

2018년 「국가교통조사 · DB시스템 운영 및 유지보수」

대중교통 정책지원 고도화를 위한 모바일 빅데이터 DB구축

13

2018년 「국가교통조사 · DB시스템 운영 및 유지보수」

대중교통 정책지원 고도화를 위한 모바일 빅데이터 DB구축

2018.12

2018년 「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 국가정보화사업 중 「2018년도 국가교통조사·DB시스템 운영 및 유지보수」의 최종보고서로 제출합니다.

2018년 12월

한국교통연구원

원장 오 재 학

**본 『2018년도 국가교통조사·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 연구위원 ◦ 박용일, 황순연, 천승훈, 장동익, 송태진, 성홍모, 김병관, 우왕희 부연구위원 ◦ 신영권, 김동호, 김규진 주임전문원, 이종우 전문연구원, 김정은 전문원 ◦ 강국수, 고두환, 김관용, 김성민, 김은미, 박미란, 박준호, 오연선, 이선아, 이선영, 이용철, 이해선, 정승환, 정승연, 조용훈, 탁지훈, 홍성표 연구원 ◦ 서유진, 노수진 연구조원
<한국해양수산개발원>	
연 구 진	<ul style="list-style-type: none"> ◦ 이호춘 부연구위원 ◦ 류희영, 반영길 연구원
<한국항공협회>	
연 구 진	<ul style="list-style-type: none"> ◦ 성인영 실장 ◦ 박수경 과장, 손병열 과장, 유인아 대리

『2018년도 국가교통조사·DB시스템 운영 및 유지보수』

보고서 구성 및 담당연구진

번 호	과 제 명	연 구 진
제 1권	요약보고서	박용일, 신영권, 박준호,
제 2권	전국여객O/D 보완갱신	조종석, 강국수, 박미란,
제 3권	빅데이터를 활용한 여객 O/D 신뢰도 제고 연구	김병관, 정승환
제 4권	항공여객 O/D조사	성인영, 박수경, 유인아
제 5권	전국화물 O/D 전수화 및 장래수요예측	성홍모, 박인기, 김정은, 조용훈 이용철
제 6권	전국해상화물 O/D 전수화 및 장래예측	이호춘, 류희영, 반영길
제 7권	빅데이터를 활용한 화물O/D 신뢰도 제고 연구	성홍모, 박인기, 김정은, 조용훈 이용철
제 8권	교통분석용 네트워크 구축	최정민, 김동호, 우왕희, 김정민 탁지훈, 이선아
제 9권	KTDB 플랫폼 기반지도 구축	최정민, 김동호, 우왕희, 김정민 탁지훈, 이선아
제10권	국가교통통계조사	황순연, 오연선, 고두환
제11권	특별교통대책기간 통행실태 조사	장동익, 김동호, 김은미
제12권	교통혼잡지도 DB구축	천승훈, 김성민, 김관용
제13권	대중교통 정책지원 고도화를 위한 모바일 빅데이 터 DB구축	송태진, 이해선, 홍성표, 이선영, 이종우
제14권	교통유발원단위 첨단조사 연구	황순연, 오연선, 고두환
제15권	국가교통DB시스템 운영 및 유지보수	김규진

『2018년도 국가교통조사·DB시스템 운영 및 유지보수』

과제별 공동참여·위탁용역 사업자

【공동사업 참여기관】

- 전국 여객O/D 현행화 공동사업 (수도권 부문)
 - 서울연구원, 경기개발연구원, 인천발전연구원
- 전국 여객O/D 현행화 공동사업 (대구광역시권 부문)
 - (재)대구경북연구원
- 전국 여객O/D 현행화 공동사업 (제주특별자치도 부문)
 - 홍익대학교산학협력단
- 전국 여객O/D 현행화 공동사업 (부산·울산권 부문)
 - 경성대학교산학협력단, (주)나우컨설턴트
- 전국 여객O/D 현행화 공동사업 (대전·충청권 부문)
 - (주)홍익기술단, 더블유비그룹코리아

【위탁용역 사업자】

- ViewT 1.0 서비스 제공을 위한 DB구축 및 시스템 개선
 - (주)큐빅웨어
- GIS기반 교통망 DB구축
 - (주)큐빅웨어
- 2018년 국가교통조사 중 특별교통통행실태조사 및 이용자 만족도 조사
 - (주)리서치랩
- 도로 및 철도 교통분석용 네트워크 보완갱신
 - (주)큐빅웨어
- 수출입 항공화물 기종점 통행량조사 위탁용역
 - (주)코리아데이터네트워크

【위탁용역 사업자】

- ViewT 2.0 서비스 제공을 위한 통신빅데이터 구축 및 기능개발
 - ㈜큐빅웨어
- 국가교통DB시스템 운영 및 유지보수 중 2018년 KTDB HW 유지보수 및 서버운영 SW 라이선스 갱신
 - ㈜휴버텍
- 모바일 빅데이터 분석 환경 구현 및 교통분석용 DB 구축
 - ㈜케이티
- 2018년도 국가교통DB Brief 발행
 - 텍스트앤드이미지
- 전국 여객 O/D 웹기반 검증프로그램 구축
 - ㈜제이에스소프트
- 국가교통통계 DB 조사관리 시스템 개선방안
 - ㈜블루와이즈
- 사용자 요구사항 분석을 위한 설문조사
 - ㈜지알아이리서치
- 국가교통빅데이터 플랫폼 아키텍처 설계
 - ㈜지음지식서비스
- 화물 O/D신뢰도 향상을 위한 검증자료 구축
 - 서울시립대학교 산학협력단
- 혼잡(불안정류)상황 교통수요 추정 방안 연구
 - 공주대학교 산학협력단

최종보고서 목차

제 1권 요약보고서

제 2권 전국여객O/D보완갱신

제 3권 빅데이터를 활용한 여객O/D 신뢰도 제고 연구

제 4권 항공여객 O/D조사

제 5권 전국화물O/D 전수화 및 장래수요예측

제 6권 전국해상화물O/D 전수화 및 장래예측

제 7권 빅데이터를 활용한 화물 O/D 신뢰도 제고 연구

제 8권 교통분석용 네트워크 구축

제 9권 KTDB 플랫폼 기반지도 구축

제 10권 국가교통통계조사

제 11권 특별교통대책기간 통행실태 조사

제 12권 교통혼잡지도 DB구축

제 13권 대중교통 정책지원 고도화를 위한 모바일 빅데이터 DB구축

제 14권 교통유발원단위 첨단조사연구

제 15권 국가교통DB시스템 운영 및 유지보수

목 차

요 약

제1장 과업의 개요	1
------------------	---

제1절 과업의 배경 및 목적 / 3

제2절 과업의 범위 및 내용 / 5

제3절 과업 추진 방법 / 8

제2장 모바일 자료 기반 DB 구축	9
---------------------------	---

제1절 모바일 자료 가공 기술 개발 / 11

제2절 모바일 자료 기반 DB 구축 / 27

제3장 모바일 분석 맵 고도화	35
------------------------	----

제1절 분석 맵 구축 방법 / 37

제2절 분석 맵 구축 결과 및 검증 / 46

제4장 모바일 자료 기반 View-T 2.0 서비스 기능 개발	51
--	----

제1절 기능 정의 및 UI 설계 / 53

제2절 기반 DB 구축 / 91

제3절 시스템 운영 및 유지보수 / 119

제5장 결론 및 향후 과제	127
----------------------	-----

제1절 결론 / 129

제2절 향후 과제 / 132

표 목 차

〈표 2- 1〉 가구통행실태조사 목적 구분	15
〈표 2- 2〉 체류지 유형	16
〈표 2- 3〉 모바일 자료 기반 통행 목적 구분	17
〈표 2- 4〉 통행목적별 통행 특성	18
〈표 2- 5〉 도착지별 로그 기록 특성	19
〈표 2- 6〉 집과 집 이외 심야시간대 주체류지 식별 기준	21
〈표 2- 7〉 집과 집 이외 심야시간대 주체류지 식별을 검토 (2017년 기준)	22
〈표 2- 8〉 통계청 기준 통근 통학 인구와 모바일 데이터 기준 통근 통학 인구 비교 (2017년 기준)	23
〈표 2- 9〉 회사와 학교 체류 식별 기준	24
〈표 2-10〉 통계청 기준 종교인구와 모바일 데이터 기준 종교인구 비교 (2017년 기준)	25
〈표 2-11〉 종교집회장소와 기타 잠재체류지 식별 기준	25
〈표 2-12〉 로그별 통행목적 추정 결과 (예시)	26
〈표 2-13〉 모바일 기반 DB 형태 (예시)	27
〈표 2-14〉 모바일 DB 테이블 정의서	28
〈표 2-15〉 체류지 식별 기준	31
〈표 3- 1〉 집계구별 기지국 개수 (예시)	38
〈표 3- 2〉 시도별 최대 교통폴리곤 형성 개수와 실제 구축된 교통폴리곤 개수 비교 ...	41
〈표 3- 3〉 주기지국 기준 변화	41
〈표 3- 4〉 최소면적 기준 변화	42
〈표 3- 5〉 시군구별 최소면적 기준	43
〈표 3- 6〉 시군구별 최소면적 기준 (계속)	44
〈표 3- 7〉 교통 폴리곤 구축 결과	46
〈표 3- 8〉 교통 폴리곤 구축 결과	48
〈표 3- 9〉 교통폴리곤과 기존 분석구 비교	49
〈표 3-10〉 기존 교통존과 교통 폴리곤 비교	50

〈표 4- 1〉 메뉴의 구성	54
〈표 4- 2〉 통행량(인구수) 검색의 출력파일 구성	88
〈표 4- 3〉 수집된 분석구 데이터의 속성정보	92
〈표 4- 4〉 MySQL 기반의 분석구 폴리곤 공간정보의 속성	93
〈표 4- 5〉 VectorTile, WMS의 지도 일반화	94
〈표 4- 6〉 GeoServer 분석구 레이어의 속성정보	95
〈표 4- 7〉 GeoServer 분석구 폴리곤 노드 레이어의 속성정보	95
〈표 4- 8〉 수집된 행정구역 데이터의 속성정보	97
〈표 4- 9〉 MySQL 기반의 행정구역 폴리곤 공간정보의 속성	98
〈표 4-10〉 GeoServer 행정구역 폴리곤 레이어의 속성정보	98
〈표 4-11〉 MySQL 기반의 행정구역 노드 공간정보의 속성	99
〈표 4-12〉 통신 빅데이터의 컬럼명 및 데이터 타입	100
〈표 4-13〉 통신 빅데이터의 연령별 코드	101
〈표 4-14〉 통신 빅데이터의 성별 코드	101
〈표 4-15〉 통신 빅데이터의 시간대 코드	102
〈표 4-16〉 통신 빅데이터의 O/D 트립타입	103
〈표 4-17〉 분석구 기준의 모바일 통행형태 통행량 컬렉션의 구성	105
〈표 4-18〉 통신 빅데이터의 도큐먼트 수의 경량화 결과	107
〈표 4-19〉 읍면동 출발 기준의 통행형태 통행량의 컬렉션의 구성	107
〈표 4-20〉 읍면동 도착 기준의 통행형태 통행량의 컬렉션의 구성	109
〈표 4-21〉 시군구 출발 기준의 통행형태 통행량의 컬렉션의 구성	110
〈표 4-22〉 시군구 도착 기준의 통행형태 통행량의 컬렉션의 구성	111
〈표 4-23〉 시도 출발 기준의 통행형태 통행량의 컬렉션의 구성	113
〈표 4-24〉 시도 도착 기준의 통행형태 통행량의 컬렉션의 구성	114
〈표 4-25〉 읍면동 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	115
〈표 4-26〉 읍면동 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	115
〈표 4-27〉 시군구 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	115
〈표 4-28〉 시군구 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	115
〈표 4-29〉 시도 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	116
〈표 4-30〉 시도 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과	116
〈표 4-31〉 Web Server 설치 정보	119
〈표 4-32〉 Web Server 설치 정보(Apache Tomcat)	120

〈표 4-33〉 데이터베이스 백업 대상 파일	121
〈표 4-34〉 응용프로그램 백업 대상 파일	122
〈표 4-35〉 백업 및 복구 절차	122
〈표 5- 1〉 모바일 기반 DB 형태 (예시)	130
〈표 5- 2〉 View-T 2.0 모바일 자료 기반 분석 서비스 개발 주요 내용	130

그림목차

〈그림 1- 1〉 모바일 데이터 기반 대중교통 최적화 시스템 (AllAboard)	3
〈그림 1- 2〉 과업 추진 절차	8
〈그림 2- 1〉 활동 목적 추정 방법	14
〈그림 2- 2〉 체류순서 부여 (개인별 이동궤적 형성)	29
〈그림 2- 3〉 이동정보와 체류정보 구분 (예시)	30
〈그림 2- 4〉 위치정보, 연령 정보 변환 (예시)	31
〈그림 2- 5〉 체류지 유형 구분 (예시)	32
〈그림 2- 6〉 통행목적 구분 (예시)	33
〈그림 2- 7〉 인구수(통행량) 집계 (예시)	34
〈그림 3- 1〉 교통 폴리곤 구축 1단계	37
〈그림 3- 2〉 셀 반경 내 집계구 병합 기준 (예시)	38
〈그림 3- 3〉 교통 폴리곤 구축 2단계	38
〈그림 3- 4〉 교통 폴리곤 구축 3단계	39
〈그림 3- 5〉 교통 폴리곤 구축 4단계	40
〈그림 3- 6〉 전국 교통폴리곤 형성 결과	47
〈그림 4- 1〉 서비스의 구성	53
〈그림 4- 2〉 통신데이터 기반의 통행량분석의 메뉴	54
〈그림 4- 3〉 유입유출 비교분석의 실행 화면	56
〈그림 4- 4〉 유입유출 비교분석 설정 UI	57
〈그림 4- 5〉 유입유출 비교분석의 출발지 및 도착지 설정	58
〈그림 4- 6〉 지도 선택을 통한 분석지역 선택	59
〈그림 4- 7〉 ID 검색을 통한 분석지역 선택	59
〈그림 4- 8〉 유입유출 비교분석의 희망선도 범례설정 UI	61
〈그림 4- 9〉 유입유출 비교분석의 분포도 범례설정 UI	62
〈그림 4-10〉 유출량 비교분석의 결과화면 예시 (출발지(1) ON, 출발지(2) ON)	63
〈그림 4-11〉 유출량 비교분석의 결과화면 예시 (출발지(1) ON, 출발지(2) OFF)	64

