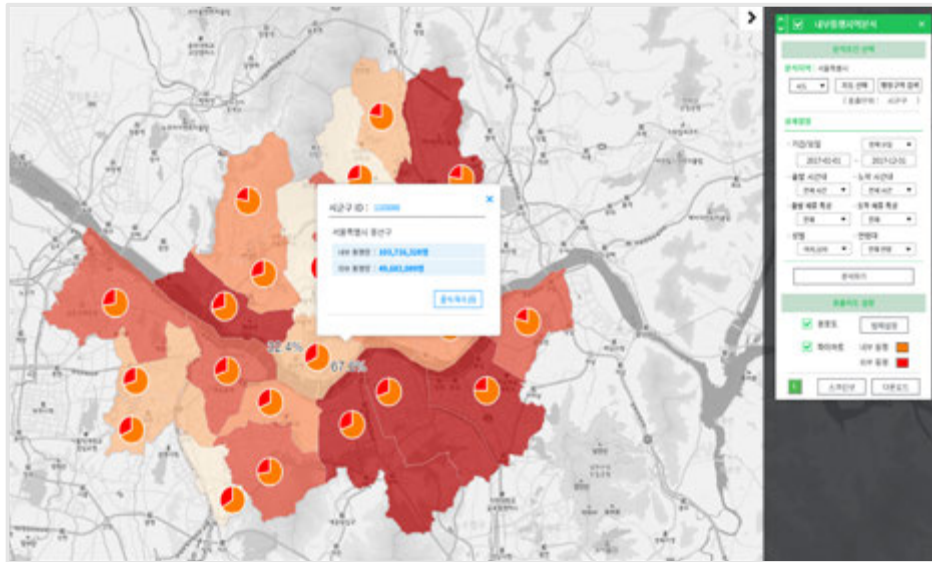
<그림 5> 부가서비스 - 유형별 통신 데이터 다운로드 UI

2) 신규 기능 개발

① 내부통행지역분석

- 내부 통행이 많은 지역을 쉽게 확인할 수 있는 분석 서비스를 개발

- 사용자가 선택한 지역의 내부 통행 비율을 주제도에서 확인할 수 있도록 기능 개발

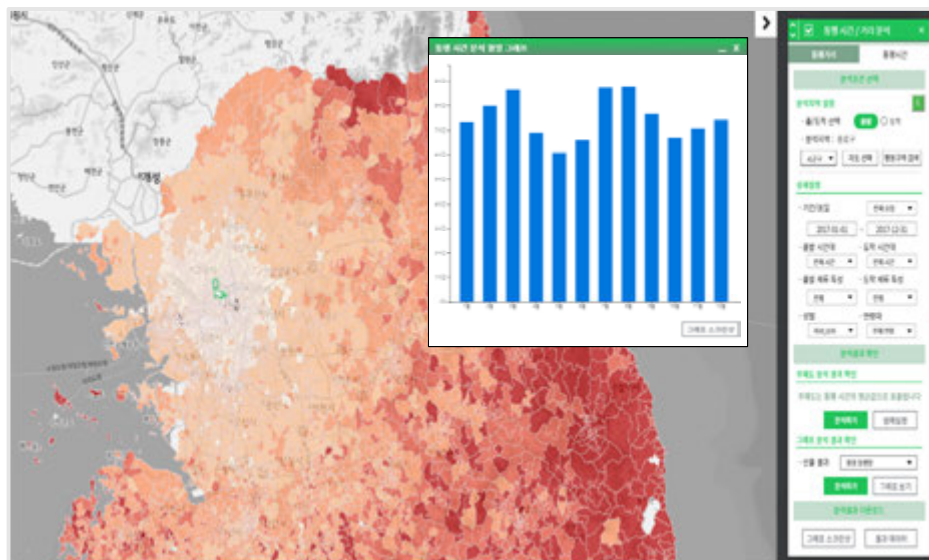


<그림 6> 내부통행분석 기능 개발 결과

② 통행시간 / 거리 분석

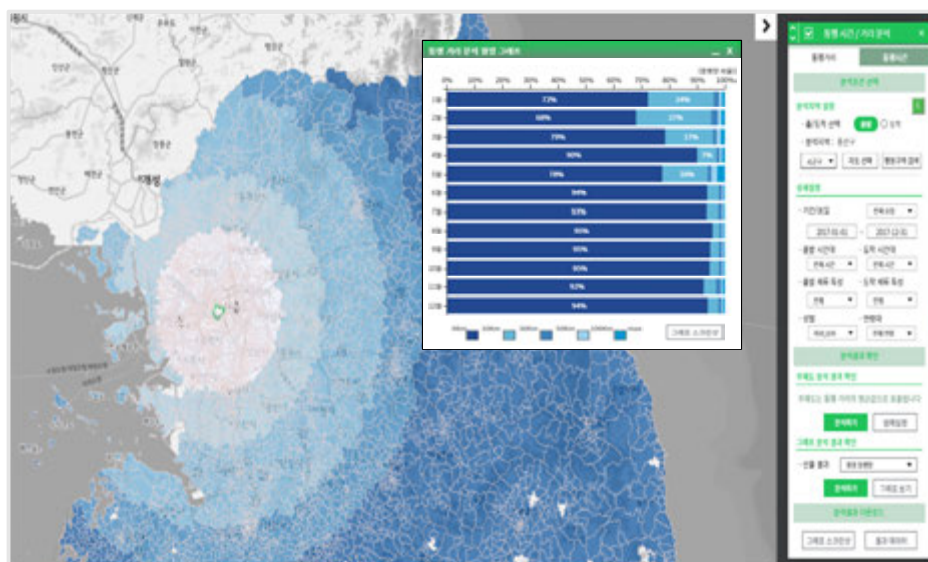
- 지역별 접근성을 쉽게 확인할 수 있는 분석 서비스를 개발

- 사용자가 선택한 지역을 기준으로 평균 통행 시간을 주제도에서 확인할 수 있도록 기능 개발



<그림 7> 통행 시간 분석 기능 개발 결과

- 사용자가 선택한 지역을 기준으로 평균 통행 거리 정보를 주제도에서 확인할 수 있도록 기능 개발



<그림 8> 통행 거리 분석의 주제도·그래프 표출 화면 예시

3) 기타(종합)

① 분석 서비스 및 메뉴 구성 변경

- 주요 분석 서비스는 기존에 개발한 주요 분석 기능 3가지와 새로 개발하는 분석 기능 2가지로 구성
 - 기존 기능에서 성격이 유사한 유입유출 지역분석 기능과 유입유출 비교분석 기능을 하나로 통합하고, 기존 분석 기능 명칭을 사용자가 이해하기 쉽도록 변경
 - 신규 분석 기능으로 내부통행지역분석, 통행거리/시간분석 기능 추가
- 메뉴 순서는 사용 빈도, 분석 기능의 복잡성에 따라 변경하고, 새로 개발하는 분석 기능을 주요 분석 기능 메뉴에 추가

② 사용방법 안내 방식 변경

- 기존 pdf 형태의 매뉴얼을 제공하는 방식에서 각 기능 메뉴 선택 시 설정 부분에 숫

자를 표기하여 기능사용 순서를 유도하는 방식으로 사용방법 안내 방식을 변경

③ 분석 조건 설정 방법 개선

- 기간 설정 시 수기 입력이 가능하도록 개선
- 요일 설정의 주중과 주말에 대한 선택 버튼을 추가하여 사용자 편의성을 높임
- 시간대 설정의 오전 첨두(6시~9시)와 오후 첨두(18시~21시)에 대한 선택 버튼을 추가하여 사용자가 자주 분석하는 시간대의 설정을 용이하게 함

④ 범례 설정 방법 개선

- 등간격, 등분위, 상위%, 직접입력 총 네가지 방식으로 범례를 설정할 수 있도록 함
- 다양한 색상의 심볼 색상 스키마를 제공하여 사용자의 편의성을 증진
- 또한 제공되는 스키마와 별개로 사용자가 설정한 시작과 종료의 색상으로 그라데이션의 스키마를 생성할 수 있게끔 함

⑤ 분석결과 유형 추가

- 주제도로 표출되는 분석 결과도 사용자가 활용할 수 있도록 스크린샷 다운로드 기능을 개발

⑥ 주제도 라벨 폰트 추가

- 주제도 표출 시 설정하는 라벨의 폰트를 추가하여 사용자 선택의 다양성을 높임

5. 결론 및 향후 과제

가. 결론

- 본 과업에서는 교통 분석에 모바일 DB가 더욱 활발히 활용될 수 있도록 '18년에 개발한 모바일 가공 알고리즘을 보완하였으며, 2018년에 생성된 모바일 빅데이터에 보완된 알고리즘을 적용하여 2018년 기준 기종점 통행량 DB를 구축하였음
- 통행자의 통행에 의해 생성된 로그가 아니라 단순 신호 이상으로 기록된 통행 왜곡 데이터(시그널 점프 데이터, 핑퐁 핸드오버 데이터)를 보정하는 방법론을 추가로 개발함
- 새로운 전처리 방법론을 적용한 데이터를 기준으로 통행자가 체류 중인지, 이동 중인지를 구분할 수 있는 기준을 다시 설정함
- 3통행 미만의 데이터가 모두 3통행으로 처리되면서 통행량이 과다하게 집계되었던 문제를 보완하기 위하여 3미만 통행 보정계수를 개발함
- 2018년 1월 1일부터 2018년 12월 31일까지 생성된 로그 기록에 새로 개발한 알고리즘을 적용하여 기존 DB를 구축하고, 적합하게 가공되었는지 기존 인력식 조사 방식의 기종점 통행량과 비교하여 검증함
- 또한 교통 모니터링·데이터 제공·분석 플랫폼(View-T)에 '18년 과업을 통해 구축한 모바일 빅데이터 기반 기종점 통행량 DB와 모바일 분석 맵 DB를 탑재하였고, 기존 개발한 사람 통행 분석 서비스를 보다 사용하기 편리하게 보완하였으며, 더 다양하고 세밀한 분석이 가능하도록 새로운 분석 서비스 2가지를 추가로 개발하였음

<표 10> View-T 2.0 사람 통행 분석 서비스 개발 주요 내용 (기반 DB 구축)

구분		내용
기반 DB 구축	분석 맵 구축	<ul style="list-style-type: none"> • 2016년 기준 교통폴리곤 (2016년 집계구와 2017년 기지국 정보를 활용하여 구축한 영역) 폴리곤/노드 구축 • 2016년 행정구역 단위의 폴리곤/노드 구축
	모바일 빅데이터 기반 기종점 통행량 DB 구축	<ul style="list-style-type: none"> • 2017년 1월 ~ 12월 기준 DB 구축 • 3통행 미만 보정계수 적용

<표 11> View-T 2.0 사람 통행 분석 서비스 개발 주요 내용 (분석 기능 개발)

구분			내용
분석 기능 개발	기존 기능 개선	주요통행 지역분석	<ul style="list-style-type: none">• 기존 유입유출 지역분석과 유입유출 비교분석기능을 하나로 통합• 기존보다 더 많은 지역을 비교 분석할 수 있도록 개선 (2개 지역 → 5개 지역)• 사용자가 선택한 분석 대상 지역의 결과를 더 쉽게 구분할 수 있도록 선택지역별 분석결과를 각기 다른 색상으로 표현할 수 있도록 함• 사용자가 분석 결과를 더 빨리 파악할 수 있도록 행정구역 명칭 ON/OFF 기능을 추가• 선택지역에 대한 분석 결과를 더 쉽게 비교할 수 있도록 상세 정보를 파이차트 형식으로 시각화하여 표현
		주요통행 구간분석	<ul style="list-style-type: none">• 지역간 통행정보도 확인할 수 있도록 개선• (지역내 최다 통행구간 확인 → 지역간 최다 통행구간 확인)• 기존보다 더 많은 지역을 분석할 수 있도록 분석 대상 지역 선택 폭 확대 (1개 지역 → 5개 지역)• 분석 결과에 대한 표출 단위를 다양하게 설정할 수 있도록 개선 (교통폴리곤 단위 → 시군구, 읍면동, 교통폴리곤 단위)
		Hot Place 분석	<ul style="list-style-type: none">• 분석 결과에 대한 표출 단위를 다양하게 설정할 수 있도록 개선 (교통폴리곤 단위 → 시군구, 읍면동, 교통폴리곤 단위)• 산출 결과를 기존보다 다양하게 확인할 수 있도록 개선 (통행량 → 통행량, 일평균 통행량, 면적대비 통행량)
		부가기능 (데이터 검색 및 다운로드)	<ul style="list-style-type: none">• 유형별 통신데이터 다운로드 기능과 지역간 OD 통행량 데이터 다운로드 기능 추가• 사용자 대기시간을 단축시키기 위해 표출화면에 표기되지 않는 데이터인 경우 분석결과를 메일로 받아볼 수 있도록 함
	신규 기능 개발	내부통행 지역분석	<ul style="list-style-type: none">• 내부 통행이 많은 지역을 쉽게 확인할 수 있는 분석 서비스를 개발
		통행시간/ 거리분석	<ul style="list-style-type: none">• 지역별 접근성을 평균통행시간과 평균통행거리 정보를 통해 쉽게 확인할 수 있도록 새로운 분석 서비스를 개발
기타(종합)			<ul style="list-style-type: none">• (사용방법 안내) 매뉴얼→ 작동시 사용 순서 유도, 설명 팁 표출• (분석조건 설정) 분석기간 수기 입력 방식 추가, 사용자가 자주 사용하는 조건 옵션 추가(주중/주말 옵션, 오전 침투시간대/오후 침투시간대 옵션)• (범례 설정) 자주 사용하는 범례 설정 방식 옵션 추가 (등간격, 등분위, 상위%), 심볼(색상)스키마 다양하게 구성• (분석결과 다운로드) 주제도 스크린샷 기능 추가• 주제로 라벨 폰트 변경 기능 추가

나. 향후 과제

- 본 과업을 통해 통행목적이 포함된 DB를 구축하였으나, 향후 궁극적으로 MaaS(Mobility as a service)를 실현하기 위해서는 통행자의 이동수단정보가 포함된 DB가 구축되어야 할 것임
- 또한 향후 모바일 DB에 대한 신뢰도를 높일 수 있도록 Prompt-recall 서비스를 실시할 필요가 있음
- 아울러, 교통 모니터링·데이터 제공·분석 플랫폼(View-T)을 통해 제공되고 있는 분석 서비스를 사용자 편의성, 활용성 측면을 고려하여 보완할 필요가 있으며, 계속해서 축적되는 DB를 시스템 내에서 효율적으로 관리할 수 있는 방안을 모색할 필요가 있음
 - 1차 분석 결과(주제도 표출)를 토대로 사용자가 분석결과를 보다 심층적으로 분석할 수 있도록 상세 분석 창을 추가로 생성할 필요가 있음
 - 차량 통행 분석 결과와 연계시켜 함께 분석할 수 있는 방안을 모색할 필요가 있음

제1장 과업의 개요

제1절 과업의 배경 및 목적

제2절 과업의 범위 및 내용

제1장 과업의 개요

제1절 과업의 배경 및 목적

1. 과업의 배경

- 현재 교통 DB는 도로 중심의 데이터가 주를 이루고 있어 대중교통 정책 지원이 어려운 실정
 - 전체 교통 DB 중에서 가구통행실태조사를 제외하고 가장 많은 양을 차지하는 데이터는 차량에 탑재된 기기(내비게이션, DTG 등)나 도로에 설치된 기기(검지기, DSRC 등)를 통해 수집된 데이터임
- 차량 외 타 교통수단의 이동을 모니터링하기 위해서는 도보, 버스, 철도 등 모든 통행자의 통행 정보가 상세하게 기록된 통신 빅데이터를 적극적으로 활용할 필요
 - 전체 사람의 이동수요 대비 대중교통 이용 변화량을 파악 가능해야 데이터 기반의 대중교통 정책 수립이 가능
- 이에 교통모니터링·분석·데이터 제공 플랫폼(View-T) 내 통신 빅데이터 DB를 구축하고, 이를 기반으로 분석할 수 있는 다양한 서비스 기능을 개발하고자 함

2. 과업의 목적

- 통행자별 이용 교통수단 정보가 포함된 통신 빅데이터 DB를 구축하고, 이를 기반으로 교통수요를 분석할 수 있는 지표, 기능 등을 개발하여 고도화된 대중교통 정책 수립 지원에 활용하고자 함

제2절 과업의 범위 및 내용

1. 과업의 범위

- 시간적 범위
 - － 모바일 빅데이터 가공 알고리즘 개발: 2018년 4월에 생성된 원천데이터
 - － 2018년도 기준 기반 DB 구축: 2018년 1월 1일 ~ 2018년 12월 31일 (1년)
 - － 웹 서비스를 위한 각종 분석도구 개발: 2017년 1월 1일 ~ 2017년 12월 31일
 - 2018년도 연구 성과물인 2017년도 기준 기반 DB 활용
- 공간적 범위: 전국

2. 과업의 내용

1) 모바일 빅데이터 가공 알고리즘 개발

- 통행 왜곡데이터 보정 방법 개발
 - － 통행 정보를 왜곡할 수 있는 Case를 제시하고, 이를 보정할 수 있는 방법론을 제시
- 체류 정보 추출 방법 개발
 - － 새로 개발한 전처리 방법론을 적용한 데이터를 기준으로 체류/이동 정보를 구분하는 방법론을 개발
- 3통행 미만 통행량 보정 방법 개발
 - － 개인정보보호법에 위배되지 않는 선에서 3미만 통행량을 보정할 수 방법을 개발

2) 모바일 빅데이터 기반 기종점 통행량 구축

- 모바일 빅데이터 기종점 통행량 구축
 - － 2018년 1~12월에 생성된 모바일 빅데이터를 대상으로 본 과업을 통해 개발한 가공 알고리즘 적용하여 2018년 기준 DB를 구축
- 구축 DB 검증

- 2018년 기준 DB에서 이상치가 존재하지 않는 지 확인, 통상적인 통행 패턴을 나타내는지 확인

3) View-T 모바일 빅데이터 기반 서비스 기능 개발

- 기존 서비스 개선 및 유지보수
 - 2018년에 개발한 분석 서비스(주요통행지역 분석, 주요통행구간 분석, Hot Place 분석, 통행량 검색 기능) 개선
- View-T 서비스 유지를 위한 기반 DB 업데이트
 - 2018년 과업 성과물인 2017년 기준 통신 빅데이터 기반 기종점 통행량 DB, 통신 분석맵 DB을 시스템에 탑재
 - 웹 서비스를 위한 데이터 정제, 웹 서비스 속도 향상을 위한 데이터 경량화
- 서비스 기능 추가 개발
 - 세분화된 체류지 구분 코드, 통행 목적 정보를 통해 구현할 수 있는 서비스 기능 개발

제2장 모바일 빅데이터 가공 알고리즘 개발

제1절 통행 왜곡데이터 보정 방법

제2절 체류 정보 추출 방법

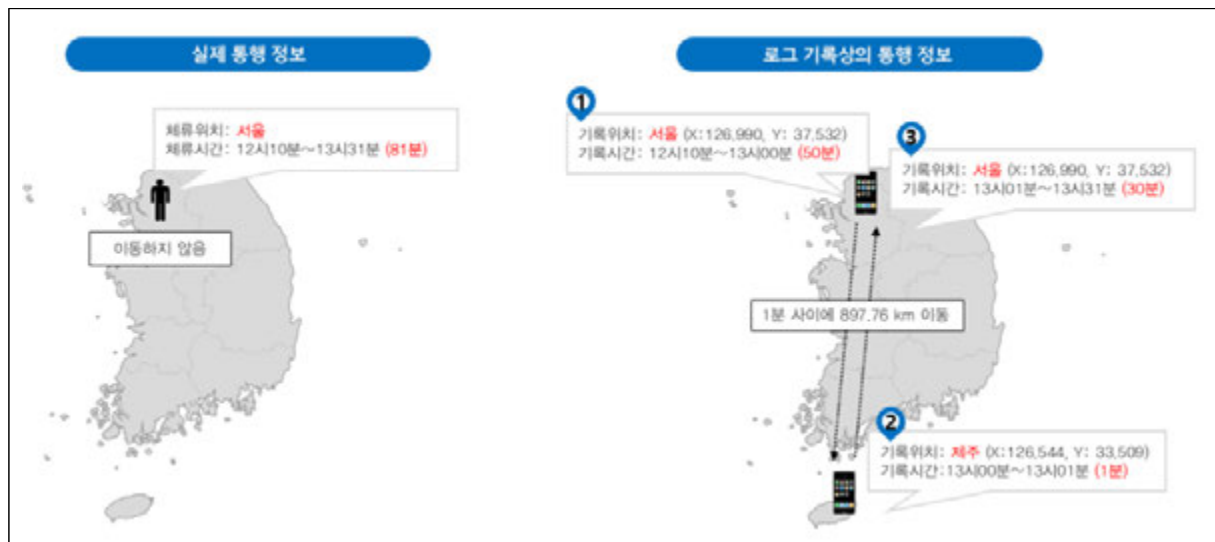
제3절 3통행 미만 통행량 보정 방법

제2장 모바일 빅데이터 가공 알고리즘 개발

제1절 통행 왜곡데이터 보정 방법

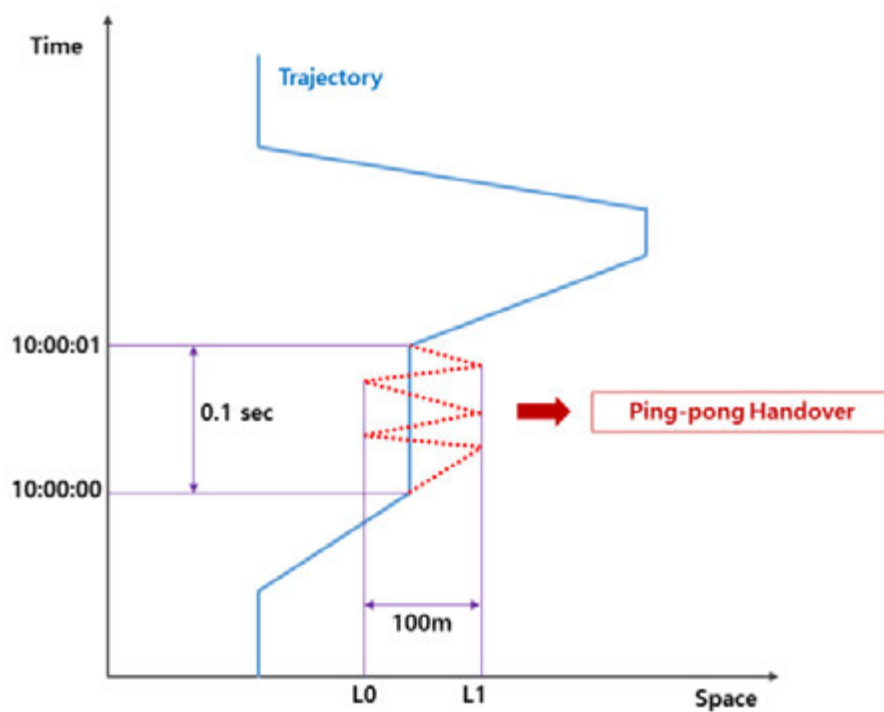
1. 통행 왜곡데이터

- 통행 왜곡 데이터란, 로그 기록에서 통행자의 움직임에 의해 기록된 데이터가 아닌 신호 이상에 의해 기록된 데이터를 의미하며, 이러한 데이터는 실제 통행 중에 기록된 데이터 사이 사이에 기록되어 있어 실제 통행정보를 왜곡시킴
 - 실제로는 이동하지 않았어도 이동한 것처럼 보이게 하거나, 매우 빠른 속도로 이동하고 있는 것처럼 보이게 함
 - 통행 왜곡 데이터를 보정하지 않으면 체류한 기록이 줄어들 수 있고, 통행자의 체류시간이 과소 추정될 수 있으며, 체류지 유형 구분에도 영향을 줄 수 있음
 - 체류지 유형은 체류 빈도와 체류시간을 기준으로 구분되기 때문
- 이러한 통행 왜곡데이터는 단말기를 연결할 기지국을 계속 변경하는 과정(핸드오버 과정)에서 다음 세 가지 경우일 때 발생할 수 있음
 - 특정 기지국의 신호 세기가 매우 센 경우
 - 기지국 간의 거리가 가까운 경우(기지국 수신범위가 작은 경우)
 - 기지국간의 신호 세기가 비슷한 경우
 - 전압조절을 하는 심야시간대에 기지국간의 신호 세기가 비슷해질 수 있음
- 통행 왜곡 데이터는 크게 두 가지로 구분할 수 있음
 - 시그널 점프 데이터 (Signal jump): 비현실적인 속도로 기지국의 위치가 변환되어 기록된 데이터



<그림 2- 1> 시그널 점프 데이터 유형

- 핑퐁 핸드오버 데이터 (Ping-pong Handover): 짧은 시간동안 복수개의 기지국이 연달아 번갈아서 기록된 데이터로, 크게 두 가지 유형으로 구분됨



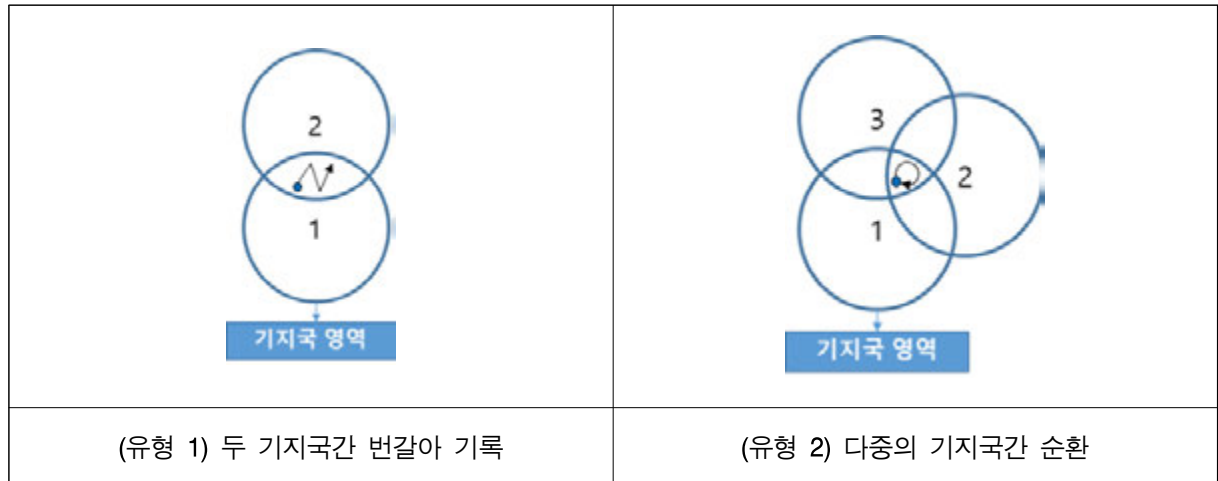
<그림 2- 2> 핑퐁 핸드오버 데이터

- 유형 1: 두 개의 기지국이 연달아 번갈아 기록된 데이터 (예: 1→2→1→2)

⇒ 신호 세기가 센 주 기지국이 2개인 경우

- 유형 2: 둘 이상의 기지국이 순환하면서 기록된 데이터 (예: 1→2→3→1)

⇒ 신호 세기가 센 주 기지국이 1개인 경우



<그림 2- 3> 핑퐁 핸드오버 데이터 유형

1. 시그널 점프 데이터 보정 방법

가. 기본방향

- 속도기반 방법(Speed-based correction method)에 의거, 허용 임계치보다 신호 기록 속도가 크게 나타날 경우 시그널 점프 데이터로 간주
 - 속도기반 방법은 본래 핑퐁 핸드오버 데이터를 탐색하는 방법 중 하나로(Iovan et al., 2013), 두 기지국간 연속적인 변동이 발생하는 패턴 중 허용 임계치(200km/h)보다 큰 기지국을 탐색하여 제외하는 방법을 말함
 - 시그널 점프 데이터를 탐색하는 데 속도기반 방법을 적용하는 이유는 시그널 점프 데이터도 해석에 따라 핑퐁 핸드오버 데이터의 일종으로 볼 수 있고, 시그널 점프 데이터의 주요한 특징이 신호 이상에 의한 기록이기 때문에 기록 속도가 비현실적으로 빠르다는 것이기 때문
- 시그널 점프가 발생한 데이터는 하나의 데이터로 행 압축
 - Iovan et al.(2013)은 탐색된 데이터를 제외하는 방법을 제시했으나, 단순 제거하는 방식으로 보정할 경우 시그널 점프 현상에 의해 분리된 데이터가 제대로 보정되지 않음
 - 따라서 시그널 점프가 발생한 기지국간의 기록을 하나의 기록으로 변환하는 방식으로 보정

할 필요가 있음

나. 보정 방법

- (1단계) 원천데이터 로그 기록을 기록 순서에 따라 나열하여 순번을 부여
 - 체류일자, 기록시작시간(이하 체류시작시간), 기록종료시간(이하 체류종료시간) 정보를 기준으로 오름차순으로 데이터를 나열하여 전체 데이터에 대한 순번을 부여
 - 고객식별번호(고객 id)를 기준으로 로그 기록을 분할하여 위와 같은 기준으로 데이터를 나열하여 고객별 기록 순번을 부여
- (2단계) 기지국간의 기록속도가 매우 빠르게 나타나는 데이터를 탐색
 - 현재 기록의 체류종료시간, 기지국 좌표 정보와 다음 순번 기록의 체류시작시간, 기지국 좌표 정보를 기준으로 기록 시간 차(time-gap)와 직선거리를 산출하여 기록 속도를 추정
 - 기록 시간 차(time-gap): 다음 순번 기록의 체류시작시간에서 현재 기록의 체류종료시간을 뺀 값
 - 직선거리: 다음 순번의 기지국 좌표와 현재 기록의 기지국 좌표를 기준으로 유클리디안 거리(Euclidean distance) 공식을 이용하여 산출한 값

<표 2- 1> 기록 간 속도 산출 방법

전체 순번	id별 순번	고객ID	체류시작 시간	체류종료 시간	기지국 좌표	시간차 (분)	거리 (Km)	속도 (km/h)
514	70	E12233	1141	1141	126.360, 33.352	1	7.2	454
515	71	E12233	1142	1142	126.406, 33.299			

- 시간 차가 0분으로 나타나 속도를 산출할 수 없거나, 산출한 속도가 300km/h이상인 경우 시그널 점프 데이터로 분류
- 속도 임계치를 300km/h로 둔 이유는 비행기 탑승시 휴대폰 전원을 끄거나 비행기 모드로 전환하기 때문에 모바일 빅데이터에서 탐지할 수 있는 최대 통행속도는 KTX 이동속도, 즉 300km/h이기 때문
- 시간 차가 0분인 경우는 분모가 0이 되어 속도를 산출할 수는 없으나 분모 값이 작을수록 속도는 빨라지기 때문에 매우 빠른 속도로 기록된 것으로 추정할 수 있으므로 보정대상에 포함시킴

- (3단계) 시그널 점프가 발생한 데이터를 하나의 로그 기록으로 압축
 - (시간 정보 보정) 시그널 점프가 발생한 데이터에서 가장 먼저 기록된 데이터의 체류시작 시간 정보와, 가장 마지막에 기록된 데이터의 체류종료시간 정보를 이용하여 다시 체류시작 시간과 체류종료시간을 업데이트
 - (위치 정보 보정) 고개별 일자별 기지국별 기록 빈도와 기지국별 총 체류시간(해당 기지국이 기록된 체류시간을 모두 더한 값)을 산출한 후, 시그널 점프가 발생한 기지국간 결과를 비교하여 빈도가 높고, 체류시간이 길게 나타난 기지국 정보로 위치 정보를 업데이트
 - 시그널 점프가 발생하면 신호 이상을 일으킨 기지국에 기록되었다가 본래 기지국으로 되돌아와 기록되기 때문에 신호 이상을 일으킨 기지국보다 본래 기지국의 기록 횟수가 더 많고, 체류시간도 길수밖에 없음

<표 2- 2> 시그널 점프 데이터 행 압축 방법

전체 순번	id별 순번	고객ID	체류시작 시간	체류종료 시간	기지국 좌표	기지국 기록 빈도	기지국별 총 체류시간(분)
514	70	E12233	1141	1141	126.360, 33.352	10	35
515	71	E12233	1142	1142	126.406, 33.299	3	5
↓							
514	70	E12233	1141	1142	126.360, 33.352		

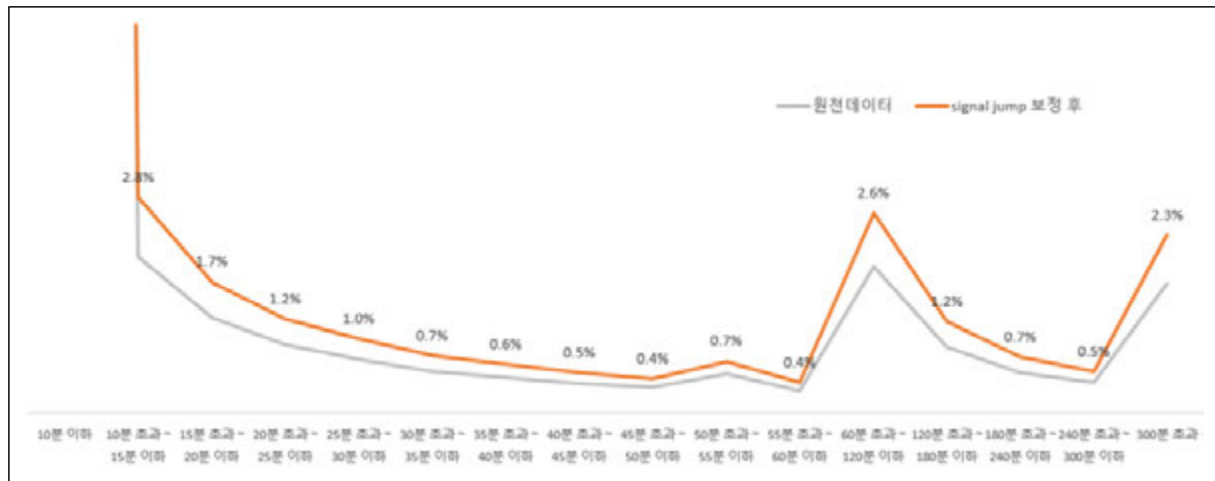
다. 보정 결과

- (주) KT에서 제공한 원천데이터(분 단위의 로그 기록)에 전술한 보정 알고리즘을 적용
 - 시간적 범위: 2018.04.16.(월)
 - 공간적 범위: 전국에 기록된 1분 단위의 로그 기록
- 시그널 점프 보정 결과, 10분 이하의 체류 기록은 줄어들고, 10분 초과인 체류 기록은 증가하는 것으로 나타남
 - Bayir et al(2010), Kim and Kwan (2003)¹⁾ 등 기존 문헌에서 10분 이상의 체류 기록을 체류했다고 간주한다는 점을 고려할 때, 통행자가 체류했을 가능성이 있는 10분 이상의 데

¹⁾ M. A. Bayir, M. Demirbas and N. Eagle, (2010). Mobility profiler: A framework for discovering mobility profiles of cell phone users. Pervasive and Mobile Computing, 6, 435-454. / H. Kim and M. Kwan, (2003). Space-time accessibility measures: A geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration. Journal of Geographical Systems, 5, 71-91.

이터가 보완된 것을 알 수 있음

- 시그널 점프 현상으로 나뉜 체류 기록이 다시 합쳐지면서 체류시간이 긴 체류기록의 빈도가 증가한 것임



<그림 2- 4> 시그널 점프 보정 전 후 비교

<표 2- 3> 시그널 점프 보정 전 후 비교

체류시간	시그널 점프 보정 전 (원천데이터)		시그널 점프 보정 후		증감	
	빈도	비율	빈도	비율	빈도	기존 대비 증감율
10분 이하	725,698,670	87.3%	506,472,180	82.7%	- 219,226,490	-30.2%
10분 초과 ~ 15분 이하	16,715,640	2.0%	16,953,440	2.8%	237,800	1.4%
15분 초과 ~ 20분 이하	10,191,930	1.2%	10,335,630	1.7%	143,700	1.4%
20분 초과 ~ 25분 이하	7,302,700	0.9%	7,377,140	1.2%	74,440	1.0%
25분 초과 ~ 30분 이하	5,802,640	0.7%	5,840,800	1.0%	38,160	0.7%
30분 초과 ~ 35분 이하	4,508,690	0.5%	4,547,590	0.7%	38,900	0.9%
35분 초과 ~ 40분 이하	3,818,520	0.5%	3,840,870	0.6%	22,350	0.6%
40분 초과 ~ 45분 이하	3,190,520	0.4%	3,214,080	0.5%	23,560	0.7%
45분 초과 ~ 50분 이하	2,725,480	0.3%	2,744,920	0.4%	19,440	0.7%
50분 초과 ~ 55분 이하	4,159,520	0.5%	4,107,140	0.7%	- 52,380	-1.3%
55분 초과 ~ 60분 이하	2,439,870	0.3%	2,456,460	0.4%	16,590	0.7%
60분 초과 ~ 120분 이하	15,643,420	1.9%	15,773,560	2.6%	130,140	0.8%
120분 초과 ~ 180분 이하	7,132,480	0.9%	7,203,670	1.2%	71,190	1.0%
180분 초과 ~ 240분 이하	4,379,770	0.5%	4,423,490	0.7%	43,720	1.0%
240분 초과 ~ 300분 이하	3,257,360	0.4%	3,288,010	0.5%	30,650	0.9%
300분 초과	13,857,580	1.7%	13,982,880	2.3%	125,300	0.9%

2. 핑퐁 핸드오버 데이터 보정 방법

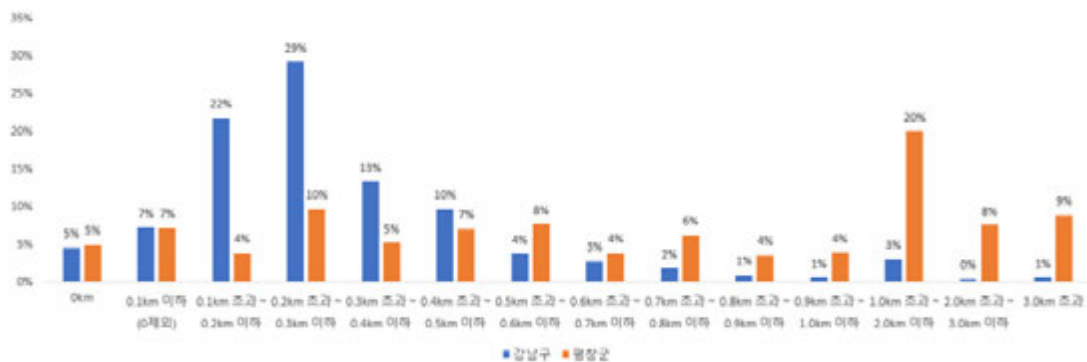
가. 기본방향

- Lee and Hou(2006)이 제안한 패턴기반 방법(Pattern-based correction method)에 의거, 두 개의 기지국간 연속적인 변동이 나타나는 패턴을 보이는 경우 핑퐁 핸드오버 데이터로 볼 수 있음
- 본 과업에서는 기지국의 수신범위를 정확히 알 수 없다는 한계점이 존재하여, 전술한 핑퐁 핸드오버의 두 가지 유형 중에 한 가지 유형만 보정 대상으로 설정함
 - 보정대상 핸드오버 유형: (유형1) 두 개의 기지국이 연달아 번갈아 기록된 데이터 (예: 1→2→1→2)
- 그러나 이러한 연속적인 변동 패턴을 보이는 모든 데이터를 핑퐁 핸드오버로 간주하고 보정할 경우 실제 통행에 의한 기록까지 보정되면서 로그 기록이 과다하게 압축될 우려가 있음
- 따라서 신호 이상에 의한 핸드오버와 실제 통행에 의한 반복 기록을 구분할 수 있는 임계값을 찾기 위해, 2018.04.16.(월) ~ 2018.04.22.(일)까지 기록된 원천데이터에서 핑퐁 핸드오버 패턴을 보이는 데이터를 추출하여 그 특징을 분석함
 - 기지국 간 거리가 가까울수록 핑퐁 핸드오버가 발생할 가능성이 높다는 이론에 근거하여, 기지국 밀도가 높은 도시부와 기지국 밀도가 낮은 지방부의 특징이 다르게 나타날 것으로 예상하고 기지국 밀도가 가장 높은 지역과 기지국 밀도가 가장 낮은 지역을 구분하여 분석
 - 기지국 밀도가 높은 서울(8.88개/km²)의 강남구과 기지국 밀도가 낮은 강원(0.42개/km²)의 평창군을 분석 대상지역으로 선정



<그림 2- 5> 핑퐁 핸드오버 데이터 분석 범위

- 분석 결과, 핑퐁 핸드오버 패턴을 보이는 데이터의 50% 이상은 지역에 상관없이 1분 이내의 시간 간격을 두고 기록된 것으로 나타났으며, 도시부는 기지국 간의 거리가 가까운 곳에서 핑퐁 핸드오버 패턴이 많이 기록된 데 반해 지방부는 비교적 거리가 먼 곳에서 핑퐁 핸드오버 패턴이 많이 기록된 것으로 나타남
- 거리 분포 분석 결과: 서울 강남구는 90%의 핑퐁 핸드오버 데이터가 0.6km 이하의 거리에서 나타나고, 0.1km 초과 0.2km 이하 또는 0.2km 초과 0.3km 이하에서 빈도가 집중적으로 높게 나타난 반면, 강원 평창군은 1.0km 이하의 거리까지는 고루 분포하다가 1.0km 초과 2.0km 이하일 때 빈도가 높게 나타남

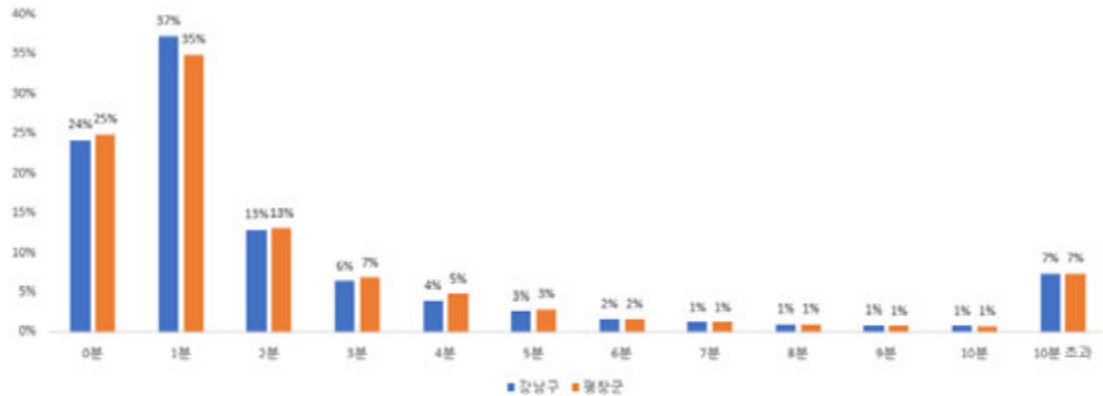


<그림 2- 6> 강남구와 평창군의 핑퐁 핸드오버 발생 특성 - 거리

<표 2- 4> 거리별 핑퐁 핸드오버 발생빈도 및 비율 (강남구 vs 평창군)

거리	강남구		평창군	
	빈도	비율	빈도	비율
0km	2,691,572	4.6%	111,511	5.0%
0km 초과 ~ 0.1km 이하	4,296,495	7.3%	159,851	7.2%
0.1km 초과 ~ 0.2km 이하	12,851,035	21.8%	85,339	3.8%
0.2km 초과 ~ 0.3km 이하	17,246,025	29.2%	216,714	9.8%
0.3km 초과 ~ 0.4km 이하	7,916,499	13.4%	115,010	5.2%
0.4km 초과 ~ 0.5km 이하	5,729,356	9.7%	157,919	7.1%
0.5km 초과 ~ 0.6km 이하	2,244,464	3.8%	170,589	7.7%
0.6km 초과 ~ 0.7km 이하	1,649,797	2.8%	86,532	3.9%
0.7km 초과 ~ 0.8km 이하	1,084,460	1.8%	138,145	6.2%
0.8km 초과 ~ 0.9km 이하	581,070	1.0%	80,372	3.6%
0.9km 초과 ~ 1.0km 이하	349,776	0.6%	88,191	4.0%
1.0km 초과 ~ 2.0km 이하	1,778,235	3.0%	445,894	20.1%
2.0km 초과 ~ 3.0km 이하	217,806	0.4%	167,825	7.6%
3.0km 초과	362,233	0.6%	197,854	8.9%

- 시간 차 분포 분석 결과: 두 지역 모두 기지국간의 기록 시간 차이가 1분일 때 핑퐁 핸드오버가 많이 발생하는 것으로 나타났고, 기지국간의 기록 시간 차이가 1분 이하일 때까지는 빈도가 증가하다가 2분일 때부터 빈도가 줄어들기 시작함



<그림 2- 7> 강남구와 평창군의 핑퐁 핸드오버 발생 특성 - 시간 차

<표 2- 5> 시간 차에 따른 핑퐁 핸드오버 발생빈도 및 비율 (강남구 vs 평창군)

시간 차	강남구		평창군	
	빈도	비율	빈도	비율
0분	14,267,741	24.2%	553,901	24.9%
1분	21,991,564	37.3%	775,923	34.9%
2분	7,599,198	12.9%	290,999	13.1%
3분	3,781,109	6.4%	152,377	6.9%
4분	2,339,543	4.0%	107,355	4.8%
5분	1,553,656	2.6%	62,396	2.8%
6분	943,191	1.6%	35,522	1.6%
7분	707,383	1.2%	25,942	1.2%
8분	566,348	1.0%	21,015	0.9%
9분	467,069	0.8%	17,020	0.8%
10분	439,350	0.7%	15,778	0.7%
10분 초과	4,342,671	7.4%	163,518	7.4%

- 따라서 본 과업에서는 핑퐁 핸드오버 패턴이 나타나는 데이터 중에서 기록 시간 차이가 1분 이내인 데이터만을 핑퐁 핸드오버 데이터로 간주하고, Lee and Hou(2006)이 제안한 방법을

본 과업에 적용하여 보정하고자 함

- 도시부와 지방부의 핑퐁 핸드오버 탐색 기준을 서울 강남구와 강원 평창군의 분석 결과로 결정하기에는 대표성의 문제가 있을 수 있으므로, 지역에 관계없이 공통적으로 나타나는 시간 차 특성을 핑퐁 핸드오버 탐색 기준으로 설정함
- Lee and Hou(2006)이 제안한대로 핑퐁 핸드오버 패턴이 관찰된 경우 가장 긴 체류시간을 가지는 기지국으로 보정

나. 보정 방법

- (1단계) 시그널 점프 데이터를 보정
 - 시그널 점프 데이터는 전술한 보정 알고리즘 적용하여 보정
- (2단계) 두 개의 기지국이 번갈아 나타나는 데이터 중에서 기록된 데이터의 시간 차이가 1분 이내인 데이터를 탐색
 - 하루 단위로(0시 0분~23시 59분), 고객별로 탐색
 - 기록 일자가 다음 날로 변경되는 경우 제외
(예: 기지국 기록 패턴은 A-B-A'-B'로 나타났지만 마지막 B'가 기록된 일자가 앞의 A-B-A'가 기록된 일자와 다른 경우)
 - 고객 ID가 다른 경우 제외
(예: 기지국 기록 패턴은 A-B-A'-B'로 나타났지만, A-B-A'의 기록은 E고객의 기록이고, 마지막 B'가 기록은 G고객의 기록인 경우)

<표 2- 6> 핑퐁 핸드오버 탐색 결과 (예시)

전체 순번	id별 순번	고객ID	체류시작시간	체류종료시간	시간 차	기지국 좌표
514	70	E12233	1141	1141	1분 1분 1분 1분	126.360, 33.352 (A)
515	71	E12233	1142	1142		126.406, 33.299 (B)
516	72	E12233	1143	1145		126.360, 33.352 (A')
517	73	E12233	1146	1146		126.406, 33.299 (B')

- (3단계) 핑퐁 핸드오버 데이터로 탐색된 데이터는 하나의 로그 기록으로 압축
 - (시간 정보 보정) 가장 먼저 기록된 A 기지국의 데이터에서 체류시작시간 정보와, 가장 마

지막에 기록된 B'기지국의 데이터에서 체류종료시간 정보를 가져와 새로운 체류시작시간과 체류종료시간을 생성

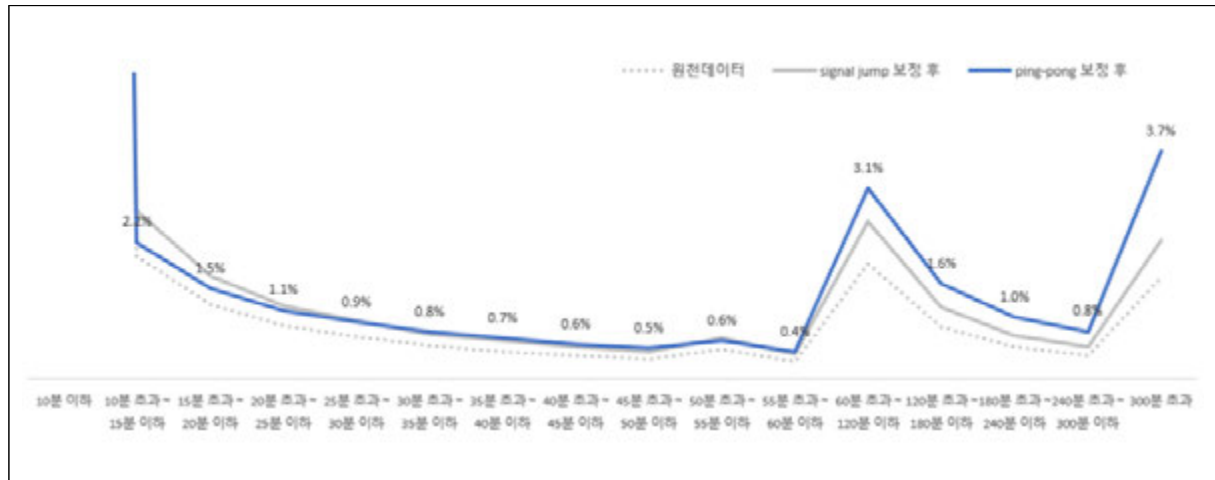
- (위치 정보 보정) 고개별 일자별 기지국별 총 체류시간(해당 기지국이 기록된 체류시간을 모두 더한 값)을 산출한 후, 핑퐁 핸드오버가 발생한 기지국(A와 B)간의 결과를 비교하여 체류시간이 길게 나타난 기지국 정보로 위치 정보를 업데이트

<표 2- 7> 핑퐁 핸드오버 데이터 행 압축 방법

전체 순번	id별 순번	고객ID	체류시작시간	체류종료시간	기지국 좌표	기지국별 총 체류시간(분)
514	70	E12233	1141	1141	126.360, 33.352 (A)	35
515	71	E12233	1142	1142	126.406, 33.299 (B)	5
516	72	E12233	1143	1145	126.360, 33.352 (A')	
517	73	E12233	1146	1146	126.406, 33.299 (B')	
↓						
514	70	E12233	1141	1146	126.360, 33.352	

다. 보정 결과

- (주) KT에서 제공한 원천데이터(분 단위의 로그 기록)에 전술한 보정 알고리즘을 적용
 - 시간적 범위: 2018.04.16.(월)
 - 공간적 범위: 전국에 기록된 1분 단위의 로그 기록
- 핑퐁 핸드오버 보정 결과, 180분 이하의 체류 기록은 줄어들고, 180분 초과인 체류 기록은 증가하는 것으로 나타남
 - 10분 이하의 체류 기록이 가장 많이 보정된 것으로 나타남 (<표 2-8> 참조)



<그림 2- 8> 핑퐁 핸드오버 데이터 보정 전 후 비교

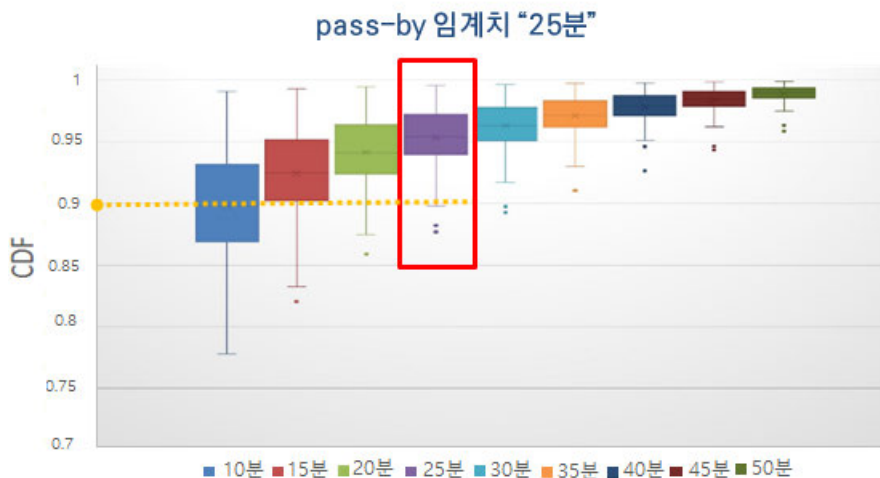
<표 2- 8> 핑퐁 핸드오버 데이터 보정 전 후 비교

체류시간	시그널 점프 보정 후		핑퐁 핸드오버 보정 후		증감	
	빈도	비율 (%)	빈도	비율 (%)	빈도	기존 대비 증감율
10분 이하	506,472,180	82.7%	360,874,110	80.4%	- 145,598,070	-28.7%
10분 초과 ~ 15분 이하	16,953,440	2.8%	10,044,630	2.2%	- 6,908,810	-40.8%
15분 초과 ~ 20분 이하	10,335,630	1.7%	6,656,820	1.5%	- 3,678,810	-35.6%
20분 초과 ~ 25분 이하	7,377,140	1.2%	5,061,760	1.1%	- 2,315,380	-31.4%
25분 초과 ~ 30분 이하	5,840,800	1.0%	4,169,080	0.9%	- 1,671,720	-28.6%
30분 초과 ~ 35분 이하	4,547,590	0.7%	3,450,570	0.8%	- 1,097,020	-24.1%
35분 초과 ~ 40분 이하	3,840,870	0.6%	2,985,550	0.7%	- 855,320	-22.3%
40분 초과 ~ 45분 이하	3,214,080	0.5%	2,567,190	0.6%	- 646,890	-20.1%
45분 초과 ~ 50분 이하	2,744,920	0.4%	2,255,150	0.5%	- 489,770	-17.8%
50분 초과 ~ 55분 이하	4,107,140	0.7%	2,820,030	0.6%	- 1,287,110	-31.3%
55분 초과 ~ 60분 이하	2,456,460	0.4%	2,014,090	0.4%	- 442,370	-18.0%
60분 초과 ~ 120분 이하	15,773,560	2.6%	13,977,620	3.1%	- 1,795,940	-11.4%
120분 초과 ~ 180분 이하	7,203,670	1.2%	6,984,520	1.6%	- 219,150	-3.0%
180분 초과 ~ 240분 이하	4,423,490	0.7%	4,526,230	1.0%	102,740	2.3%
240분 초과 ~ 300분 이하	3,288,010	0.5%	3,494,710	0.8%	206,700	6.3%
300분 초과	13,982,880	2.3%	16,716,920	3.7%	2,734,040	19.6%

제2절 체류 정보 추출 방법

1. 기존 방법 검토 및 기준 변경의 필요성

- 기존에는 체류(Stay)와 이동(Pass-by)을 구분하는 기준으로 누적분포함수 값(CDF)을 활용
 - 일부 기지국(서울 마포구, 세종시 200개의 기지국)에 체류했던 이용자들의 체류시간(1~60분)을 기준으로 누적 분포함수 값을 산출하여, 모든 누적분포함수 값이 90%를 넘는 체류시간을 체류와 이동을 구분하는 기준으로 둠



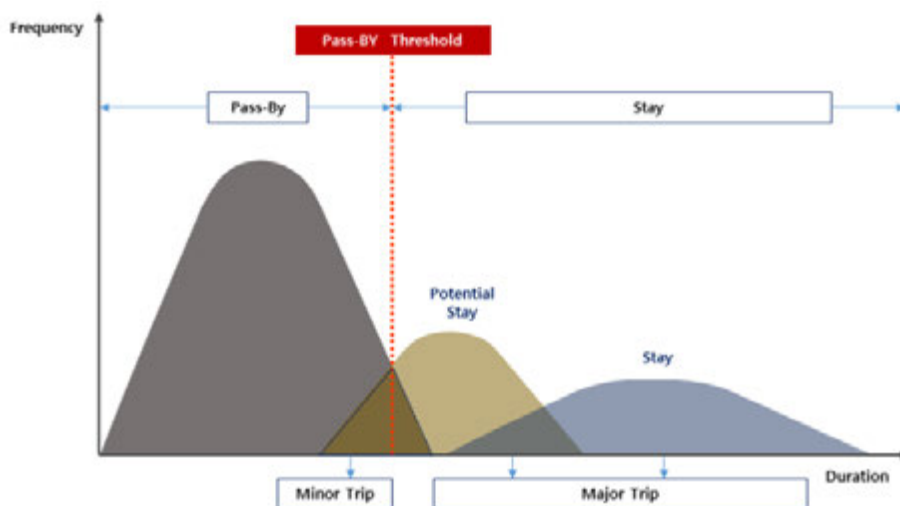
자료: 한국교통연구원 (2017), 2017년 국가교통조사

<그림 2- 9> 2017년 기준 이동/체류 기준 설정 방법

- 그러나 누적분포함수 값만으로는 체류와 이동을 구분하는데 한계가 있음
 - 누적분포함수값은 체류시간 분포를 설명하는 값으로 대체로 통행자가 체류 중이라고 판단할 수 있는 근거가 될 수는 있지만, 체류와 이동을 구분하는 임계치 역할로는 부적합
- 또한 기존에는 일부 지역만 분석한 자료이기 때문에 모든 데이터를 설명한다고 보기 어려우며, 올해 개발한 전처리 방법이 적용되지 않았기 때문에 기존 방법에서 산출한 기준을 그대로 적용하는 것은 바람직하지 않음

2. 개선방안

- 로그 기록 빈도를 체류시간별로 출력해보면 파동처럼 굴곡을 띤 그래프가 그려지는데, 세 가지의 체류 특성이 반영된 것으로 볼 수 있음
 - 이동(Pass-by): 이동 중에 기지국 수신범위 내에 인식된 로그 → 이동 중에는 기지국 수신영역을 빠르게 지나가 체류시간이 대체로 짧게 기록됨
 - 잠재체류(Potential stay): 일시적으로 머무르는 중에 기지국 수신범위에 인식된 로그 → 이동할 때보다 체류시간은 길게 나타나지만 체류시간이 특정 범위 내에서 기록됨
 - 주 체류(stay): 출발지 또는 목적지로 추정되는 기지국 수신범위 내에 인식된 로그 → 체류시간이 대체로 길게 나타남
- <그림 2-2>는 실제 그래프를 토대로 굴곡을 극대화하여 다시 도식화한 것임
- 각 체류 특성별 그래프, 3개의 그래프가 겹쳐져 마치 파동처럼 보여지는 것으로 볼 수 있으며, 그래프에서 3개의 마루는 곧 세 가지 체류 특성이 두드러지게 나타나는 지점이라고 해석할 수 있음
- 따라서 본 과업에서는 이동과 체류가 중첩되는 1차 변곡점을 체류와 이동을 구분하는 임계치로 설정하고자 함
 - 원천데이터에 전처리 알고리즘을 적용한 후 체류시간별 빈도를 추출한 결과, 체류시간 15분에서 1차 변곡점이 나타나는 것으로 확인



<그림 2- 10> 경로 체류시간 임계치 설정

제3절 3통행 미만 통행량 보정 방법

1. 개발 필요성

- 기 과업을 통해 구축한 모바일 빅데이터 기반 기종점 통행량 DB는 개인정보보호법에 저촉되지 않도록 3통행 미만의 통행을 모두 3통행으로 처리한 것으로, 실제 통행보다 통행량이 높게 나타날 수밖에 없음

<표 2- 9> 기존 3통행 미만 통행 처리 방식

출발 일자	출발 시간	출발 위치	출발 체류지 유형	도착 일자	도착 시간	도착 위치	도착 체류지 유형	성	연령	통행량	실제 값
171121	13	482703	집	171123	14	123909	기타	M	20	1	
↓											
출발 일자	출발 시간	출발 위치	출발 체류지 유형	도착 일자	도착 시간	도착 위치	도착 체류지 유형	성	연령	통행량	보정 값
171121	13	482703	집	171123	14	123909	기타	M	20	3	

2. 3통행 미만 보정 방법

1) 기존 문헌 고찰

- 모바일 빅데이터를 활용하는 타 과제에서는 최대한 3통행 미만이 나오지 않도록 DB를 설계하거나, 3통행 미만인 경우 실제 통행량을 정확하게 알 수 없도록 비식별화 처리하고 있음
 - 최대한 3통행 미만이 나오지 않도록 DB를 설계한다는 것은 필드 구분을 최소화하여 현재 필드의 구분으로 통행량이 분리된 것을 다시 합친다는 것을 의미함
 - 비식별화 처리 방식은 3통행 미만인 경우 별표 표시(*)로 실제 통행량 정보를 공개하지 않

는 다는 것을 의미함

<표 2- 10> 타 과제에서의 3통행 미만 처리 방법 - 비식별화 처리 방법

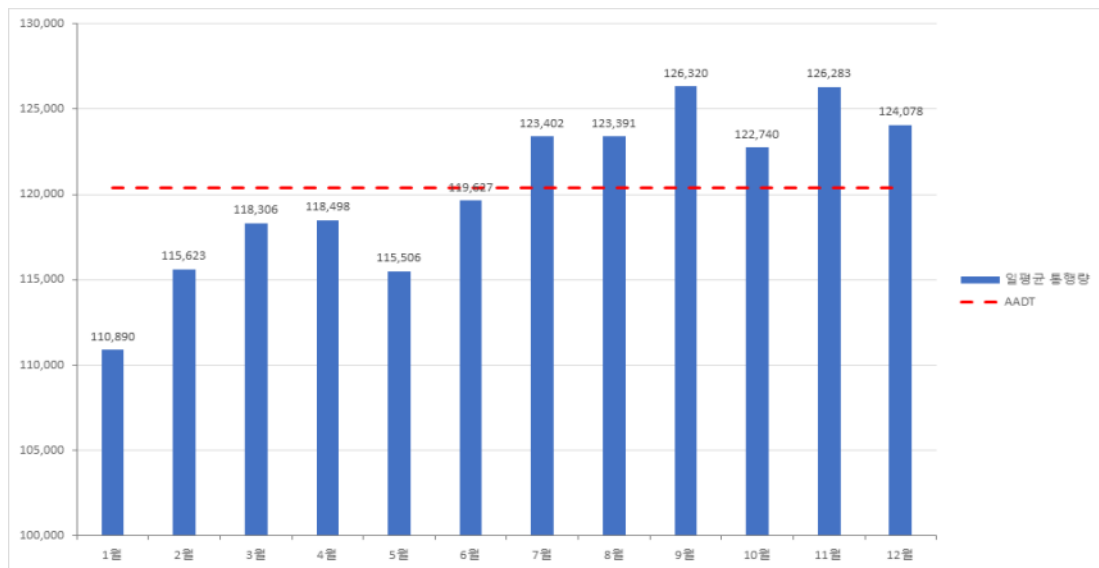
기준일시 (yymmddhh)	출발지	목적지	총 이동 (실제 통행량)	M0105	M0610	...	M1115
18060423	111414	111325	* (1)	*	*	...	* (1)
18060410	111123	111178	6	4	* (1)		* (1)

자료: 서울시·한국교통연구원·KT(2019. 08. 28.), 통근통학인구 데이터 공동개발 전문가 자문회의 자료, P. 25.