〈그림 4-12〉 유출량 비교분석의 결과화면 예시 (희망선도 ON, 분포도 OFF)	64
〈그림 4-13〉 유출량 비교분석의 결과화면 예시 (희망선도 OFF, 분포도 ON)	65
〈그림 4-14〉 유입유출 지역분석의 실행 화면	66
〈그림 4-15〉 유입유출 지역분석의 설정 UI	67
〈그림 4-16〉 유입유출 지역분석의 표출지도 ZOOM 단계 설정 UI	68
〈그림 4-17〉 유입유출 지역분석의 표출지도 범례설정 UI	70
〈그림 4-18〉 ZOOM 단계에 따른 범례설정의 변경 예시	71
〈그림 4-19〉 범례설정의 표출단위 선택에 따른 표출지도 변경 예시	72
〈그림 4-20〉 유입유출 지역분석의 분석구단위 유출량 분석의 화면 예시	73
〈그림 4-21〉 유입유출 지역분석의 읍면동단위 유출량 분석의 화면 예시	73
〈그림 4-22〉 유입유출 지역분석의 시군구단위 유출량 분석의 화면 예시	74
〈그림 4-23〉 유입유출 지역분석의 시도단위 유출량 분석의 화면 예시	74
〈그림 4-24〉 O/D 분석의 설정 UI	76
〈그림 4-25〉 O/D 분석의 희망선도 범례설정 UI	78
〈그림 4-26〉 O/D 분석결과 화면 예시(상위 10% 강조)	79
〈그림 4-27〉 O/D 분석결과 화면 예시(상위 30% 강조)	80
〈그림 4-28〉 Hot Place 분석의 실행 화면	81
〈그림 4-29〉 Hot Place 분석의 설정 UI	82
〈그림 4-30〉 Hot Place 분석의 분포도 범례설정 UI	84
〈그림 4-31〉 HeatMap의 반경과 흐림	85
〈그림 4-32〉 유입기준 Hot Place 분석결과 화면 예시(순위표출개수 : 10)	86
〈그림 4-33〉 유입기준 Hot Place 분석결과 HeatMap 화면 예시	86
〈그림 4-34〉 통행량(인구수) 검색의 설정 UI	87
〈그림 4-35〉 범례정보 저장 및 불러오기	89
〈그림 4-36〉 팝업 기능을 통한 분석결과 정보제공	89
〈그림 4-37〉 분석결과에 대한 통신데이터의 파일출력 예시	90
〈그림 4-38〉 데이터 구축 프로세스	91
〈그림 4-39〉 분석구 레이어의 구축 절차	92
〈그림 4-40〉 GeoServer 기반의 분석구 폴리곤 레이어	94
〈그림 4-41〉 GeoServer 기반의 분석구 노드 레이어	96
〈그림 4-42〉 GeoServer 기반의 행정구역 폴리곤 레이어	98
〈그림 4-43〉 GeoServer 기반의 행정구역 노드 레이어	99

〈그림 4-44〉 수집된 통신 빅데이터 형태	100
〈그림 4-45〉 모바일 통행형태 기반 DB의 구축 절차	103
〈그림 4-46〉 분석구와 읍면동 행정구역 코드의 매칭정보	104
〈그림 4-47〉 1차 가공데이터의 MongoDB 입력데이터 예시	105
〈그림 4-48〉 MongoDB Sharding의 구성	118
〈그림 4-49〉 모바일 통행형태 기반 DB의 Sharding 구성 결과	118
〈그림 4-50〉 유지보수 범위 및 조직	124
〈그림 4-51〉 장애관리 절차	125
〈그림 4-52〉 장애처리 절차	126
 〈그림 5- 1〉 향후 모바일 DB 설계 방안	 132

요약

요 약

1. 과업의 개요

가. 과업의 배경 및 목적

- 전 세계적으로 대중교통 및 비동력 수단 모니터링을 위한 플랫폼 개발이 활발히 이루어지고 있으며, 이를 기반으로 고도화된 대중교통 정책을 지원 중이나, 현재 국내에는 표준화된 대중교통 및 비동력 수단 모니터링 시스템 부재한 실정
 - 플랫폼을 활용하여 버스 노선을 탄력적으로 운영하거나, 연령별, 성별 등 인적 특성 기반 통행패턴을 모니터링하여 이용자 맞춤형 대중교통 서비스를 제공 중
 - 교통은 파생수요라는 측면에서, 신(新) 교통수단 등 다양한 수단을 고려한 교통정책을 지원하기 위해서는 ‘사람의 이동’을 파악하는 것이 중요하나 기존의 모니터링 시스템은 차량 중심으로 구축됨
- 이에 본 과업을 통해 사람의 이동정보가 담긴 통신자료를 활용해 대중교통 정책 지원에 직접적으로 활용 가능한 대중교통 및 비동력 수단 모니터링 시스템의 기반을 구축하고자 함
 - 향후 대중교통 모니터링 시스템으로 확장할 수 있도록 교통카드 자료 연계를 고려하여 지리정보체계 기반 통신자료 가공 기술을 개발하여 통신자료 기반 DB를 구축하고, 대중교통 및 비동력 수단 이동 모니터링 및 분석이 가능하도록 핵심 통계 지표 및 분석 기술을 개발하여 기 개발된 플랫폼(View-T)에 구축하고자 함

나. 과업의 범위 및 내용

1) 과업의 범위

- 시간적 범위
 - 모바일 자료 가공 기술 개발 및 2017년도 기준 기반 DB 구축: 2017년 1월 1일 ~ 2017년 12월 31일 (1년)
 - 2017년 4월 한 달간의 원천데이터를 활용하여 가공 알고리즘을 개발한 후, 2017년에 생성된 시그널 데이터에 개발한 가공 알고리즘을 적용

- 웹 서비스를 위한 각종 분석도구 개발: 2016년 3월 7일 ~ 2016년 12월 31일
 - 2017년도 연구 성과물인 2016년도 기준 기반 DB 활용
- 공간적 범위: 전국

2) 과업의 내용

① 모바일 자료 기반 DB 구축

- 이동인구 이동계적 자료 전처리 기술 개발
 - 통신 원시자료에서 지역별, 연령대별, 성별 등 다양한 특성을 고려한 이동계적 자료 전처리 방법론 개발 및 적용 방안 제시
- 통행목적 구분 로직 개발
 - 이동인구 통행 패턴 특성 분석을 통해 개인정보보호법에 위반되지 않는 범위 내에서 목적 O/D 정보 추정 및 제공 방안 개발
- 기반 DB 설계 및 '17년 기준 기반 DB 구축
 - 분석구 Polygon 네트워크와 통행사슬 정보기반의 O/D DB 산출
 - 이용자 맞춤형 특성별(주중·주말, 첨두·비첨두) 쿼리형 O/D 정보 제공

② 모바일 자료 분석 맵 구축

- 분석 맵 구축 알고리즘 고도화
- '17년 기준 분석 맵 구축

③ View-T 모바일 자료 기반 분석 서비스 개발

- GIS 서비스 레이어 구축
 - 웹 서비스를 위한 분석구 데이터 경량화 작업
 - 웹 서비스를 위한 분석구 단위의 노드 구축
 - 분석구 단위의 검색 POI 구축
- 웹 서비스를 위한 모바일 통행행태 기반 DB 구축
 - 데이터 분석 및 DB 설계

- DB 설계 기반의 데이터 구축을 위한 가공 시스템 개발
- 웹서비스용 데이터 구축 및 업로드
- 데이터 검증
- 통행행태 기반 데이터 압축
 - 웹 서비스를 위한 데이터 정제
 - 웹 서비스 속도 향상을 위한 데이터 압축(경량화) (검색속도 향상을 위한 Mongo DB 기반의 데이터 구축)
 - 데이터 검증
- 분석 서비스 기능 개발
 - 기능 정의 및 UI 설계
 - O/D 서비스 기능 개발
 - 이동생활패턴 분석 서비스 기능 개발
 - 시간대별 대중교통·유동인구 인적특성별 Hot Place 분석 서비스 기능 개발

2. 모바일 자료 가공 기술 개발

가. 전처리 기술 개발

- 야간에 주로 발생하는 핑퐁 핸드오버 데이터 특성을 반영하여 핑퐁 현상이 나타난 로그 기록을 추출하는 기준을 1일 단위(0시 1분부터 23시 59분)에서 일주일 단위(예: 1일 0시 1분부터 7일 23시 59분)로 변경
- 기존 핑퐁 핸드오버 데이터 보정 알고리즘은 핑퐁 현상을 1일 단위로 식별하도록 설계되어 있어 일자가 변경되는 때 발생한 핑퐁 현상을 보정하지 못함
- 최대한 많은 핑퐁 현상을 보정하기 위해서는 최초로 기록된 로그부터 가능한 한 가장 최근에 기록된 로그를 모두 연결한 상태에서 핑퐁 현상을 검열해야 하나, 축적된 원천데이터의 양이 방대하기 때문에 가공 효율성이 떨어질 수 있고, 핑퐁 현상 탐색 기준과 체류지 식별 기준이 다를 경우 체류지 식별이 불가하므로 최대 기간을 일주일로 결정함
- 핑퐁현상 발생에 대한 탐색 기준과 체류지 식별 기준이 상이할 경우 체류지 식별이 불가능하기 때문에 기준이 최대한 매칭이 되도록 추출 기준을 설정

- 주체류지는 주 단위의 식별 기준에 의해 식별되므로 한 달 기준으로 핑퐁 현상을 가공할 경우에는 일주일 기간이 넘는 핑퐁 데이터들이 보완되면서 기존에 주체류지로 식별되었던 기지국이 주체류지에서 누락될 가능성이 존재함

나. 통행 목적 구분 알고리즘 개발

- 본 과업에서는 기지국 기록 패턴을 활용하여 체류지(기점과 종점)의 활동을 추론하는 방식으로 출근 / 등교 / 여가 / 귀가 / 기타 총 5가지의 통행 목적을 구분
 - 주로 방문하는 시간대, 평균 체류 시간, 방문 빈도 등을 고려하여 체류지를 구분
- 추론 대상(체류지) 유형은 기존 가구통행실태조사의 통행 목적별 기종점 유형을 참조하여 집 / 집 이외 심야시간대 주체류지 / 회사 / 학교 / 종교집회장소 / 기타 (여가 / 오락/쇼핑/친교 장소)로 구분

<표 1> 체류지 유형

17년도 기준	'18년도 기준
심야시간대 주체류지	집
	집 이외 심야시간대 주체류지 (예: 친지 집)
낮시간대 주체류지	회사
	학교
잠재체류지	교회, 성당 등 종교집회장소
	기타

- 통행목적별 통행목적별 기종점 체류지 유형을 정리하면 <표 13-2>와 같음

<표 2> 모바일 자료 기반 통행 목적 구분

목적 구분		출발지	도착지
출근		집	회사
		집 이외의 심야시간대 주체류지	
등교		집	학교
		집 이외의 심야시간대 주체류지	
여가		집	기타
		집 이외의 심야시간대 주체류지	
		회사	
		학교	
		기타	
귀가	퇴근	회사	집
			집 이외의 심야시간대 주체류지
	하교	학교	집
			집 이외의 심야시간대 주체류지
	귀가	종교집회장소	집
			집 이외의 심야시간대 주체류지
기타		집	종교집회장소
		집 이외의 심야시간대 주체류지	
		회사	
		학교	
		기타	
	친지 방문	집	집 이외의 심야시간대 주체류지
	기타	종교집회장소	회사
			학교
		기타	회사
			학교

- 체류지 식별 기준은 <표 13-3>과 같이 통행목적에 따라 다르게 나타나는 통행 시간대, 통행 계층 등 통행 특성을 참고하여 설정함
- 통계청 인구 정보 등 실제 기준 값을 통해 식별 기준의 타당성을 검토함

<표 3> 통행목적별 통행 특성

통행목적		특성
출근		<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오전 6시 ~ 오전 9시) 주로 발생¹⁾ • 특정 계층에 해당(사회에 진출한 20세 이상의 통행자)
등교		<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오전 6시 ~ 오전 9시) 주로 발생¹⁾ • 특정 계층에 해당(취학 대상의 연령층, 대체로 20세 미만)
여가		<ul style="list-style-type: none"> • 비반복 통행
귀가	퇴근	<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오후 4시 ~ 오후 23시) 주로 발생²⁾ • 특정 계층에 해당(사회에 진출한 20세 이상의 통행자)
	하교	<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오후 4시 ~ 오후 23시) 주로 발생²⁾ • 특정 계층에 해당(취학 대상의 연령층, 대체로 20세 미만)
	귀가	<ul style="list-style-type: none"> • 반복 통행 • 특정 시간대(오후 7시 ~ 오후 10시) 주로 발생³⁾ • 요일에 영향을 받지 않음
기타	종교활동	<ul style="list-style-type: none"> • 반복 통행 • 주로 주말에 발생 • 특정 시간대(예배 시간) 주로 발생
	친지 방문	<ul style="list-style-type: none"> • 비반복 통행
	기타	<ul style="list-style-type: none"> • 특징 없음

주 1) 2016년 3월 7일 ~ 2017년 3월 31일에 기록된 통신 데이터를 분석한 결과 심야시간대 주체류지(예: 집)에서 낮 시간대 주체류지(예: 회사, 학교 등)로 주로 오전 6시에서 오전 8시 사이에 출발하여 오전 7시에서 오전 9시 사이에 도착하는 것으로 나타남(모바일 Mobility report, p.55).