- 하지만 위와 같은 방식은 활용성이 떨어질 우려가 있고, 무엇보다 비식별화 처리 방안은 모두 3통행으로 처리하는 기존 방식과 반대로 오히려 통행량이 과소하게 집계될 수 있기 때문에 적절하지 않다고 판단됨
 - 3통행 미만의 통행이 발생하는 경우가 없도록 필드 구분을 최소화하려면 통행을 해석하는데 기본적으로 필요한 통행 정보(출발일자, 출발지, 출발시간, 도착일자, 도착지, 도착시간)를 제외하고 성, 연령, 체류지 유형 필드를 하나의 필드로 통합하거나 시간, 연령, 체류지 유형 필드와 같이 범위에 따라 값을 재분류할 수 있는 필드를 기준으로 데이터를 나눌 범위를 다시 설정해야 하지만, 이처럼 필드를 통합하거나 데이터를 분류하는 범위를 크게 설정할 경우 모바일 빅데이터의 최대 장점 중 하나인 상세성을 활용할 수 없게 됨
 - 또한 전체 3통행을 넘지 않는 경우가 많은 경우 비식별화 된 상태에서는 집계가 불가능하기 때문에 실제 통행량보다 매우 낮게 통행량이 집계될 수 있음
 - 아울러 두 가지 방법 모두 가능한 3통행 미만의 값을 갖지 않도록 DB를 설계하는 것이 중요한데, 매년 이에 맞춰 DB를 재설계하기에는 어려움이 따름
- 본 과업에서는 기존 과제에서 제시하고 있는 방법과는 다르게 후(後) 보정 방법으로 3통행 미만 통행을 보정하고자 함
 - 타 과제에서 제시하고 있는 방법들은 원천데이터에서 실제 통행량(참 값)을 어떻게 처리할지를 논의하는 것이라면, 본 과업에서 제시하는 방법은 이미 설계한 DB에 맞춰 모두 3통행으로 처리된 3미만의 통행량을 원래 값과 유사하도록 다시 되돌리는 방법임

2) 3통행 미만 보정 방법

- 원천데이터에서 본원에서 설계한 DB 형태로 가공할 때 곱해진 값을 반대로 그대로 나눌 경우 실제 값으로 변환할 수 있으나, 이 경우 완전한 참 값이 되어 개인정보보호법에 위배됨
- 따라서 보정한 통행량이 가장 참 값과 유사한 값을 가질 수 있도록 오차율이 낮은 보정계수를 개발해야 함
- 본 과업에서는 오차율이 낮은 보정계수를 찾기 위해서 먼저 보정계수를 산출할 기준 달을 설정하였음
 - 2018년도 과업을 통해 구축한 2017년 1월부터 12월까지의 모바일 빅데이터 기반 기종점 통행량 DB에서 연평균 일통행량(AADT)을 산출한 후 이와 가장 유사한 일평균 통행량 값을 갖는 달을 기준 달로 설정함
 - 연평균 일통행량: 120,402,573 통행
 - 기준달(6월) 일통행량: 119,626,502 통행



<그림 2- 11> 2017년 기준 DB의 AADT와 월별 일평균 통행량

- 그 다음 출발 정보와 성, 연령을 기준 필드로 두고 원천데이터에서 집계한 통행량을 기준 DB에서 같은 조건으로 집계한 통행량으로 나누어 보정계수를 산출함

<표 2- 11> 3통행 미만 통행량 보정계수 산출 기준 표

출발 시도	출발 요일	출발 시간대	출발 체류지 유형	성	연령	가공 DB상 집계 통행량	원천데이터상 집계 통행량	보정계수
서울	월	1	D	남	20	34500	15000	0.4348

- 기준 달의 보정계수로 보정할 수 없는 경우에는 기 산출한 보정계수에서 연령 구분 필드를 제외하고 보정계수를 재산출
 - 기준 달에는 존재하지 않는 경우의 수가 다른 달에 존재하는 경우에는 기 산출한 보정계수 적용 불가
 - 이러한 경우에는 연령을 제외하고 출발 시도, 출발 요일, 출발 시간대, 출발지 체류 특성, 통행자 성별 정보를 기준으로 기 산출한 보정계수의 평균값을 보정계수로 활용

<표 2- 12> 기준 달에서 누락된 보정계수 재산출 방법

출발 시도	출발 요일	출발 시간대	출발 체류지 유형	성	연령	가공 DB상 집계 통행량	원천데이터상 집계 통행량	보정계수
서울	월	1	D	남	20	34500	15000	0.43
서울	월	1	D	남	10	21240	18100	0.85
서울	월	1	D	남	50	10300	8910	0.86
평균								0.71

3. 2017년 기준 3통행 미만 보정계수

- 기준 달에서 총 265,220개의 보정계수를 산출
 - 평균 보정계수 값: 약 0.3497
- 기준 달에서 누락된 보정계수는 총 28,689개였으며, 위에서 제시한 방법을 적용하여 재산출한 평균 보정계수 22,154개로 4,898,639건의 데이터를 보정함
- 본 과업에서 개발한 보정계수를 2017년 DB에 적용한 결과, 보정 전 DB에서 산출한 연평균

일통행량보다 보정 후 DB에서 산출한 연평균 일통행량이 약 78,843천 통행 낮게 나타남

<표 2- 13> 3통행 미만 통행량 보정계수 적용 전후 비교

구분	보정계수 적용 전	보정계수 적용 후
연평균 일통행량 (AADT)	120,402,573	41,558,998
KT 시장점유율 적용 결과 ¹⁾	427.353.477	147.508.329
인당 일통행량 ²⁾	8.26	2.85

주 1: 2017. 12 기준 KT 시장점유율 28% 적용

주 2: 시장점유율을 적용한 연평균 일통행량을 대한민국 인구(51,709,098명)로 나눈 값

제3장 모바일 빅데이터 기반 기종점 통행량 DB 구축

제1절 기준년도 DB 구축

제2절 구축 DB 검증

제3장 모바일 빅데이터 기반 기종점 통행량 DB 구축

제1절 기준년도 DB 구축

1. DB 설계

- 「개인정보보호법」과 「위치정보의 보호 및 이용 등에 관한 법률」에 따라 개인의 위치와 이동계적을 추적할 수 없도록 기반 DB는 다음 <표 3-1>과 같이 집계된 형태로 구축하고자 함

<표 3- 1> 모바일 기반 DB 형태

출발				도착				성별 코드	연령 코드	통행량	통행 시간 (분)
일자	시간	교통 폴리곤 ID	체류지 유형	일자	시간	교통 폴리곤 ID	체류지 유형				
20181121	13	48270310	집	20181123	14	12390981	기타	M	20	10	34
20180501	06	48270310	집	20180501	09	32789014	회사	F	50	25	39
20180311	17	48270320	집	20180312	18	45608912	학교	M	10	15	25

- 개인의 통행 정보가 노출되지 않도록 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)를 기준으로 통행 정보가 같은 인구를 집계
- 일자 정보는 1일 단위로 구축하도록 함
- 출발 시간 및 도착 시간 정보는 집계된 인구가 3명 이하가 될 가능성을 줄이면서, DB 사용자에게 가능한 섬세한 통행 정보를 제공할 수 있도록 한 시간 단위로 입력하도록 함
- * 출발 시간은 출발지에서의 체류종료시간을, 도착 시간은 도착지에서의 체류시작시간을 의미
- 출발지와 도착지 위치는 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 기지국 좌표 그대로 노출하지 않고 매칭 되는 교통폴리곤(교통폴리곤은 기지국의 수신범위를 고려하여 구축한 모바일 기반 DB 분석 단위) ID로 변환하도록 함
- 출발 및 도착 체류지 유형은 평소 통행자가 해당 위치에 얼마나 자주, 오래 머무르는지를 추정할 수 있는 정보를 제공하기 위한 것으로, 「2018년 국가교통조사」에서 설정한 체류지 식별 기준에 따라 출발지(출발 기지국 좌표)와 도착지(도착 기지국 좌표)의 특성을 확인한 후 각각 체류지 유형을 구분하여 입력하도록 함

- * 「2018년 국가교통조사」에서 구분하고 있는 체류지 유형 중에서 집 이외 심야시간대 주체류지는 제외하고, 총 5가지 유형으로 구분 (집, 회사, 학교, 정기적 잠재체류지, 비정기적 잠재체류지)
- 통행자에 대한 정보를 확인할 수 있도록 성, 연령 필드 구성하고, 통행 정보가 지나치게 세분되지 않도록 가능한 통행 특성이 유사한 계층을 묶어 10세 단위로 연령을 구분하여 입력하도록 함
- 통행량 필드에는 전술한 기준에 의해 집계된 인구(단말기 수)를 입력하되, 집계된 통행량이 3 이하의 값을 갖는 경우, 2장에서 개발한 3통행 미만 보정계수를 적용하여 추정된 통행량을 입력하도록 함
- 새로운 지표 개발을 위해 통행시간 필드를 추가
- 통행시간 필드에는 집계된 통행 정보를 기준으로 산출한 평균통행시간 값을 입력

<표 3- 2> 모바일 DB 테이블 정의서

컬럼 ID	컬럼명	Type	비고
o_base_ymd	출발 일자	string	yyyymmdd
o_timezn_cd	출발시간대	string	00~23 (1시간 단위)
o_polygon	출발폴리곤ID	string	-
o_trip_type	출발 체류지 유형	string	H: 집 C: 회사 S: 학교 R: 종교집회장소 X: 기타 잠재체류지
d_base_ymd	도착 일자	string	yyyymmdd
d_timezn_cd	도착시간대	string	00~23 (1시간 단위)
d_polygon	도착폴리곤ID	string	-
d_trip_type	도착 체류지 유형	string	H: 집 C: 회사 S: 학교 R: 종교집회장소 X: 기타 잠재체류지 -: 이동정보 없음
age_itg_cd	연령	integer	0~110 (예: 10: 10세 이상 20세 미만)
sex_type_itg_cd	성	string	M: 남성, F: 여성
total	통행량	integer	보정계수 적용 후 일의 자리에서 올림 처리
travel_time	통행시간	integer	1분 단위

2. DB 구축 방법

1) 원천데이터 수집 및 이상치 제거

- 2018년 1월 1일부터 2018년 12월 31일까지 기록된 로그를 수집
- 기지국 좌표가 Null값이거나 행정동 정보가 매칭 되지 않는 데이터를 제거

2) 선분 이력으로 변환

- 포인트 단위(밀리세컨 단위) 로그를 선분 단위(분 단위) 로그로 변환
 - 닷(dot) 형태로 기록된 로그를 기록된 시간 순서대로 연결한 후, 최초 기록된 시간(이하 ‘체류시작시간’)과 마지막으로 기록된 시간(이하 ‘체류종료시간’) 정보만 추출하여 단말기가 해당 기지국에 식별된 시간(이하 ‘체류시간’)을 산출
 - 체류시간은 체류종료시간에서 체류시작시간을 뺀 값이며, 초(sec) 단위는 생략



주: 자료를 참고하여 도식화 함

자료: 박미율·주은정 (2017), LTE 시그널 정보를 이용한 위치정보가공 및 유동인구집계 방법, 한국통신학회 2017년도 동계종합학술발표회, p. 285.

<그림 3- 1> 선분 이력으로의 변환

3) 통행 정보 왜곡 데이터 보완

- 데이터 보완 범위를 한정하기 위해 로그 기록을 한 달 단위로 구분
 - 통행 정보 왜곡 데이터를 탐색하고 보완하는 범위를 1개월로 한정 (예: 4월 1일부터 4월 30일까지의 로그 기록을 연결하여 탐색하고 보완)
- 통행 정보를 왜곡할 수 있는 ‘필요 이상의 핸드오버 데이터(Unnecessary Handover)’, ‘핑

퐁 핸드오버(Ping-pong Handover)로 인한 데이터'를 본 과업에서 개발한 전처리 기술을 통해 보정함

4) 체류순서 부여

- 로그 기록 일자, 체류시작시간, 체류종료시간을 기준으로 로그 기록 순서에 따라 개인별 데이터를 정렬하여 체류 순서를 구분
 - 고객식별번호(단말기 구분코드)와 데이터 생성일자를 기준으로 KEY 값을 갖는 필드를 형성한 후('기준일-고객식별번호'), 같은 KEY 값을 갖는 데이터 내에서 정렬

ID	그룹 번호	KEY 합형 (기준일- 고객식별번호)	가상기차국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간
			x좌표	y좌표				
1	1	20180520-17232311	126.914	37.545	1400	1750	33	240
2	2	20180520-17232311	126.914	37.549	1751	1752	33	1
3	3	20180520-17232311	126.916	37.548	1754	1809	33	15
4	4	20180520-17232311	126.914	37.549	1810	1813	33	3
5	5	20180520-17232311	126.915	37.551	1813	1814	33	1
6	6	20180520-17232311	126.914	37.549	1814	1815	33	1
		⋮	⋮	⋮				
18	18	20180520-17232311	126.815	37.538	1920	2009	33	279
		⋮	⋮	⋮				
453	1	20180520-184000	126.934	37.556	0805	1041	25	156
454	2	20180520-184000	126.936	37.556	1041	1042	25	1
455	3	20180520-184000	126.935	37.554	1043	1044	25	1
456	4	20180520-184000	126.933	37.552	1045	1050	25	5
457	5	20180520-184000	126.929	37.547	1051	1052	25	1
458	6	20180520-184000	126.923	37.547	1053	1055	25	2
459	7	20180520-184000	126.918	37.547	1056	1059	25	3
460	8	20180520-184000	126.914	37.545	1100	1200	25	60
		⋮	⋮	⋮				
679	227	20180522-184000	126.934	37.556	0803	0830	25	507
680	228	20180522-184000	126.936	37.556	0831	0833	25	2
		⋮	⋮	⋮				
711	259	20180522-184000	126.815	37.551	0850	1100	25	130

<그림 3- 2> 체류순서 부여 (개인별 이동궤적 형성)

5) 체류 정보 추출

- 이동 중에 기록된 로그 기록을 제외하고 체류 중에 기록된 로그 기록만을 추출
 - 체류시간이 15분을 초과하는 로그 기록만 추출
 - 트립타임 필드를 추가하여 체류시간이 15분 이하인 경우는 '경로(Pass-by)'로, 15분 초과인 경우는 '체류(Stay)'로 구분

ID	그룹 번호	KEY 칼럼 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입
			X좌표	Y좌표					
1	1	20160520-17232311	126.914	37.545	1400	1750	33	240	체류
2	2	20160520-17232311	126.914	37.549	1751	1752	33	1	경로
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로
4	4	20160520-17232311	126.914	37.549	1810	1813	33	3	경로
5	5	20160520-17232311	126.915	37.551	1813	1814	33	1	경로
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로
		⋮	⋮	⋮					
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	체류
		⋮	⋮	⋮					
453	1	20160520-1843030	126.934	37.556	0805	1041	25	156	체류
454	2	20160520-1843030	126.936	37.556	1041	1042	25	1	경로
455	3	20160520-1843030	126.935	37.554	1043	1044	25	1	경로
456	4	20160520-1843030	126.933	37.552	1045	1050	25	5	경로
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	체류
		⋮	⋮	⋮					
679	227	20160522-1843030	126.934	37.556	0003	0830	25	507	체류
680	228	20160522-1843030	126.936	37.556	0831	0833	25	2	경로
		⋮	⋮	⋮					
711	259	20160522-1843030	126.815	37.561	0850	1100	25	130	체류

<그림 3- 3> 이동정보와 체류정보 구분 (예시)

6) 체류지 유형 구분

- 체류지 식별 기준(<표 3-3> 참조)에 따라 체류지 유형을 구분하여 필드 추가

<표 3- 3> 체류지 식별 기준

체류지 유형			식별 기준				
			체류특성			통행자 연령	비고
			체류시작시간 ~ 체류종료시간	체류 시간	체류빈도		
주 상 주 지 역	심 야 시 간 대	집	오후 9시 ~ 오전 7시	3시간 이상	주 3회 이상 기록된 주가 월 3회 이상	-	-

잠재 활동 지역	역						
	낮 시 간 대	회사	오전 9시 ~ 오후 6시	3시간 이상	주 2회 이상(주말 제외) 기록된 주가 월 2회 이상	20세 이상	20세 이상 25세 미만 연령의 경우 방학기간 (7-8월, 1-2월)에도 위 기준이 충족될 경우 회사로 간주
		주 상 주 지 역	학교	오전 9시 ~ 오후 6시	3시간 이상	주 2회 이상(주말 제외) 기록된 주가 월 2회 이상	25세 미만
잠재 활동 지역	정기적	-	-	주 3회 이상	-	-	
	비정기적	집, 회사, 학교, 정기적 잠재활동지역으로 구분되지 않은 그 외의 체류지					

ID	그룹 번호	KEY 칼럼 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입	체류지 유형
			X좌표	Y좌표						
1	1	20160520-17232311	126.914	37.545	1400	1750	33	240	체류	집
2	2	20160520-17232311	126.914	37.549	1751	1752	33	1	경로	-
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로	-
4	4	20160520-17232311	126.914	37.549	1810	1813	33	3	경로	-
5	5	20160520-17232311	126.915	37.551	1813	1814	33	1	경로	-
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로	-
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	체류	회사
453	1	20160520-1843030	126.934	37.556	0805	1041	25	156	체류	집
454	2	20160520-1843030	126.936	37.556	1041	1042	25	1	경로	-
455	3	20160520-1843030	126.935	37.554	1043	1044	25	1	경로	-
456	4	20160520-1843030	126.933	37.552	1045	1050	25	5	경로	-
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로	-
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로	-
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로	-
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	체류	기타
679	227	20160522-1843030	126.934	37.556	0003	0830	25	507	체류	집
680	228	20160522-1843030	126.936	37.556	0831	0833	25	2	경로	-
711	259	20160522-1843030	126.815	37.561	0850	1100	25	130	체류	종교집회 장소

<그림 3- 4> 체류지 유형 구분 (예시)

7) 위치정보, 시간정보, 연령정보 변환

- 개인의 이동궤적을 추적할 수 없도록 원천데이터에서 출발지와 도착지의 위치 정보를 분석 맵의 교통폴리곤 ID로 변경한 후, 시간 정보(체류시작시간, 체류종료시간)에서 분 정보를 생략한 시간대 필드를 생성하고, 1세 단위의 연령 정보를 10세 단위로 변경함

ID	그룹 번호	KEY 칼럼 (기준일- 고객식별번호)	가상기차국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입	체류지 유형	위치 정보 변환 (교통폴리곤 ID)	체류시작 시간대	체류종료 시간대	연령대 변환
			X좌표	Y좌표										
1	1	20160520-17232311	126.914	37.545	1400	1750	33	240	체류	집	210001	14	17	30
2	2	20160520-17232311	126.914	37.549	1751	1752	33	1	경로	-	-	-	-	-
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로	-	-	-	-	-
4	4	20160520-17232311	126.914	37.549	1810	1813	33	3	경로	-	-	-	-	-
5	5	20160520-17232311	126.915	37.561	1813	1814	33	1	경로	-	-	-	-	-
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로	-	-	-	-	-
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	체류	회사	230017	19	23	30
453	1	20160520-1843030	126.934	37.566	0805	1041	25	156	체류	집	311014	08	10	20
454	2	20160520-1843030	126.936	37.566	1041	1042	25	1	경로	-	-	-	-	-
455	3	20160520-1843030	126.935	37.564	1043	1044	25	1	경로	-	-	-	-	-
456	4	20160520-1843030	126.933	37.562	1045	1050	25	5	경로	-	-	-	-	-
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로	-	-	-	-	-
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로	-	-	-	-	-
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로	-	-	-	-	-
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	체류	기타	312123	11	12	20
679	227	20160522-1843030	126.934	37.566	0003	0830	25	507	체류	집	110016	00	08	20
680	228	20160522-1843030	126.936	37.566	0831	0833	25	2	경로	-	-	-	-	20
711	259	20160522-1843030	126.815	37.561	0850	1100	25	130	체류	종교집회 장소	210003	08	11	20

<그림 3- 5> 위치정보, 시간정보, 연령정보 변환 (예시)

8) 통행량 집계

- 앞서 부여한 체류순서에 따라 출발과 도착을 구분한 다음, 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)가 동일한 로그 기록을 집계함
 - 마지막 체류지가 출발 정보로 구분되고, 도착 정보가 없는 경우 이동하지 않은 것으로 보고 ‘무통행’으로 간주함

ID	그룹 번호	KEY 칼럼 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입	체류지 유형	위치 정보 변환 (교통플리근 ID)	체류시작 시간대	체류종료 시간대	연령대 변환	
			X좌표	Y좌표											
1	1	20160520-17232311	126.914	37.546	1400	1750	33	240	체류	집	210001	14	17	30	출발
2	2	20160520-17232311	126.914	37.549	1751	1752	33	1	경로	-	-	-	-	-	1통행 ↓
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로	-	-	-	-	-	
4	4	20160520-17232311	126.914	37.549	1810	1813	33	3	경로	-	-	-	-	-	
5	5	20160520-17232311	126.915	37.551	1813	1814	33	1	경로	-	-	-	-	-	
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로	-	-	-	-	-	
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	체류	회사	230017	19	23	30	도착
453	1	20160520-1843030	126.934	37.556	0805	1041	25	156	체류	집	311014	08	10	20	출발
454	2	20160520-1843030	126.936	37.556	1041	1042	25	1	경로	-	-	-	-	-	1통행 ↓
455	3	20160520-1843030	126.935	37.554	1043	1044	25	1	경로	-	-	-	-	-	
456	4	20160520-1843030	126.933	37.552	1045	1050	25	5	경로	-	-	-	-	-	
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로	-	-	-	-	-	
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로	-	-	-	-	-	
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로	-	-	-	-	-	
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	체류	기타	312123	11	12	20	도착
679	227	20160522-1843030	126.934	37.556	0003	0830	25	507	체류	집	110016	00	08	20	출발
680	228	20160522-1843030	126.936	37.556	0831	0833	25	2	경로	-	-	-	-	20	1통행 ↓
711	259	20160522-1843030	126.815	37.551	0850	1100	25	130	체류	종교집회 장소	210003	08	11	20	도착

<그림 3- 6> 출발/도착 구분 및 통행량 집계 (예시)

9) 통행량 보정

- 통행량이 3 이하(0~3통행)인 경우 모두 3통행으로 변환한 후, 3통행 미만 보정계수(2장 참고)를 적용하여 통행량을 보정
 - 개인정보보호법에 의거, 3 통행 이하의 값은 실제 값이 아닌, 추정 값을 입력
- 실제 통행 분석에 활용할 때에는 통신업체의 시장점유율을 기준으로 전수화 하여 사용
 - 과학기술정보통신부에 고시되고 있는 무선통신서비스 가입자 통계 정보 중 LTE 가입 정보를 활용하여 KT 시장점유율 산출 가능

<표 3- 4> KT 시장점유율

구분		가입자 수	비율
LTE	SKT	24, 205, 134	43.9%
	KT	15, 259, 806	27.7%
	LGU+	12, 647, 000	22.9%
	MVNO	3, 021, 741	5.5%
	소계	55, 133, 681	100.0%

자료 : 과학기술정보통신부 “무선 통신서비스 통계 현황 (2018년 12월)”, <https://www.msit.go.kr/> (20.03.25).

10) 평균통행시간 산출

- 각 개인의 통행정보에서 통행시간(도착시간-출발시간, 분 단위)을 산출한 다음, 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령)가 동일한 로그 기록의 통행시간을 집계한 후 집계한 통행시간을 산술평균하여 산출

ID	그룹 번호	KEY 값임 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입	체류지 유형	위치 정보 변환 (교통플라그 ID)	체류시작 시간대	체류종료 시간대	연령대 변환	
			X좌표	Y좌표											
1	1	20160520-1223211	126.984	37.546	1400	1750	33	240	체류	집	210008	14	17	30	출발
2	2	20160520-1223211	126.984	37.549	1751	1752	33	1	경로	-	-	-	-	-	1시간 30분 ↓
3	3	20160520-1223211	126.986	37.548	1754	1809	33	15	경로	-	-	-	-	-	
4	4	20160520-1223211	126.984	37.549	1810	1813	33	3	경로	-	-	-	-	-	
5	5	20160520-1223211	126.985	37.561	1813	1814	33	1	경로	-	-	-	-	-	
6	6	20160520-1223211	126.984	37.549	1814	1815	33	1	경로	-	-	-	-	-	
18	18	20160520-1223211	126.985	37.538	1820	2319	33	279	체류	회사	230017	19	23	30	도착
453	1	20160520-1843030	126.984	37.556	0805	1041	25	156	체류	집	311014	08	10	20	출발
454	2	20160520-1843030	126.986	37.556	1041	1042	25	1	경로	-	-	-	-	-	19분 ↓
455	3	20160520-1843030	126.985	37.554	1043	1044	25	1	경로	-	-	-	-	-	
456	4	20160520-1843030	126.983	37.552	1045	1050	25	5	경로	-	-	-	-	-	
457	5	20160520-1843030	126.989	37.547	1051	1052	25	1	경로	-	-	-	-	-	
458	6	20160520-1843030	126.983	37.547	1053	1055	25	2	경로	-	-	-	-	-	
459	7	20160520-1843030	126.988	37.547	1056	1059	25	3	경로	-	-	-	-	-	
460	8	20160520-1843030	126.984	37.546	1100	1200	25	60	체류	기타	312123	11	12	20	도착
659	227	20160522-1843030	126.984	37.556	0803	0800	25	507	체류	집	110016	08	08	20	출발
660	228	20160522-1843030	126.986	37.556	0831	0833	25	2	경로	-	-	-	-	20	20분 ↓
711	259	20160522-1843030	126.985	37.561	0850	1100	25	130	체류	종교장회 장소	210003	08	11	20	도착

<그림 3- 7> 통행시간 산출 (예시)

제2절 구축 DB 검증

1. DB 검증 방법

가. 기존 방식의 기·종점 통행량과 단순 비교

- 2016년 기준 가구통행실태조사를 통해 추정된 기종점 통행량과 비교
 - － 지역간 통행량, 지역내 통행량 특징을 비교하여 최다 통행 지역과 최소 통행 지역이 유사하게 나타나는지 확인
 - － 지역간, 지역내 통행량 차이가 크게 나타나는 지역이 있는지 확인

나. 파라미터 변화에 따른 통행 수요 변화

- 시간대별, 요일별, 성별, 연령별 통행량 변화를 확인하여 통상적인 통행 패턴을 벗어나는지 확인
 - － 시간대별 통행량 변화: 심야시간대에는 주간시간대보다 통행량이 낮게 나타남
 - － 요일별 통행량 변화: 평일에는 주기적으로 발생하는 통근·통학 통행으로 인해 주말에 비해 평일의 통행량이 높게 나타남
 - － 성별 통행량 변화: 남/여 통행량은 차이가 없음
 - － 연령별 통행량 변화: 경제활동인구(20세 이상 ~ 60세 미만)의 통행량이 고령인구(60세 이상)에 비해 통행량이 높게 나타남

2. DB 검증 결과

가. 기존 방식의 기·종점 통행량과 단순 비교 결과

1) 가구통행실태조사 기반 기종점 통행량

○ 일평균 1인당 1.7회 정도 통행하는 것으로 나타남

- 지역간 일평균 통행량은 약 49,484통행이며, 경기에서 서울로 이동한 통행이 가장 많고, 울산에서 세종으로 이동한 통행이 가장 적은 것으로 분석됨
- 지역 내 일평균 통행량은 약 4,318,207통행이며, 서울이 가장 많고 세종이 가장 적은 것으로 분석됨

<표 3- 5> 인력식 조사 기준 전국 기·종점 통행량 (시·도)

(단위: 통행/일)

시·도	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	계
서울	14,260,257	15,858	13,159	470,969	9,236	21,983	4,796	8,167	2,942,056	68,467	42,714	65,245	16,641	10,421	18,927	14,629	13,289	17,996,813
부산	14,453	5,935,035	12,463	2,571	2,448	3,032	58,672	360	7,469	1,208	1,186	1,661	3,847	9,175	20,239	380,445	3,273	6,457,538
대구	13,004	11,154	3,821,731	3,731	1,138	4,934	6,029	492	8,781	2,622	3,840	2,800	1,611	3,205	415,096	31,179	1,710	4,333,058
인천	434,152	2,527	3,317	3,389,696	2,177	4,571	690	870	486,457	7,053	7,863	15,893	3,402	2,305	4,252	2,176	2,916	4,370,318
광주	8,706	2,592	1,303	2,474	2,713,585	2,607	230	437	6,843	370	1,198	4,132	28,041	339,032	1,094	3,855	1,212	3,117,712
대전	21,761	3,141	5,045	5,338	2,516	2,700,392	1,672	87,505	23,720	3,265	87,509	117,909	14,223	2,448	10,027	4,349	1,040	3,091,862
울산	4,787	65,574	6,089	786	250	1,729	1,910,807	115	2,606	3,393	744	1,001	1,062	2,216	65,920	50,149	504	2,117,733
세종	8,630	399	510	1,053	466	87,441	116	210,597	7,071	238	37,600	32,559	1,288	486	620	648	144	389,864
경기	2,963,440	8,396	8,802	501,634	7,143	23,996	2,731	7,109	14,782,458	83,802	64,685	159,034	15,931	7,070	17,683	8,781	5,834	18,668,530
강원	64,540	1,159	2,325	7,492	351	3,021	3,213	324	85,024	2,514,529	19,538	1,365	628	186	8,937	746	639	2,714,015
충북	43,599	1,339	3,828	9,461	1,313	93,874	728	39,330	67,243	17,838	2,462,759	33,717	6,868	989	15,837	3,486	1,262	2,803,473
충남	63,897	1,821	2,723	20,928	3,792	121,285	932	32,927	156,200	1,294	36,453	3,392,040	26,381	3,693	3,948	2,778	770	3,871,861
전북	16,007	3,539	1,844	4,448	27,762	15,543	1,159	1,300	15,465	616	6,856	25,584	2,780,320	29,596	2,375	9,362	664	2,942,438
전남	9,789	8,835	3,472	2,580	164,253	2,520	1,379	464	6,514	189	888	3,517	30,457	2,349,236	2,733	17,282	2,175	2,606,282
경북	17,901	18,200	408,978	4,448	1,126	9,430	67,044	574	16,469	9,120	15,504	3,700	2,334	3,007	3,823,717	28,842	807	4,431,201
경남	13,770	375,540	33,341	2,388	3,852	4,119	48,831	860	8,277	648	2,908	2,625	9,784	18,085	30,021	4,740,344	1,763	5,297,159
제주	13,186	3,182	1,673	3,015	1,167	1,010	490	140	5,787	625	1,230	752	635	2,029	792	1,715	1,622,016	1,659,446
계	17,971,878	6,458,293	4,330,603	4,433,014	2,942,575	3,101,487	2,109,519	391,571	18,628,440	2,715,277	2,793,476	3,863,535	2,943,452	2,783,178	4,442,218	5,300,768	1,660,019	86,869,303

주 : 2016년 기준, 승용차(택시), 버스, 철도(일반/고속/지하철), 항공, 해운수단에 대한 목적통행량임.

2) 모바일 빅데이터 기반 기종점 통행량

○ 일평균 1인당 3.9회 정도 통행하는 것으로 나타남

- 지역간 일평균 통행량은 약 33,810통행이며, 경기에서 서울로 이동한 통행이 가장 많고, 세종에서 제주로 이동한 통행이 가장 적은 것으로 분석됨

- 지역 내 일평균 통행량은 약 11,180,592통행이며, 서울이 가장 많고 세종이 가장 적은 것으로 분석됨

<표 3- 6> 모바일 데이터 기준 전국 기·종점 통행량 (시·도) - (주) KT 데이터 기준

(단위: 통행/일)

시·도 \ 기·종점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	계
서울	11,054,814	2,726	2,139	58,182	1,312	4,158	677	1,453	489,023	11,531	7,150	12,586	2,630	1,598	3,739	1,407	4,481	11,659,606
부산	2,455	4,009,033	3,016	335	216	539	13,147	77	1,204	204	629	394	407	1,140	8,594	71,550	1,166	4,114,106
대구	2,009	3,123	2,548,695	262	217	861	1,338	112	1,408	417	1,391	563	631	470	75,316	8,727	583	2,646,123
인천	58,754	239	181	2,592,057	147	544	42	166	109,647	1,322	1,389	3,614	491	206	504	161	248	2,769,712
광주	1,416	232	220	211	1,530,433	500	36	97	966	68	252	1,581	6,182	45,956	183	672	350	1,589,355
대전	496	577	851	702	454	1,688,227	174	10,993	4,881	496	14,323	23,231	3,774	541	1,829	843	6	1,752,398
울산	747	14,225	1,568	78	36	203	1,318,895	26	394	110	227	170	94	193	13,833	9,672	82	1,360,553
세종	1,450	66	112	173	99	11,417	17	207,370	1,703	89	7,116	7,084	509	112	300	94	4	237,715
경기	506,944	1,244	1,584	112,615	1,002	4,843	346	1,714	11,907,536	27,048	20,149	35,747	3,880	1,417	5,659	1,329	367	12,633,424
강원	11,657	190	427	1,470	69	472	89	91	26,338	1,792,170	6,289	708	197	70	4,190	187	44	1,844,658
충북	7,103	758	1,571	1,451	306	14,930	231	6,943	19,955	6,063	1,612,704	8,541	1,978	465	8,504	1,251	651	1,693,405
충남	12,378	379	601	3,724	1,743	24,512	138	6,957	35,351	666	8,471	2,147,417	15,468	2,719	1,896	1,748	30	2,264,198
전북	2,816	395	656	636	6,282	4,061	90	499	4,061	205	1,836	18,486	1,778,315	9,551	709	3,235	192	1,832,025
전남	1,399	1,193	468	260	50,311	632	200	113	1,393	73	400	2,337	9,950	1,598,627	414	6,751	1,435	1,675,956
경북	3,955	8,515	81,131	634	192	2,001	13,608	294	5,646	4,285	8,503	1,797	676	402	2,560,909	8,429	36	2,701,013
경남	1,383	73,118	9,367	205	755	1,098	9,448	99	1,318	198	1,125	1,504	3,373	6,905	8,821	3,320,816	118	3,439,651
제주	3,161	805	399	311	375	62	114	17	908	79	529	97	215	1,304	154	290	939,250	948,070
계	11,672,937	4,116,818	2,652,986	2,773,306	1,593,949	1,759,060	1,358,590	237,021	12,611,732	1,845,024	1,692,483	2,265,857	1,828,770	1,671,676	2,695,554	3,437,162	949,043	55,161,968

주 : 2018년 1월 1일부터 2018년 12월 31일까지의 데이터를 활용함.

<표 3- 7> 모바일 데이터 기준 전국 기·종점 통행량 (시·도) - 전수화

(단위:통행/일)

시·도 \ 기·종점	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	계
서울	39,941,044	9,849	7,728	210,212	4,740	15,023	2,446	5,250	1,766,840	41,662	25,833	45,473	9,502	5,774	13,509	5,083	16,190	42,126,158
부산	8,870	14,484,637	10,897	1,210	780	1,947	47,500	278	4,350	737	2,273	1,424	1,470	4,119	31,050	258,510	4,213	14,864,265
대구	7,259	11,283	9,208,435	947	784	3,111	4,834	405	5,087	1,507	5,026	2,034	2,280	1,698	272,117	31,531	2,106	9,560,443
인천	212,278	864	654	9,365,102	531	1,965	152	600	396,155	4,776	5,018	13,057	1,774	744	1,821	582	896	10,006,970
광주	5,116	838	795	762	5,529,455	1,807	130	350	3,490	246	910	5,712	22,336	166,039	661	2,428	1,265	5,742,340
대전	1,792	2,085	3,075	2,536	1,640	6,099,564	629	39,718	17,635	1,792	51,749	83,934	13,635	1,955	6,608	3,046	22	6,331,414
울산	2,699	51,395	5,665	282	130	733	4,765,168	94	1,424	397	820	614	340	697	49,979	34,945	296	4,915,678
세종	5,239	238	405	625	358	41,250	61	749,228	6,153	322	25,710	25,594	1,839	405	1,084	340	14	858,864
경기	1,831,589	4,495	5,723	406,878	3,620	17,498	1,250	6,193	43,021,929	97,724	72,798	129,154	14,018	5,120	20,446	4,802	1,326	45,644,563
강원	42,117	686	1,543	5,311	249	1,705	322	329	95,159	6,475,110	22,722	2,558	712	253	15,138	676	159	6,664,750
충북	25,663	2,739	5,676	5,242	1,106	53,942	835	25,085	72,097	21,906	5,826,700	30,859	7,147	1,680	30,725	4,520	2,352	6,118,272
충남	44,722	1,369	2,171	13,455	6,297	88,562	499	25,136	127,723	2,406	30,606	7,758,618	55,886	9,824	6,850	6,316	108	8,180,548
전북	10,174	1,427	2,370	2,298	22,697	14,672	325	1,803	14,672	741	6,633	66,790	6,425,052	34,508	2,562	11,688	694	6,619,107
전남	5,055	4,310	1,691	939	181,774	2,283	723	408	5,033	264	1,445	8,444	35,949	5,775,840	1,496	24,391	5,185	6,055,229
경북	14,289	30,765	293,126	2,291	694	7,230	49,166	1,062	20,399	15,482	30,721	6,493	2,442	1,452	9,252,565	30,454	130	9,758,760
경남	4,997	264,175	33,843	741	2,728	3,967	34,136	358	4,762	715	4,065	5,434	12,187	24,948	31,870	11,998,109	426	12,427,459
제주	11,421	2,908	1,442	1,124	1,355	224	412	61	3,281	285	1,911	350	777	4,711	556	1,048	3,393,510	3,425,377
계	42,174,323	14,874,064	9,585,239	10,019,955	5,758,938	6,355,484	4,908,586	856,357	45,566,189	6,666,072	6,114,941	8,186,542	6,607,346	6,039,766	9,739,037	12,418,467	3,428,892	199,300,197

주 : (주)KT 시장점유율을 반영하여 전수화한 값임.

3) 비교 분석 결과

- 모바일 데이터 기반 인당 일평균 통행횟수는 기존 방식으로 산출한 것보다 약 2~3배 높음
 - － 모바일 데이터에서는 기·종점이 일정 시간 이상 머물렀는가, 즉 체류 여부로 결정되어 기존 방식보다 인당 일평균 통행횟수가 더 클 수밖에 없음
 - 기존 방식에서는 출근, 등교, 업무, 쇼핑, 귀가 등 통행 목적에 따라 기·종점 통행량을 집계하기 때문에, 목적이 없는 이동은 통행에서 제외됨
- 시도별 통행 분포는 거의 일치하고, 서울과 경기지역에서만 통행 비율이 약 2~5% 차이나는 것으로 분석됨

<표 3- 8> 전국 기·종점 통행 비율 (비교 결과)

(단위: %)

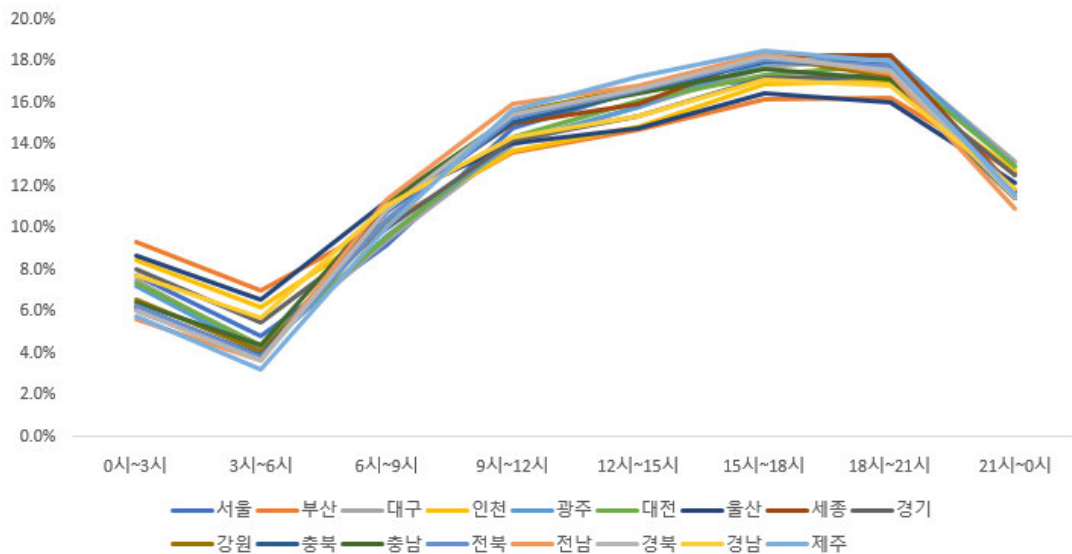
○ \ D	서울	부산	대구	인천	광주	대전	울산	세종	경기	강원	충북	충남	전북	전남	경북	경남	제주	계
서울	3.62	-0.01	-0.01	-0.44	-0.01	-0.02	0.00	-0.01	-2.50	-0.06	-0.04	-0.05	-0.01	-0.01	-0.02	-0.01	-0.01	0.42
부산	-0.01	0.44	-0.01	0.00	0.00	0.00	-0.04	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01	-0.01	-0.31	0.00	0.02
대구	-0.01	-0.01	0.22	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	-0.34	-0.02	0.00	-0.19
인천	-0.39	0.00	0.00	0.80	0.00	0.00	0.00	0.00	-0.36	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.00	-0.01
광주	-0.01	0.00	0.00	0.00	-0.35	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.02	-0.31	0.00	0.00	0.00	-0.71
대전	-0.02	0.00	0.00	0.00	0.00	-0.05	0.00	-0.08	-0.02	0.00	-0.07	-0.09	-0.01	0.00	-0.01	0.00	0.00	-0.38
울산	0.00	-0.05	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.05	-0.04	0.00	0.03
세종	-0.01	0.00	0.00	0.00	0.00	-0.08	0.00	0.13	-0.01	0.00	-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	-0.02
경기	-2.49	-0.01	-0.01	-0.37	-0.01	-0.02	0.00	-0.01	4.57	-0.05	-0.04	-0.12	-0.01	-0.01	-0.01	-0.01	-0.01	1.41
강원	-0.05	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	-0.05	0.35	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.22
충북	-0.04	0.00	0.00	-0.01	0.00	-0.08	0.00	-0.03	-0.04	-0.01	0.09	-0.02	0.00	0.00	0.00	0.00	0.00	-0.16
충남	-0.05	0.00	0.00	-0.02	0.00	-0.10	0.00	-0.03	-0.12	0.00	-0.03	-0.01	0.00	0.00	0.00	0.00	0.00	-0.35
전북	-0.01	0.00	0.00	0.00	-0.02	-0.01	0.00	0.00	-0.01	0.00	0.00	0.00	0.02	-0.02	0.00	0.00	0.00	-0.07
전남	-0.01	-0.01	0.00	0.00	-0.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	0.19	0.00	-0.01	0.00	0.04
경북	-0.01	-0.01	-0.32	0.00	0.00	-0.01	-0.05	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.24	-0.02	0.00	-0.20
경남	-0.01	-0.30	-0.02	0.00	0.00	0.00	-0.04	0.00	-0.01	0.00	0.00	0.00	-0.01	-0.01	-0.02	0.56	0.00	0.14
제주	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.16	-0.19
계	0.47	0.03	-0.18	-0.08	-0.50	-0.38	0.03	-0.02	1.42	0.22	-0.15	-0.34	-0.07	-0.17	-0.23	0.13	-0.19	0.00

주 : 모바일 자료 기반 통행 비율에서 KTDB 통행 비율을 뺀 값임.

나. 파라미터 변화에 따른 통행 수요 변화 결과

1) 시간대별 통행량 변화

- 심야시간대(0~6시) 통행량은 주간시간대(12~18시) 통행량의 1/3 수준인 것으로 분석됨
 - 가장 통행이 낮은 시간대는 3~6시였으며, 6시를 기점으로 통행량이 점차 증가하다가 18~21시에서 가장 높은 통행량을 보이고, 21시 이후부터 다시 통행이 줄어드는 것으로 나타남
 - 17개 시도 모두 같은 패턴을 보였으며, 타 지역에 비해 심야시간대 통행량이 높은 지역은 부산, 주간시간대 통행량이 높은 지역은 제주인 것으로 분석됨



<그림 3- 8> 시간대별 통행량 변화 (유입 기준)

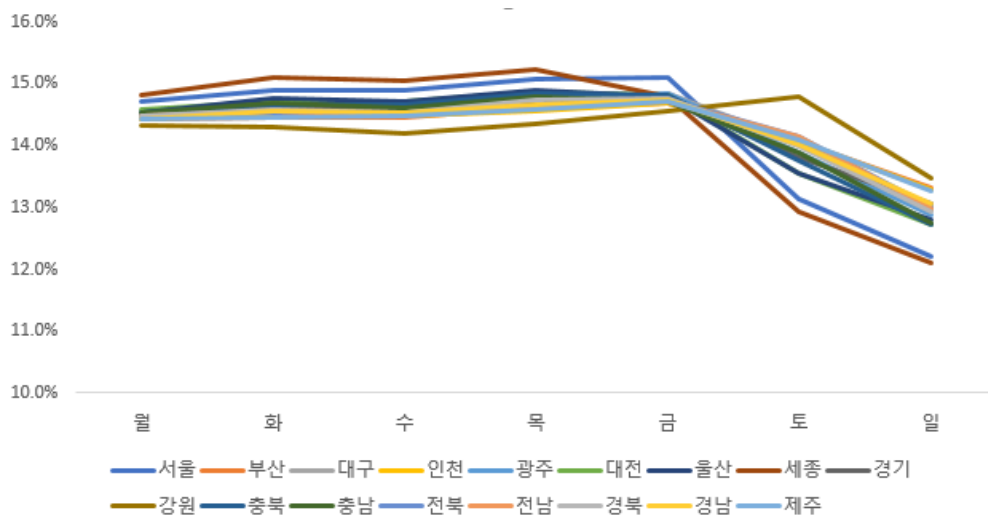
<표 3- 9> 시간대별 통행량 변화 - (주) KT 데이터 기준

(단위: 통행/일)

구분		새벽		오전		오후		저녁	
		0시~3시	3시~6시	6시~9시	9시~12시	12시~15시	15시~18시	18시~21시	21시~0시
수도권	서울	1,156,685	722,902	1,375,011	2,234,255	2,532,430	2,621,642	2,638,889	1,895,804
	부산	486,725	368,234	538,795	710,382	764,856	838,301	846,084	660,865
	대구	266,706	150,948	333,932	492,107	545,913	611,893	640,259	470,873
	인천	312,143	228,842	374,004	496,271	533,121	604,259	615,011	467,532
	광주	154,937	88,479	205,436	299,102	327,720	370,142	387,484	278,710
	대전	172,048	103,544	222,375	330,955	365,750	394,699	411,783	300,969
	울산	150,180	113,564	196,161	242,851	254,841	282,863	274,723	209,324
	세종	19,528	11,484	35,943	48,742	51,874	58,820	58,942	38,463
	경기	1,346,006	930,177	1,688,075	2,365,164	2,549,163	2,822,546	2,852,129	2,093,032
	강원	157,061	97,282	2,444,464	374,787	408,975	444,494	417,245	274,159
	충북	139,072	85,818	240,590	338,807	372,270	405,585	399,402	261,047
	충남	187,269	126,138	328,131	453,132	490,402	527,330	504,324	333,267
	전북	152,179	92,451	249,975	369,008	402,510	437,390	429,865	284,604
	전남	122,802	78,843	247,849	351,416	378,486	413,587	382,940	236,790
	경북	212,986	127,779	384,719	551,410	605,671	661,495	627,861	403,797
	경남	342,953	253,123	494,676	640,378	688,342	764,656	748,309	527,212
	제주	71,747	40,212	125,170	195,934	215,485	231,380	225,287	144,999
외곽	서울	1,173,818	737,726	1,399,830	2,240,547	2,525,036	2,593,488	2,609,355	1,910,196
	부산	485,362	365,299	537,528	711,400	768,745	841,418	845,073	659,055
	대구	262,574	147,725	330,560	491,877	553,703	621,476	637,673	463,687
	인천	308,177	224,634	368,517	496,299	539,311	612,896	619,462	462,035
	광주	152,129	86,353	202,404	297,559	333,709	377,695	386,025	273,914
	대전	170,428	101,837	220,548	330,213	369,911	399,375	410,784	298,410
	울산	150,105	113,482	194,887	241,985	254,713	283,220	276,251	209,676
	세종	19,508	11,985	35,502	48,381	51,583	59,070	59,088	38,115
	경기	1,329,137	914,503	1,660,703	2,358,772	2,558,627	2,852,199	2,877,743	2,078,874
	강원	158,424	98,859	248,562	375,934	405,217	439,558	417,148	276,309
	충북	140,012	87,181	242,107	338,637	370,321	403,137	399,262	262,547
	충남	190,196	129,023	331,737	453,367	485,178	520,644	504,684	337,656
	전북	152,478	92,945	251,620	369,413	401,145	435,833	429,821	285,130
	전남	125,422	80,907	251,163	352,783	372,329	406,238	384,637	241,352
	경북	217,569	132,274	388,639	551,457	597,873	651,611	628,946	410,734
	경남	343,799	254,495	495,330	640,036	685,396	762,243	749,745	528,944
	제주	71,886	40,589	125,668	196,041	215,615	230,980	224,841	144,812

2) 요일별 통행량 변화

- 월요일부터 금요일까지 평일동안 일정 수준의 통행량이 유지되다가 토요일부터 통행이 줄어들어 일요일에는 통행량이 가장 낮게 나타남
- 17개 시도 중에서 강원은 예외적으로 주말 중 토요일 통행량이 평일의 통행량보다 다소 높게 나타남
- 서울과 세종은 타 지역에 비해 평일과 주말 통행량 차이가 크게 나타나는 것이 특징임



<그림 3- 9> 요일별 통행량 변화 (유입 기준)

<표 3- 10> 요일별 통행량 변화 - (주) KT 데이터 기준

(단위: 통행/일)

구분		월	화	수	목	금	토	일
유출	서울	16, 237, 322	16, 437, 754	16, 447, 908	16, 639, 753	16, 729, 257	14, 519, 098	13, 383, 700
	부산	5, 424, 377	5, 438, 295	5, 429, 064	5, 487, 035	5, 513, 197	5, 290, 878	5, 013, 208
	대구	3, 687, 494	3, 697, 918	3, 699, 457	3, 731, 696	3, 772, 903	3, 581, 707	3, 298, 884
	인천	3, 819, 476	3, 836, 429	3, 836, 382	3, 855, 007	3, 874, 630	3, 715, 382	3, 532, 110
	광주	2, 225, 446	2, 232, 551	2, 235, 368	2, 257, 101	2, 268, 964	2, 130, 132	1, 969, 212
	대전	2, 434, 690	2, 454, 335	2, 451, 931	2, 481, 706	2, 471, 601	2, 263, 844	2, 117, 834
	울산	1, 806, 922	1, 834, 619	1, 829, 610	1, 851, 990	1, 845, 409	1, 688, 787	1, 581, 927
	세종	352, 969	360, 494	359, 429	365, 056	356, 231	309, 819	284, 599
	경기	17, 556, 203	17, 711, 937	17, 721, 159	17, 836, 338	17, 898, 328	16, 757, 263	15, 851, 212
	강원	2, 532, 132	2, 530, 100	2, 507, 415	2, 536, 474	2, 552, 766	2, 601, 800	2, 412, 408
	충북	2, 382, 154	2, 406, 647	2, 407, 193	2, 435, 291	2, 432, 313	2, 254, 897	2, 087, 334
	충남	3, 123, 104	3, 164, 153	3, 144, 348	3, 183, 495	3, 160, 976	2, 985, 402	2, 751, 562
	전북	2, 550, 396	2, 564, 103	2, 551, 518	2, 580, 153	2, 593, 717	2, 487, 947	2, 306, 789
	전남	2, 350, 506	2, 366, 062	2, 349, 396	2, 373, 111	2, 387, 148	2, 288, 864	2, 119, 514
	경북	3, 795, 586	3, 822, 893	3, 804, 027	3, 862, 418	3, 867, 382	3, 654, 112	3, 393, 498
	경남	4, 675, 039	4, 717, 109	4, 706, 062	4, 750, 440	4, 779, 697	4, 540, 495	4, 239, 191
	제주	1, 308, 802	1, 311, 329	1, 315, 238	1, 322, 767	1, 335, 319	1, 278, 187	1, 205, 494
유입	서울	16, 236, 647	16, 440, 664	16, 449, 502	16, 638, 352	16, 663, 575	14, 490, 450	13, 477, 903
	부산	5, 421, 197	5, 438, 099	5, 429, 630	5, 488, 171	5, 518, 552	5, 292, 151	5, 009, 052
	대구	3, 684, 430	3, 698, 749	3, 699, 994	3, 732, 136	3, 780, 091	3, 574, 902	3, 299, 681
	인천	3, 819, 310	3, 838, 757	3, 837, 864	3, 857, 733	3, 887, 469	3, 723, 377	3, 517, 813
	광주	2, 223, 574	2, 233, 334	2, 235, 962	2, 256, 182	2, 272, 237	2, 125, 262	1, 972, 041
	대전	2, 433, 935	2, 455, 010	2, 452, 051	2, 480, 678	2, 470, 343	2, 260, 686	2, 122, 939
	울산	1, 807, 842	1, 834, 996	1, 829, 428	1, 851, 165	1, 839, 767	1, 684, 785	1, 590, 868
	세종	353, 925	360, 644	359, 464	363, 802	353, 158	308, 733	288, 895
	경기	17, 558, 995	17, 715, 451	17, 715, 009	17, 830, 908	17, 920, 034	16, 751, 405	15, 829, 733
	강원	2, 530, 962	2, 527, 303	2, 507, 882	2, 538, 073	2, 571, 297	2, 615, 692	2, 381, 162
	충북	2, 384, 202	2, 406, 734	2, 407, 240	2, 435, 185	2, 429, 832	2, 255, 069	2, 087, 259
	충남	3, 127, 111	3, 162, 164	3, 144, 621	3, 183, 294	3, 163, 157	2, 988, 480	2, 743, 750
	전북	2, 549, 131	2, 562, 879	2, 551, 837	2, 581, 367	2, 600, 162	2, 493, 186	2, 295, 462
	전남	2, 350, 910	2, 363, 923	2, 349, 469	2, 374, 853	2, 388, 032	2, 298, 270	2, 108, 579
	경북	3, 796, 266	3, 820, 796	3, 804, 363	3, 863, 554	3, 866, 766	3, 663, 050	3, 384, 534
	경남	4, 675, 524	4, 715, 944	4, 705, 658	4, 751, 113	4, 780, 221	4, 544, 642	4, 233, 858
	제주	1, 308, 656	1, 311, 282	1, 315, 530	1, 323, 266	1, 335, 146	1, 278, 475	1, 204, 947

3) 성별 통행량 변화

- 전체 통행량에서 남성과 여성의 통행량은 평균적으로 53.8%, 46.2%를 차지하는 것으로 나타났으며, 남성과 여성의 인구 성비가 50.1% 49.9%인 것을 고려할 때 통행량에는 거의 차이가 없다고 볼 수 있음

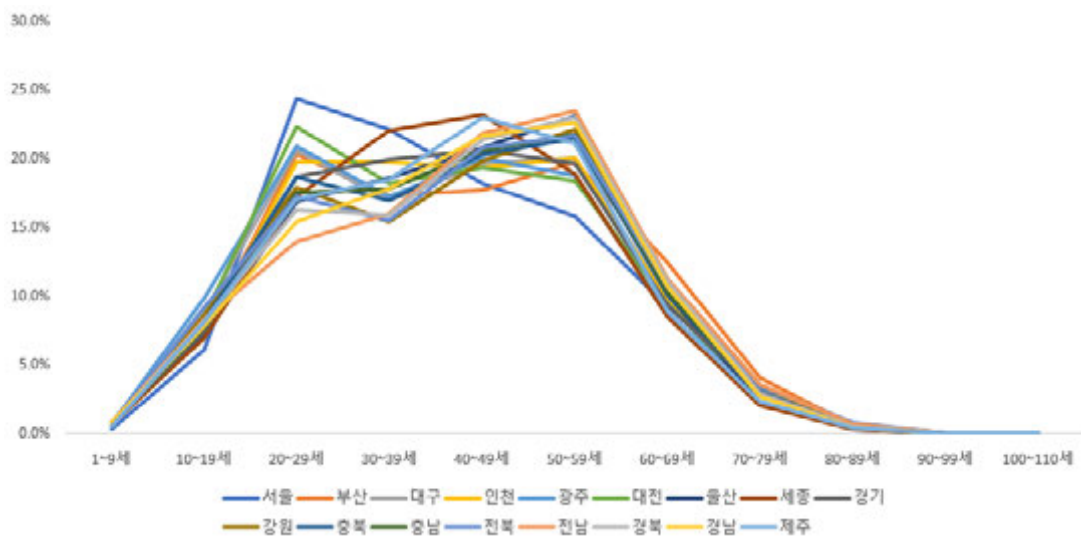
<표 3- 11> 성별 통행량 변화 - (주) KT 데이터 기준

(단위: 통행/일, %)

구분		남성		여성	
		일평균 통행량	비율	일평균 통행량	비율
유출	서울	7,874,652	49.9%	7,897,311	50.1%
	부산	2,688,315	50.1%	2,682,696	49.9%
	대구	1,849,747	50.8%	1,788,966	49.2%
	인천	1,966,371	52.0%	1,815,078	48.0%
	광주	1,122,776	51.3%	1,065,721	48.7%
	대전	1,242,292	52.1%	1,140,128	47.9%
	울산	980,087	55.2%	797,032	44.8%
	세종	194,207	56.9%	147,053	43.1%
	경기	9,176,108	52.9%	8,157,708	47.1%
	강원	1,391,022	55.1%	1,133,725	44.9%
	충북	1,308,981	55.8%	1,034,814	44.2%
	충남	1,770,963	57.6%	1,302,464	42.4%
	전북	1,363,999	54.1%	1,155,318	45.9%
	전남	1,317,930	56.8%	1,001,384	43.2%
	경북	2,115,574	56.5%	1,627,416	43.5%
	경남	2,521,203	54.5%	2,108,640	45.5%
	제주	695,479	53.6%	601,287	46.4%
유입	서울	7,874,792	49.9%	7,897,496	50.1%
	부산	2,688,363	50.1%	2,682,753	49.9%
	대구	1,849,731	50.8%	1,788,963	49.2%
	인천	1,967,284	52.0%	1,816,004	48.0%
	광주	1,122,762	51.3%	1,065,704	48.7%
	대전	1,242,271	52.1%	1,140,105	47.9%
	울산	980,058	55.2%	797,005	44.8%
	세종	194,209	56.9%	147,057	43.1%
	경기	9,175,321	52.9%	8,156,949	47.1%
	강원	1,390,982	55.1%	1,133,659	44.9%
	충북	1,308,965	55.8%	1,034,791	44.2%
	충남	1,770,954	57.6%	1,302,419	42.4%
	전북	1,363,957	54.1%	1,155,271	45.9%
	전남	1,317,891	56.8%	1,001,344	43.2%
	경북	2,115,543	56.5%	1,627,365	43.5%
	경남	2,521,134	54.5%	2,108,557	45.5%
	제주	695,490	53.6%	601,300	46.4%

4) 연령별 통행량 변화

- 경제활동인구의 통행량이 가장 많고, 60세 이상부터 점차 줄어드는 것으로 나타남
 - － 서울, 대전, 광주, 부산 등 특별시·광역시에서는 경제활동인구 중 20대의 통행 비율이 두드러지게 높게 나타났으나 그 외 지역에서는 50대의 통행 비율이 높게 나타남
 - 세종은 30대와 40대 통행이 높게 나타나는 예외적인 패턴을 보임



<그림 3- 10> 연령별 통행량 변화 (유입 기준)

<표 3- 12> 연령별 통행량 변화 - (주) KT 데이터 기준

(단위: 통행/일)

구분		1~9세	10 ~19세	20 ~29세	30 ~39세	40 ~49세	50 ~59세	60 ~69세	70 ~79세	80 ~89세	90 ~99세	100 ~110세
광역시 · 도	서울	54,320	972,555	3,845,269	3,497,795	2,881,476	2,506,488	1,460,139	510,465	84,757	4,864	141
	부산	31,844	387,777	1,092,387	933,982	952,250	1,064,713	667,405	219,032	34,434	1,811	32
	대구	22,090	320,672	751,104	618,813	716,945	729,574	357,861	110,103	21,308	1,234	36
	인천	30,686	272,100	750,164	750,793	733,386	760,076	366,214	110,723	19,644	1,155	14
	광주	15,595	215,451	459,115	379,308	439,860	411,223	190,137	68,892	14,472	1,024	16
	대전	13,341	202,701	533,491	435,206	461,741	438,920	225,365	65,523	12,309	765	9
	울산	13,435	127,585	299,929	330,464	370,934	411,718	181,081	40,567	5,826	346	6
	세종	2,370	24,038	58,940	75,613	79,474	64,546	29,062	7,136	1,270	72	1
	경기	126,543	1,298,845	3,249,756	3,473,531	3,579,437	3,415,743	1,622,016	523,166	93,874	5,034	115
	강원	19,170	218,519	454,353	389,778	503,364	559,631	281,903	88,728	15,846	1,029	25
	충북	19,700	191,146	439,945	397,867	478,516	505,334	238,537	67,030	12,178	781	13
	충남	20,828	239,449	540,334	549,997	636,072	670,282	318,690	88,550	17,242	1,033	20
	전북	16,485	232,748	435,629	392,197	524,496	546,271	269,563	88,037	19,836	1,602	55
	전남	18,203	193,931	324,536	371,789	507,534	546,567	262,866	83,309	16,561	1,241	9
	경북	25,652	306,998	611,326	596,970	803,174	863,504	415,950	109,392	20,261	1,350	29
	경남	36,852	362,620	717,605	827,112	1,004,207	1,051,431	494,743	127,089	20,595	1,191	17
	제주	6,892	105,916	221,391	240,057	299,590	274,559	115,871	29,977	5,851	483	9
광역시 · 도	서울	53,745	968,373	3,834,652	3,488,501	2,872,828	2,499,014	1,456,561	509,017	84,592	4,858	140
	부산	31,575	386,262	1,089,819	931,434	949,475	1,061,792	665,998	218,531	34,385	1,810	32
	대구	21,874	319,313	749,170	616,948	714,615	727,459	356,988	109,785	21,270	1,232	36
	인천	30,388	270,906	747,893	748,497	731,030	757,872	365,477	110,443	19,609	1,154	14
	광주	15,440	214,534	457,955	378,138	438,461	410,072	189,679	68,695	14,445	1,023	16
	대전	13,203	201,846	532,180	433,957	460,305	437,642	224,833	65,345	12,286	764	9
	울산	13,317	127,062	299,142	329,554	369,857	410,713	180,756	40,487	5,818	345	6
	세종	2,332	23,906	58,748	75,293	79,195	64,344	28,990	7,114	1,267	72	1
	경기	125,256	1,293,054	3,239,534	3,462,369	3,567,679	3,405,654	1,618,240	521,645	93,686	5,027	114
	강원	18,968	217,558	453,246	388,565	501,741	557,961	281,269	88,460	15,814	1,028	25
	충북	19,489	190,290	438,764	396,617	476,959	503,873	237,988	66,826	12,154	780	13
	충남	20,609	238,409	538,918	548,338	634,114	668,445	317,971	88,303	17,208	1,032	20
	전북	16,315	231,736	434,520	390,961	522,802	544,737	268,929	87,773	19,795	1,600	55
	전남	18,005	193,041	323,627	370,560	505,917	545,046	262,211	83,050	16,525	1,240	9
	경북	25,371	305,566	609,702	595,040	800,590	861,051	414,926	109,062	20,218	1,348	29
	경남	36,496	361,036	715,633	824,553	1,001,103	1,048,621	493,707	126,771	20,560	1,190	17
	제주	6,816	105,436	220,812	239,326	298,682	273,857	115,633	29,901	5,839	482	9

제4장 모바일 빅데이터 기반 통행 분석 서비스 개발

제1절 개요

제2절 UI 설계 및 기능 개발

제3절 기반 DB 구축





제3절 운영 및 유지보수

제4장 모바일 빅데이터 기반 통행 분석 서비스 개발

제1절 개요

1. 모바일 빅데이터 기반 통행 분석 서비스 개발 현황

- 기 개발된 교통 모니터링·데이터 제공·분석 플랫폼(View-T)에 모바일 빅데이터를 추가로 탑재하여 2019년 1월부터 사람 통행 분석 서비스를 제공하기 시작
 - 기존에는 내비게이션 데이터 기반의 차량 통행 분석 서비스만 제공
- 사람 통행 분석 서비스를 제공하기 위해 2018년 총 4개의 주요 분석 기능과 1개의 부가 기능을 개발함

	
유입유출 지역분석	유입유출 비교분석
	
O/D 분석	Hot Place 분석

<그림 4- 1> 2018년 개발한 주요 분석 기능 4가지

– 주요 분석 기능

- 유입유출 지역분석: 특정 지역(사용자가 분석하기 원하는 지역, 이하 ‘선택지역’)으로 유입되는 통행량과 선택지역에서 유출되는 통행량의 분포를 시도, 시군구, 읍면동,

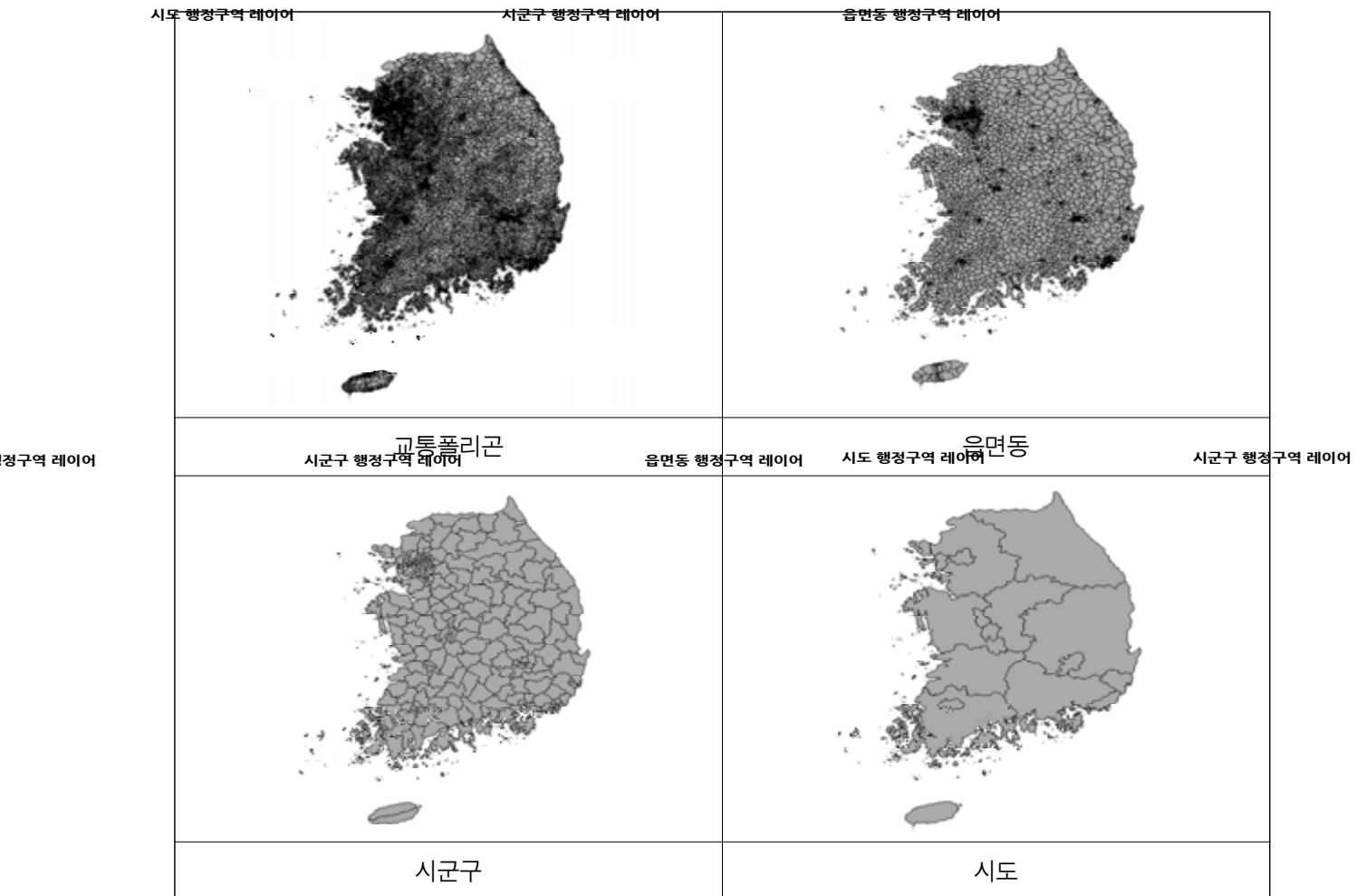
교통폴리곤 단위로 확인할 수 있는 기능

- 유입유출 비교분석: 선택지역으로 유입되는 통행량과 선택지역에서 유출되는 통행량의 분포를 비교 분석할 수 있는 기능 (최대 2개 지역 비교 분석 가능)
- O/D 분석: 선택지역 내 통행량이 높은 구간을 확인할 수 있는 기능
- Hot Place 분석: 선택지역에서 유출량이 많은 지점 또는 유입량이 많은 지점을 확인할 수 있는 기능
- 부가 기능: 사용자가 조건을 설정하여 해당되는 통행량을 검색하고 이를 파일 형태로 출력할 수 있는 기능
- 또한, 분석 재료가 되는 모바일 빅데이터 기종점 통행량 DB와 분석결과를 시각화 할 수 있는 모바일 분석 맵을 View-T에 탑재함
- 모바일 빅데이터 기종점 통행량 DB: 2016년 3월 7일 ~ 2016년 12월 31일에 기록된 로그 기록을 아래와 같은 형태로 가공한 DB

<표 4-1> 통신 빅데이터의 컬럼명 및 데이터 타입

NO	필드명	설명	데이터 타입
1	o_polygon	출발폴리곤	STRING
2	o_base_ymd	출발 일자	STRING
3	o_timezn_cd	출발시간대	STRING
4	o_trip_type	출발지 OD트립타입	STRING
5	d_polygon	도착폴리곤	STRING
6	d_base_ymd	도착 일자	STRING
7	d_timezn_cd	도착시간대	STRING
8	d_trip_type	도착지 OD트립타입	STRING
9	age_itg_cd	연령대 정보	INT
10	sex_type_itg_cd	성별 정보	STRING
11	total	이동 인구수	INT
12	base_ymd	기준일	STRING

- 모바일 분석 맵: 전국 15,937개의 교통폴리곤, 교통폴리곤의 속성정보가 되는 행정구역(읍면동, 시군구, 시도)
- 교통폴리곤: 2015년 집계구와 2016년 기지국 정보를 활용하여 구축한 영역
- 행정구역(읍면동, 시군구, 시도): 2015년 기준



<그림 4-2> 모바일 분석 맵 4가지 (2018년 구축 공간정보)

2. 2019년 개발 방향 및 범위

가. 기존 분석 서비스 고도화

1) 주요통행지역분석

- 유입유출 지역분석과 유입유출 비교 분석을 하나로 통합
 - 유입유출 지역분석과 유입유출 비교 분석은 하나의 분석 대상 지역을 선택할 것인지, 여러 개의 분석 대상 지역을 선택할 것인지에 따라 분리된 것일 뿐, 분석 성격·결과는 동일
- 기존보다 더 많은 지역을 비교 분석할 수 있도록 개선
- 사용자가 선택한 분석 대상 지역의 결과를 더 쉽게 구분할 수 있도록 표출 방법(시각화 방법)을 개선
- 사용자가 분석결과를 더 빨리 파악할 수 있도록 행정구역 명칭 ON/OFF 기능을 추가
- 선택지역에 대한 분석결과를 더 쉽게 비교할 수 있도록 상세 정보도 시각화하여 표현

<표 4- 2> 주요통행지역분석

구분	기존	개선(안)
분석 대상 지역	- 2개 지역까지 선택 가능	- 5개 지역까지 선택 가능 - 권한에 따라 선택 개수 제한
표출 방법	- 선택지역의 분석결과가 동일한 색상으로 표현	- 선택지역별 분석결과를 각기 다른 색상으로 표현
기타	- 주제도 클릭시 상세결과를 텍스트 형식으로 확인	- 행정구역 ON/OFF 기능 추가 - 주제도 클릭시 상세결과를 파이차트 형식으로 확인 가능

2) 주요통행구간분석 (전 O/D 분석)

- 지역간 통행 정보도 확인할 수 있도록 개선
 - 기존에는 사용자가 선택한 지역을 기준으로 지역 내 주요 통행 구간만 확인 가능
- 기존보다 더 많은 지역을 분석할 수 있도록 분석 대상 지역 선택 폭 확대

- 분석결과에 대한 표출단위를 다양하게 설정할 수 있도록 개선

<표 4- 3> 주요통행구간분석 개선 방향

구분	기존	개선(안)
분석 대상 지역	- 1개 지역만 선택 가능	- 5개 지역까지 선택 가능 - 권한에 따라 선택 개수 제한
분석 결과	- 지역 내 통행 정보 확인 가능	- 지역 간 통행 정보 확인 가능
표출 단위	- 교통폴리곤 단위	- 시군구 단위 - 읍면동 단위 - 교통폴리곤 단위