주 2) 2016년 3월 7일 ~ 2017년 3월 31일에 기록된 통신 데이터를 분석한 결과 낮 시간대 주체류지(예: 회사, 학교 등)에서 심야시간대 주체류지(예: 집)로 주로 오후 4시에서 오후 7시 사이에 출발하여 오후 5시에서

오후 11시 사이에 도착하는 것으로 나타남(모바일 Mobility report, p. 56).

주 3) 2016년 3월 7일 ~ 2017년 3월 31일에 기록된 통신 데이터를 분석한 결과 잠재활동지(예: 식당, 백화점, 공원 등)에서 심야시간대 주체류지(예: 집)으로 주로 오후 7시에서 오후 9시 사이에 출발하여 오후 9시에서 오후 10시 사이에 도착하는 것으로 나타남(모바일 Mobility report, p. 57).

자료: KTDB 정책자료집 『모바일 Mobility Report』, p. 55~57.

○ 체류지 유형별 식별 기준을 정리하면 다음 <표 13-4>와 같음

<표 4> 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	-	-
		집 이외	오후 9시 ~ 오전 7시	3시간 이상	주 1회	-	집과의 거리가 5km, 이상인 경우, 모두 집 이외 주체류지로 식별, 5km미만인 경우 심야시간대 마지막에 등장한 기지국을 집 이외 주체류지로 식별
	낮 시 간 대	회사	오전 9시 ~ 오후 6시	3시간 이상	주 2회 이상(주말 제외) 기록된 주가 월 3회 이상	20세 이상	20세 이상 25세 미만 연령의 경우 방학기간 (7~8월, 1~2월)에도 위 기준이 충족될 경우 회사로 간주
		학교	오전 9시 ~ 오후 6시	3시간 이상	주 2회 이상(주말 제외) 기록된 주가 월 3회 이상	25세 미만	20세 이상 25세 미만 연령의 경우 방학기간 (7~8월, 1~2월)에 위 기준이 충족되지 않는 경우 회사로 간주
	잠재 활동 지역	종교 집회장소	오전 8시 ~ 오후 14시 / 오전 9시 ~ 오후 16시	1시간 이상	주 1회 이상 기록된 주가 월 2회 이상	-	‘토요일 또는 일요일’에 기록된 로그에 한 함
		기타	집, 집 이외 심야시간대 주체류지, 회사, 학교, 종교집회장소로 구분되지 않은 그 외의 체류지				

다. 모바일 자료 기반 DB 구축

- 개인별 로그 기록은 「개인정보보호법」과 「위치정보의 보호 및 이용 등에 관한 법률」 등으로 인해 공개될 수 없으므로 기반 DB는 개인의 위치와 이동궤적이 추정되지 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), 통행목적 정보가 동일한 인구를 집계하여 <표 13-5>와 같은 형태로 DB를 구축함
- 2017년 1월 1일부터 2017년 12월 31일까지 전국에서 발생한 로그 기록을 가공하여 총 약 4.3TB의 기반 DB가 생성됨(1일 기준 약 12GB)

<표 5> 모바일 기반 DB 형태 (예시)

출발				도착				통행 목적	성 별 코 드	연 령 코 드	유동 인구 수
일자	시 간	교통 폴리곤 ID	체류지 유형	일자	시 간	교통 폴리곤 ID	체류지 유형				
20171121	13	48270310	집	20171123	14	12390981	기타	여가	M	20	10
20170501	06	48270310	집	20170501	09	32789014	회사	출근	F	50	25
20170311	17	48270320	집	20170312	18	45608912	학교	하교	M	10	15

- DB에서 제공되는 일자, 시간, 위치, 트립타입, 통행목적, 성, 연령 정보를 활용하여 분석 조건에 따라 다양한 형태로 기종점 통행량 산출 가능

3. 모바일 분석 맵 고도화

가. 기존 방식의 문제점 및 개선방향

1) 주기지국 안전성 개선

- 주기지국을 한 달 기준으로 선별할 경우 특정 기간에만 기록된 기지국이 주기지국 정보에 포함되어 있을 수 있으므로, 주 기지국 선정 기준을 1년 기준으로 365일 이상 신호가 기록된 기지국으로 변경함

<표 6> 주기지국 기준 변화

'17년 주기지국	'18년 주기지국
2016년 5월 한 달 동안 25일 이상 신호가 기록된 기지국	1년 기준으로 365일 이상 신호가 기록된 기지국

2) 교통폴리곤 최소면적 기준 변경

- '시도별 폴리곤 면적 누적분포함수 값 5% 이하를 최소면적으로 적용할 경우, 세종시 등 일부 지역은 큰 규모의 교통폴리곤이 형성될 수밖에 없으므로, 최소면적 기준을 시군구 단위로 변경함

<표 7> 최소면적 기준 변화

'17년 최소면적 기준	'18년 최소면적 기준
시도별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하	시군구별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하

나. 모바일 분석 맵 구축 개선안

1) 1단계: 기지국별 가상의 영역(셀 반경)을 설정

- 365일 이상 기록된 기지국을 (이하 '주 기지국') 기준으로 보로노이 기법을 적용하여 기지국별 영역

형성

2) 2단계: 주기지국 수신범위를 기준으로 1차 병합

- 집계구의 중심점을 생성한 후 기지국 수신범위와 중첩시켜 동일한 수신범위에 있는 집계구를 동일 행정(읍면동) 내에서 병합
 - 기지국이 매핑 되지 않은 집계구는 동일 행정동(읍면동) 내 맞닿아있는 경계선(border)이 많은 폴리곤과 병합

3) 3단계: 부적합한 교통폴리곤 보정

- 지나치게 작게 형성된 폴리곤, Multi-part 폴리곤 보정
 - 시군구별 집계구 면적 누적분포함수 값(Cumulative distribution function, CDF) 5%를 폴리곤 면적 최소 기준으로 두고(〈표 3-5〉, 〈표 3-6〉 참조), 이보다 작게 형성된 폴리곤, 하나의 폴리곤으로 인식되나, 실제로는 떨어져 위치하여 하나의 폴리곤으로 보기 어려운 폴리곤을 동일 행정동(읍면동) 내 맞닿아있는 경계선(border)이 많은 폴리곤과 병합

4) 4단계: 비(非)매칭 교통 폴리곤 보정

- 기지국이 '0'개 매핑된 폴리곤 중에서 2단계에서 병합되지 않은 폴리곤을 동일 읍면동 내에서 가까운 폴리곤과 병합하여 보정
 - 인근 폴리곤과 중심점간의 거리 기준으로 병합 대상 탐색

다. 모바일 분석 맵 구축 결과

- 전국 총 16,335개의 교통폴리곤이 형성됨
 - '17년도 교통폴리곤 대비 '18년도 교통폴리곤은 약 2.5% 증가하였으며, 시도별 교통폴리곤 면적은 평균 약 6.8% 감소함
 - 기존 시군구 대비 약 1/65, 읍면동 대비 약 1/5 수준의 규모로 구축됨

<표 8> 교통 폴리곤 구축 결과

구분	'17년도 교통폴리곤		'18년도 교통폴리곤		증감율(%)	
	개수	평균 면적 (㎡)	개수	평균 면적 (㎡)	개수	면적
서울특별시	1,617	375,320	1,664	363,731	2.9	-3.1
부산광역시	789	1,006,100	764	1,026,162	-3.2	2.0
대구광역시	555	1,585,891	582	1,511,316	4.9	-4.7
인천광역시	440	2,655,597	441	2,492,872	0.2	-6.1
광주광역시	422	1,181,355	449	1,109,139	6.4	-6.1
대전광역시	437	1,234,709	469	1,150,260	7.3	-6.8
울산광역시	205	5,218,281	253	4,195,228	23.4	-19.6
세종특별자치시	70	6,653,727	89	5,222,970	27.1	-21.5
경기도	3,204	3,248,864	2,762	3,711,108	-13.8	14.2
강원도	776	21,832,096	941	17,863,176	21.3	-18.2
충청북도	726	10,211,441	821	9,023,623	13.1	-11.6
충청남도	1,218	6,839,509	1,289	6,410,660	5.8	-6.3
전라북도	1,055	7,622,155	1,190	6,798,629	12.8	-10.8
전라남도	1,389	9,119,564	1,469	8,455,669	5.8	-7.3
경상북도	1,433	13,292,880	1,496	12,720,036	4.4	-4.3
경상남도	1,272	8,350,587	1,328	7,946,745	4.4	-4.8
제주특별자치도	329	5,745,843	328	5,685,216	-0.3	-1.1
계	15,937	-	16,335	-		



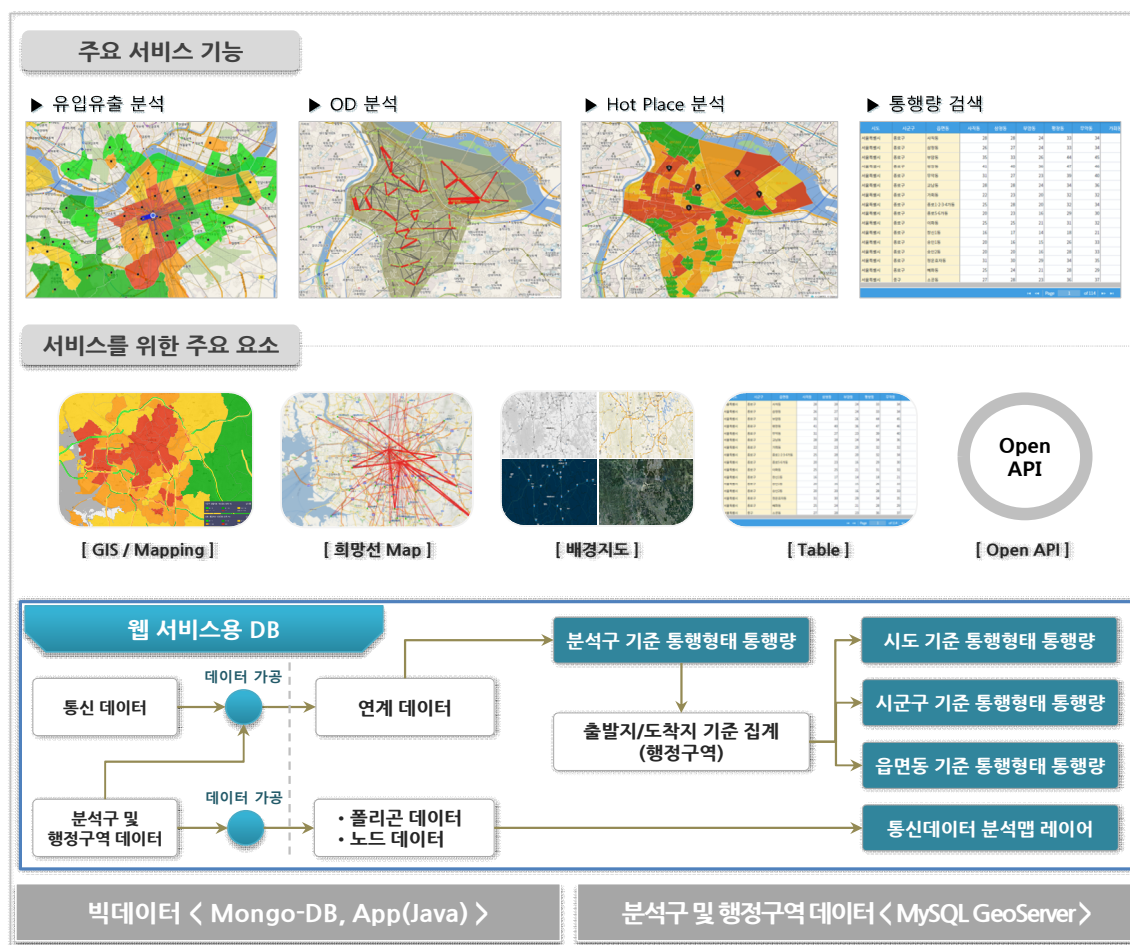
<그림 2> 전국 교통폴리곤 형성 결과

4. 모바일 자료 기반 View-T 2.0 서비스 기능 개발

가. 기능 정의 및 UI 설계

1) 개요

- GIS 기반으로 유입유출 지역분석, 유입유출 비교분석, OD 분석, Hot Place 분석을 할 수 있는 서비스와 사용자가 조건을 설정하여 해당되는 통행량(이동인구수)를 검색하고 이를 파일형태로 출력할 수 있는 서비스를 개발함

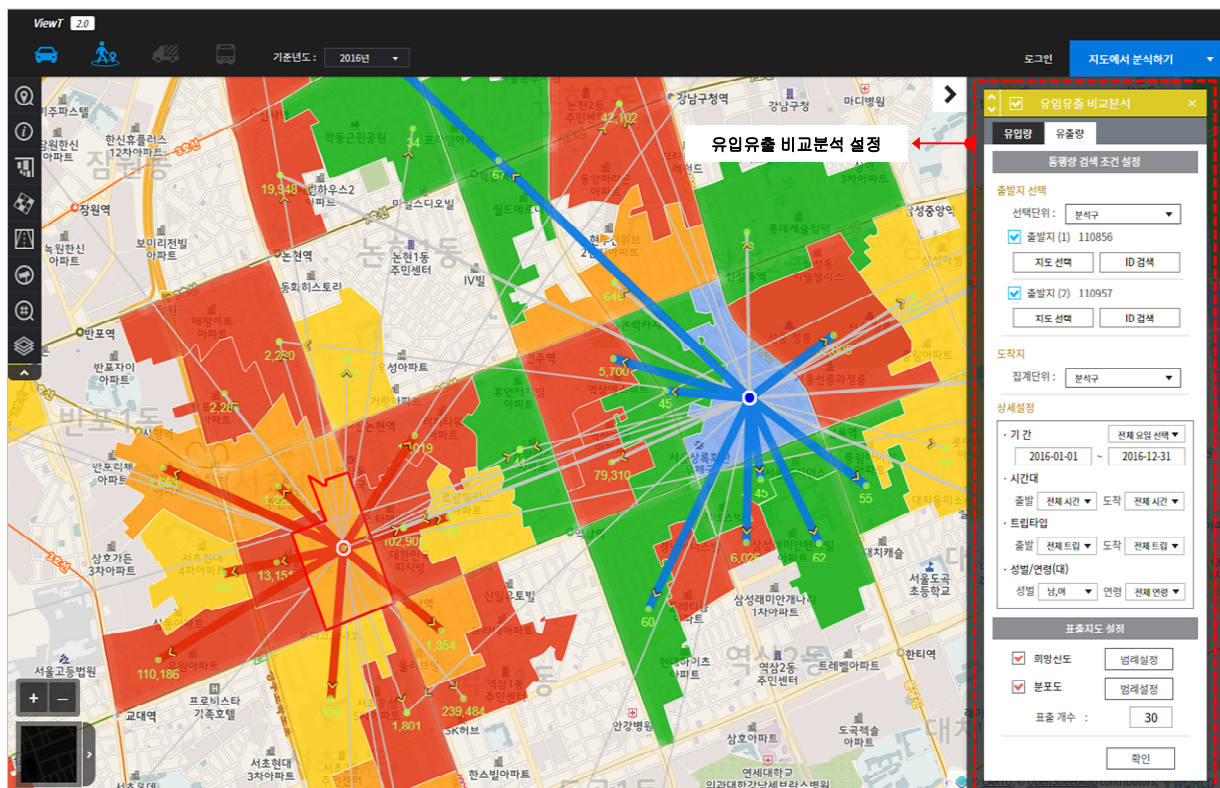


<그림 3> 서비스의 구성

2) 주요기능

① 유입유출 비교분석

- 시도, 시군구, 읍면동, 분석구 단위로 복수개 지역을 선택하여 선택한 지역으로 유입되는 통행량 또는 선택한 지역서 유출되는 통행량을 비교분석할 수 있도록 개발함
 - 최대 5개의 지역 비교 분석 가능
- 유입량 분석은 출발지의 집계단위, 유출량 분석은 도착지의 집계단위를 시도, 시군구, 읍면동, 분석구 단위로 설정할 수 있도록 개발함
- 기간, 요일, 출발 및 도착시간대, 출발 및 도착 트립타입, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 분석결과를 희망선도, 분포도로 지도에 표출하고 표출지도의 범례를 설정할 수 있도록 개발함
- 유입유출 비교분석에 활용된 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 4> 유입유출 비교분석의 실행 화면

- 선택한 지역에 대한 유입량, 유출량 비교분석이 가능하도록 UI를 구성함
- 분석결과를 희망선도, 분포도로 지도에 표출하고 표출지도의 범례를 설정할 수 있도록 개발함

유입량 분석 시 설정 UI

유입량 분석 시 설정 UI

유입량 유출량

동행량 검색 조건 설정

도착지 선택

선택단위 : 분석구

☒ 도착지 (1) 110856

지도 선택 ID 검색

☒ 도착지 (2) 110957

지도 선택 ID 검색

출발지

집계단위 : 분석구

상세설정

기간 : 2016-01-01 ~ 2016-12-31

시간대 : 출발 전체 시간 도착 전체 시간

트립타입 : 출발 전체 트립 도착 전체 트립

성별/연령(대) : 성별 남,여 연령 전체 연령

표출지도 설정

☒ 희망선도 범례설정

☒ 분포도 범례설정

표출 개수 : 30

출력 확인

① 출발지 및 도착지 설정

- 2개의 비교지역 설정
 - 유입 : 도착지
 - 유출 : 출발지
 - 레이어 ON/OFF 설정

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

③ 표출지도 설정

- 희망선도 범례 설정
- 분포도 범례 설정
- 표출지도 ON/OFF
- 표출개수 설정

유출량 분석 시 설정 UI

유출량 분석 시 설정 UI

유입량 유출량

동행량 검색 조건 설정

출발지 선택

선택단위 : 분석구

☒ 출발지 (1) 선택안됨

지도 선택 ID 검색

☒ 출발지 (2) 선택안됨

지도 선택 ID 검색

도착지

집계단위 : 분석구

상세설정

기간 : 2016-01-01 ~ 2016-12-31

시간대 : 출발 전체 시간 도착 전체 시간

트립타입 : 출발 전체 트립 도착 전체 트립

성별/연령(대) : 성별 남,여 연령 전체 연령

표출지도 설정

☒ 희망선도 범례설정

☒ 분포도 범례설정

표출 개수 : 30

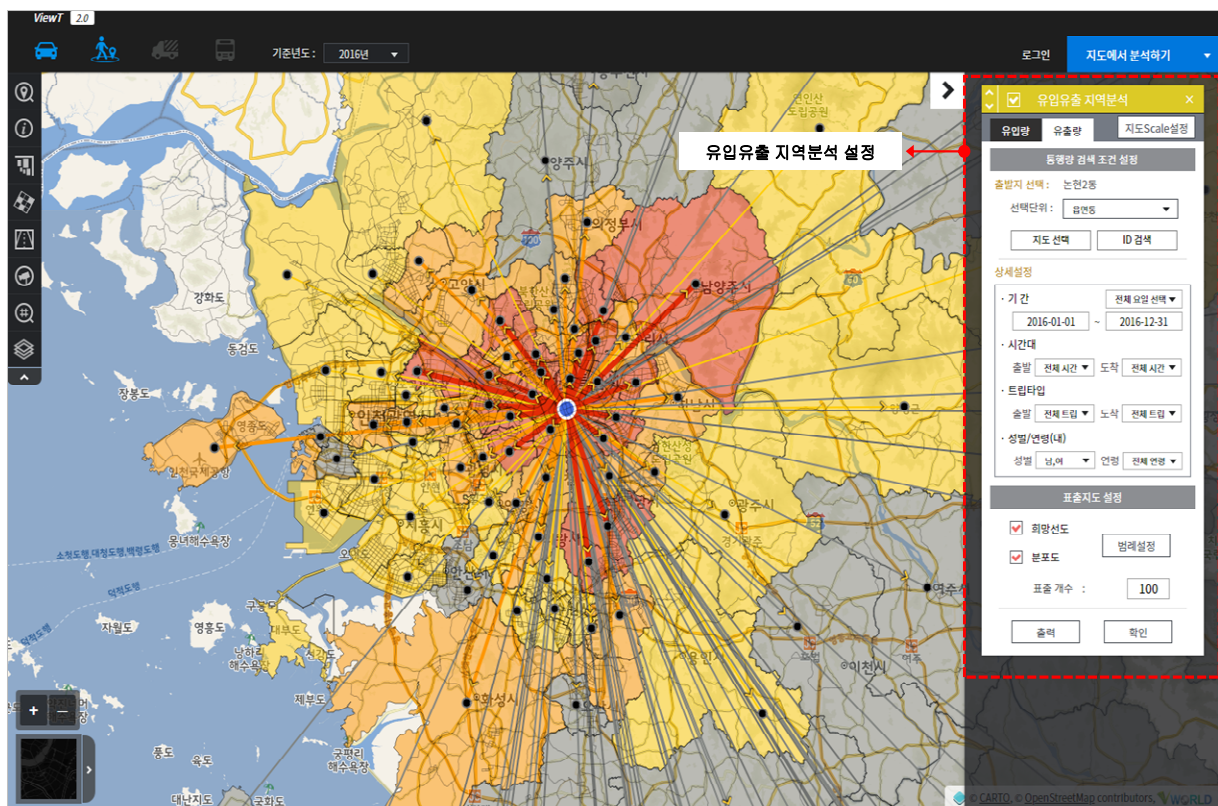
출력 확인

<그림 5> 유입유출 비교분석 설정 UI

- 사용자가 분석결과를 쉽게 확인할 수 있도록 다양한 ON/OFF 기능을 제공함
 - 선택 지역에 대한 레이어 ON/OFF 기능을 개발함
 - 희망선도 및 분포도에 대한 ON/OFF 기능을 개발함
- 지도에 표출되는 분포도, 희망선의 개수를 설정할 수 있도록 개발함

② 유입유출 지역분석

- 전국의 시도, 시군구, 읍면동, 분석구에서 특정지역으로 유입되는 통행량과 특정지역에서 전국의 시도, 시군구, 읍면동, 분석구로 유출되는 통행량을 분석할 수 있도록 개발함
- 시도, 시군구, 읍면동, 분석구 단위로 집계된 유입 통행량, 유출 통행량을 지도의 ZOOM 단계에 따라 지도에 표출되도록 개발함
- 유입 통행량, 유출 통행량의 분석결과에 대한 시도, 시군구, 읍면동, 분석구 단위의 표출지도가 표현되는 ZOOM 단계를 사용자가 설정할 수 있도록 개발함
- 기간, 요일, 출발 및 도착시간대, 출발 및 도착 트립타입, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 분석결과를 희망선도, 분포도로 표출할 수 있도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 유입유출 지역분석에 활용된 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 6> 유입유출 지역분석의 실행 화면

- 선택단위를 통해서 선택된 시도, 시군구, 읍면동, 분석구에 대하여 유입량, 유출량 분석이 가능하도록 UI를 구성함
- 분석결과가 반영된 시도, 시군구, 읍면동, 분석구 단위의 지도가 표출되는 ZOOM 단계를 사용자가 설정할 수 있도록 개발함
- 검색조건 설정을 통해서 분석대상 통신데이터를 설정할 수 있도록 개발함
- 분포도와 희망선도로 표출되는 지도의 범례를 설정할 수 있도록 개발함
- 지도에 표출되는 분포도, 희망선의 개수를 설정할 수 있도록 개발함

■ 유입량 분석 시 설정 UI

The UI for Inflow Analysis Settings is divided into several sections. At the top, there are tabs for '유입량' (Inflow), '유출량' (Outflow), and '지도Scale설정' (Map Scale Setting), with '유입량' selected. Below this is a '동행량 검색 조건 설정' (Companion Search Condition Setting) section with a '도착지 선택' (Destination Selection) dropdown set to '선택안됨' (Not Selected) and a '선택단위' (Selection Unit) dropdown set to '분석구' (Analysis Area). There are buttons for '지도 선택' (Map Selection) and 'ID 검색' (ID Search). The '상세설정' (Detailed Setting) section includes a '기간' (Period) with dates '2016-01-01' and '2016-12-31', a '시간대' (Time Zone) dropdown, and a '출발' (Departure) dropdown set to '전체 시간' (All Time). It also has a '도착' (Arrival) dropdown set to '전체 시간' (All Time), a '트립타입' (Trip Type) section with '출발' (Departure) and '도착' (Arrival) dropdowns set to '전체 트립' (All Trip), and a '성별/연령(대)' (Gender/Age Group) section with '성별' (Gender) set to '남,여' (Male, Female) and '연령' (Age) set to '전체 연령' (All Age). The '표출지도 설정' (Map Display Setting) section has checkboxes for '희망선도' (Desired Line Map) and '분포도' (Distribution Map), both checked, and a '범례설정' (Legend Setting) button. Below this is a '표출 개수' (Display Count) input field set to '100'. At the bottom are '출력' (Output) and '확인' (Confirm) buttons.