3) Hot Place 분석

- 분석결과에 대한 표출단위를 다양하게 설정할 수 있도록 개선
- 산출 결과를 기존보다 다양하게 확인할 수 있도록 개선

<표 4- 4> Hot place 분석 개선 방향

구분	기존	개선(안)
표출 단위	- 교통폴리곤 단위	- 시군구 단위 - 읍면동 단위 - 교통폴리곤 단위
표출 결과	- 통행량 (총 합계)	- 통행량 (총 합계) - 일평균 통행량 (통행량/분석일) - 면적대비 통행량 (통행량/분석단위 면적)

4) 부가 서비스 (데이터 검색 및 다운로드)

- 다양한 형태로 통신데이터를 집계해서 다운로드할 수 있도록 유형별 통신데이터 다운로드 기능과 지역간 OD 통행량 데이터 다운로드 기능 추가

- 기존에 사용자가 원천데이터를 다운 받은 후 필요에 따라 집계해서 사용해야 했던 불필요한 단계를 단축
- 사용자 대기시간을 단축시키기 위해 표출화면에 표기되지 않는 데이터인 경우 분석결과를 메일로 받아볼 수 있도록 함

<표 4- 5> 부가 서비스 개선 방향

구분	기존	개선(안)
출력(통행량 집계) 조건 설정	- 설정 불가	- 사용자가 설정한 행 조건(가로축), 열 조건(세로축)에 따라 다양하게 출력(집계) 가능 - 지역간 통행량 출력(집계) 가능
기타	-	- 분석 결과를 메일로 송부

나. 신규 분석 서비스 개발

1) 내부통행지역분석

- 내부 통행이 많은 지역을 쉽게 확인할 수 있는 분석 서비스를 개발
 - 사용자가 선택한 지역의 내부 통행 비율을 주제도에서 확인할 수 있도록 기능 개발

2) 통행시간 / 거리 분석

- 지역별 접근성을 쉽게 확인할 수 있는 분석 서비스를 개발
 - 사용자가 선택한 지역을 기준으로 평균 통행시간, 평균 통행 거리 정보를 주제도에서 확인할 수 있도록 기능 개발

다. 기반 DB 업데이트

1) 모바일 분석 맵 업데이트

- 2018년에 개발한 교통폴리곤 영역과 2016년 기준 행정구역 정보를 구축

<표 4- 6> 분석 맵 업데이트

구분	기존 구축 DB	신규 구축 DB
교통폴리곤	- 2015년 집계구와 2016년 기지국 정보를 활용하여 구축한 영역 (2017년 개발)	- 2016년 집계구와 2017년 기지국 정보를 활용하여 구축한 영역 (2018년 개발)
행정구역	- 2015년 기준 읍면동, 시군구, 시도 영역	- 2016년 기준 읍면동, 시군구, 시도 영역

2) 모바일 빅데이터 기종점 통행량 업데이트

- 2017년에 생성된 로그 기록을 활용하여 구축한 기종점 통행량 DB 구축
- 3통행 미만 보정계수를 적용하여 과다하게 집계되던 통행량을 보정
 - 원천데이터에서 기종점 통행량 정보로 가공할 때 개인정보보호법에 따라 3통행 미만인 경우에는 모두 3통행으로 처리하여 실제 통행량보다 높게 집계되는 문제 발생
 - 본 과업에서 개발한 3통행 미만 보정계수를 적용하여 부풀려진 통행량을 보정한 후 시스템에 탑재

<표 4- 7> 모바일 빅데이터 기종점 통행량 업데이트

구분	기존 구축 DB	신규 구축 DB
시간적 범위	2016년 3월 ~ 2016년 12월	2017년 1월 ~ 2017년 12월
공간적 범위	전국	좌동
기타	-	3통행 미만 보정계수 적용

라. 사용자 편의 개선

1) 사용방법 안내

- 매뉴얼을 찾아보지 않더라도 사용자가 분석 기능을 사용하는 데 어려움을 느끼지 않도록 사용 순서와 설명을 분석 기능 동작 시 표시

2) 분석조건 설정

- 분석 기간 설정은 수기로 입력 가능하도록 개선
 - 시작 일자와 종료 일자를 전체 기간에서 찾아서 마우스로 클릭하는 기존 방식은 분석 기간 설정에 번거로움을 유발
- 사용자가 자주 선택하는 조건을 옵션으로 추가
 - 요일 설정 부분에 주중/주말 옵션 추가
 - 시간대 설정 부분에 오전 첨두시간대(6시~9시) / 오후 첨두시간대(18시~21시) 옵션 추가

3) 범례 설정

- 범례를 쉽게 설정할 수 있도록 자주 사용하는 범례 설정 방식(등간격, 등분위, 상위%)을 옵션에 추가
- 사용자가 색상 선택을 편하게 설정할 수 있도록 심볼(색상) 스키마를 다양하게 구성

4) 분석결과 다운로드

- 분석 결과(주제도)를 사용자가 자유롭게 활용할 수 있도록 스크린 샷 기능을 추가

5) 기타

- 주제도에 표출되는 라벨의 폰트를 변경할 수 있도록 개선

<표 4- 8> 사용자 편의 개선

구분	기존	개선(안)
사용방법 안내	- PDF 방식의 매뉴얼 탑재	- 사용 순서와 설명을 분석 기능 동작시(마우스 오버) 표출 - PDF 방식의 매뉴얼 탑재
분석조건 설정	- 분석기간을 마우스로 클릭하여 설정 - 분석요일을 마우스로 클릭 - 1시간 단위의 분석시간대를 마우스로 클릭	- 분석기간을 수기로 입력 (마우스로 클릭하는 기존 방식도 유지) - 주중/주말 옵션 추가 - 오전 첨두 시간대/오후 첨두 시간대 옵션 추가
범례 설정	- 수기 입력 - 단계별 심볼 색상 직접 선택	- 등간격, 등분위, 상위%, 수기입력 가능 - 사전 정의된 스키마로 심볼 색상 자동 선택 가능
분석결과 다운로드	- 분석결과를 텍스트 형태(CSV 파일)로 다운로드	- 분석결과를 텍스트 형태(CSV 파일), 그림 형태(PNG, JPEG, BMP, GIF)로 다운로드 가능
기타	- 단일 라벨 폰트 사용	- 라벨 폰트 변경 가능

제2절 II 설계 및 기능 개발



1. 총괄


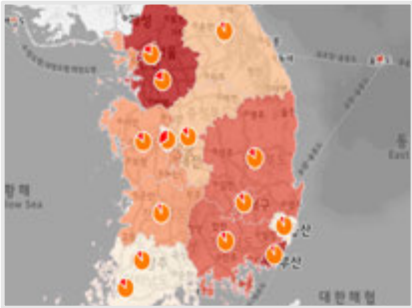
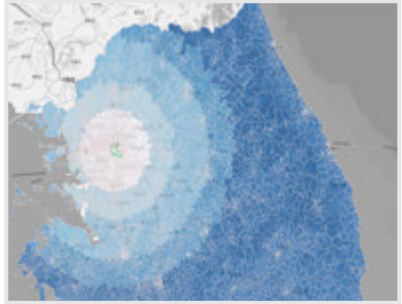
가. 분석 서비스 및 메뉴 구성 변경

1) 분석 서비스 구성

- 기존 주요 분석 기능 3가지와 새로 개발하는 분석 기능 2가지로 구성
 - 기존 기능에서 성격이 유사한 유입유출 지역분석 기능과 유입유출 비교분석 기능을 하나로 통합하고, 기존 분석 기능 명칭을 사용자가 이해하기 쉽도록 변경
 - 신규 분석 기능으로 내부통행지역분석, 통행거리/시간분석 기능 추가

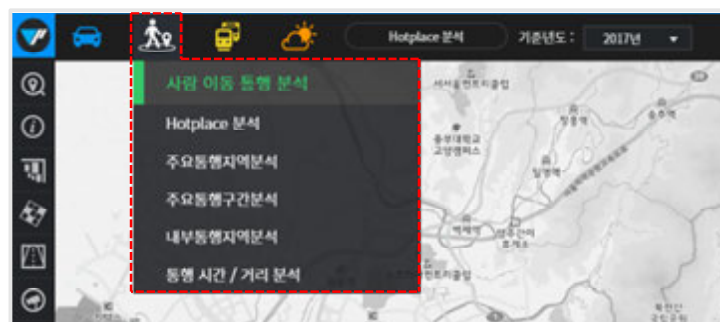
<표 4- 9> 사용자 편의 개선

기능명	설명	실행화면
Hot Place 분석	- 통행수요가 높은 출발지 또는 도착지를 확인할 수 있는 기능	
주요통행 지역분석	- 선택 지역에 대한 주요 출발지와 주요 도착지를 확인·비교할 수 있는 기능(출발지, 도착지의 분포를 확인할 수 있는 기능) - 기존 '유입유출 지역분석'과 '유입유출비교분석'을 통합	

주요통행 구간분석	- 통행이 많은 구간을 확인할 수 있는 기능	
내부통행 지역분석	- 내부통행량이 많은 지역을 확인할 수 있는 기능	
통행시간/ 거리분석	- 지역별 접근성을 평균통행시간, 평균통행거리로 확인할 수 있는 기능	

2) 메뉴 구성

- 사용 빈도, 분석 기능의 복잡성에 따라 메뉴 순서 변경
 - 가장 마지막에 배치되어 있었던 Hot Place 분석 기능을 맨 앞으로 배치
- 새로 개발하는 분석 기능을 주요분석기능 메뉴에 추가



<그림 4- 3> 사람 이동 통행 분석 메뉴

- 메뉴의 상세 구성은 다음과 같음

<표 4- 10> 메뉴의 구성

페이지메뉴	상세 메뉴	기능	설명
사람 이동 통행 분석	Hotplace 분석	통행 기준 선택	유입기준 또는 유출기준 선택
		분석지역 선택	시도, 시군구, 읍면동 단위로 선택
		표출단위 선택	선택 분석지역의 하위 단위 선택
		상세설정	기간, 요일, 시간대, 체류특성, 성별, 연령대 선택
		산출결과 선택	총 유동인구, 일평균 유동인구, 면적대비 유동인구 선택
		표출지도 설정	분석결과에 대한 주제도 범례설정 분석결과에 대한 순위 표출 설정
		다운로드	표출지도 이미지 다운로드, 분석 결과 파일출력(csv)
	주요통행 지역분석	통행 기준 선택	유입량 또는 유출량 선택
		비교지역 선택	비교지역에 대한 출발지(유출), 도착지(유입) 선택 권한에 따라 2-5개까지의 비교지역 선택
사람 이동 통행 분석	주요통행 지역분석	표출지역 선택	통행량 집계지 출발지(유입), 도착지(유출) 선택 선택 지역의 하위 행정구역 단위로 표출 단위 선택
		상세설정	기간, 요일, 시간대, 체류특성, 성별, 연령대 선택
		산출결과 선택	총 유동인구, 일평균 유동인구, 면적대비 유동인구 선택
		표출지도 설정	분석결과에 대한 희망선, 주제도 범례설정
		다운로드	표출지도 이미지 다운로드, 분석 결과 파일출력(csv)
	주요통행 구간분석	분석지역 선택	시군구, 읍면동 단위의 5개의 분석지역 선택 권한에 따라 단일 시도 선택
		표출단위 선택	선택 분석지역의 하위 단위 선택
		상세설정	기간, 요일, 시간대, 체류특성, 성별, 연령대 선택
		산출결과 선택	OD 통행량, OD + DO 통행량, 일평균 OD 통행량, 일평균 OD + DO 통행량 선택
		표출지도 설정	분석결과에 대한 희망선 범례설정
		다운로드	표출지도 이미지 다운로드, 분석 결과 파일출력(csv)
	내부통행 지역분석	분석지역 선택	전국, 시도, 시군구, 읍면동 단위 선택
		상세설정	기간, 요일, 시간대, 체류특성, 성별, 연령대 선택

		표출지도 설정	분석결과에 대한 주제도, 파이차트 범례설정
		다운로드	표출지도 이미지 다운로드, 분석 결과 파일출력(csv)
	통행 거리/시간 분석	통행 거리/시간 선택	통행 거리 또는 통행 시간 선택
		분석지역 선택	시군구, 읍면동 단위 선택 가능 출/도착 기준 선택
		상세설정	기간, 요일, 시간대, 체류특성, 성별, 연령대 선택
		그래프 산출 결과 선택	일자별, 요일별, 평일/주말별, 월별 통행량 선택
		그래프 표출	그래프 산출 유형의 분석 결과로 그래프 확인
		표출지도 설정	분석결과에 대한 주제도 범례설정
		다운로드	표출지도 이미지 다운로드, 분석 결과 파일출력(csv)

나. 사용방법 안내 방식 변경

- 기존 pdf 형태의 매뉴얼을 제공하는 방식에서 각 기능 메뉴 선택 시 설정 부분에 숫자를 표기하여 기능사용 순서를 유도하는 방식으로 사용방법 안내 방식을 변경
 - 해당 숫자의 마우스 오버 시 설정에 대한 내용을 보여줌으로써 사용 방법의 어려움을 개선
 - 더 이상 보지 않기 버튼을 클릭하면 하루 동안 사용자 기능 유도 메시지 표시안함



<그림 4- 4> 사용자 기능 유도 표출 예시

다. 분석 조건 설정 방법 개선

- 기간 설정 시 수기 입력이 가능하도록 개선
- 요일 설정의 주중과 주말에 대한 선택 버튼을 추가하여 사용자 편의성을 높임
- 시간대 설정의 오전 첨두(6시~9시)와 오후 첨두(18시~21시)에 대한 선택 버튼을 추가하여 사용자가 자주 분석하는 시간대의 설정을 용이하게 함

상세설정

기간/요일: 2016-03-07 ~ 2016-12-31

시간대: 출발 전제 시간, 도착 전제 시간

주요 체류 특성: 출발 전제, 도착 전제

성별/연령(대): 성별 여/남자, 연령 전제 연령

▶ 사용자 수기 입력 가능

▶ 주중/주말 선택

▶ 오전 첨두/오후 첨두 선택

<그림 4- 5> 상세설정 UI

라. 범례 설정 방법 개선

- 등간격, 등분위, 상위%, 직접입력 총 네 가지 방식으로 범례를 설정할 수 있도록 함

범례 설정

범례 설정 방법: 직접입력, 등간격, 등분위, 상위%

범례 설정 값: 최소, 최대, 상위%

범례 설정 방법: 직접입력, 등간격, 등분위, 상위%

범례 설정 값: 최소, 최대, 상위%

범례 설정 방법: 직접입력, 등간격, 등분위, 상위%

범례 설정 값: 최소, 최대, 상위%

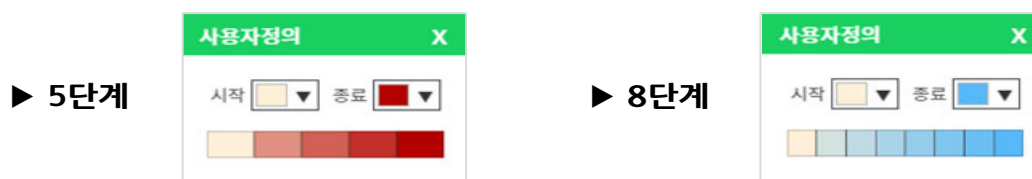
<그림 4- 6> 범례설정 UI

- 다양한 색상의 심볼 색상 스키마를 제공하여 사용자의 편의성을 증진
 - 단계에 따라 색상 스키마가 변경되며, 특정 스키마 선택 시 테이블의 심볼 색상이 변경됨



<그림 4- 7> 심볼 스키마 예시

- 또한 제공되는 스키마와 별개로 사용자가 설정한 시작과 종료의 색상으로 그라데이션의 스키마를 생성할 수 있게끔 함
 - 단계에 따라 색상 스키마가 변경되며, 특정 스키마 선택 시 테이블의 심볼 색상이 변경됨

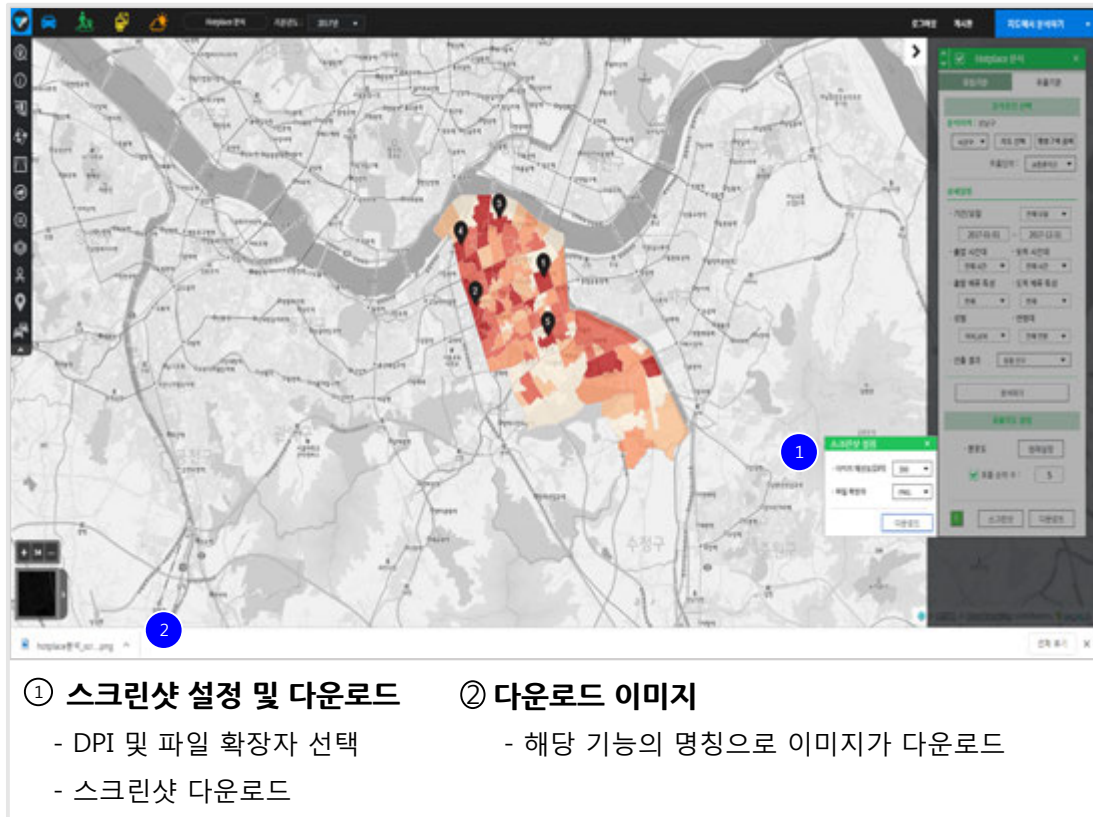


<그림 4- 8> 심볼 색상 사용자정의 예시

마. 분석결과 유형 추가

- 주제도로 표출되는 분석결과도 사용자가 활용할 수 있도록 스크린샷 다운로드 기능을 개발
 - 이미지 해상도(DPI) : 72, 100, 200, 250, 300, 350, 400, 450, 500

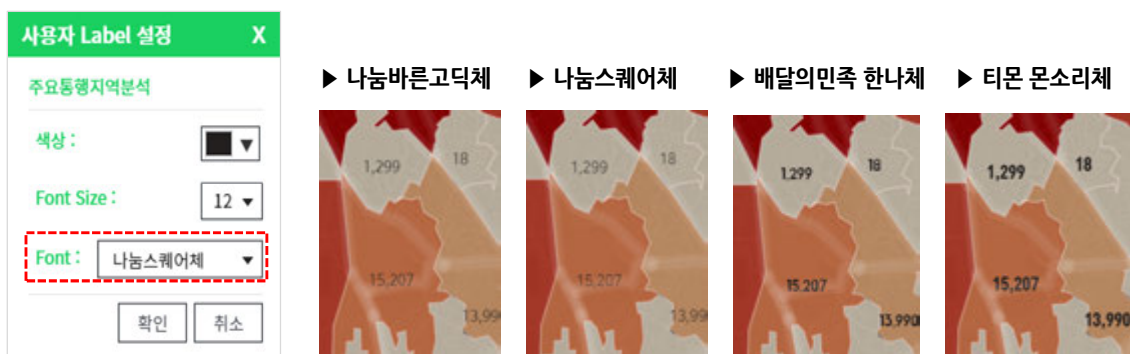
– 파일 확장자 : PNG, JPEG, BMP, GIF



<그림 4- 9> 스크린샷 UI

바. 주제도 라벨 폰트 추가

- 주제도 표출 시 설정하는 라벨의 폰트를 추가하여 사용자 선택의 다양성을 높임



<그림 4- 10> 라벨 폰트 표출 예시

2. 분석 기능별 UI 설계 및 개발 결과

가. Hot Place 분석 기능

- 사용자가 선택한 특정지역 범위 내의 사람이 가장 많이 몰리는 곳, 외부로 가장 많이 나가는 곳을 분석할 수 있도록 개발함
- 지도에 표출되는 분석결과에 대한 표출 단위를 사용자가 설정할 수 있도록 개선함
 - 시군구, 읍면동, 교통폴리곤 단위로 선택 가능 (단, 사용자가 선택한 분석지역의 하위 단위로만 선택 가능)



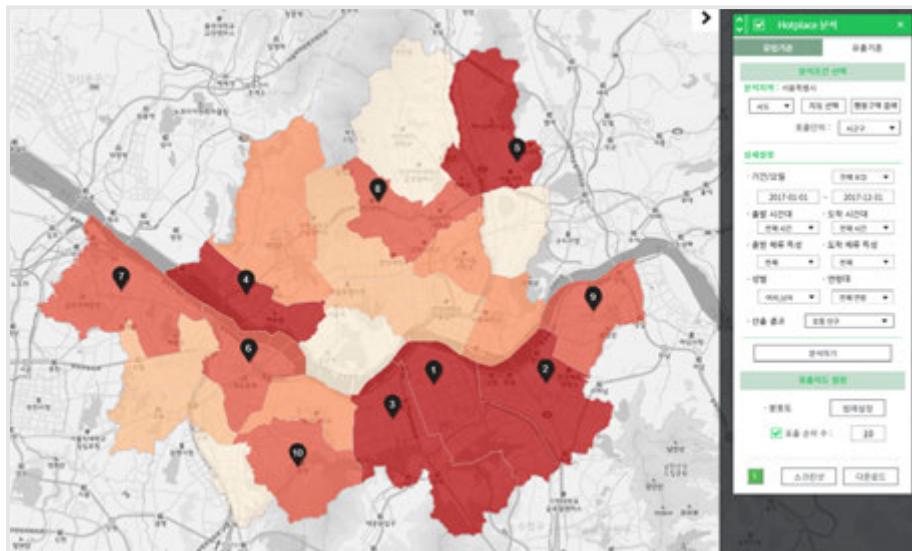
<그림 4- 11> Hotplace 분석 개선된 UI

- 사용자가 산출 결과를 선택하여 유동인구(합계), 유동인구/분석일(평균), 면적대비 유동인구의 다양한 분석이 가능하도록 구현함
- 분석결과는 주제도로 표출하고 표출지도에 대한 범례를 설정할 수 있도록 구현함
- 통행이 많은 지역을 쉽게 확인할 수 있도록 표출지도에 순위를 함께 표출함
- 분석결과를 파일형태(csv)로, 표출지도를 스크린샷으로 다운로드할 수 있도록 개발함

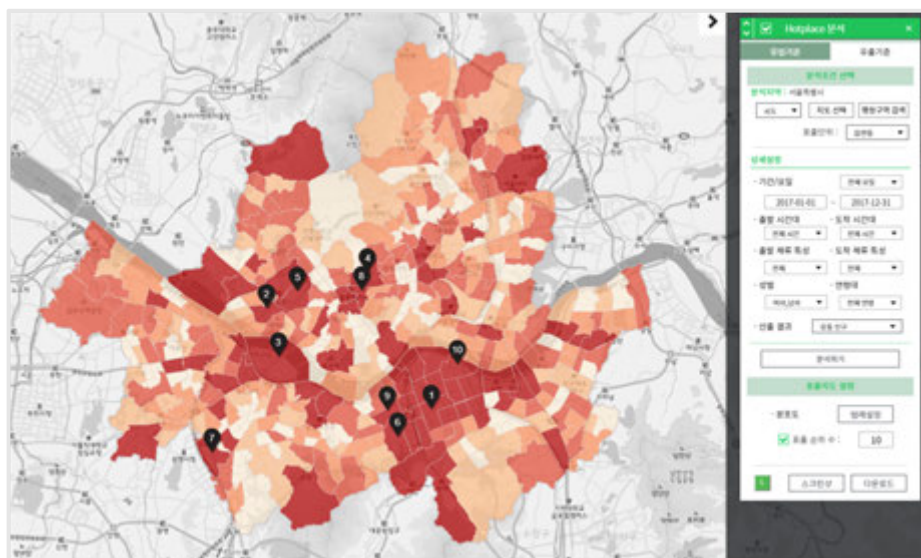
※ Hot Place 분석 예시

○ 분석 조건

- 통행유형 : 유출기준 통행량 / 분석지역 : 서울특별시
- 표출단위 : 시군구 / 읍면동
- 기간 : 2017.01.01. ~ 2017.12.31. / 요일 : 전체 요일 / 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성 / 성별 : 남, 여 / 연령대 : 전체 연령대



<그림 4- 12> Hot Place 분석 시군구 단위 표출 예시



<그림 4- 13> Hot Place 분석 읍면동 단위 표출 예시

나. 주요통행지역분석 기능

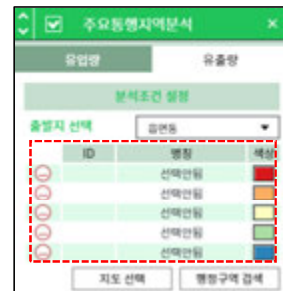
- 시도, 시군구, 읍면동, 교통폴리곤 단위로 다중 지역을 선택하여 선택지역으로 유입되는 통행량 또는 유출되는 통행량을 비교 분석할 수 있도록 개발함
- 권한에 따라 2 - 5개까지의 지역을 선택할 수 있도록 개선함

▶ 개선 전



▶ 개선 후

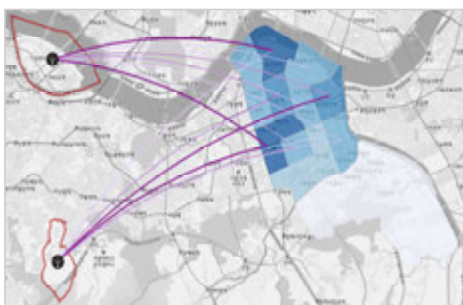
· 고급사용자 권한 미만 (최대2개) · 고급사용자 권한 이상 (최대5개)



<그림 4- 14> 주요통행지역분석 권한에 따른 선택 지역 UI

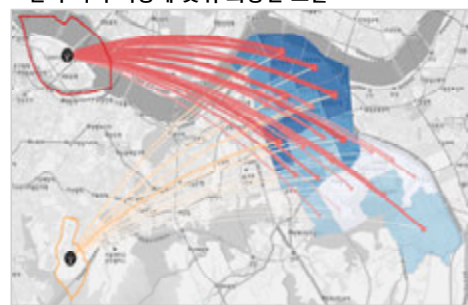
- 선택지역에 대한 분석결과를 더 쉽게 구분할 수 있도록 개선함
 - － 선택지역별 대표 색상을 선택할 수 있도록 함
 - － 통행량(유출량, 유입량) 정도에 따라 희망선의 굵기와 농도가 자동 조정되게끔 함

▶ 개선 전



▶ 개선 후

· 선택 지역 색상에 맞춰 희망선 표출

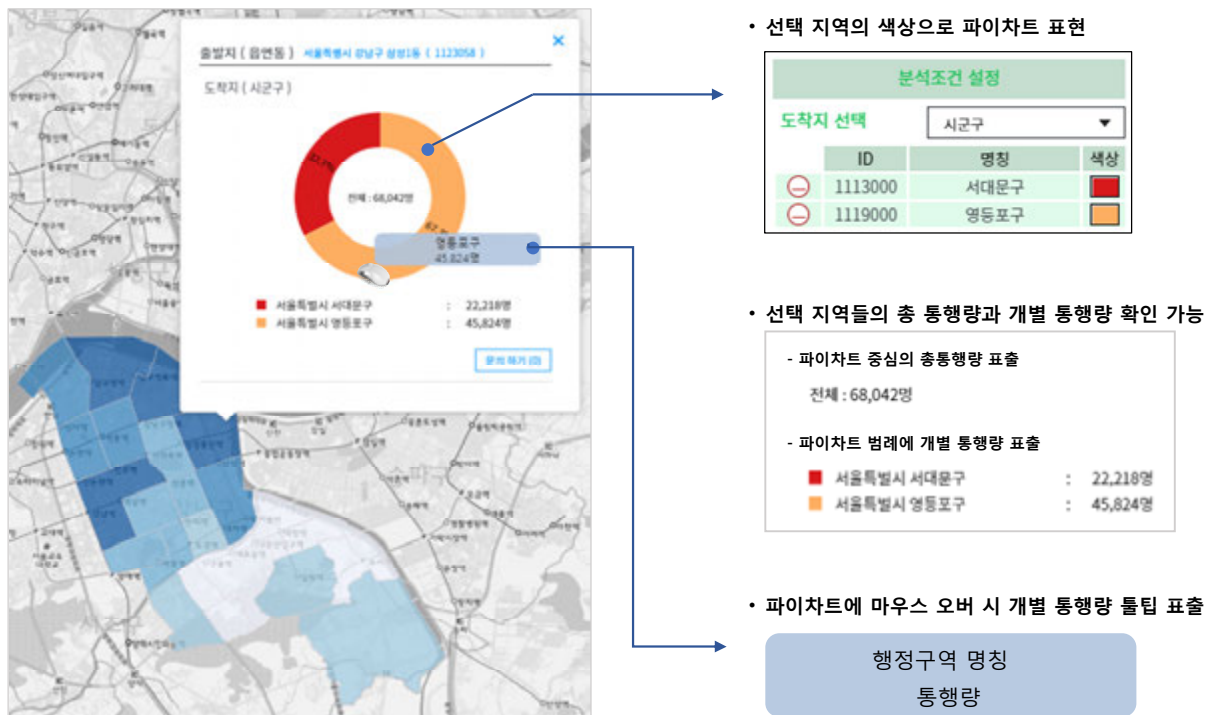


<그림 4- 15> 주요통행지역분석 희망선 색상 구분

- 유입량 분석은 다중의 도착지역 선택과 단일의 출발지역 선택, 유출량 분석은 다중의 출발지역 선택과 단일의 도착지역 선택이 가능하도록 구현함
- 유입량 분석은 출발지 집계단위, 유출량 분석은 도착지의 집계단위를 시도, 시군구, 읍면동,

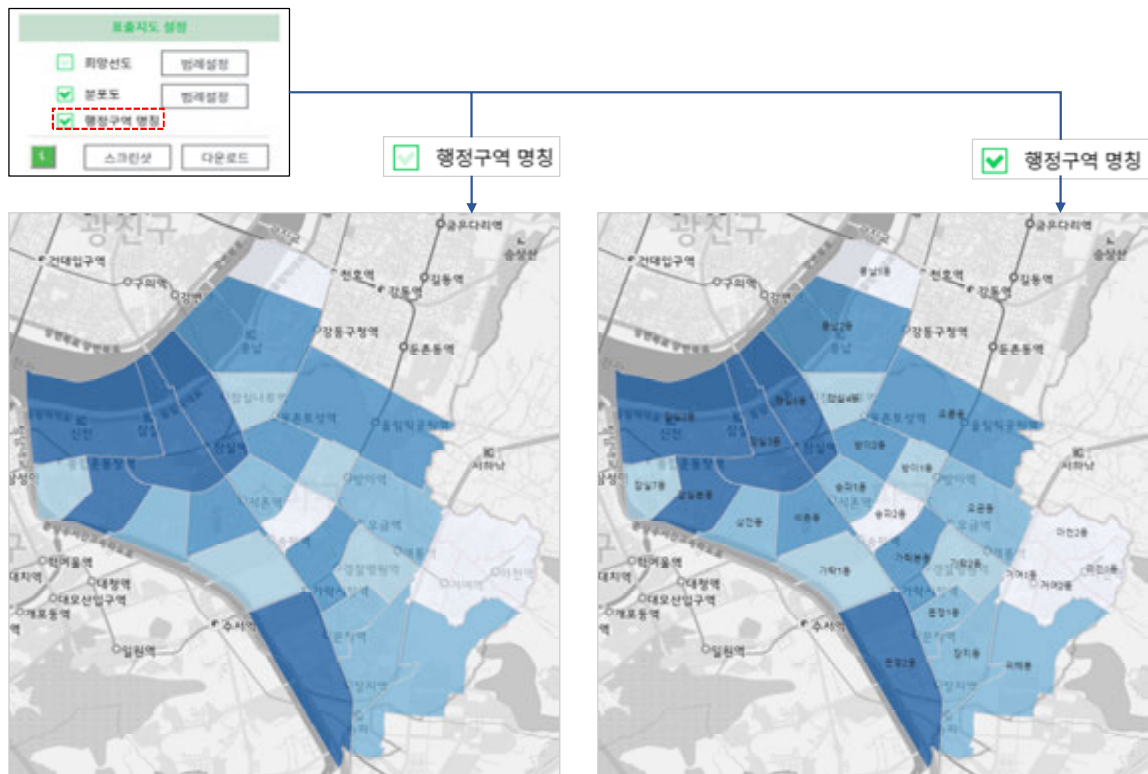
교통폴리곤 단위로 설정할 수 있도록 구현함

- 사용자가 산출 결과를 선택하여 유동인구(합계), 유동인구/분석일(평균), 면적대비 유동인구의 다양한 분석이 가능하도록 개발함
- 분석결과를 희망선과 주제도로 표출하고 표출지도의 범례를 설정할 수 있도록 구현함
- 주제도로 표출된 분석결과에 대한 상세 설명을 파이차트 형식으로 제공하여 선택지역 간의 통행량을 보다 쉽게 비교할 수 있도록 함
- 상세 설명은 표출된 주제도를 클릭할 때 팝업창으로 표출되며, 행정구역 명칭, 행정구역 ID, 파이차트(전체 통행량 대비 각 선택지역의 통행량이 차지하는 비율 표기), 파이차트의 범례로 구성됨
- 파이차트에 마우스 오버 시 선택지역의 행정구역 명칭과 통행량 값 확인 가능