① 표출지도 ZOOM 단계 설정

- 시도, 시군구, 읍면동, 분석구 단위의 표출지도

② 출발지 및 도착지 설정

- 유입 : 도착지
- 유출 : 출발지

③ 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

④ 표출지도 설정

- 희망선도 범례 설정
- 분포도 범례 설정

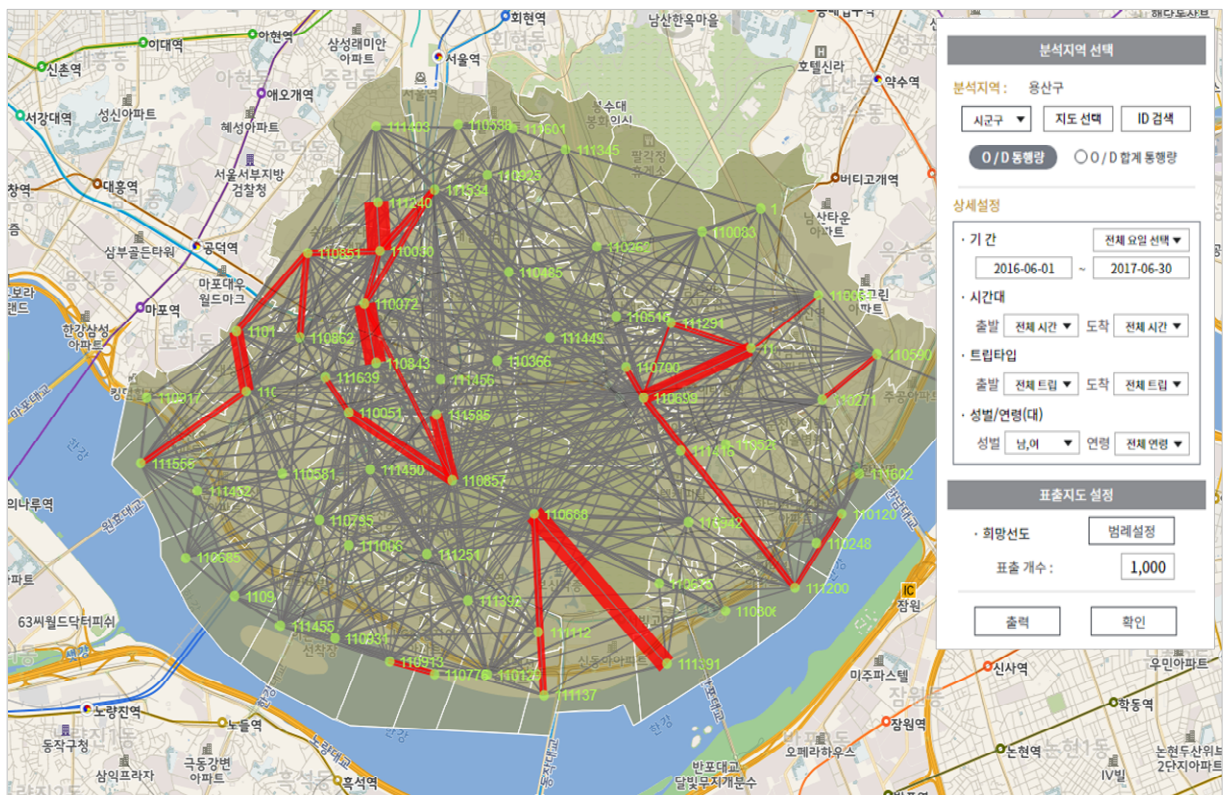
■ 유출량 분석 시 설정 UI

The UI for Outflow Analysis Settings is similar to the Inflow UI. It has tabs for '유입량' (Inflow), '유출량' (Outflow), and '지도Scale설정' (Map Scale Setting), with '유출량' selected. The '동행량 검색 조건 설정' (Companion Search Condition Setting) section has a '출발지 선택' (Departure Selection) dropdown set to '110580' and a '선택단위' (Selection Unit) dropdown set to '분석구' (Analysis Area). There are buttons for '지도 선택' (Map Selection) and 'ID 검색' (ID Search). The '상세설정' (Detailed Setting) section includes a '기간' (Period) with dates '2016-01-01' and '2016-12-31', a '시간대' (Time Zone) dropdown, and a '출발' (Departure) dropdown set to '전체 시간' (All Time). It also has a '도착' (Arrival) dropdown set to '전체 시간' (All Time), a '트립타입' (Trip Type) section with '출발' (Departure) and '도착' (Arrival) dropdowns set to '전체 트립' (All Trip), and a '성별/연령(대)' (Gender/Age Group) section with '성별' (Gender) set to '남,여' (Male, Female) and '연령' (Age) set to '전체 연령' (All Age). The '표출지도 설정' (Map Display Setting) section has checkboxes for '희망선도' (Desired Line Map) and '분포도' (Distribution Map), both checked, and a '범례설정' (Legend Setting) button. Below this is a '표출 개수' (Display Count) input field set to '50'. At the bottom are '출력' (Output) and '확인' (Confirm) buttons.

<그림 7> 유입유출 지역분석의 설정 UI

③ O/D분석

- 사용자가 분석지역을 선택하여 분석구간의 통행량을 분석할 수 있도록 개발함
- 분석지역은 행정구역 단위로 선택할 수 있도록 개발함
- 분석구 간의 단방향, 양방향 통행량을 분석할 수 있도록 개발함
- 표출지도는 분석구 노드를 직선으로 연결한 희망선으로 표출되도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 검색조건에 해당되는 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 8> O/D 분석결과의 화면 예시(상위 10% 강조)

- 분석지역을 시도, 시군구, 읍면동 단위로 선택하여 O/D 분석을 할 수 있도록 개발함
- 분석지역에 포함된 분석구 간의 O/D 통행량을 분석할 수 있도록 개발함
- 단방향 통행량, 양방향 통행량을 사용자 선택해서 분석할 수 있도록 개발함
- 기간, 요일, 출발 및 도착시간대, 성별, 연령대에 대하여 사용자가 검색조건을 설정

할 수 있도록 개발함

- 출발 분석구 노드와 도착 분석구 노드를 직선으로 연결한 희망선으로 분석결과를 지도에 표출할 수 있도록 개발함
- 희망선도의 범례를 설정할 수 있도록 개발함

분석지역 선택

분석지역 : 성남시분당구

시군구 ▼ 지도 선택 ID 검색

① ☒ O/D동행량 ☐ O/D합계동행량

상세설정

· 기간 ~ 전체 요일 선택 ▼

· 시간대 ② 출발 ▼ 도착 ▼

· 트립타입 출발 ▼ 도착 ▼

· 성별/연령(대) 성별 ▼ 연령 ▼

표출지도 설정

③ · 희망선도

표출 개수 :

① 분석지역 선택 및 통행량 산출방법 설정

- 분석지역 : 시도, 시군구, 읍면동 선택
- O/D 통행량 : 단방향 통행량 산출
- O/D 합계 통행량 : 양방향 통행량 산출

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

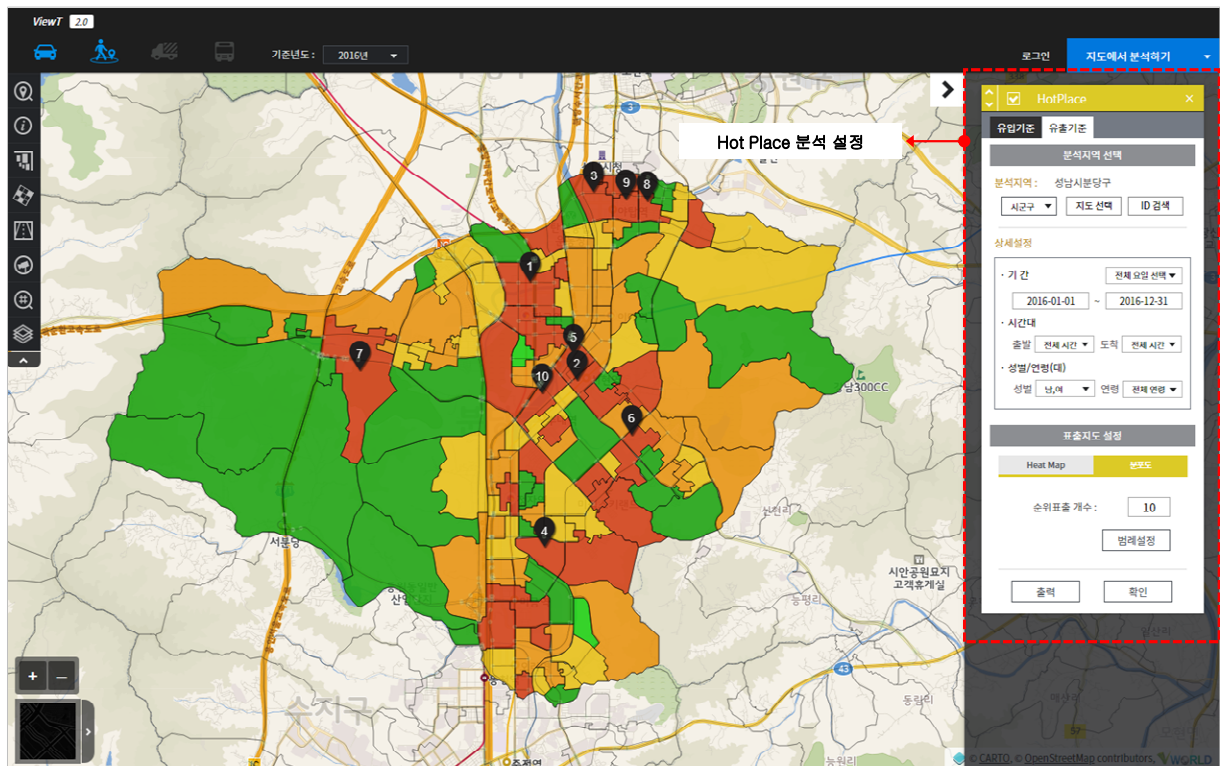
④ 표출지도 설정

- 희망선도 범례 설정
- 희망선도 표출 개수 설정

<그림 9> O/D 분석의 설정 UI

④ Hot Place 분석

- 사용자가 선택한 특정지역의 범위 내에서 사람이 가장 많이 몰리는 곳, 외부로 가장 많이 나가는 곳을 분석할 수 있도록 개발함
- 분석구 단위로 분석결과를 지도에 표출할 수 있도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 검색조건에 해당되는 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 10> Hot Place 분석의 실행 화면

- 분석지역을 시도, 시군구, 읍면동, 분석구로 선택하여 Hot Place 분석을 할 수 있도록 개발함
- 유입기준 및 유출기준의 Hot Place 분석이 가능하도록 개발함
- 기간, 요일, 출발 및 도착시간대, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 분석결과를 Heat Map, 분포도로 지도에 표출하고 표출되는 지도의 형태를 사용자가 설정할 수 있도록 개발함

유입기준 분석 시 설정 UI

유입기준

유출기준

분석지역 선택

분석지역: 선택안됨

① 시군구 ▼

지도 선택

ID 검색

상세설정

· 기간

전체 요일 선택 ▼

2016-01-01 ~ 2016-12-31

· 시간대

출발 전체 시간 ▼ 도착 전체 시간 ▼

· 성별/연령(대)

성별 남,여 ▼ 연령 전체 연령 ▼

표출지도 설정

Heat Map

분포도

③ 반경 1 — 25 — 50

흐림 1 — 30 — 50

출력

확인

① 분석지역 선택

- 시도, 시군구, 읍면동

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

③ 표출지도 설정

- HeatMap 범례 설정
- 분포도 범례 설정

유출기준 분석 시 설정 UI

유입기준

유출기준

분석지역 선택

분석지역: 서울특별시

시도 ▼

지도 선택

ID 검색

상세설정

· 기간

전체 요일 선택 ▼

2016-01-01 ~ 2016-12-31

· 시간대

출발 전체 시간 ▼ 도착 전체 시간 ▼

· 성별/연령(대)

성별 남,여 ▼ 연령 전체 연령 ▼

표출지도 설정

Heat Map

분포도

순위표출 개수: 10

범례설정

출력

확인

<그림 11> Hot Place 분석의 설정 UI

⑤ 통행량(인구수) 검색

- 본 연구에서는 사용자가 사용목적에 따라 통신데이터를 직접 가공할 수 있도록 통신 데이터를 검색할 수 있는 기능을 개발함
- 검색된 결과를 파일형태(csv)로 출력할 수 있도록 개발함
- 사용자가 기간, 요일, 출발시간대, 도착시간대, 성별, 연령대 조건을 설정하여 통신 데이터를 검색하고 출력할 수 있도록 개발함
- 요일, 출발시간대, 도착시간대, 성별, 연령대의 설정은 다중선택이 가능하도록 개발함

<그림 12> 통행량(인구수) 검색의 설정 UI

나. 기반 DB 구축

1) 통신데이터 분석 맵 레이어 구축

① 분석구 레이어 구축

- 지역 : 전국
- 형상 : 폴리곤
- 개체 수 : 15,937
- 좌표계 : WGS84 경위도 좌표계
- 용량 : 약 108 MB
- 데이터형식 : shp

<표 9> 수집된 분석구 데이터의 속성정보

NO	필드명	설명	데이터 타입	비고
1	FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	KEY	교통폴리곤 ID	DOUBLE	-
4	행정동 ID	읍면동 코드	STRING	-



- 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 13> 분석구 레이어의 구축 절차

- GeoServer, Openlayers 기반의 웹 GIS 형태로 서비스를 개발하였음
 - GeoServer : GIS 서버 개발에 적용
 - Openlayer : 클라이언트 서비스 개발에 적용

- 수집된 분석구 데이터를 활용하여 GeoServer 형식의 분석구 폴리곤 레이어와 분석구 노드 레이어를 구축함
- 분석구 폴리곤 레이어는 유입유출 지역분석 및 비교분석, OD 분석 등에서 분석결과에 대한 분포도 생성의 기반 공간정보로 활용됨
- 분석구 노드 레이어는 희망선 지도표출, Hot Place 분석결과와 지표표출 등에서 기반 공간정보로 활용됨
- 분석구 노드는 분석구 폴리곤의 대표 포인트(노드)를 의미함

② 행정구역 레이어 구축

- 지역 : 전국
- 형상 : 폴리곤
- 개체수 : 3,771
- 좌표계 : Korean 1985 Katech 좌표계
- 용량 : 약 84.3 MB
- 데이터형식 : shp

<표 10> 수집된 행정구역 데이터의 속성정보

NO	필드명	설명	데이터 타입	비고
1	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 분석구 레이어 구축과 동일한 절차에 따라 구축하였음
- 수집된 행정구역 데이터를 활용하여 GeoServer 형식의 폴리곤 레이어와 노드 레이어를 구축함
- 시도, 시군구, 읍면동에 대한 각각의 폴리곤 레이어와 노드 레이어를 구축함
- 행정구역 폴리곤 레이어는 분석구 폴리곤 레이어와 같이 유입유출 지역분석 및 비교 분석, OD 분석 등에서 분석결과에 대한 분포도 생성의 기반 공간정보로 활용됨
- 행정구역 노드 레이어 또한, 희망선 지도표출, Hot Place 분석결과의 지표표출 등에서 기반 공간정보로 활용됨

2) 모바일 통행행태 기반 DB 구축

- 지역 : 전국
- 기간 : 2016. 03. 07. ~ 2016. 12. 31.
- 용량 : 약 555 GB
- 일별로 구분된 통신빅데이터를 수집하였으며 데이터 형태는 csv로 되어 있음

<표 11> 통신빅데이터의 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



<그림 14> 모바일 통행형태 기반 DB의 구축 절차

- 본 연구에서는 데이터의 검색 및 처리속도, 빅데이터 관리의 원활함을 위해서 수집된 데이터를 가공하여 MongoDB 기반의 빅데이터 DB를 구축함
 - MongoDB의 데이터셋을 다수의 데이터베이스에 분산 저장하는 Sharding 기법을 적용하여 구축함
- 수집된 데이터의 건수(도큐먼트 수)가 매우 커서 가공 없이 DB를 구축할 경우 데이터 검색 및 처리 속도의 저하로 정상적인 서비스를 제공하는데 한계가 있으므로, 본 연구에서는 수집된 데이터의 도큐먼트 수를 경량화 시킨 분석구 기준의 모바일 통행형태 기반 DB를 구축하여 정상적인 서비스를 제공할 수 있도록 함
- 또한 이동인구수의 분석을 행정구역 단위로 하는 경우 행정구역에 포함되는 분석구의 이동인구수를 검색하고 처리하는데 많은 시간이 소요되므로, 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 복합 인덱스를 생성함(시도, 시군구, 읍면동, 분석구)

다. 시스템 운영 및 유지보수

1) 시스템 운영

- Web Server는 Apache HTTP Server를 사용하며, Web Application Server는 Apache Tomcat을 사용함
 - Apache HTTP Server와 Apache Tomcat는 윈도우즈 서비스로 등록되어 시스템 부팅 시 자동으로 실행되나 프로그램의 오류 등으로 중지되어 부득이하게 수동으로 실행 시에는 수동 실행해야 함
- 시스템 운영 중 발생할 수 있는 장애에 대응하기 위해서 DB 데이터 파일, 로그파일을 별도의 드라이브로 복사하여 백업하거나 DB관리 툴의 백업 기능을 이용하여 백업함
 - 시스템의 응용프로그램 백업 대상인 프로그램 실행파일, 환경설정 파일을 포함한 데이터베이스의 백업, 복구 절차는 다음 <표 13-12>와 같음

<표 12> 백업 및 복구 절차

범위	수행단계	설명
백업	백업 수행 및 모니터링	• 백업 담당자는 계획된 백업에 대해 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 해결 한 후 백업을 재수행
	백업관리대장 반영	• 백업 담당자는 백업 결과를 관리대장에 반영하고, 백업 중 에러가 발생한 경우 해당 내용에 대한 조치경과를 기록
백업	백업 종료	• 백업 종료 후 백업 결과를 확인하고 현황을 관련자에게 통보
복구	복구요청 접수	• 데이터의 복구가 필요한 경우 백업담당자에게 복구요청을 신청
	복구계획 수립	• 백업 담당자는 복구 요청을 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 이상유무를 확인하고 결과를 관련자에게 통보
	후속조치	• 백업담당자 및 관리자는 검증결과에 이상이 있을 경우 복구나 백업장비 교체 등의 후속조치를 수행

2) 시스템 유지보수

- 본 사업 완료 후 유지보수 기간 동안 도입 시스템의 안정적이고 효율적인 시스템 운영, 신속한 장애처리 서비스를 위하여 개발 참여인력들로 운영지원팀을 구성하여 안정적인 유지보수 서비스를 제공함
- 장애발생에 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 문제발생 시 신속히 대처함

5. 결론 및 향후과제

가. 결론

본 과업에서는 교통 분석에 모바일 DB가 더욱 활발히 활용될 수 있도록 '17년 개발한 모바일 가공 알고리즘과 모바일 분석 맵 구축 알고리즘을 보완하였으며, 이를 기반으로 17년 기준 모바일 DB를 구축하였음

- 핑퐁 핸드오버 데이터의 주요 발생원인(야간시간대 전압 조절로 인한 신호 탐색)을 고려하여 하루를 넘어가면서 발생하는 핑퐁 데이터가 최대한 보정될 수 있도록 탐색 기준을 일 단위에서 일주일 단위로 변경함
- 사람의 이동특성을 파악하는데 필수 요소인 '통행목적'정보를 '17년에 개발한 DB에 추가하여 모바일 DB를 구축함
 - 모바일 DB를 통해 통행목적을 분류하려면 기점과 종점의 기준이 되는 체류지 유형을 기존보다 구체화 시킬 필요가 있어, 심야시간대 주체류지, 낮 시간대 주체류지, 잠재체류지 세 가지로 구분되었던 유형을 각각 나누어 낮 시간대 주체류지는 집과 집 이외 심야시간대 주체류지로, 낮 시간대 주체류지는 회사와 학교로 잠재체류지는 종교집회장소과 기타로 세분류함
 - 통행 로그에서 각 체류지를 구분하기 위해 공간적 특성이 개인의 통행과 활동에 영향을 줄 수 있다는 가정 하에 주로 방문하는 시간대, 평균 체류 시간, 방문 빈도, 주요 통행자(연령 특성 반영), 관심지점정보 위치 여부를 기준으로 체류지를 구분하는 기준을 개발함
 - 이를 통해 추출할 수 있는 통행목적 유형은 총 5가지(출근 / 등교 / 여가 / 귀가 / 기타)이며, 도착지에 따라 귀가 목적 통행은 '퇴근', '하교', '귀가'로 기타

- 목적 통행은 ‘종교 활동’, ‘친지 방문’ ‘기타’로 세분 가능함
- 통행 목적이 구분된 각 개인별 로그 기록은 「개인정보보호」, 「위치정보의 보호 및 이용 등에 관한 법률」에 저촉되지 않도록 집계한 형태의 DB를 설계하였으며, 이를 기준으로 2017년 1월 1일부터 2017년 12월 31일까지 발생한 로그 기록을 가공하여 DB를 구축함
 - 전국 약 15,937개로 나뉘었던 교통폴리곤(‘17년 기준)은 주·기지국 기준을 변경하고, 교통폴리곤 최소면적 기준을 강화하여 16,335개로 세분화 되었으며, 이로 인해 시도별 교통폴리곤 면적은 평균 약 6.8% 감소하였음
 - 주·기지국 기준을 한 달 동안 25일 이상 신호가 기록된 기지국에서 1년 365일 이상 신호가 기록된 기지국으로 변경하고, 교통폴리곤 최소면적을 시도별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하에서 시군구별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하로 변경함
 - 부산광역시, 경기도, 제주특별자치시를 제외하고 나머지 지역은 모두 ‘17년도 보다 세분되었으며, 특히 세종특별자치시, 울산광역시, 강원도, 충청북도, 전라북도 지역이 많이 개선된 것으로 나타남
 - 또한 ‘17년 개발된 교통 모니터링·데이터 제공·분석 플랫폼(View-T)를 통해 모바일 자료 기반의 분석 서비스가 제공될 수 있도록 기반 정보 DB를 설계하고, 모바일 DB 기반 분석 기능(통행자 기반의 O/D 분석, 유입유출 분석, Hot Place 분석 등)을 개발하였음

<표 13> View-T 2.0 모바일 자료 기반 분석 서비스 개발 주요 내용

구분		내용
데이터구축	통신데이터 분석 맵 레이어	<ul style="list-style-type: none"> · 분석구(교통폴리곤) 구축 · 분석구 단위의 노드 구축 · 행정구역 단위의 폴리곤/노드 구축
	모바일 통행형태 기반 DB 구축	<ul style="list-style-type: none"> · 서비스 기능 구현 및 확장성을 고려한 DB 설계 · 검색속도 향상을 위한 MongoDB 기반의 DB 설계 · 통신 빅데이터 기반의 통행량 DB 구축

통신데이터 기반의 서비스 기능 개발	O/D 분석	<ul style="list-style-type: none"> • 분석지역에 대한 OD 분석 • 행정구역(시도, 시군구, 읍면동) 단위로 분석지역 설정 • 분석구 단위로 OD 분석 • 단방향, 양방향 통행량 분석 • OD 희망선 표출지도 생성 및 표출지도 범례 설정 등
	유입유출 지역분석	<ul style="list-style-type: none"> • 특정지역에 대한 유입유출 통행량 분석 • 시도, 시군구, 읍면동, 분석구 단위로 특정지역 설정 • 전국 시도, 시군구, 읍면동, 분석구에서 특정지역으로 유입되는 통행량을 분석 • 특정지역에서 전국의 시도, 시군구, 읍면동, 분석구로 유출되는 통행량 분석 • 분석결과에 대한 분포도(폴리곤) 생성 • 도착지-출발지(유입), 출발지-도착지(유출) 희망선 표출지도 생성 • 분포도, 희망선 표출지도 범례 설정 등
	유입유출 비교분석	<ul style="list-style-type: none"> • 2지역의 유입유출 통행량 분석 • 시도, 시군구, 읍면동, 분석구 단위로 2지역으로 유입되는 통행량을 비교분석 • 2지역에서 시도, 시군구, 읍면동, 분석구 단위로 유출되는 통행량을 비교분석 • 분포도, 희망선 표출지도 생성 및 표출지도 범례 설정 등
	Hot Place 분석	<ul style="list-style-type: none"> • 시도, 시군구, 읍면동을 대상으로 특정조건에서 유입유출되는 통행량의 분포(도) 분석 • 기간, 연령별, 성별, 시간대 등에 따른 유입유출 통행량을 분석구 단위로 분석 • 분석결과에 대한 분포 및 순위 표출지도 생성 • 표출지도에 대한 범례설정 등
통행량(인구수) 데이터 검색		<ul style="list-style-type: none"> • 검색조건에 따른 일별 통행량 데이터 검색 • 검색조건 : 기간, 연령별, 성별, 시간대, 트립타입 • 검색결과에 대한 파일 출력 등

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 야간시간대 발생하는 핑퐁 핸드오버 데이터를 추가적으로 보완할 방안을 모색해야 할 것이며, 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, 금년도에 개발된 모바일 DB를 기준으로 교통 모니터링·데이터 제공·분석 플랫폼(View-T)에 제공할 분석 서비스가 추가적으로 개발될 필요가 있으며, 서비스 만족도를 높이기 위해 검색, 표출 속도를 개선할 필요가 있음
 - 새로 추가된 통행목적 정보를 기준으로 통행목적별 유출/유입 분석, 통행목적별 주요 통행구간(O/D) 분석 등 새로운 분석 기능을 개발할 수 있음

제1장 과업의 개요

제1절 과업의 배경 및 목적

제2절 과업의 범위 및 내용

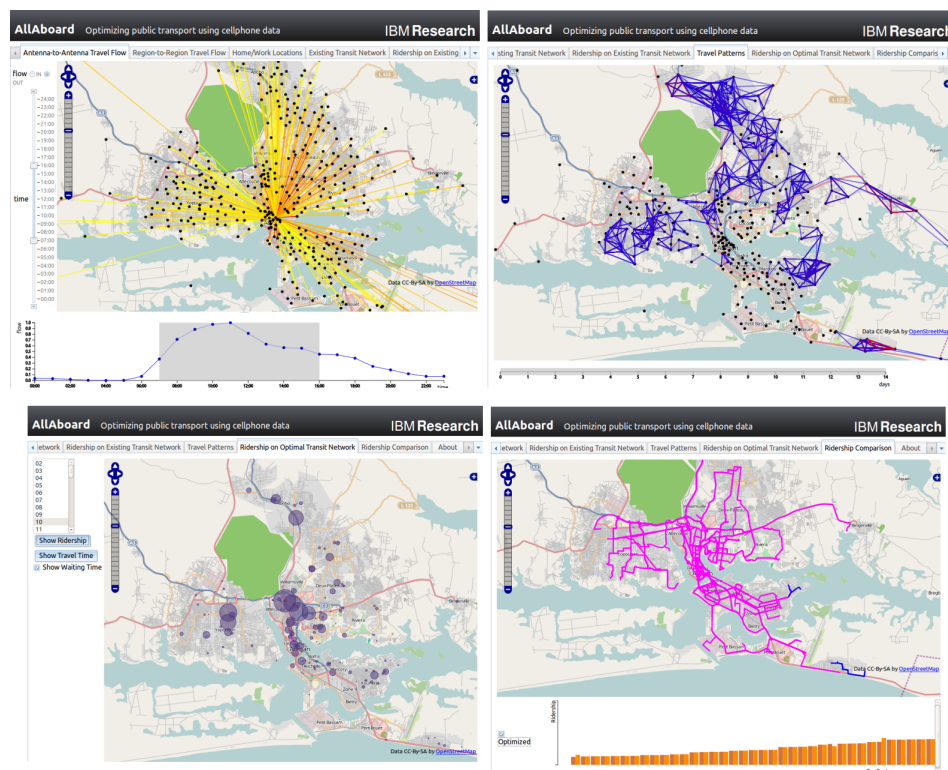
제3절 과업 추진 방법

제1장 과업의 개요

제1절 과업의 배경 및 목적

1. 과업의 배경

- 전 세계적으로 대중교통 및 비동력 수단 모니터링을 위한 플랫폼 개발이 활발히 이루어지고 있으며, 이를 기반으로 고도화된 대중교통 정책을 지원 중임
- － 플랫폼을 활용하여 버스 노선을 탄력적으로 운영하거나, 연령별, 성별 등 인적 특성 기반 통행패턴을 모니터링하여 이용자 맞춤형 대중교통 서비스를 제공 중



자료: Michele Berlingiero et al., AllAboard: a system for exploring urban mobility and optimizing public transport using cellphone data. p. 3-4.