<그림 4- 16> 주요통행지역분석 파이차트 정보표출

- 행정구역 명칭 on/off 기능으로 표출된 주제도의 위치 확인을 용이하게 개선함
- 교통폴리곤이 속한 읍면동 명칭과 교통폴리곤의 ID가 같이 표출됨
- 교통폴리곤 이외의 행정구역은 행정구역 명칭만 표출됨



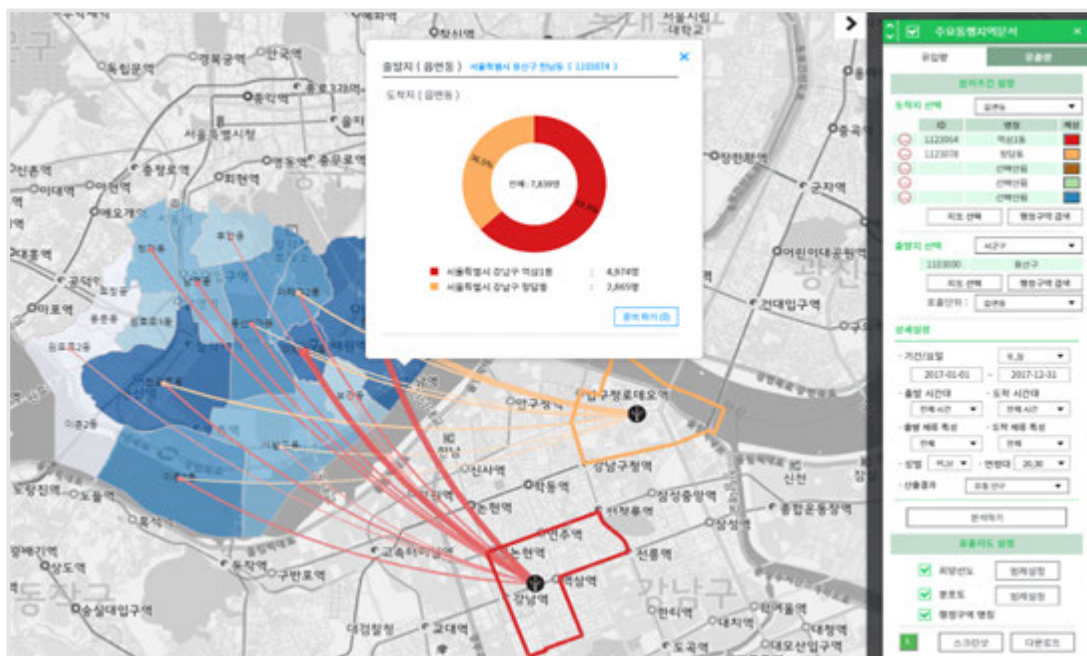
<그림 4- 17> 주요통행지역분석 주제도의 행정구역 명칭 제어

- 분석결과를 파일형태(csv)로, 표출지도를 스크린샷으로 다운로드할 수 있도록 구현함

※ 주요통행지역분석 예시

○ 분석 조건

- 통행유형 : 유입량
- 분석지역
 - 출발지 : 서울특별시 강남구 역삼1동, 서울특별시 강남구 청담동
 - 도착지 : 서울특별시 용산구
- 표출단위 : 읍면동
- 기간 : 2017.01.01. ~ 2017.12.31. / 요일 : 주말
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 20, 30대



<그림 4- 18> 주요통행지역분석 분석 화면 예시(행정구역 명칭 ON / 정보표출 ON)

다. 주요통행구간분석 기능

- 사용자가 선택한 분석지역 내의 표출단위 행정구역간의 통행량을 분석할 수 있도록 개발함
- 시도를 제외한 시군구, 읍면동 단위는 분석지역을 최대 5개까지 선택할 수 있도록 개선하였으며, 표출단위는 사용자가 표출단위를 자유롭게 선택할 수 있도록 개선함
 - 시군구, 읍면동, 교통폴리곤 단위의 표출단위 선택 가능

▶ 개선 전

개선

- 최대 5개의 분석지역 선택
- 표출단위 선택

▶ 개선 후

<그림 4- 19> 주요통행구간분석 개선된 UI

- 표출된 행정구역간의 단방향, 양방향 통행량을 분석할 수 있도록 구현함
- 표출지도는 분석지역 내 표출단위의 행정구역간 노드를 직선으로 연결한 희망선으로 표출되도록 구현함
- 분석결과는 희망선으로 표출하고 표출지도에 대한 범례를 설정할 수 있도록 구현함
- 분석결과의 행정구역들을 표출지도에 표현하고 행정구역 명칭이 보이도록 구현함
- 표출단위가 교통폴리곤일 시 행정구역 명칭은 읍면동의 명칭으로 표현하고 교통폴리곤의 ID가 같이 보이도록 구현함

- 분석결과를 파일형태(csv)로, 표출지도를 스크린샷으로 다운로드할 수 있도록 구현함

※ 주요통행구간분석 예시

○ 분석 조건

- 분석지역: 서울특별시 강서구, 서울특별시 마포구, 서울특별시 영등포구, 서울특별시 양천구, 서울특별시 구로구
- 표출단위 : 읍면동
- 기간 : 2017.01.01. ~ 2017.12.31.
- 요일 : 전체 요일
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 전체 연령대



<그림 4- 20> 주요통행구간분석 읍면동 단위 표출 예시

라. 내부통행지역분석 기능

- 사용자가 선택한 특정 지역 범위 내의 내부 통행량과 외부 통행량을 비교 분석할 수 있도록 개발함
 - 내부통행량: 출발지와 도착지의 행정구역이 동일한 데이터의 통행량을 집계
 - 외부통행량: 출발지와 도착지의 행정구역이 다른 데이터의 통행량을 집계
- 분석지역은 지도에서 바로 선택하거나, 행정구역을 검색하여 선택할 수 있도록 개발하였으며, 분석지역을 전국, 시도, 시군구, 읍면동 단위로 선택할 수 있도록 개발함
- 표출단위는 분석지역 행정구역 단위의 한 단계 하위 단위로 선택되도록 구현함
- 사용자가 분석조건(기간, 요일, 출발 및 도착 시간대, 출발 및 도착 체류특성, 성별, 연령대)을 자유롭게 설정할 수 있도록 개발함
 - 기간 : 일 단위 날짜 설정
 - 요일 : 월, 화, 수, 목, 금, 토, 일
 - 출발 및 도착 시간대 : 00시 ~ 23시
 - 출발 및 도착 체류특성 : 집, 집 이외 주체류지, 회사, 학교, 종교집회장소, 기타
 - 성별 : 남, 여
 - 연령대 : 0 ~ 9세, 10 ~ 19세, 20 ~ 29세, 30 ~ 39세, 40 ~ 49세, 50 ~ 59세, 60 ~ 69세, 70 ~ 79세, 80 ~ 89세, 90 ~ 99세, 100 ~ 109세, 110 ~ 119세
- 분석결과는 사용자가 선택한 분석조건에 맞는 데이터를 검색하여 산출한 내부통행량과 외부 통행량을 이용하여 표출하도록 개발함
 - 내부통행량으로 주제도를 표출하고, 내부통행량과 외부통행량을 이용해 파이차트가 표출되도록 개발함
 - 분석결과의 주제도와 파이차트는 체크 박스로 제어할 수 있도록 구현함

내부통행지역분석

분석조건 선택

① **분석지역** : 선택안됨

시도 ▼ 지도 선택 행정구역 검색
(표출단위 : 시군구)

② **상세설정**

· 기간/요일 전체 요일 ▼
2017-01-01 ~ 2017-12-31

· 출발 시간대 전체 시간 ▼ · 도착 시간대 전체 시간 ▼



· 출발 체류 특성 전체 ▼ · 도착 체류 특성 전체 ▼

· 성별 여자,남자 ▼ · 연령대 전체 연령 ▼

분석하기

표출지도 설정

③ ☒ 분포도 범례설정

☒ 파이차트 내부 통행  외부 통행 

스크린샷 다운로드

① 분석지역 선택

- 전국, 시도, 시군구, 읍면동, 선택
- 분석지역의 한단계 하위 행정구역 단위로 표출단위 선택됨

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 체류특성, 성별, 연령대

③ 표출지도 설정

- 주제도 on/off 설정 및 범례 설정
- 파이차트 on/off 설정 및 색상 설정

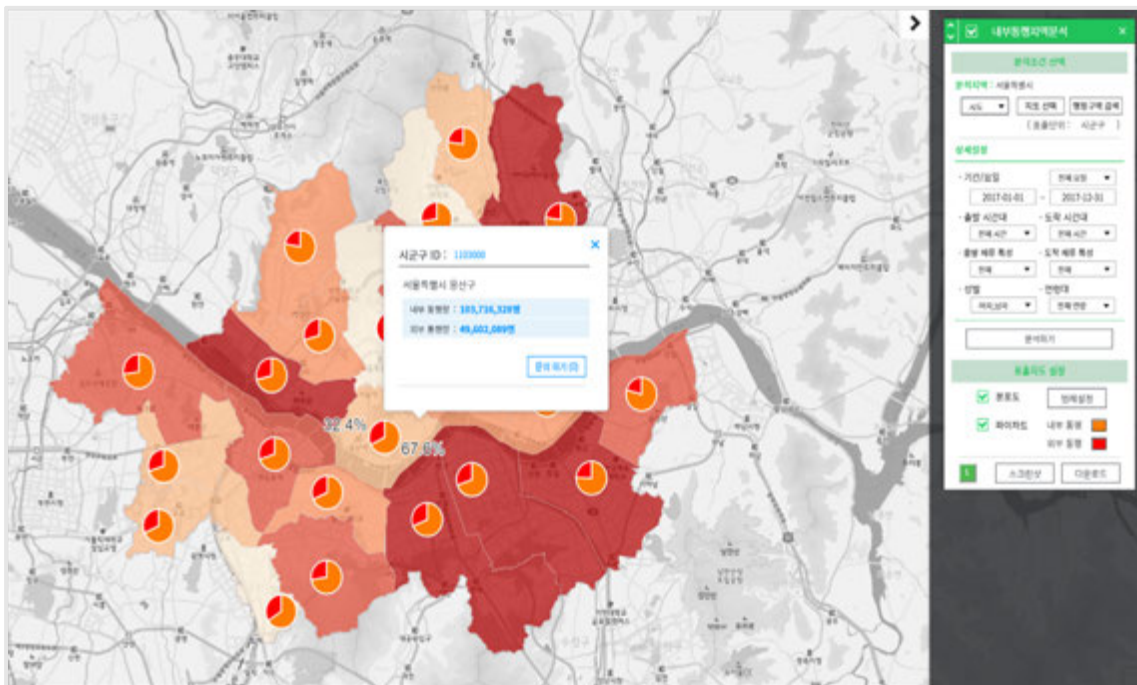
<그림 4- 21> 내부통행지역분석의 설정 UI

- 주제도와 파이차트에 대한 범례를 설정할 수 있도록 구현함
- 분석결과를 파일형태(csv)로, 표출지도를 스크린샷으로 다운로드할 수 있도록 구현함

※ 내부통행분석 예시

○ 분석 조건

- 분석지역 : 서울특별시
- 표출단위 : 시군구
- 기간 : 2017.01.01. ~ 2017.12.31.
- 요일 : 전체 요일
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 전체 연령대



<그림 4- 22> 내부통행분석 분석 화면 예시(파이차트ON / 정보표출 ON)

마. 통행시간/거리분석 기능

- 사용자가 선택한 지역을 기준으로 평균통행시간 또는 평균통행거리를 확인할 수 있도록 개발함
 - － 평균통행시간: 총 통행시간(도착시간에서 출발시간을 뺀 시간과 통행량을 곱한 값을 합산한 값)을 전체 통행량으로 나눈 값
 - － 평균통행거리: 총 통행거리(출발지와 도착지간의 직선거리와 통행량을 곱한 값을 합산한 값)를 전체 통행량으로 나눈 값
- 접근성을 어떤 지표로 확인할 것인지 선택할 수 있도록 개발함
 - － 사용자가 통행시간 분석, 통행 거리 분석을 선택할 수 있도록 개발함
- 출발지를 기준으로 분석할지, 도착지를 기준으로 분석할지 사용자가 선택할 수 있도록 함
- 분석지역은 지도에서 바로 선택하거나, 행정구역을 검색하여 선택할 수 있도록 개발하였으며, 분석지역을 시도, 시군구, 읍면동 단위로 선택할 수 있도록 개발함
- 표출단위는 전국의 교통폴리곤으로 표출되도록 구현함
- 사용자가 분석조건(기간, 요일, 출발 및 도착 시간대, 출발 및 도착 체류특성, 성별, 연령대)을 자유롭게 설정할 수 있도록 개발함
 - － 기간 : 일 단위 날짜 설정
 - － 요일 : 월, 화, 수, 목, 금, 토, 일
 - － 출발 및 도착 시간대 : 00시 ~ 23시
 - － 출발 및 도착 체류특성 : 집, 집 이외 주체류지, 회사, 학교, 종교집회장소, 기타
 - － 성별 : 남, 여
 - － 연령대 : 0 ~ 9세, 10 ~ 19세, 20 ~ 29세, 30 ~ 39세, 40 ~ 49세, 50 ~ 59세, 60 ~ 69세, 70 ~ 79세, 80 ~ 89세, 90 ~ 99세, 100 ~ 109세, 110 ~ 119세
- 분석결과는 주제도와 그래프 두 가지 방식으로 확인할 수 있도록 개발함
 - － 그래프는 사용자가 주제도를 보다 풍부하게 해석할 수 있도록 도와주는 보조 역할을 담당
 - 주제도에는 사용자가 선택한 분석조건에 해당하는 값들의 평균값이 표출되며, 그래프에는 사용자가 선택한 분석조건에 해당하는 값들이 단순 집계되어 표출됨
 - － 그래프는 주제도를 클릭하거나, 그래프 분석결과를 클릭하여 확인 가능
 - 주제도의 특정 지역을 클릭했을 때 표출되는 그래프는 해당 지역에 대한 상세 정보(

통행시간별, 통행거리별 통행량 정보)를 사용자에게 제공

▶ 통행 거리 분석의 정보표출



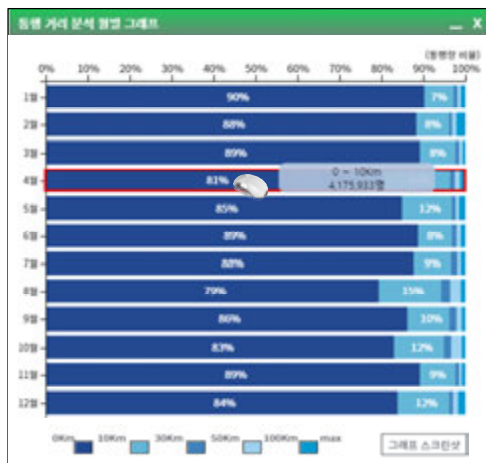
▶ 통행 시간 분석의 정보표출



<그림 4- 23> 주제도 클릭시 나타나는 그래프 UI

- 그래프 분석결과를 클릭했을 때 표출되는 그래프는 주제도에 표출된 모든 지역에 대한 상세 정보(일자별, 평일/주말별, 요일별, 월별 통행시간, 통행거리 정보)를 사용자에게 제공

▶ 통행 거리 분석의 월별 그래프



▶ 통행 시간 분석의 월별 그래프



<그림 4- 24> 그래프 분석하기 클릭시 나타나는 그래프 UI

- 사용자가 분석하기 버튼을 클릭하면 우선 주제도만 표출되며, 그래프는 사용자가 그래프 보기 버튼을 눌렀을 때에만 표출됨

- 주제도와 그래프 표출 처리 과정을 분리하여 기 표출된 분석결과에 영향을 미치지 않도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 분석결과는 기본적으로 파일형태(csv)로 출력할 수 있도록 하였으며, 표출지도 및 그래프는 스크린샷으로 다운로드할 수 있도록 개발함

① 분석유형 선택

- 통행 거리, 통행 시간 분석 유형 선택

② 분석지역 설정

- 출도착 설정
- 시도, 시군구, 읍면동 선택

③ 통신타이머 검색조건 설정

- 기간, 요일, 시간대
- 체류특성, 성별, 연령대

④ 분석결과 확인

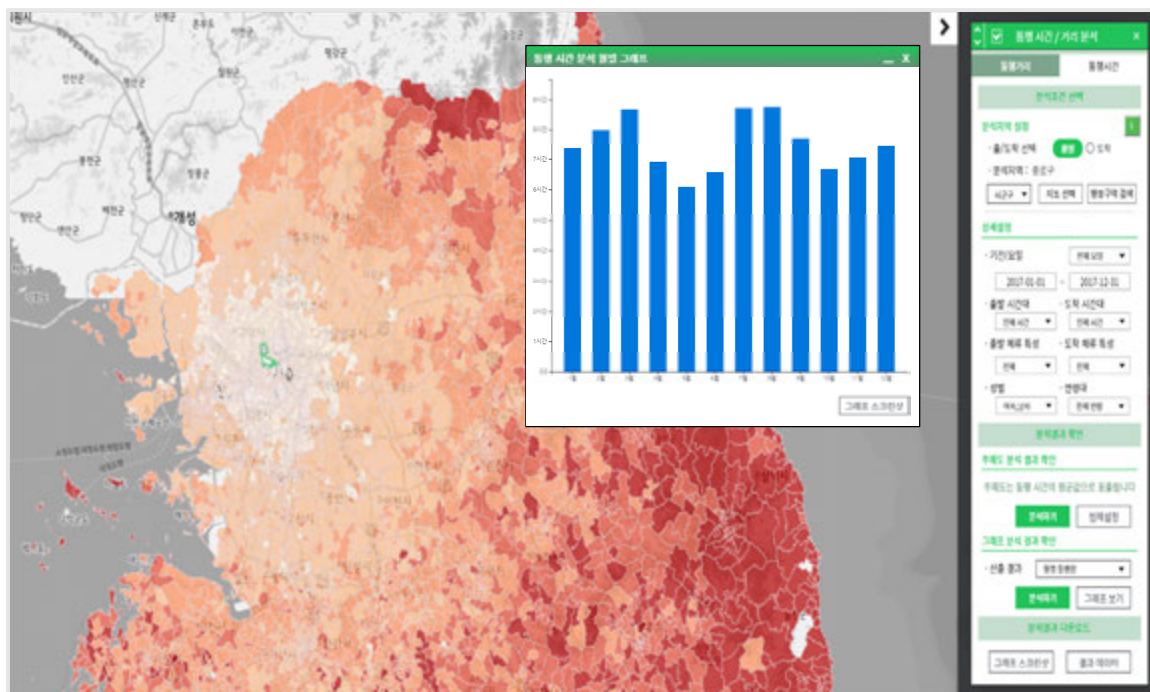
- 주제도 분석, 그래프 분석 개별 구성
- 주제도는 범례 설정 가능
- 그래프는 산출결과를 통해 유형 (일자, 평일/주말, 요일, 월별) 선택 가능

<그림 4- 25> 통행 시간/거리 분석의 설정 UI

※ 통행시간분석 예시

○ 분석 조건

- 분석지역 : 서울특별시 종로구
- 표출단위 : 전국의 교통폴리곤
- 기간 : 2017.01.01. ~ 2017.12.31.
- 요일 : 전체 요일
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 전체 연령대
- 그래프 산출결과 : 월별 통행량

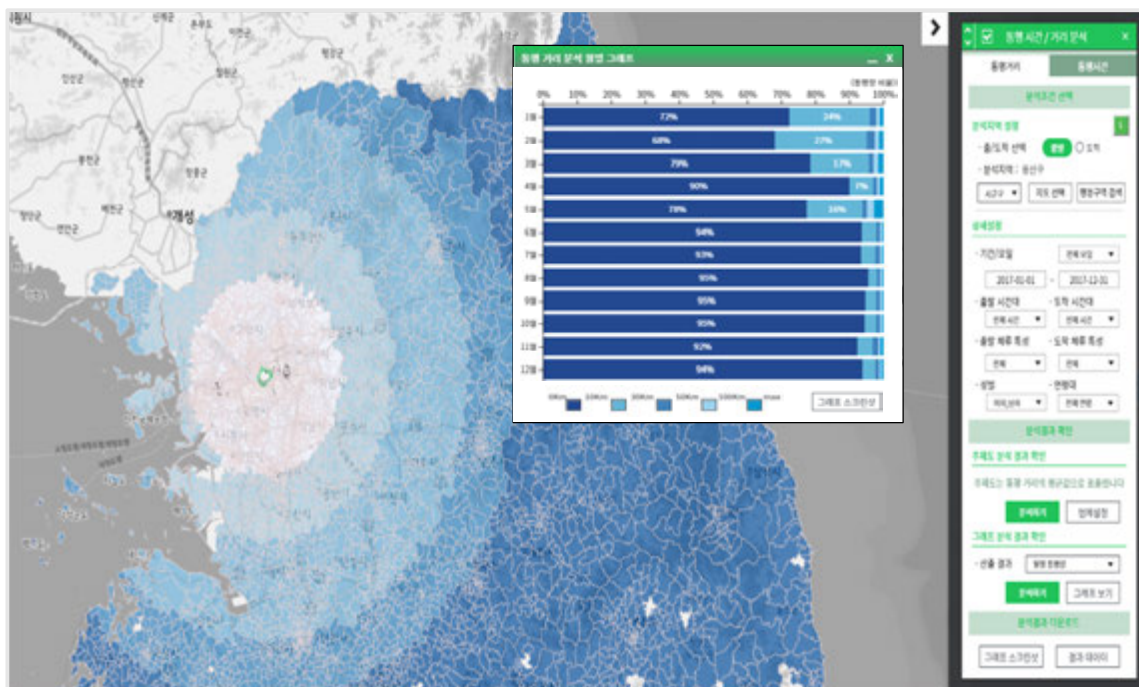


<그림 4- 26> 통행 시간 분석의 주제도·그래프 표출 화면 예시

※ 통행거리분석 예시

○ 분석 조건

- 분석지역 : 서울특별시 용산구
- 표출단위 : 전국의 교통폴리곤
- 기간 : 2017.01.01. ~ 2017.12.31.
- 요일 : 전체 요일
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 전체 연령대
- 그래프 산출결과 : 월별 통행량



<그림 4- 27> 통행 거리 분석의 주제도·그래프 표출 화면 예시

바. 부가 기능

1) 유형별 통신데이터 다운로드

- 사용자가 설정한 데이터 조건과 테이블 형식에 따라 다양한 형태의 통신 집계 데이터를 다운로드 할 수 있도록 개발함
 - 데이터 조건 설정: 사용자가 기준년도, 기간, 요일, 분석지역 공간 범위, 출발 및 도착 시간대, 출발 및 도착 체류특성, 성별, 연령대 설정을 할 수 있도록 함
 - 테이블 형식 설정: 테이블의 열(세로축)과 행(가로축)을 사용자가 정의
 - 세로축과 가로축 설정 옵션: 행정구역별, 월별, 요일별, 일별, 계절별, 출발 시간대, 도착 시간대, 출발 체류특성, 도착 체류특성, 성별, 연령대별

① 데이터 설정

- 기준년도, 기간, 요일 설정
- 분석지역 공간 범위 설정
- 시간대, 체류특성, 성별, 연령대 설정

② 테이블 설정

- 가로축/세로축 별 다양한 집계유형 조건 설정

③ 다운로드

- 데이터 설정과 테이블 설정의 조건으로 분석한 결과를 사용자의 E-Mail로 발송

④ 테이블 설정에 따른 예시

- 사용자가 설정한 조건을 테이블 형태로 예시를 나타냄

<그림 4- 28> 유형별 통신데이터 다운로드 설정 UI

- 분석결과는 사용자의 E-mail로 발송되도록 개발하였으며, 사용자가 E-mail의 다운로드 링크를 클릭하였을 때에만 다운로드할 수 있도록 개발함
 - 다운로드 전 사용자가 설정한 조건을 다시 확인할 수 있도록 메일 내에 설정 사항을 정리하여 안내함
 - 해당 데이터는 7일간 다운로드 가능함

ViewT
ViewT 2.0은 사용자에게 효율성에 대한 새로운 시각을 제공합니다.

님이 다운로드하신 통신 유형별 데이터를 발송하였습니다.
 데이터는 7일간 다운로드 가능 합니다.

설정 사항

기간	20170101 - 20171231
요일	월 / 화 / 수 / 목 / 금 / 토 / 일
분석지역 공간 범위	2900000 (시도 : 세종특별자치시)
출발 시간대	0 / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 15 / 16 / 17 / 18 / 19 / 20 / 21 / 22 / 23 시
도착 시간대	0 / 1 / 2 / 3 / 4 / 5 / 6 / 7 / 8 / 9 / 10 / 11 / 12 / 13 / 14 / 15 / 16 / 17 / 18 / 19 / 20 / 21 / 22 / 23 시
출발 주요 체류 특성	집 / 집 이외 주거류지 / 회사 / 학교 / 종교집회장소 / 기타
도착 주요 체류 특성	집 / 집 이외 주거류지 / 회사 / 학교 / 종교집회장소 / 기타
성별	여자 / 남자
연령대	20 / 30 대
데이터 유형	혼합동행-요일별 데이터

다운로드 링크 클릭

(우)30147 세종특별자치시 시청대로 370 세종국책연구단지 과학프라자 TEL : 044-211-3114 FAX : 044-211-3222
 Copyright© 2016 THE KORLA TRANSPORT INSTITUTE All rights reserved. 이용약관개인정보처리방침

<그림 4- 29> 유형별 통신데이터 다운로드 링크 메일

※ 부가 기능 (유형별 통신데이터 다운로드) 예시

○ 분석 조건

- 기간 : 2017.01.01 ~ 2017.12.31
- 요일 : 전체 요일
- 분석지역 공간 범위 : 세종특별자치시
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 20대, 30대
- 테이블 세로축 : 읍면동
- 테이블 가로축 : 요일별

	A	B	C	D	E	F	G	H
1	읍면동별-요일별 테이블							
2	읍면동별	요일별						
3		일	월	화	수	목	금	토
4	세종특별자치시 세종시 조치원읍	254661	394117	388298	397537	322987	288198	250999
5	세종특별자치시 세종시 연기면	54112	80739	83092	81077	66528	63590	54185
6	세종특별자치시 세종시 연동면	12887	20706	21047	21476	17634	16970	14088
7	세종특별자치시 세종시 부강면	11294	26118	28346	28333	23505	20469	15008
8	세종특별자치시 세종시 금남면	32659	42997	47701	46247	38076	36731	35844
9	세종특별자치시 세종시 장군면	66651	87273	88947	90541	73715	73372	69582
10	세종특별자치시 세종시 연서면	41256	61049	60313	61439	49134	48537	42820
11	세종특별자치시 세종시 전외면	20705	36511	38788	40681	32309	29504	24207
12	세종특별자치시 세종시 전동면	7679	11619	12378	12102	9523	9594	8891
13	세종특별자치시 세종시 소정면	1788	2530	2724	2670	2303	2180	1726
14	세종특별자치시 세종시 한솔동	139278	171686	180799	177202	142880	143095	143176
15	세종특별자치시 세종시 도담동	450421	587165	612120	599051	485927	478576	460749
16	세종특별자치시 세종시 아름동	3115	4156	4147	4441	3150	3079	2636

<그림 4- 30> 사용자가 설정한 조건에 해당하는 유형별 통신 엑셀 데이터

2) 지역 간 OD 통행량 데이터 다운로드

- 지역 간 OD 통행량 분석의 광범위한 자료들을 효과적으로 활용할 수 있도록 OD Matrix 형식의 데이터 다운로드 기능을 개발함
 - － 사용자가 기준 년도, 기간, 요일, 분석지역 단위, 분석지역 공간 범위, 출발 및 도착 시간대, 출발 및 도착 체류 특성, 성별, 연령대 설정을 할 수 있도록 함
 - － 사용자가 분석지역 단위와 분석지역 공간 범위 조건을 변경할 때마다 조건에 맞는 지역 간 OD 테이블 예시를 나타내어 사용자의 편의성을 높임

지역 간 OD 통행량 데이터

데이터 설정

기준년도: 2017년

기간설정: 2017.01.01 ~ 2017.12.31

요일설정: * * * * *

분석지역 단위: ☐ 시/도 ☒ 시/군/구 ☐ 읍/면/동

분석지역 공간 범위: 서울특별시 > 시/군/구 > 읍/면/동

시간대 설정

출발: 전체 시간 도착: 전체 시간

주요 체류 특성 설정

출발: 집, 집 하역, 회사, 학교, 종교/취미/스포츠, 기타

도착: 집, 집 하역, 회사, 학교, 종교/취미/스포츠, 기타

성별 설정: 여자, 남자

연령대 설정: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100

다운로드 다운로드 시 E-Mail로 발송합니다.

데이터 설정에 따른 예시

시/군/구명	서울특별시 용문구	서울특별시 용문구	서울특별시 용문구	서울특별시 용문구	서울특별시 용문구	서울특별시 용문구
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						
서울특별시 용문구						

예시 테이블에는 설정에 해당되어지는 값이 나오지 않습니다. 값을 확인 하시려면 다운로드 해주세요.