<그림 1- 1> 모바일 데이터 기반 대중교통 최적화 시스템 (AllAboard)

- 그러나 현재 국내에는 표준화된 대중교통 및 비동력 수단 모니터링 시스템 부재한 실정

- 교통은 파생수요라는 측면에서, 신(新) 교통수단 등 다양한 수단을 고려한 교통정책을 지원하기 위해서는 ‘사람의 이동’을 파악하는 것이 중요하나 기존의 모니터링 시스템은 차량 중심으로 구축됨
- View-T도 내비게이션 통행패턴 데이터와 교통량 조사 데이터를 기반으로 차량 통행 분석 도구만 제공

2. 과업의 목적

- 본 과업의 목적은 사람의 이동정보가 담긴 통신자료를 활용해 대중교통 정책 지원에 직접적으로 활용 가능한 대중교통 및 비동력 수단 모니터링 시스템의 기반을 구축하는 것임
- 향후 대중교통 모니터링 시스템으로 확장할 수 있도록 교통카드 자료 연계를 고려하여 지리정보체계 기반 통신자료 가공 기술을 개발하여 통신자료 기반 DB를 구축하고, 대중교통 및 비동력 수단 이동 모니터링 및 분석이 가능하도록 핵심 통계 지표 및 분석 기술을 개발하여 기 개발된 플랫폼(View-T)에 구축하고자 함

제2절 과업의 범위 및 내용

1. 과업의 범위

- 시간적 범위
 - － 모바일 자료 가공 기술 개발 및 2017년도 기준 기반 DB 구축: 2017년 1월 1일 ~ 2017년 12월 31일 (1년)
 - 2017년 4월 한 달간의 원천데이터를 활용하여 가공 알고리즘을 개발한 후, 2017년에 생성된 시그널 데이터에 개발한 가공 알고리즘을 적용
 - － 웹 서비스를 위한 각종 분석도구 개발: 2016년 3월 7일 ~ 2016년 12월 31일
 - 2017년도 연구 성과물인 2016년도 기준 기반 DB 활용
- 공간적 범위: 전국

2. 과업의 내용

1) 모바일 자료 기반 DB 구축

- 유동인구 이동궤적 자료 전처리 기술 개발
 - － 통신 원시자료에서 지역별, 연령대별, 성별 등 다양한 특성을 고려한 이동궤적 자료 전처리 방법론 개발 및 적용 방안 제시
- 통행목적 구분 로직 개발
 - － 유동인구 통행 패턴 특성 분석을 통해 개인정보보호법에 위반되지 않는 범위 내에서 목적 O/D 정보 추정 및 제공 방안 개발
- 기반 DB 설계 및 '17년 기준 기반 DB 구축
 - － 분석구 Polygon 네트워크와 통행사슬 정보기반의 O/D DB 산출
 - － 이용자 맞춤형 특성별(주중·주말, 침두·비침두) 쿼리형 O/D 정보 제공

2) 모바일 자료 분석 맵 구축

- 분석 맵 구축 알고리즘 고도화
- '17년 기준 분석 맵 구축

3) View-T 모바일 자료 기반 분석 서비스 개발

- GIS 서비스 레이어 구축
 - 웹 서비스를 위한 분석구 데이터 경량화 작업
 - 웹 서비스를 위한 분석구 단위의 노드 구축
 - 분석구 단위의 검색 POI 구축
- 웹 서비스를 위한 모바일 통행행태 기반 DB 구축
 - 데이터 분석 및 DB 설계
 - 웹 서비스 기능 분석 및 이를 기반으로 한 DB 설계
 - 주중·주말, 시간단위 서비스를 고려한 DB 설계
 - 년 단위 서비스를 위한 DB 설계
 - DB 설계 기반의 데이터 구축을 위한 가공 시스템 개발
 - 1일 단위 데이터를 월 단위/주중·주말·특송/시간단위 데이터로 구축
 - 년 단위 데이터 병합
 - 분석구 단위의 시간대별 O/D 구축
 - 분석구 단위의 통행사슬 구축
 - 웹서비스용 데이터 구축 및 업로드
 - 데이터 검증
- 통행행태 기반 데이터 압축
 - 웹 서비스를 위한 데이터 정제
 - 웹 서비스 속도 향상을 위한 데이터 압축(경량화)(검색속도 향상을 위한 Mongo DB 기반의 데이터 구축)
 - 데이터 검증
- 분석 서비스 기능 개발
 - 기능 정의 및 UI 설계

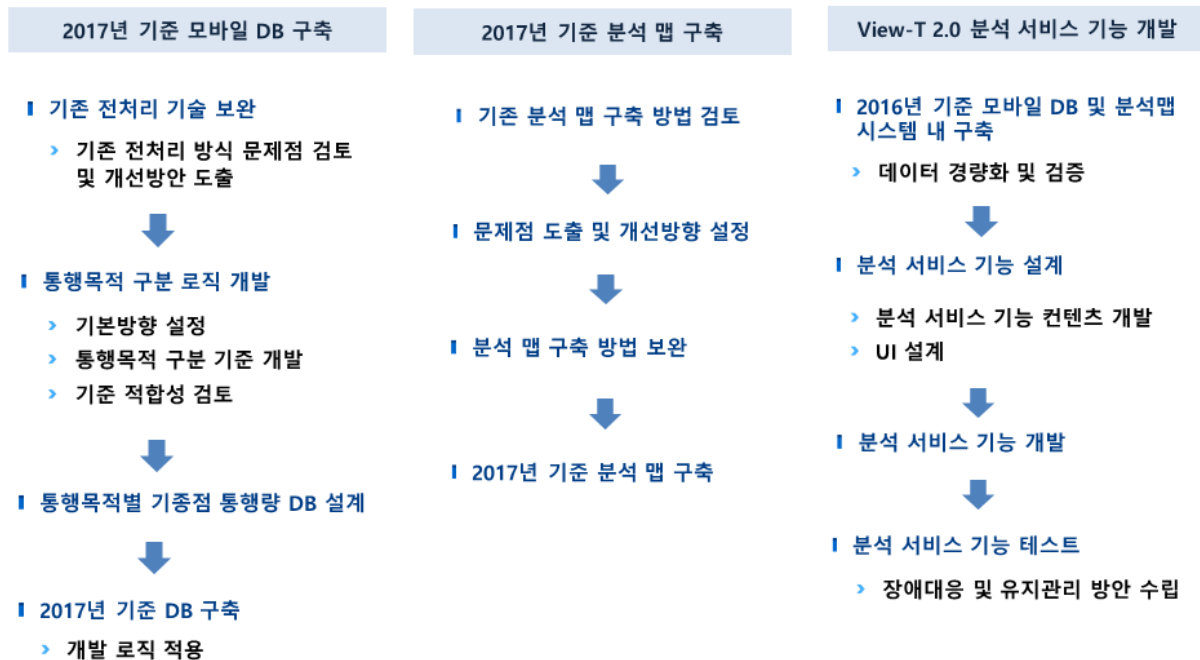
- O/D 서비스 기능 개발
 - O/D에 대한 분석구/읍면동/시군구/시도 단위 데이터 표출 기능 구현
 - 시공간/성별/연령대별 표출 기능 구현
 - 범례 설정 기능 구현
 - 출력 기능 구현
- 이동생활패턴 분석 서비스 기능 개발
 - 이동생활패턴에 대한 분석구/읍면동/시군구/시도 단위 데이터 표출 기능 구현
 - 시공간/성별/연령대별 표출 기능 구현
 - 범례 설정 기능 구현
 - 출력 기능 구현
- 시간대별 대중교통·유동인구 인적특성별 Hot Place 분석 서비스 기능 개발

제3절 과업 추진 방법

1. 과업 추진 방법

- 모바일 원천데이터 보유 기관에서 구축한 분석 환경 내에서 모바일 데이터 가공 로직을 개발한 후 전국 단위로 해당 로직을 적용하여 DB 구축
 - － 통신 분야 전문가와 교통 분야 전문가가 함께 로직 개발에 참여
 - 모바일 데이터를 가공하기 위해서는 모바일 데이터 생성원리, 특성 등 모바일 데이터에 대한 이해도가 높은 통신 분야 전문가 참여 필요
- 구축된 모바일 DB 기반으로 특성(연령별, 성별, 기간별, 분석구별 분석 가능)을 부각시킬 수 있는 방안을 모색하여 웹 서비스 기능 개발
 - － 가공·분석·정제 등 플랫폼 내 빅데이터 처리에 대해 지식을 갖춘 전문가 참여

2. 과업 추진 절차



<그림 1- 2> 과업 추진 절차

제2장 모바일 자료 기반 DB 구축

제1절 모바일 자료 가공 기술 개발

제2절 모바일 자료 기반 DB 구축

제2장 모바일 자료 기반 DB 구축

제1절 모바일 자료 가공 기술 개발

1. 전처리 기술 개발

가. 기존¹⁾ 전처리 기술

1) 이상치 제거

- 기지국 특성상 좌표 칼럼에 \N으로 표기되는 NULL 값
- 기지국 좌표에 행정동 코드가 매칭 되지 않는 데이터
 - 통신업체에서 행정동 코드를 부여하는 영역 밖에 위치한 기지국의 경우 행정동이 NULL값으로 입력됨
 - 행정동 코드가 부여되는 영역: 최소값 $x=124.529$, $y=33.144$, 최대값 $x=130.942$, $y=38.616$

2) 체류시간 보정

- 포인트 이력을 선 이력으로 변환하는 과정에서 분리된 데이터 통합
 - 실제로 같은 위치에서 기록된 데이터지만 밀리세컨 단위로 기록된 포인트 정보를 1분 단위의 기록으로 잇는 과정에서 기지국 속성이 반영되지 않아 분리된 데이터를 연결
 - 연달아 발생한 데이터 중에서 위치 정보(기지국 좌표 정보)가 동일하지만 체류시간정보가 나뉜 경우 가장 마지막에 기록된 로그를 기반으로 체류종료시간을 업데이트 한 후, 이를 기준으로 총 체류시간(체류종료시간-체류시작시간) 정보를 업데이트 함

3) 왜곡 데이터 보정

- 통행자가 목적지로 이동하면서 기록되는 로그 이외에 신호 탐색과정에서 축적된 '핸드오버

1) 『2017년 국가교통조사』 중 제3권 모바일 자료 기반 통행수요 추정 및 교통지표 발굴 과업에서 제시한 전처리 기술을 의미함.

데이터'를 제거함

- 끊임 없는 통신서비스를 제공하기 위해 단말기는 신호 세기가 강한 주변 기지국을 계속 탐색하고, 적합한 기지국을 찾아 신호를 전환하는 핸드오버(Handover) 과정이 이루어지는데, 기지국간의 신호 세기가 비슷하거나 기지국 간의 거리가 가까운 경우 단말기에 실제 통행이력과 다른 로그가 기록됨
- 필요 이상의 핸드오버(Unnecessary Handover) 데이터로 인한 통행정보 왜곡: 속도가 빠른 교통수단을 탑승하여 이동경로 상의 두 기지국(또는 2개 이상의 기지국)이 동시에 기록되면서 이동경로가 하나로 형성되지 않음
- 핑퐁 핸드오버(Ping-pong Handover) 데이터로 인한 통행정보 왜곡: 야간 시간대 기지국 신호 강도가 조정되어 기지국간 신호 세기가 비슷한 경우 실제 움직이지 않더라도 이동한 것처럼 서로 다른 기지국 좌표가 연달아 발생함
- 하나의 단말기의 로그 기록 시작시간과 로그 종료시간이 동일하게 나타나는 필요 이상의 핸드오버 데이터는 최단거리로 통행하는 통행자 특성을 반영하여 기지국간의 거리가 최소가 되는 로그 기록만 제외하고 나머지 데이터는 제거함
- 기지국 거리가 동일한 경우 기지국 기록 빈도를 기준으로 하여 최다 기록 기지국만 제외하고 나머지 데이터를 제거함
- 1분 간격으로 기지국 좌표가 번갈아 나타나는 핑퐁 핸드오버 데이터는 기지국 기록 빈도와 체류시간 기준으로 핑퐁현상이 발생한 로그 기록을 추출하여 하나의 기록으로 보정함
- 핑퐁 현상은 로그 기록에서 기지국 좌표가 A-B(1분)-A' 형태로 나타나므로 이전 기지국과 다음 기지국이 동일하고,中间的 기지국에서의 체류시간이 1분인 경우를 추출하여 이러한 현상이 나타나는 시작 시점과 마지막 지점 정보를 종합하되 기지국 좌표는 가장 많이 기록된 기지국의 좌표를 활용함

나. 기존 전처리 방식의 한계점 및 개선방안

1) 기존 전처리 방식의 한계점

- 기존 핑퐁 핸드오버 데이터 보정 알고리즘은 핑퐁 현상을 1일 단위로 식별하도록 설계되어 있어 일자가 변경되는 때 발생한 핑퐁 현상을 보정하지 못함
- 기지국 데이터가 1일 단위로 갱신되고 있어 제한된 보정 알고리즘을 개발한 것임
- 핑퐁 현상 중 일부는 야간시간대 전압 조절로 인해 기지국간의 신호 강도가 비슷할 때 발생하므로 완벽하게 핑퐁 현상을 보정하기 위해서는 일자가 변경되는 야간시간대에 발생한 핑

풍 현상을 식별할 수 있는 기준이 필요함

- 기존 보정 알고리즘은 기지국의 위치가 서로 가까워 발생하는 핑퐁 현상을 보정하는 데에는 탁월하나, 야간시간대 발생한 핑퐁 현상을 식별하는 데에는 한계가 있음