<그림 4- 31> 지역 간 OD 통행량 데이터 다운로드 설정 UI

- 분석결과는 사용자의 E-mail로 발송되도록 개발하였으며, 사용자가 E-mail의 다운로드 링크를 클릭하였을 때에만 다운로드할 수 있도록 개발함
 - － 다운로드 전 사용자가 설정한 조건을 다시 확인할 수 있도록 메일 내에 설정 사항을 정리 하여 안내함
 - － 해당 데이터는 7일간 다운로드 가능함

※ 부가 기능 (지역간 OD통행량 데이터 다운로드) 예시

○ 분석조건

- 기간 : 2017.01.01 ~ 2017.12.31
- 요일 : 전체 요일
- 분석지역 단위 : 시군구
- 분석지역 공간 범위 : 서울특별시
- 시간대 : 출발, 도착 전체 시간대
- 체류특성 : 출발, 도착 전체 체류특성
- 성별 : 남, 여
- 연령대 : 20대, 30대, 40대, 50대

	A	B	C	D	E	F
1	서울특별시(시군구별) 데이볼					
2	시군구별	서울특별시 강도구	서울특별시 중구	서울특별시 용산구	서울특별시 성동구	서울특별시 광진구
3	서울특별시 강도구	13205349	1652368	354327	329537	209837
4	서울특별시 중구	1692520	13452269	1034132	990115	260735
5	서울특별시 용산구	335769	1010235	14462558	288706	119384
6	서울특별시 성동구	327316	982931	287352	17233338	1153407
7	서울특별시 광진구	201845	254556	117147	1139663	19349724
8	서울특별시 동대문구	494490	410673	128335	1162364	462053
9	서울특별시 중랑구	169945	193892	75708	214655	544208
10	서울특별시 성북구	1425298	536433	175404	215123	139349
11	서울특별시 강북구	293111	274092	78893	99684	70509
12	서울특별시 도봉구	203883	205449	63709	78063	72582
13	서울특별시 노원구	297724	313305	118245	168426	224922
14	서울특별시 은평구	555390	344776	142484	75840	46008
15	서울특별시 서대문구	841415	629830	228371	108000	83713
16	서울특별시 마포구	481042	552966	1268176	171120	145914

<그림 4- 32> 사용자가 설정한 조건에 해당하는 지역 간 OD 엑셀 데이터

제3절 기반 DB 구축

1. 모바일 분석맵 구축

가. 교통폴리곤 구축

- 2018년에 개발한 교통폴리곤(2016년 집계구와 2017년 기지국 정보를 활용하여 구축한 영역)
 - 지역 : 전국
 - 형상 : 폴리곤
 - 개체수 : 16,335
 - 좌표계 : GRS80 TM 좌표계
 - 용량 : 약 180MB
 - 데이터형식 : Shp

<표 4- 11> 수집된 교통폴리곤 데이터의 속성정보

NO	필드명	설명	데이터 타입	비고
1	OBJECTID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	sido	시도 코드	STRING	-
4	emd	읍면동 코드	STRING	-
5	polycode	교통폴리곤 코드	STRING	-
6	area	교통폴리곤 면적	DOUBLE	-



- 데이터 수집
- 분석구 데이터 (폴리곤)
- 형상정보 가공
- 분석구 폴리곤의 노드 데이터 생성
- 속성정보 가공
- 컬럼명/타입 수정
- 부가정보 추가
- 데이터 변환
- MySQL 기반
- 공정정보 DB 구축
- GeoServer 레이어 구축
- 분석구 폴리곤 레이어
- 분석구 노드 레이어

<그림 4- 33> 교통폴리곤 구축 절차

- 교통폴리곤을 이용 GeoServer, Openlayers 기반의 GIS 형태로 서비스를 구축함

- GeoServer : 공간 Data를 인터넷 GIS 인터페이스로 공급하는 서버 프로그램
- Openlayers : 웹에서 지도데이터를 표현하고 조작할 수 있는 자바스크립트 라이브러리
- 교통폴리곤에 대한 폴리곤 레이어와 노드 레이어를 구축함
 - 폴리곤 레이어: 각 기능의 분석결과에 대한 분포도 생성의 기반 공간정보로 활용
 - 노드 레이어: 교통폴리곤의 대표 포인트를 의미하며, 희망선 표출에 대한 중심점 및 라벨 생성의 기반 공간정보로 활용

1) 교통폴리곤 - 폴리곤 레이어



- 수집된 교통폴리곤 데이터의 컬럼명 및 타입을 다음과 같이 변경하여 MySQL기반의 공간정보로 변환함

<표 4- 12> MySQL 기반의 교통폴리곤 공간정보의 속성

NO	필드명	설명	데이터 타입	비고
1	OGR_FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	poly_id	교통폴리곤 코드	INT	-
4	zone_id	읍면동 코드	INT	-
5	areavalue	교통폴리곤 면적	DOUBLE	-
6	name	읍면동 명칭	STRING	-
7	full_name	상위행정구역 명칭 + 읍면동 명칭	STRING	-

- MySQL 교통폴리곤 공간정보를 활용하여 GeoServer 기반의 교통폴리곤 레이어를 구축함
- 전국은 16,335개의 교통폴리곤으로 구성되어 있으며 분석결과를 교통폴리곤 단위의 지도로 표출할 경우 사용자의 컴퓨터 환경에서 지도 표출이 원활하여야 함
- GeoServer의 VectorTile, WMS 서비스는 지도의 뷰 레벨에 따라 최적화된 형상의 이미지를 제공함 (지도 일반화)

<표 4- 13> VectorTile, WMS의 지도 일반화

구분	일반화 적용 전	일반화 적용 후
교통폴리곤 (Polygon)		
비고	일반화 적용 전에는 곡선 부분 등에서 절점이 중복되거나 가까이 위치하여 데이터의 크기 증가	데이터의 형상은 최대한으로 유지하면서 절점의 개수를 줄여 데이터 크기 감소 및 그리기 속도 향상

- 본 연구에서는 VectorTile, WMS 서비스가 가능한 형태로 교통폴리곤 레이어를 구축함



Scale = 1 : 4M

TMOBILE_TPOLYGON_SHP_2017

fid	poly_id	zone_id	areavalue	lon	lat	name	full_name
TMOBILE_TPOLYGON_SHP_2017.5950	311239	3111056	4579153.87291	127.026597038	37.421207662	문원 동	경기도 과천시 문원동
TMOBILE_TPOLYGON_SHP_2017.6205	311494	3117056	7944749.62323	127.021545839	37.4026728412	청계 동	경기도 의왕시 청계동

<그림 4- 34> GeoServer 기반의 교통폴리곤 레이어

<표 4- 14> GeoServer 교통폴리곤 레이어의 속성정보

NO	필드명	설명	데이터 타입	비고
1	fid	ID	Object ID	-
2	poly_id	교통폴리곤 코드	INT	-
3	zone_id	읍면동 코드	INT	-
4	areavalue	교통폴리곤 면적	DOUBLE	-
5	lon	X 좌표	DOUBLE	-
6	lat	Y 좌표	DOUBLE	-
7	name	읍면동 명칭	STRING	-
8	full_name	상위행정구역 명칭 + 읍면동 명칭	STRING	-
9	geom	형상정보	MULTIPOLYGON	-

2) 교통폴리곤 - 노드 레이어

- 희망선 지도 표출, 분석결과의 라벨표출 시 사용되는 교통폴리곤의 대표 포인트(노드)를 구축함
- 상용 GIS툴(ARCGIS)을 기반으로 교통폴리곤 데이터를 활용하여 노드 레이어를 구축함
- MySQL 기반의 공간정보를 구축하고 이를 활용하여 GeoServer 기반의 교통폴리곤 노드 레이어를 구축함
- VectorTile, WMS 서비스가 가능한 형태로 교통폴리곤 노드 레이어를 구축함

<표 4- 15> GeoServer 교통폴리곤 노드 레이어의 속성정보

NO	필드명	설명	데이터 타입	비고
1	fid	ID	Object ID	-
2	poly_id	교통폴리곤 코드	INT	-
3	zone_id	읍면동 코드	INT	-
4	lon	X 좌표	DOUBLE	-
5	lat	Y 좌표	DOUBLE	-
6	name	읍면동 명칭	STRING	-
7	full_name	상위행정구역 명칭 + 읍면동 명칭	STRING	-
8	geom	형상정보	POINT	-



ZONE_MOBILE_TPOLYGON_SHP_P_2017

fid	poly_id	zone_id	lon	lat	name	full_name
ZONE_MOBILE_TPOLYGON_SHP_P_2017.119	110119	1103059	126.956879642	37.5386464203	영문 영문	서울특별시 용산구 영문동
ZONE_MOBILE_TPOLYGON_SHP_P_2017.820	110820	1114075	126.951620226	37.539140658	도화 도화	서울특별시 마포구 도화동

<그림 4- 35> GeoServer 기반의 교통폴리곤 노드 레이어

나. 행정구역 구축

- 2016년 기준 읍면동, 시군구, 시도 영역
 - 지역 : 전국
 - 형상 : 폴리곤
 - 개체수 : 3,770
 - 좌표계 : GRS 1980 ITRF 2000 UTM K 좌표계
 - 용량 : 약 88.4 MB
 - 데이터형식 : Shp
 - 속성정보

<표 4- 16> 수집된 행정구역 데이터의 속성정보

NO	필드명	설명	데이터 타입	비고
1	OGR_FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	DISTRICT_I	행정구역 코드	STRING	-
4	DISTRICT_N	행정구역 명칭	STRING	-
5	DISTRICT_T	행정구역 구분 코드 (시도 : 2, 시군구 :3, 읍면동 : 4)	STRING	-
6	X_COORDINA	X 좌표	DOUBLE	-
7	Y_COORDINA	Y 좌표	DOUBLE	-
8	AREA	면적	DOUBLE	-

- 행정구역 레이어는 교통폴리곤 레이어 구축과 동일한 절차에 따라 구축하였음
- 시도, 시군구, 읍면동에 대해 GeoServer 형식의 폴리곤 레이어와 노드 레이어를 구축함
 - 폴리곤 레이어: 교통폴리곤의 폴리곤 레이어와 같이 Hotplace분석, 주요통행지역분석 등에서 분석결과에 대한 분포도 생성의 기반 공간정보로 활용
 - 노드 레이어: 희망선 표출에 대한 중심점 및 라벨 생성의 기반 공간정보로 활용

1) 행정구역 - 폴리곤 레이어

- 수집된 행정구역 폴리곤 데이터의 컬럼명 및 타입을 다음과 같이 변경하여 MySQL기반의 공간정보로 변환함(시도, 시군구, 읍면동 동일함)

<표 4- 17> MySQL 기반의 행정구역 폴리곤 공간정보의 속성

NO	필드명	설명	데이터 타입	비고
1	OGR_FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	zone_id	해당 행정구역 코드	INT	-
4	lon	해당 행정구역 X 좌표	DOUBLE	-
5	lat	해당 행정구역 Y 좌표	DOUBLE	-
6	name	해당 행정구역 명칭	STRING	-
7	full_name	상위행정구역 명칭 + 해당 행정구역 명칭	STRING	-
8	areavalue	해당 행정구역 면적	DOUBLE	-

- 교통폴리곤의 폴리곤 레이어와 동일한 절차에 따라 GeoServer의 행정구역 폴리곤 레이어를 구축함(시도, 시군구, 읍면동 GeoServer 폴리곤 레이어의 속성정보 모두 동일함)

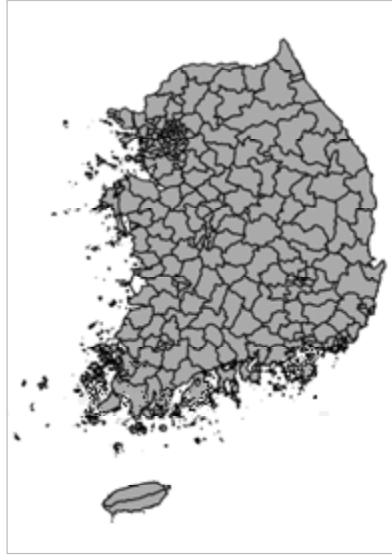
<표 4- 18> GeoServer 행정구역 폴리곤 레이어의 속성정보

NO	필드명	설명	데이터 타입	비고
1	fid	ID	Object ID	-
2	zone_id	해당 행정구역 코드	INT	-
3	name	해당 행정구역 명칭	STRING	-
4	lon	해당 행정구역 X 좌표	DOUBLE	-
5	lat	해당 행정구역 Y 좌표	DOUBLE	-
6	areavalue	해당 행정구역 면적	DOUBLE	-
7	full_name	상위행정구역 명칭 + 해당 행정구역 명칭	STRING	-
8	geom	형상정보	MULTIPOLYGON	-

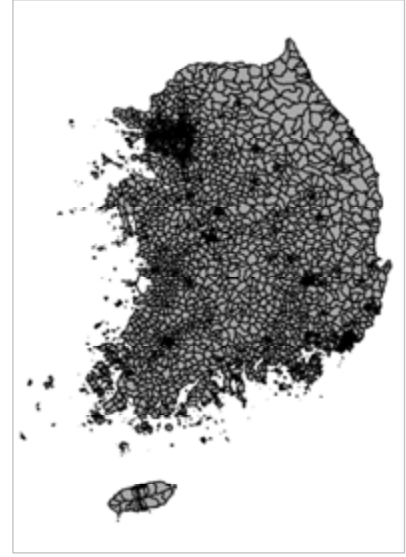
시도 행정구역 레이어



시군구 행정구역 레이어



읍면동 행정구역 레이어



<그림 4- 36> GeoServer 기반의 행정구역 폴리곤 레이어

2) 행정구역 - 노드 레이어

- 희망선 지도표출, 라벨표출 시 사용되는 행정구역의 대표 포인트(노드)를 구축함
(시도, 시군구, 읍면동 동일함)

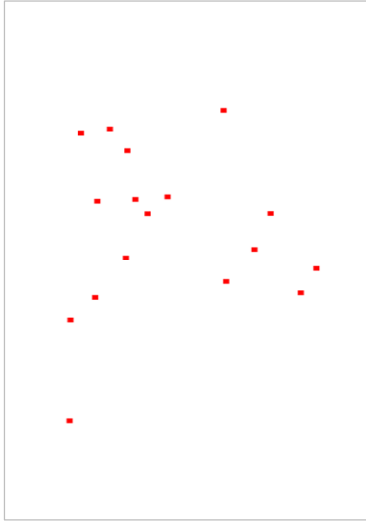
<표 4- 19> MySQL 기반의 행정구역 노드 공간정보의 속성

NO	필드명	설명	데이터 타입	비고
1	OGR_FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	zone_id	해당 행정구역 코드	INT	-
4	lon	해당 행정구역 X 좌표	DOUBLE	-
5	lat	해당 행정구역 Y 좌표	DOUBLE	-
6	name	해당 행정구역 명칭	STRING	-
7	full_name	상위행정구역 명칭 + 해당 행정구역 명칭	STRING	-
8	areavalue	해당 행정구역 면적	DOUBLE	-

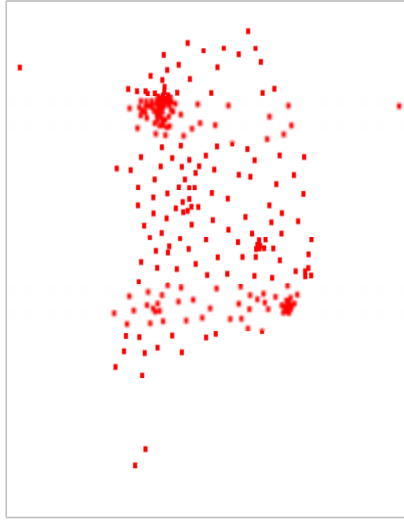
- 교통폴리곤의 노드 레이어와 동일한 절차에 따라 GeoServer의 행정구역 노드 레이어를 구축함

(시도, 시군구, 읍면동 GeoServer 노드 레이어의 속성정보 모두 동일함)

■ 시도 행정구역 노드 레이어



■ 시군구 행정구역 노드 레이어



■ 읍면동 행정구역 노드 레이어



<그림 4- 37> GeoServer 기반의 행정구역 노드 레이어

2. 모바일 빅데이터 기종점 통행량 DB 구축

가. 모바일 빅데이터 기종점 통행량 DB 개요

- 2017년에 생성된 로그 기록을 활용하여 가공한 기종점 통행량 DB 구축
 - 지역 : 전국
 - 기간 : 2017.01.01. ~ 2017.12.31.
 - 용량 : 약 675GB
 - 형태: 일단위 텍스트 파일 (csv 파일)

<표 4- 20> 통신빅데이터의 컬럼명 및 데이터 타입

NO	필드명	설명	데이터 타입
1	o_polygon	출발지 (교통폴리곤 코드)	STRING
2	o_base_ymd	출발일자	STRING
3	o_timezn_cd	출발시간대	STRING
4	o_trip_type	출발체류특성	STRING
5	d_polygon	도착지 (교통폴리곤 코드)	STRING
6	d_base_ymd	도착일자	STRING
7	d_timezn_cd	도착시간대	STRING
8	d_trip_type	도착체류특성	STRING
9	age_itg_cd	연령대	INT
10	sex_trip_itg_cd	성별	STRING
11	total	이동 인구수	INT

```
110002,20170307,00,N,230346,20170307,01,N,040,F,3
110086,20170307,00,C,110084,20170307,00,H,040,M,3
110101,20170307,00,H,110100,20170307,01,R,010,F,3
```

<그림 4- 38> 모바일 빅데이터 기종점 통행량 DB 형태

- 연령별, 성별, 출발 및 도착 시간대, 출발지 및 도착지 체류특성 정보는 코드로 기록되어 있으며 코드 정보는 다음과 같음

<표 4- 21> 모바일 빅데이터 기종점 통행량 DB의 연령별 코드

구분	코드	내용	비고
연령별	0	0세 ~ 9세	-
	10	10세 ~ 19세	-
	20	20세 ~ 29세	-
	30	30세 ~ 39세	-
	40	40세 ~ 49세	-
	50	50세 ~ 59세	-
	60	60세 ~ 69세	-
	70	70세 ~ 79세	-
	80	80세 ~ 89세	-
	90	90세 ~ 99세	-
	100	100세 ~ 109세	-
	110	110세 ~ 119세	-

<표 4- 22> 모바일 빅데이터 기종점 통행량 DB의 성별 코드

구분	코드	내용	비고
성별	F	여자	-
	M	남자	-

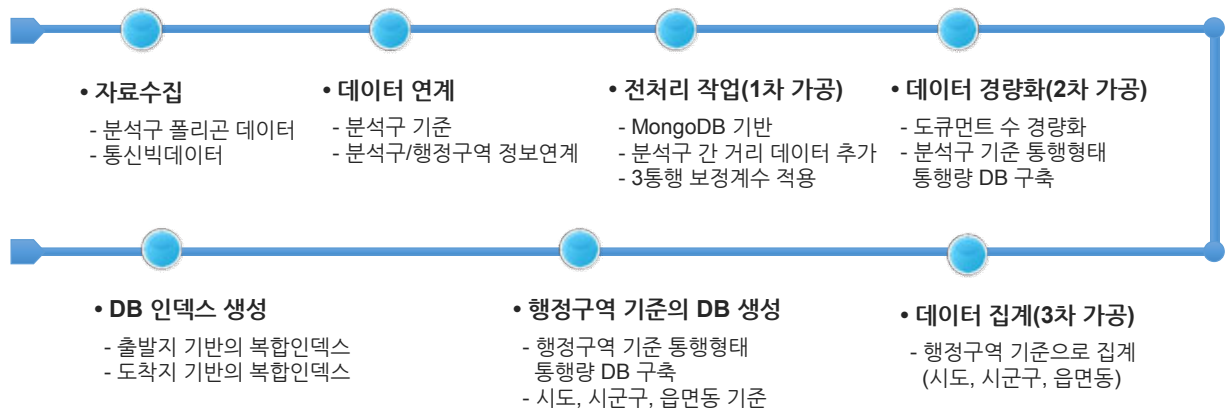
<표 4- 23> 모바일 빅데이터 기종점 통행량 DB의 시간대 코드

구분	코드	내용	비고
시간대별	0	0시 ~ 1시	-
	1	1시 ~ 2시	-
	2	2시 ~ 3시	-
	3	3시 ~ 4시	-
	4	4시 ~ 5시	-
	5	5시 ~ 6시	-
	6	6시 ~ 7시	-
	7	7시 ~ 8시	-
	8	8시 ~ 9시	-
	9	9시 ~ 10시	-
	10	10시 ~ 11시	-
	11	11시 ~ 12시	-
	12	12시 ~ 13시	-
	13	13시 ~ 14시	-
	14	14시 ~ 15시	-
	15	15시 ~ 16시	-
	16	16시 ~ 17시	-
	17	17시 ~ 18시	-
	18	18시 ~ 19시	-
	19	19시 ~ 20시	-
	20	20시 ~ 21시	-
	21	21시 ~ 22시	-
	22	22시 ~ 23시	-
	23	23시 ~ 0시	-

<표 4- 24> 모바일 빅데이터 기종점 통행량 DB의 체류특성별 코드

구분	코드	내용	비고
체류특성별	H	집	-
	N	집 이외 주체류지	-
	C	회사	-
	S	학교	-
	R	종교집회장소	-
	X	기타	-

나. 구축 절차



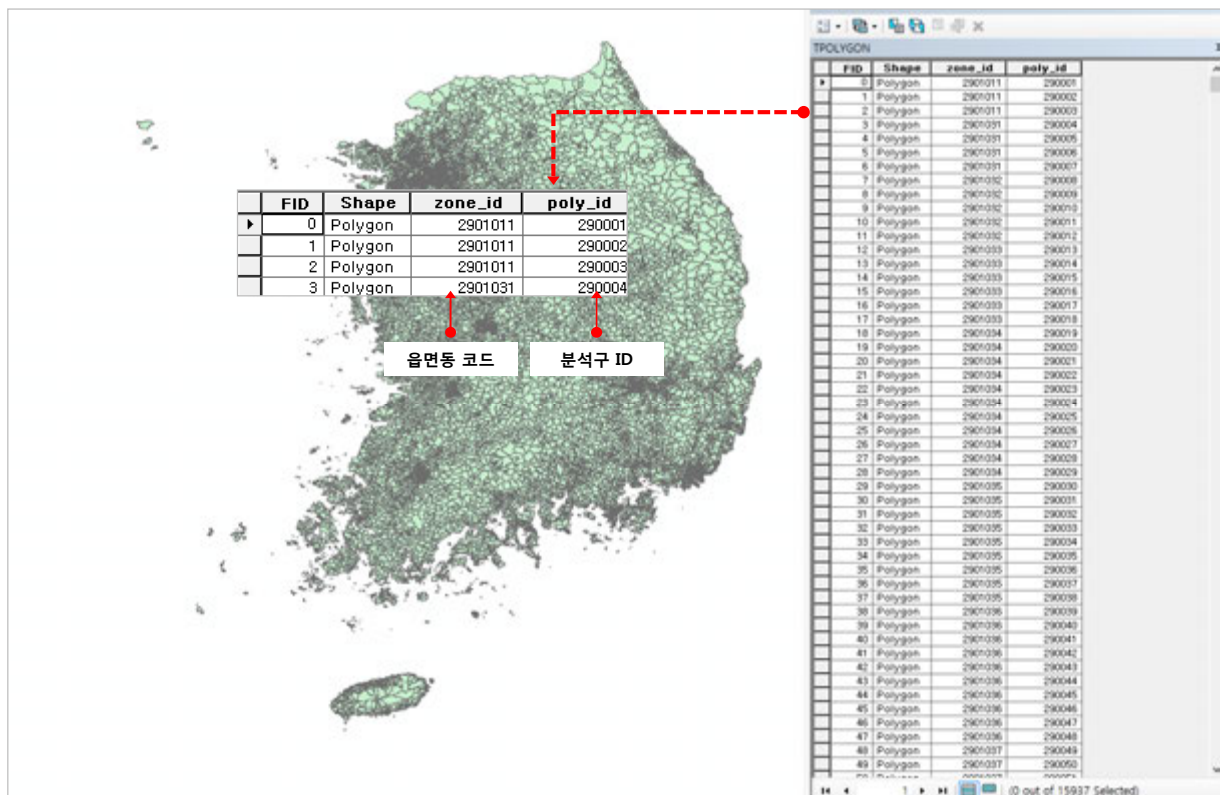
<그림 4- 39> 모바일 빅데이터 기종점 통행량 DB의 구축 절차

- 데이터의 검색 및 처리속도, 빅데이터 관리의 원활함을 위해서 MongoDB 기반의 빅데이터 DB를 구축함
- 데이터의 건수(도큐먼트 수)가 매우 커서 가공 없이 DB를 구축할 경우 데이터 검색 및 처리속도의 저하로 정상적인 서비스를 제공하는 데 한계가 있음
 - 1년 데이터 건수 : 약 140억 건(용량 : 약 675GB)
 - (1일 데이터 건수 : 약 4,200만 건(용량 : 약 2GB))
- 따라서 데이터의 도큐먼트 수를 경량화시킨 교통폴리곤 단위의 모바일 통행형태 기반 DB를 구축하여 정상적인 서비스를 제공할 수 있도록 함

- 이동 인구수의 분석을 행정구역 단위로 하는 경우 행정구역에 포함되는 교통폴리곤의 이동 인구수를 검색하고 처리하는데 많은 시간이 소요됨
- 출발지 또는 도착지가 행정구역(시도, 시군구, 읍면동)인 경우 서비스 속도향상을 위하여 행정구역 기준의 모바일 통행형태 기반 DB를 구축함
- 입력된 DB에 대한 인덱스를 생성하여 검색속도를 향상시킴
 - 출발지 및 도착지 정보를 활용한 복합인덱스를 생성함(시도, 시군구, 읍면동, 교통폴리곤 ID)

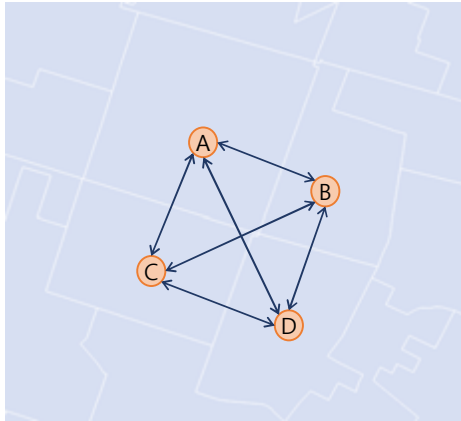
1) 전처리 작업

- 수집된 통신 빅데이터에 행정구역 코드 추가, 교통폴리곤간 거리 추가, 보정계수 적용 및 문자형 타입 데이터를 숫자형으로 변환을 하여 MongoDB 입력 파일 형태로 데이터를 1차 가공함
- 출발지, 도착지 교통폴리곤에 대한 시도, 시군구, 읍면동 행정구역 코드를 추가함
- 교통폴리곤과 행정구역 매칭정보는 수집된 교통폴리곤 공간정보의 속성정보를 활용함



<그림 4- 40> 교통폴리곤과 읍면동 행정구역 코드의 매칭정보

- 통행 거리에 따른 데이터 분석 기능을 구현하기 위해 출·도착 교통폴리곤 간의 거리 데이터를 추가함
- 교통폴리곤의 노드 공간정보에서 대표 포인트(노드)의 X 좌표와 Y 좌표를 참조하여 통신 빅데이터의 출발 교통폴리곤과 도착 교통폴리곤 간의 직선거리(Km)를 생성하여 추가함



출발 분석구	도착 분석구	거리(Km)
A	B	10
A	C	15
A	D	20
B	C	25
B	D	20
C	D	15

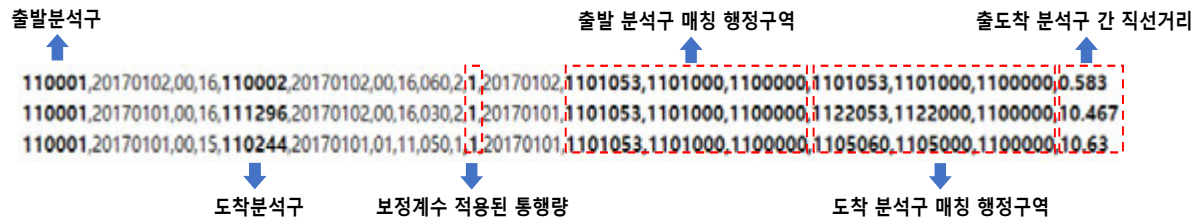
<그림 4- 41> 출·도착 교통폴리곤 간 거리 산출 예시

- 통신 빅데이터는 개인정보보호법과 위치 정보의 보호 및 이동 등에 관한 법률로 인해 최소 이동 인원수가 3인으로 생성됨
- 실제 통행과 유사한 데이터를 생성하기 위해 3인 통행 데이터에 3통행 보정계수를 적용하여 1인, 2인 통행에 대한 데이터를 생성함
- 3통행 보정계수에는 시도 코드, 출발 일자의 요일, 출발 시간대, 성별, 연령대, 출발 체류특성의 조건을 가지고 있음

시도 코드	요일 코드	시간대	성별	연령대	체류특성	보정계수
11	1	1	M	40	C	0.336612
11	2	1	M	30	X	0.360702
11	3	0	F	80	H	0.335057
11	4	0	F	0	X	0.347685
11	5	0	M	80	N	0.333333
11	6	0	M	30	R	0.348798
11	7	1	F	80	R	0.333333

<그림 4- 42> 3통행 보정계수 데이터 예시

- 가공된 MongoDB 입력 파일 형태는 다음과 같음



<그림 4- 43> 1차 가공데이터의 MongoDB 입력데이터 예시

2) 교통폴리곤 단위의 통행량 DB 구축

- 교통폴리곤 단위의 통행량 DB는 서비스의 속도를 결정하는 DB이며 행정구역 단위의 모바일 통행형태 통행량 DB를 구축하는데 활용되는 기반 DB임
- 전처리 작업 후 경량화 작업을 거쳐 최종적으로 DB를 구축하며, 데이터 경량화 작업은 다음과 같이 이루어짐
 - object 타입으로 ID를 정의하고 ID를 기준으로 1차 가공 DB를 재분류함
 - ID : 출발지코드(교통폴리곤 ID), 도착지코드(교통폴리곤 ID), 성별, 연령대, 출발 체류특성, 도착 체류특성의 복합정보
 - ID별로 출발 및 도착 시간대 인구수를 출/도착 일자별로 구분하여 DB를 구축함
- 행정구역 기준의 통행량 DB 구축 시 활용되는 행정구역정보와 교통폴리곤 간 거리 데이터를 추가하여 교통폴리곤 단위의 모바일 통행형태 통행량 DB를 구축함
- 컬렉션(MongoDB의 테이블)의 구성은 다음과 같음

<표 4- 25> 교통폴리곤 단위의 모바일 통행형태 통행량 컬렉션의 구성

컬렉션명	TMOBILE_2017			
컬렉션 설명	교통폴리곤 단위의 모바일 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A” : 출발교통폴리곤코드, “E” : 도착교통폴리곤코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A, E, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : { { “C” : 출발시간대, “G” : 도착시간대, “K” : { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, } “B”:출발날짜, “F”:도착날짜, “V”:인구수, } } }	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	“A1” : 출발읍면동코드	
도착읍면동	E1	INT	“E1” : 도착읍면동코드	
출발시군구	A2	INT	“A2” : 출발시군구코드	
도착시군구	E2	INT	“E2” : 도착시군구코드	
출발시도	A3	INT	“A3” : 출발시도코드	
도착시도	E3	INT	“E3” : 도착시도코드	
교통폴리곤 간 거리	S	DOUBLE	“S” : 교통폴리곤 간 거리	

- 경량화의 결과는 다음과 같음

<표 4- 26> 통신빅데이터의 문서 수의 경량화 결과

구분	수집데이터	경량화 (교통폴리곤 단위의 통행형태 통행량)	감소율
문서 수	약 140억 건	약 7억 건	약 95% 감소

3) 행정구역 단위의 모바일 통행형태 통행량 DB 구축

- 교통폴리곤 단위의 모바일 통행형태 통행량 DB를 기반으로 행정구역 ↔ 행정구역, 행정구역 ↔ 교통폴리곤의 통행량을 분석할 경우 행정구역에 포함되는 교통폴리곤의 이동 인구수를 검색하고 처리하는데 많은 시간이 소요됨
- 행정구역 ↔ 행정구역, 행정구역 ↔ 교통폴리곤의 통행량 분석을 위해서는 교통폴리곤 단위의 모바일 통행형태 통행량 DB의 문서 수를 경량화시킨 별도의 통행형태 통행량 DB가 필요함
- 따라서 행정구역 기준의 모바일 통행형태 통행량 DB를 추가로 구축함
 - 시도, 시군구, 읍면동에서 출발하는 통행량을 분석하기 위한 DB를 구축함
 - 시도, 시군구, 읍면동에서 도착하는 통행량을 분석하기 위한 DB를 구축함
- 출발지 및 도착지의 행정구역 단위로 인구수를 각각 집계하여 MongoDB 기반의 행정구역 기준 통행형태 통행량 DB를 구축함

① 행정구역 기준의 통행형태 통행량 컬렉션의 구성

<표 4- 27> 읍면동 출발 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_UMD_POLY_2017			
컬렉션 설명	읍면동 출발 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : {	- A1, E, D, H, J, I : integer

			“A1” : 출발읍면동코드, “E” : 도착교통폴리곤코드, “J” : 성별, “I” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : { { “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }] } }	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인구수
도착읍면동	E1	INT	“E1” : 도착읍면동코드	
출발시군구	A2	INT	“A2” : 출발시군구코드	
도착시군구	E2	INT	“E2” : 도착시군구코드	
출발시도	A3	INT	“A3” : 출발시도코드	
도착시도	E3	INT	“E3” : 도착시도코드	

<표 4- 28> 읍면동 도착 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_POLY_UMD_2017			
컬렉션 설명	읍면동 도착 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A” : 출발교통폴리곤코드, “E1” : 도착읍면동코드, }	- A, E1, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13,

			“J” : 성별, “I” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : { { “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }] } }	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	“A1” : 도착읍면동코드	
출발시군구	A2	INT	“A2” : 출발시군구코드	
도착시군구	E2	INT	“E2” : 도착시군구코드	
출발시도	A3	INT	“A3” : 출발시도코드	
도착시도	E3	INT	“E3” : 도착시도코드	

<표 4- 29> 시군구 출발 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_SGG_POLY_2017			
컬렉션 설명	시군구 출발 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A2” : 출발시군구코드, “E” : 도착교통폴리곤코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A2, E, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : [{ “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }]	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인구수
도착읍면동	E1	INT	“E1” : 도착읍면동코드	
도착시군구	E2	INT	“E2” : 도착시군구코드	
출발시도	A3	INT	“A3” : 출발시도코드	
도착시도	E3	INT	“E3” : 도착시도코드	

<표 4- 30> 시군구 도착 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_POLY_SGG_2017			
컬렉션 설명	시군구 도착 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A” : 출발교통폴리곤코드, “E2” : 도착교통폴리곤코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A, E2, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : [{ “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }] }]	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인 구수
출발읍면동	A1	INT	“A1” : 도착읍면동코드	
출발시군구	A2	INT	“A2” : 출발시군구코드	
출발시도	A3	INT	“A3” : 출발시도코드	
도착시도	E3	INT	“E3” : 도착시도코드	

<표 4- 31> 시도 출발 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_SIDO_POLY_2017			
컬렉션 설명	시도 출발 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A3” : 출발시도코드, “E” : 도착교통폴리곤코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A3, E, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : [{ “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }] }] }	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인구수
도착읍면동	E1	INT	“E1” : 도착읍면동코드	
도착시군구	E2	INT	“E2” : 도착시군구코드	
도착시도	E3	INT	“E3” : 도착시도코드	

<표 4- 32> 시도 도착 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_POLY_SIDO_2017			
컬렉션 설명	시도 도착 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A” : 출발교통폴리곤코드, “E3” : 도착시도코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A, E3, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : [{ “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }] }]	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	“A1” : 도착읍면동코드	
출발시군구	A2	INT	“A2” : 출발시군구코드	
출발시도	A3	INT	“A3” : 출발시도코드	

<표 4- 33> 읍면동 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_UMD_2017			
컬렉션 설명	기준 출도착 행정구역이 읍면동인 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	<pre> _id : { "A1" : 출발읍면동코드, "E1" : 도착읍면동코드, "J" : 성별, "I" : 연령대, "D" : 출발체류특성, "H" : 도착체류특성 } </pre>	<ul style="list-style-type: none"> - A1, E1, D, H, J, I : integer - 체류특성 코드 <ul style="list-style-type: none"> ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 <ul style="list-style-type: none"> ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	<pre> CG : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "B":출발날짜, "F":도착날짜, "V":인구수, }, { "B":출발날짜, "F":도착날짜, "V":인구수, }] }] </pre>	<ul style="list-style-type: none"> - C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당 하는 원시데이터의 인구수
출발시군구	A2	INT	"A2" : 출발시군구코드	
도착시군구	E2	INT	"E2" : 도착시군구코드	
출발시도	A3	INT	"A3" : 출발시도코드	
도착시도	E3	INT	"E3" : 도착시도코드	

<표 4- 34> 시군구 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_SGG_2017			
컬렉션 설명	기준 출·도착 행정구역이 시군구인 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A2” : 출발시군구코드, “E2” : 도착시군구코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A2, E2, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : [{ “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }] }]	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당 하는 원시데이터의 인구수
출발시도	A3	INT	“A3” : 출발시도코드	
도착시도	E3	INT	“E3” : 도착시도코드	

<표 4- 35> 시도 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_SIDO_2017			
컬렉션 설명	기준 출·도착 행정구역이 시도인 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	_id	Object	_id : { “A3” : 출발시도코드, “E3” : 도착시도코드, “J” : 성별, “T” : 연령대, “D” : 출발체류특성, “H” : 도착체류특성 }	- A3, E3, D, H, J, I : integer - 체류특성 코드 ·H:11, N:12, C:13, S:14, R:15, X: 16 - 성별 코드 ·F:1, M:2
출·도착시간대 인구수	CG	Object 배열	CG : [{ “C” : 출발시간대, “G” : 도착시간대, “K” : [{ “B”:출발날짜, “F”:도착날짜, “V”:인구수, }, { “B”:출발날짜, “F”:도착날짜, “V”:인구수, }]	- C, G, B, F, V : integer - K : object 배열 - V : _id, C, G, B, F에 해당 하는 원시데이터의 인구수

② 행정구역 기준의 통행형태 통행량의 경량화 결과

<표 4-36> 읍면동 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
약 7억 건	약 4억 건	약 50% 감소

<표 4-37> 읍면동 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	읍면동 도착 기준의 통행형태 통행량	감소율
약 7억 건	약 4억 건	약 50% 감소

<표 4-38> 시군구 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	시군구 출발 기준의 통행형태 통행량	감소율
약 7억 건	약 2억 건	약 70% 감소

<표 4-39> 시군구 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	시군구 도착 기준의 통행형태 통행량	감소율
약 7억 건	약 2억 건	약 70% 감소

<표 4-40> 시도 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	시도 출발 기준의 통행형태 통행량	감소율
약 7억 건	약 3,000 만 건	약 95% 감소

<표 4-41> 시도 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	시도 도착 기준의 통행형태 통행량	감소율
약 7억 건	약 3,000 만 건	약 95% 감소

<표 4-42> 읍면동 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	읍면동 기준의 통행형태 통행량	감소율
약 7억 건	약 3억 건	약 40% 감소

<표 4- 43> 시군구 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	시군구 기준의 통행형태 통행량	감소율
약 7억 건	약 1억 건	약 80% 감소

<표 4- 44> 시도 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

교통폴리곤 단위의 통행형태 통행량	시도 기준의 통행형태 통행량	감소율
약 7억 건	약 2,000 만 건	약 97% 감소

3. Sharding 기반의 DB 구축

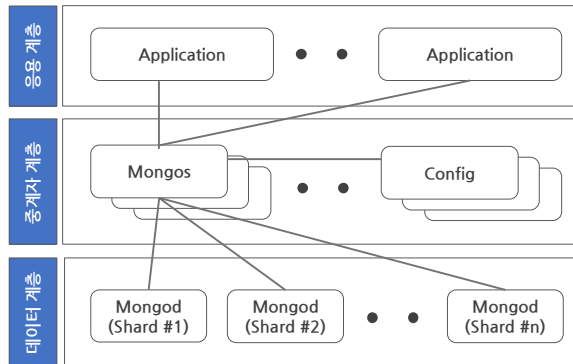
가. Sharding

- Sharding이란 데이터 셋이 단일 데이터베이스에 저장하기에 매우 큰 경우, 데이터 셋을 다수의 데이터베이스에 분산 저장하는 것을 말함
- 본 과업에서는 MongoDB의 Sharding 기법을 적용하여 모바일 통행형태 기반 DB를 구성함

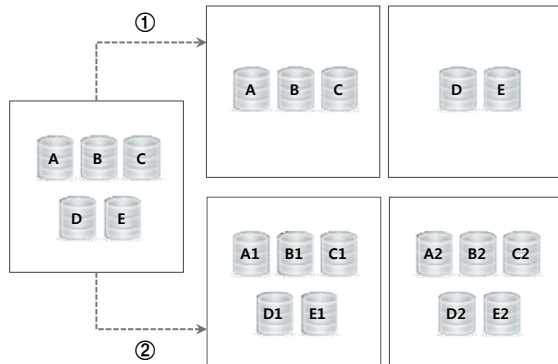
나. Sharding의 구성

- Config 서버, Mongos 서버, Shard 서버로 Sharding을 구성함
 - Config 서버 : 데이터가 어느 Shard에 저장되어 있는지 등의 Shard Meta 정보 저장
 - Mongos 서버 : 중계자 역할 수행, 응용계층의 질의와 데이터 계층의 Shard 서버의 응답을 중계하는 역할을 수행함
 - Shard 서버 : 실제 데이터가 저장되는 서버
- 데이터 셋의 분산저장 방식은 테이블 단위로 분리하는 방법(①)과 테이블 자체를 분할 하는 방법(②)이 있음
- 본 연구에서는 테이블 자체를 분할 하는 방법(②)을 적용함
 - 데이터가 한쪽으로 몰리는 단점이 있는 분리방법(①)을 보완한 방법
 - Shard Key 필드의 값에 해쉬 함수를 적용하여 데이터를 균등하게 배분
 - 다음과 같이 복합정보로 구성된 _id를 Shard Key로 설정함
 - 출발지 정보(시도, 시군구, 읍면동, 교통폴리곤)
 - 도착지 정보(시도, 시군구, 읍면동, 교통폴리곤)
 - 성별 정보
 - 연령대 정보
 - 출발체류특성 정보
 - 도착체류특성 정보

■ MongoDB의 Sharding 서버 구성



■ MongoDB의 분산저장 방식



<그림 4- 44> MongoDB Sharding의 구성

다. 모바일 통행형태 기반 DB의 분산저장 결과

■ 교통플리곤 단위의 통행형태 통행량 DB의 분산저장 결과

```
data : 442.67GiB docs : 480983056 chunks : 7868
Shard shard0000 contains 8.82% data, 8.83% docs in cluster, avg obj size on shard : 987B
Shard shard0001 contains 7.73% data, 7.74% docs in cluster, avg obj size on shard : 987B
Shard shard0002 contains 7.23% data, 7.22% docs in cluster, avg obj size on shard : 988B
Shard shard0003 contains 8.75% data, 8.75% docs in cluster, avg obj size on shard : 987B
Shard shard0004 contains 8.53% data, 8.52% docs in cluster, avg obj size on shard : 989B
Shard shard0005 contains 8.84% data, 8.83% docs in cluster, avg obj size on shard : 988B
Shard shard0006 contains 8.4% data, 8.4% docs in cluster, avg obj size on shard : 987B
Shard shard0007 contains 8.69% data, 8.69% docs in cluster, avg obj size on shard : 988B
Shard shard0008 contains 8.66% data, 8.65% docs in cluster, avg obj size on shard : 989B
Shard shard0009 contains 8.1% data, 8.09% docs in cluster, avg obj size on shard : 988B
Shard shard0010 contains 8.17% data, 8.17% docs in cluster, avg obj size on shard : 988B
Shard shard0011 contains 8.03% data, 8.04% docs in cluster, avg obj size on shard : 986B
```

■ 읍면동 출발 기준의 통행형태 통행량 DB의 분산저장 결과

```
data : 347.87GiB docs : 278093705 chunks : 5035
Shard shard0000 contains 8.53% data, 8.53% docs in cluster, avg obj size on shard : 1KiB
Shard shard0001 contains 8.36% data, 8.37% docs in cluster, avg obj size on shard : 1KiB
Shard shard0002 contains 8.07% data, 8.06% docs in cluster, avg obj size on shard : 1KiB
Shard shard0003 contains 7.93% data, 7.92% docs in cluster, avg obj size on shard : 1KiB
Shard shard0004 contains 7.76% data, 7.76% docs in cluster, avg obj size on shard : 1KiB
Shard shard0005 contains 8.5% data, 8.5% docs in cluster, avg obj size on shard : 1KiB
Shard shard0006 contains 8.83% data, 8.84% docs in cluster, avg obj size on shard : 1KiB
Shard shard0007 contains 8.21% data, 8.2% docs in cluster, avg obj size on shard : 1KiB
Shard shard0008 contains 8.47% data, 8.47% docs in cluster, avg obj size on shard : 1KiB
Shard shard0009 contains 8.83% data, 8.83% docs in cluster, avg obj size on shard : 1KiB
Shard shard0010 contains 7.86% data, 7.88% docs in cluster, avg obj size on shard : 1KiB
Shard shard0011 contains 8.59% data, 8.58% docs in cluster, avg obj size on shard : 1KiB
```

<그림 4- 45> 모바일 통행형태 기반 DB의 Sharding 구성 결과

- 12개의 Shard 서버에 분산 저장함
- 서버당 평균 약 8%의 데이터를 분산 저장함

제4절 운영 및 유지보수

1. 시스템 운영 및 관리

가. 시스템 운영

1) Web Server

① 설치 및 설정

- Web Server는 Apache HTTP Server를 사용하며 설치 정보는 다음과 같음

<표 4- 45> Web Server 설치 정보

구분	설명
제품	Apache HTTP Server
버전	2.2
설치경로	D:\Apache2.2

- Apache HTTP Server의 설정 정보는 D:\Apache2.2\conf\httpd 파일에 저장되어 있으며 특히 Tomcat과의 연동을 위해 Tomcat Connector(mod_jk)에 대한 설정이 추가돼있음

② 실행 및 중지

- Apache HTTP Server는 윈도우즈 서비스로 등록되어 시스템 부팅시 자동으로 실행되나 프로그램의 오류 등으로 중지되어 부득이하게 수동으로 실행시에는 수동 실행해야 함
 - 윈도우즈 서비스에서 실행
 1. 윈도우즈 제어판 -> 관리도구 -> 서비스를 실행
 2. 서비스 목록 중 Apache2.2를 선택한 후 시작 명령을 클릭
 - Apache Service Monitor에서 실행
 1. 윈도우즈 작업표시줄에서 Apache Service Monitor 아이콘을 더블클릭
 2. Service 리스트에서 대상 서비스명을 선택하고 우측의 Start 버튼을 클릭

- Web Server를 중지하는 절차

- 윈도우즈 서비스 이용

1. 윈도우즈 제어판 -> 관리도구 -> 서비스를 실행
2. 서비스 목록 중 Apache2.2를 선택한 후 중지 명령을 클릭

- Apache Service Monitor 이용

1. 윈도우즈 작업표시줄에서 Apache Service Monitor 아이콘을 더블클릭
2. Service 리스트에서 대상 서비스명을 선택하고 우측의 Stop 버튼을 클릭

2) WAS(Web Application Server)

① 설치 및 설정

- Web Application Server는 Apache Tomcat을 사용하고 원활한 서비스를 위하여 다수의 프로그램을 설치하였으며 설치 정보는 다음과 같음

<표 4- 46> Web Server 설치 정보(Apache Tomcat)

구분	설명
제품	Apache Tomcat
버전	7.0.82
설치경로	apache-tomcat-7.0.82
JDK 버전	JDK 1.8.0.181

② 실행 및 중지

- Apache Tomcat는 윈도우즈 서비스로 등록되어 시스템 부팅시 자동으로 실행되나 프로그램의 오류 등으로 중지되어 부득이하게 수동으로 실행시에는 수동 실행해야 함
- 윈도우즈 서비스에서 실행
 1. 윈도우즈 제어판 -> 관리도구 -> 서비스를 실행
 2. 서비스 목록 중 tomcat을 선택한 후 시작 명령을 클릭
- Apache Tomcat를 중지하는 절차

– 윈도우즈 서비스에서 이용

1. 윈도우즈 제어판 -> 관리도구 -> 서비스를 실행
2. 서비스 목록 중 tomcat을 선택한 후 중지 명령을 클릭

③ 웹 어플리케이션 정보

- Apache Tomcat에서 웹 서비스가 설치되는 경로는 [Tomcat 설치경로]\webapps 디렉토리의 아래에 설치되며 본 사업에서 설치된 웹 애플리케이션은 다음과 같음
 - ViewT 서비스 : [Tomcat 설치경로]\webapps\cong
 - GeoServer : [Tomcat 설치경로]\webapps\geoserver

나. 백업 및 복구

- 시스템 운영 중 발생할 수 있는 장애에 대응하기 위해 평시 백업과 복구에 대한 절차 필요

1) 백업 대상 및 방법

- 데이터베이스 : DB데이터 파일, 로그파일

<표 4- 47> 데이터베이스 백업 대상 파일

구분	대상 디렉토리 / 파일
DBMS 데이터 파일	시스템 데이터 파일 : C:\ProgramData\MySQL Server 5.7\Data
DBMS 로그 파일	시스템 로그 : C:\ProgramData\MySQL\MySQL Server 5.7\Data
Big Data의 데이터 파일	몽고 DB 데이터 파일 : D:\mongodb\shard

- 대상 파일을 별도의 드라이브로 복사하여 백업하거나 DB관리 툴(ex: Microsoft SQL Server management Studio 이용)의 백업 기능을 이용하여 백업을 실행함
- 응용프로그램 파일 : 프로그램 실행 파일, 환경설정 파일
 - 시스템의 응용프로그램 백업 대상은 Apache HTTP Server와 Apache Tomcat Server를 대상으로 함

<표 4- 48> 응용프로그램 백업 대상 파일

구분	대상 디렉토리 / 파일
Apache HTTP Server (Web Server)	D:\Apache2.2 아래의 모든 파일
Apache Tomcat (Web Application Server)	D:\apache-tomcat-7.0.82 모든 파일
MongoDB (Big Data)	D:\mongodb\sherd 아래의 파일 중 데이터를 제외한 파일

2) 백업 및 복구 절차

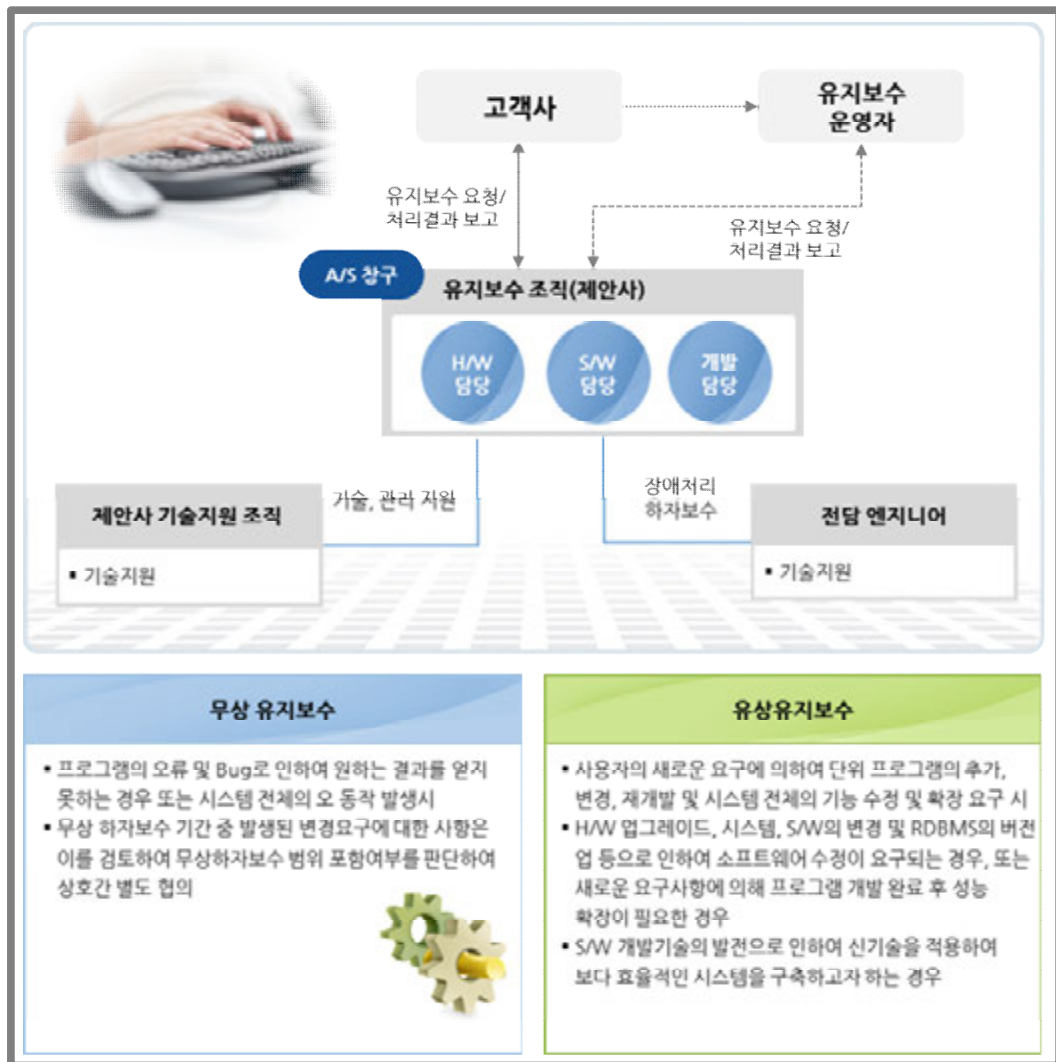
<표 4- 49> 백업 및 복구 절차

구분	수행단계	설명
백업	백업 수행 및 모니터링	백업 담당자는 계획된 백업에 대한 직접 또는 스케줄링 된 백업에 대한 모니터링 실시
	백업실패 조치	백업 담당자는 백업 중 발생한 에러 유형을 확인하고 조치
	백업 재수행	백업 중 발생한 에러를 해결한 후 백업을 재수행
	백업관리대장 반영	백업 담당자는 백업 결과를 관리대장에 반영하고, 백업 중 에러가 발생한 경우 해당 내용에 대한 조치 경과를 기록
	백업 종료	백업 종료 후 백업 결과를 확인하고 현황을 관련자에게 보고
복구	복구요청 접수	데이터의 복구가 필요한 경우 백업 담당자에게 복구요청을 신청
	복구계획 수립	백업 담당자는 복구요청을 분석하고 복구계획을 수립
	복구계획 승인	백업 관리자는 백업 담당자가 수립한 복구계획을 확인하고 이상이 없으면 복구를 승인
	복구 실행	백업 담당자는 복구계획에 따라 백업 데이터를 복구함
	복구 검증	백업 데이터에 대한 검증과 응용프로그램 실행의 이상 유무를 확인하고 결과를 관련자에게 통보
	후속 조치	백업 담당자 및 관리자는 검증결과에 이상이 있을 경우 복구나 백업장비 교체 등의 후속 조치를 수행

2. 유지보수

가. 유지보수 범위 및 조직

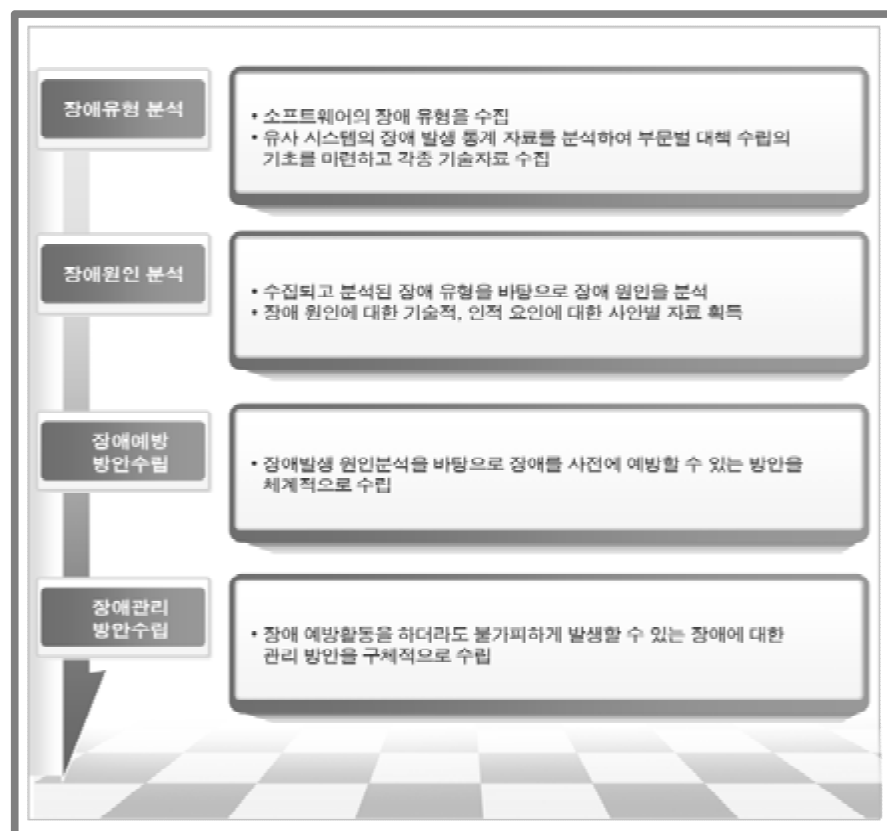
- 본 사업 완료 후 유지보수 기간동안 도입 시스템의 안정적인고 효율적인 시스템 운영, 신속한 장애처리 서비스를 위하여 개발 참여 인력들로 운영지원팀을 구성하여 안정적인 유지보수 서비스를 제공함
- 장애 발생에 대한 신속한 대응과 문제에 대한 근본적인 원인을 종합적으로 분석하여 장애 재발을 최소화할 수 있도록 고객사와 대화창구 단일화를 위한 유지보수총괄조직을 구성하고, 각 분야별 유지보수 팀을 운영하여 문제 발생 시 신속히 대처함



<그림 4- 46> 유지보수 범위 및 조직

나. 장애관리

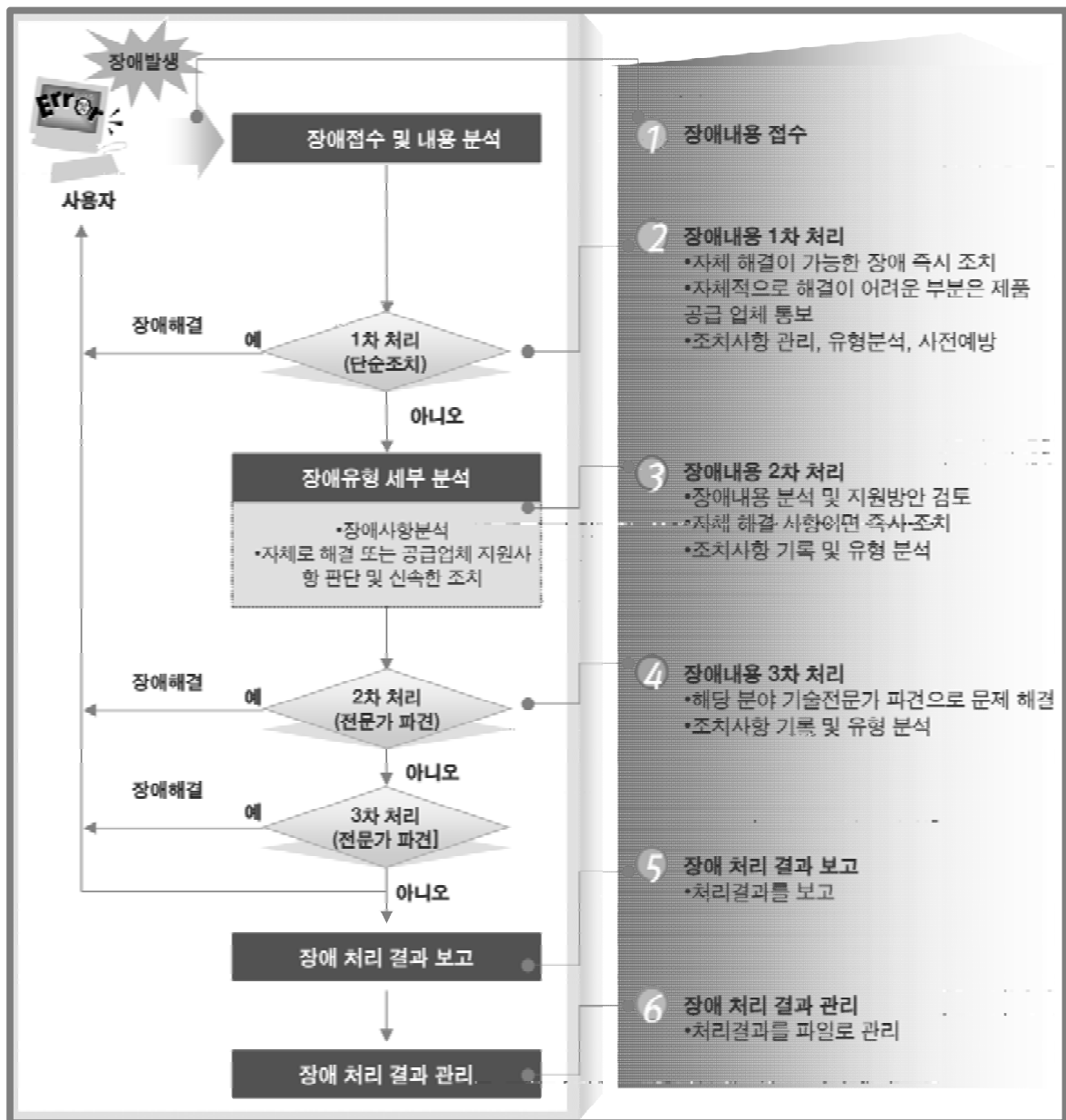
- 시스템의 원활한 서비스와 효율적인 운영을 위해서는 시스템의 안정성, 신뢰성, 유연성 및 장애 대책이 확보되어야 함. 따라서 각 시스템 자원에 대한 원활한 운영 및 장애 발생에 대한 조속한 대응이 가능하도록 중단 없는 서비스, 신속한 시스템 복구, 시스템 장애의 Zero화를 장애 대책의 최우선 목표로 함
- 예상되는 데이터 오류 등의 장애 요인을 파악하고 대처 방안을 마련하여 즉각적인 조치가 가능하도록 함
- 사전에 장애 및 오류 사항에 대해 원인을 분석하여 해당 장애 발생 시 신속하고 효율적인 장애처리가 가능하도록 함
- 시스템을 원활하고 지속적으로 운영하기 위해서 시스템 환경에 적합하고 효율적인 장애관리 방안을 수립하여 체계적으로 관리함
- 장애관리 방안을 수립하기 위해 필요한 단계별 절차는 장애유형분석, 원인분석, 예방방안, 장애 대응 방안 수립으로 정의됨



<그림 4-47> 장애관리 절차

다. 장애 대응

- 장애 발생 시 신속 정확한 장애 원인분석 및 조치로 장애 시간을 단축하고, 장애 발생의 통계 및 관리를 통하여 장애 발생을 예측, 분석하여 원활한 시스템 운영이 가능하도록 지원함
- 장애가 발생하면 접수, 현장 선 조치, 장애 해결, 장애처리 결과보고 순서에 따라 처리하고, 문제 해결을 위해 관련 전문가와 즉시 공조하여 장애 시간을 최소화함



<그림 4- 48> 장애처리 절차

제5장 결론 및 향후 과제

제1절 결론

제2절 향후 과제

제5장 결론 및 향후 과제

제1절 결론

- 본 과업에서는 교통 분석에 모바일 DB가 더욱 활발히 활용될 수 있도록 '18년에 개발한 모바일 가공 알고리즘을 보완하였으며, 이를 기반으로 18년 기준 모바일 DB를 구축하였음
- 통행자의 통행에 의해 생성된 로그가 아니라 단순 신호 이상으로 기록된 로그를 보정하는 방법론을 추가로 개발함
 - 시그널 점프 현상에 의한 로그 기록 보정 방법: 로그 기록 속도가 300km/h가 넘는 경우(로그 기록의 시간차와 기지국간의 거리정보를 이용하여 로그 기록 속도 산출) 하나의 로그 기록으로 변환 (기록 빈도가 높게 나타난 기지국, 체류시간이 높게 기록된 기지국 중심으로 통행 정보 업데이트)
 - 평풍 핸드오버 현상에 의한 로그 기록 보정 방법: 두 개의 기지국이 번갈아 연속해서 번갈아 나타나면서 (A-B-A'-B') 시간 차가 1분 이내인 경우 하나의 로그 기록으로 변환 (체류시간이 높게 기록된 기지국 중심으로 통행 정보 업데이트)
- 새로운 전처리 방법론을 적용한 데이터를 기준으로 통행자가 체류 중인지, 이동 중인지를 구분할 수 있는 기준을 다시 설정함
 - 기존에는 통행자의 체류시간이 25분 이상이면 체류 중으로 판단하고 25분 미만이면 이동 중으로 판단, 즉 '체류시간 25분'이 체류와 이동을 구분할 수 있는 기준이었음
 - 본 과업을 통해 재탐색한 체류/이동 구분 기준은 기존에 비해 10분 더 짧은 15분이었음
 - 이는 전처리 단계에서 신호 이상으로 인해 분리되어 있었던 1분 내외의 매우 짧은 체류의 로그 기록이 하나의 기록으로 합쳐졌기 때문
- 3통행 미만의 데이터가 모두 3통행으로 처리되면서 통행량이 과다하게 집계되었던 문제를 보완하기 위하여 3미만 통행 보정계수를 개발함
 - 그동안 개인정보보호법에 의거 개개인의 통행정보를 추적할 수 없도록 3미만의 통행은 3통행으로 가공하여 DB를 구축하였으나, 체류지 유형을 세분화할수록 3통행 미만의 데이터가 늘어나 실제 통행량 정보와 차이가 크게 벌어지는 현상 발생
 - '17년 기준 DB에서 연평균 일통행량(AADT)과 가장 유사한 달인 6월을 기준 달로 설정한 후, 3통행 보정 후 집계한 통행량을 보정 전 집계한 통행량으로 나누어 3미만 통행을 보정할 수 있는 보정계수를 개발함

- 2018년 1월 1일부터 2018년 12월 31일까지 생성된 로그 기록에 새로 개발한 알고리즘을 적용하여 기존 DB를 구축함
- 체류지 유형은 '18년에 개발한 체류지 구분 알고리즘을 그대로 적용하여 구분하였으나, 집 이외의 심야시간대 주체류지(N)는 활용성이 낮다는 판단 하에 제외하고, 총 5가지의 체류 유형으로 구분함
- 각 개인별 로그 기록은 「개인정보보호」, 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 <표 5-1>과 같이 집계한 형태의 DB를 설계하였으며, 이를 기준으로 2018년 1월 1일부터 2018년 12월 31일까지 발생한 로그 기록을 가공하여 DB를 구축함

<표 5- 1> 모바일 기반 DB 형태

출발				도착				성별 코드	연령 코드	통 행 량	통행 시간 (분)
일자	시 간	교통 폴리곤 ID	체류지 유형	일자	시 간	교통 폴리곤 ID	체류지 유형				
20181121	13	48270310	집	20181123	14	12390981	기타	M	20	10	34
20180501	06	48270310	집	20180501	09	32789014	회사	F	50	25	39
20180311	17	48270320	집	20180312	18	45608912	학교	M	10	15	25

- 또한 교통 모니터링·데이터 제공·분석 플랫폼(View-T)에 '18년 과업을 통해 구축한 모바일 빅데이터 기반 기종점 통행량 DB와 모바일 분석맵 DB를 탑재하였고, 기 개발한 사람 통행 분석 서비스를 보다 사용하기 편리하게 보완하였으며, 더 다양하고 세밀한 분석이 가능하도록 새로운 분석 서비스 2가지를 추가로 개발하였음

<표 5- 2> View-T 2.0 사람 통행 분석 서비스 개발 주요 내용 (기반 DB 구축)

구분		내용
기반 DB 구축	분석 맵 구축	<ul style="list-style-type: none"> • 2016년 기준 교통폴리곤 (2016년 집계구와 2017년 기지국 정보를 활용하여 구축한 영역) 구축 • 2016년 교통폴리곤 단위의 노드 구축 • 2016년 행정구역 단위의 폴리곤/노드 구축
	모바일 빅데이터 기반 기종점 통행량 DB 구축	<ul style="list-style-type: none"> • 2017년 1월 ~ 12월 기준 DB 구축 • 3통행 미만 보정계수 적용

<표 5- 3> View-T 2.0 사람 통행 분석 서비스 개발 주요 내용 (분석 기능 개발)

구분			내용
분석 기능 개발	기존 기능 개선	주요통행 지역분석	<ul style="list-style-type: none">• 기존 유입유출 지역분석과 유입유출 비교분석기능을 하나로 통합• 기존보다 더 많은 지역을 비교 분석할 수 있도록 개선 (2개 지역 → 5개 지역)• 사용자가 선택한 분석 대상 지역의 결과를 더 쉽게 구분할 수 있도록 선택지역별 분석결과를 각기 다른 색상으로 표현할 수 있도록 함• 사용자가 분석 결과를 더 빨리 파악할 수 있도록 행정구역 명칭 ON/OFF 기능을 추가• 선택지역에 대한 분석 결과를 더 쉽게 비교할 수 있도록 상세 정보를 파이차트 형식으로 시각화하여 표현
		주요통행 구간분석	<ul style="list-style-type: none">• 지역간 통행정보도 확인할 수 있도록 개선• (지역내 최다 통행구간 확인 → 지역간 최다 통행구간 확인)• 기존보다 더 많은 지역을 분석할 수 있도록 분석 대상 지역 선택 폭 확대 (1개 지역 → 5개 지역)• 분석 결과에 대한 표출 단위를 다양하게 설정할 수 있도록 개선 (교통폴리곤 단위 → 시군구, 읍면동, 교통폴리곤 단위)
		Hot Place 분석	<ul style="list-style-type: none">• 분석 결과에 대한 표출 단위를 다양하게 설정할 수 있도록 개선 (교통폴리곤 단위 → 시군구, 읍면동, 교통폴리곤 단위)• 산출 결과를 기존보다 다양하게 확인할 수 있도록 개선 (통행량 → 통행량, 일평균 통행량, 면적대비 통행량)
		부가기능 (데이터 검색 및 다운로드)	<ul style="list-style-type: none">• 유형별 통신데이터 다운로드 기능과 지역간 OD 통행량 데이터 다운로드 기능 추가• 사용자 대기시간을 단축시키기 위해 표출화면에 표기되지 않는 데이터인 경우 분석결과를 메일로 받아볼 수 있도록 함
	신규 기능 개발	내부통행 지역분석	<ul style="list-style-type: none">• 내부 통행이 많은 지역을 쉽게 확인할 수 있는 분석 서비스를 개발
		통행시간/ 거리분석	<ul style="list-style-type: none">• 지역별 접근성을 평균통행시간과 평균통행거리 정보를 통해 쉽게 확인할 수 있도록 새로운 분석 서비스를 개발
	기타(종합)		

제2절 향후 과제

- 본 과업을 통해 통행목적이 포함된 DB를 구축하였으나, 향후 궁극적으로 MaaS(Mobility as a service)를 실현하기 위해서는 <그림 5-1>과 같이 통행자의 이동수단정보가 포함된 DB가 구축되어야 할 것임

고객 식별 번호	기저국 X좌표	기저국 Y좌표	체류 시작 시간	체류 종료 시간	일시	연령	성별	체류 시간(분)	1단계	2단계	3단계	4단계
									체류 특성	체류 순서	체류 목적	수단
184302	127,13	36,47	1247	1250	2017 0520	20	남	3	경로	1	-	버스
184302	127,13	36,47	1300	1335	2017 0520	20	남	35	상주	2	학교	-
184302	127,13	36,47	1530	1600	2017 0520	20	남	30	잠재 상주	3	쇼핑	-
:	:	:	:	:	:	:	:	:	:	:	:	:

<그림 5- 1> 향후 모바일 DB 설계 방안

- 또한 향후 모바일 DB에 대한 신뢰도를 높일 수 있도록 Prompt-recall 서비스를 실시할 필요가 있음
 - Prompt-recall 서비스란 모바일 이용자의 시간대별 활동(통행, 출발시간, 도착시간, 체류시간, 활동목적, 활동위치 등)을 조사하는 것을 의미함
- 아울러, 교통 모니터링·데이터 제공·분석 플랫폼(View-T)을 통해 제공되고 있는 분석 서비스를 사용자 편의성, 활용성 측면을 고려하여 보완할 필요가 있으며, 계속해서 축적되는 DB를 시스템 내에서 효율적으로 관리할 수 있는 방안을 모색할 필요가 있음
 - 1차 분석 결과(주제도 표출)를 토대로 사용자가 분석결과를 보다 심층적으로 분석할 수 있도록 상세 분석 창을 추가로 생성할 필요가 있음
 - 차량 통행 분석 결과와 연계시켜 함께 분석할 수 있는 방안을 모색할 필요가 있음