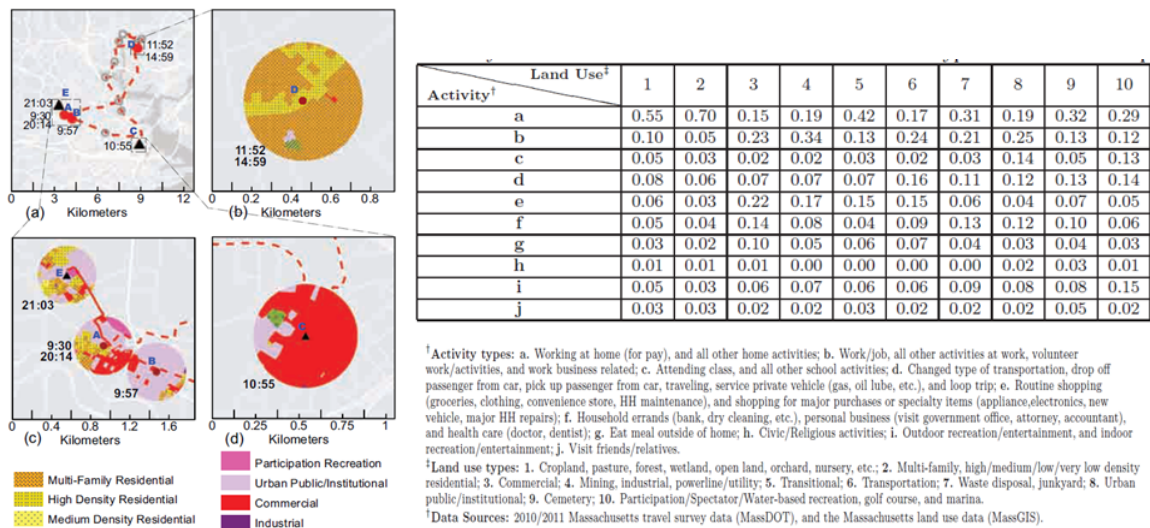
2) 개선방안

- 로그 기록을 추출하는 기준을 1일 단위(0시 1분부터 23시 59분)에서 일주일 단위(예: 1일 0시 1분부터 7일 23시 59분)로 변경
 - 야간에 발생하는 핑퐁 핸드오버 데이터 특성 반영이 가능해야 함
 - 핑퐁 핸드오버 보정 비율을 높이기 위해서는 기지국 데이터 로그가 지속적으로 연결되어야 함
 - 가공 효율성을 확보한 추출 기준을 확보해야 함
 - 축적된 원천 데이터 양이 방대하므로 기준을 넓게 한다면 가공 효율성이 떨어짐
 - 핑퐁현상 발생에 대한 탐색 기준과 체류지 식별 기준이 상이할 경우 체류지 식별이 불가능하기 때문에 기준이 최대한 매칭이 되도록 추출 기준을 설정
 - 주체류지는 주 단위의 식별 기준에 의해 식별되므로 한달 기준으로 핑퐁 현상을 가공할 경우에는 일주일 기간이 넘는 핑퐁 데이터들이 보완되면서 기준에 주체류지로 식별되었던 기지국이 주체류지에서 누락될 가능성이 존재함
 - 예) 4월 1일부터 4월 15일까지(약 2주간) 집에 머무른 고객의 데이터에 한 달 기준의 핑퐁 현상 보정 알고리즘을 적용할 경우 4월 1일부터 4월 15일까지의 데이터가 하나로 합쳐지면서 탐색조건(일주일, 예: 4월 1일부터 4월 7일까지)에 필터링 되지 않음

2. 통행목적 구분 추출 알고리즘 개발

가. 기본방향

- 모바일 자료를 기반으로 통행 목적을 추정하는 것은 기점과 종점의 활동을 추론하는 것에서 부터 시작되며, 기점과 종점의 활동은 <그림 2-1>과 같이 공간정보를 활용하여 추론하거나 기지국 기록 패턴을 활용하여 추론할 수 있음
- Jiang et al(2013)은 기지국 수신영역에 매칭 되는 토지용도를 기준으로 기지국이 기록된 지점에서의 활동을 추론
- 기지국 수신범위 내에 여러 유형의 토지용도가 함께 존재하는 경우, <그림 2-2>에 제시된 표처럼 토지용도별 활동 유형을 추정할 수 있는 값(토지용도별 활동 유형별 확률)²⁾을 토대로 추정하였으며, 토지용도 외에 관심지점정보(Point of Interest, POI)를 활용하는 방법도 있음



자료: Jiang, et al. (2013, August).

<그림 2-1> 활동 목적 추정 방법

- 공간정보를 활용하여 추론하는 방식은 다양한 건물이 밀집해 있는 도시부에는 적용하는데 한계점이 존재하므로 본 과업에서는 기지국 기록 패턴을 활용하여 체류지(기점과 종점)의 활동을 추론하는 방식으로 통행 목적을 구분하고자 함

2) 활동 유형과 목적지 정보가 포함된 조사 데이터를 기반으로 수행한 다항로지 회귀분석 결과임.

- 주로 방문하는 시간대, 평균 체류 시간, 방문 빈도 등을 고려하여 체류지를 구분
 - 통근, 통학과 같이 통행대상이 명확한 경우는 통행자 특성(연령)을 구분하여 기지국 기록 패턴을 파악함
- 추론 대상(체류지) 유형은 기존 가구통행실태조사의 통행 목적별 기종점 유형을 참조하여 재정의하고자 함

나. 체류지 유형

- 가구통행실태조사에서는 통행목적을 출근 / 등교 / 업무 / 쇼핑 / 학원 / 여가 / 기타 / 귀가 총 8가지로 구분하고 있음
- 통행목적별 출발지와 도착지를 정리해보면 출발지 또는 도착지가 될 수 있는 체류지 유형은 크게 집 / 집 이외의 휴식 장소(예: 친지 집) / 회사 / 학교 / 출장지 / 쇼핑몰 / 학원 / 여가·오락·친교 장소 / 기타(병원, 종교집회장소 등) 총 9가지라는 것을 알 수 있음

<표 2- 1> 가구통행실태조사 목적 구분

목적 구분	출발지	도착지
출근	집	회사
	집 이외의 휴식 장소*	
등교	집	학교
	집 이외의 휴식 장소	
업무	회사	출장지
	기타	
쇼핑	집	쇼핑몰
	집 이외의 휴식 장소	
	기타	
학원	집	학원
	집 이외의 휴식 장소	
	기타	
여가	집	여가/오락/친교 장소
	집 이외의 휴식 장소	
	기타	
기타	집	병원, 종교집회장소 등
	집 이외의 휴식 장소	
	기타	
귀가	기타	집
		집 이외의 출발지

*: 친지 집, 호텔 등.

- 모바일 자료 기반의 통행목적별 구분하기 위해서는 전술한 체류지가 식별되어야 함
- 본 과업에서는 작년 사업 내 심야시간대 주체류지, 낮시간대 주체류지, 잠재체류지로 총 3가지로 구분되었던 체류지 유형을 다음 <표 2-2>와 같이 6가지 유형으로 세분화하고자 함
 - 기존 체류지 유형으로는 출근과 등교 통행을 구분할 수 없어 통행수요가 가장 높은 시간대인 출퇴근 시간대 통행을 심층 분석하기 어렵기 때문

<표 2- 2> 체류지 유형

'17년도 기준	'18년도 기준
심야시간대 주체류지	집
	집 이외 심야시간대 주체류지 (예: 친지 집)
낮시간대 주체류지	회사
	학교
잠재체류지	교회, 성당 등 종교집회장소
	기타

- 세분화 된 체류지를 기준으로 기종점 체류지 유형을 조합하여 추출할 수 있는 통행 목적 유형은 출근 / 등교 / 여가 / 귀가 / 기타 총 5가지이며, 통행목적별 기종점 체류지 유형을 정리하면 <표 2-3>과 같음
 - 쇼핑과 여가 통행 모두 비정기적인 통행으로 통행 패턴이 유사하며 단순 토지이용 정보 등 공간정보를 추가하여 이들을 구분하기에는 한계가 존재하므로 이들을 구분하지 않았음
 - 학원 통행은 타 통행 대비 통행 패턴이 명확하지 않기 때문에 이를 제외함
 - 귀가 통행의 경우 도착지에 따라 '퇴근', '하교', '귀가'로 세분 가능하며, 기타 통행도 마찬가지로 도착지에 따라 '종교 활동', '친지 방문' '기타'로 세분 가능함
 - 체류지 구분 기준으로 개인별 인적특성(연령)이 반영될 것을 고려하여 학교와 회사가 기점과 종점으로 엮이는 경우는 제외함

<표 2- 3> 모바일 자료 기반 통행 목적 구분

목적 구분		출발지	도착지
출근		집	회사
		집 이외의 심야시간대 주체류지	
등교		집	학교
		집 이외의 심야시간대 주체류지	
여가		집	기타
		집 이외의 심야시간대 주체류지	
		회사	
		학교	
		기타	
귀가	퇴근	회사	집
			집 이외의 심야시간대 주체류지
	하교	학교	집
			집 이외의 심야시간대 주체류지
	귀가	종교집회장소	집
			집 이외의 심야시간대 주체류지
기타		집	종교집회장소
		집 이외의 심야시간대 주체류지	
		회사	
		학교	
		기타	
	친지 방문	집	집 이외의 심야시간대 주체류지
	기타	종교집회장소	회사
			학교
		기타	회사
			학교

다. 체류지 식별 기준

- <표 2-4>와 같이 통행목적에 따라 다르게 나타나는 통행 시간대, 통행 계층 등 통행 특성을 참고하여 체류지 식별 기준의 틀을 설정하고자 함
- － 통행목적별 통행 특성은 주로 목적지(도착지)에 따라 달라지므로, 통행목적별 통행 특성을 도착지 구분 기준으로 두고자 함

<표 2- 4> 통행목적별 통행 특성

통행목적		특성
출근		<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오전 6시 ~ 오전 9시) 주로 발생¹⁾ • 특정 계층에 해당(사회에 진출한 20세 이상의 통행자)
등교		<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오전 6시 ~ 오전 9시) 주로 발생¹⁾ • 특정 계층에 해당(취학 대상의 연령층, 대체로 20세 미만)
여가		<ul style="list-style-type: none"> • 비반복 통행
귀가	퇴근	<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오후 4시 ~ 오후 23시) 주로 발생²⁾ • 특정 계층에 해당(사회에 진출한 20세 이상의 통행자)
	하교	<ul style="list-style-type: none"> • 반복 통행 • 주로 주중에 발생 • 특정 시간대(오후 4시 ~ 오후 23시) 주로 발생²⁾ • 특정 계층에 해당(취학 대상의 연령층, 대체로 20세 미만)
	귀가	<ul style="list-style-type: none"> • 반복 통행 • 특정 시간대(오후 7시 ~ 오후 10시) 주로 발생³⁾ • 요일에 영향을 받지 않음
기타	종교활동	<ul style="list-style-type: none"> • 반복 통행 • 주로 주말에 발생 • 특정 시간대(예배 시간) 주로 발생
	친지 방문	<ul style="list-style-type: none"> • 비반복 통행
	기타	<ul style="list-style-type: none"> • 특징 없음

주 1) 2016년 3월 7일 ~ 2017년 3월 31일에 기록된 통신 데이터를 분석한 결과 심야시간대 주체류지(예: 집)에서 낮 시간대 주체류지(예: 회사, 학교 등)로 주로 오전 6시에서 오전 8시 사이에 출발하여 오전 7시에서

오전 9시 사이에 도착하는 것으로 나타남(모바일 Mobility report, p.55).

주 2) 2016년 3월 7일 ~ 2017년 3월 31일에 기록된 통신 데이터를 분석한 결과 낮 시간대 주체류지(예: 회사, 학교 등)에서 심야시간대 주체류지(예: 집)로 주로 오후 4시에서 오후 7시 사이에 출발하여 오후 5시에서 오후 11시 사이에 도착하는 것으로 나타남(모바일 Mobility report, p.56).

주 3) 2016년 3월 7일 ~ 2017년 3월 31일에 기록된 통신 데이터를 분석한 결과 잠재활동지(예: 식당, 백화점, 공원 등)에서 심야시간대 주체류지(예: 집)으로 주로 오후 7시에서 오후 9시 사이에 출발하여 오후 9시에서 오후 10시 사이에 도착하는 것으로 나타남(모바일 Mobility report, p.57).

자료: KTDB 정책자료집 『모바일 Mobility Report』, p.55-57.

○ 통행목적별 통행특성을 기준으로 도착지에서의 로그 기록 특성을 정리하면 다음 <표 2-5>와 같음

<표 2- 5> 도착지별 로그 기록 특성

통행목적		목적지(도착지)	특성
출근		회사	<ul style="list-style-type: none"> 반복적으로 기록됨 주로 주중에 기록됨 주로 특정 시간대(낮 시간대) 기록됨 특정 계층에 한해서만 기록됨(20대 이상의 휴대폰 소지자)
등교		학교	<ul style="list-style-type: none"> 반복적으로 기록됨 주로 주중에 기록됨 주로 특정 시간대(낮 시간대) 기록됨 특정 계층에 한해서만 기록됨(20세 미만의 휴대폰 소지자)
여가		기타	<ul style="list-style-type: none"> 특징 없음
귀가		집 또는 집 이외의 심야시간 주체류지	<ul style="list-style-type: none"> 요일에 관계없이 반복적으로 기록됨 주로 특정시간대(심야시간대) 기록됨
기타	종교활동	종교집회장소	<ul style="list-style-type: none"> 반복적으로 기록됨 주로 주말에 기록됨 주로 특정시간대(예배 시간대) 기록됨
	친지 방문	집 이외의 심야시간대 주체류지	<ul style="list-style-type: none"> 주로 특정 시간대(심야시간대) 기록됨 반복적으로 기록되지 않음
	기타	회사 또는 학교	<ul style="list-style-type: none"> 특징 없음

- 체류지를 정확하게 구분하기 위해서는 정확한 식별 임계치가 필요하므로, 전술한 로그 기록 특성을 토대로 체류지별 로그 기록 시간대, 총 기록 시간, 기록 횟수 기준 값을 정할 필요가 있음
- 본 과업에서는 『2017년도 국가교통조사』에서 정의한 체류지 식별 기준에 식별 기준을 추가하여 체류지를 세분하고자 하였으며, 식별 기준이 타당한 지는 통계청 인구 정보 등 실제 기준 값을 통해 검토하고자 함

1) 집과 집 이외의 심야시간대 주체류지 식별 기준

- 집과 집 이외의 심야시간대 주체류지는 주로 심야시간대 기록되는 특성이 있으므로, 다른 체류지와 구분하는 기준으로 『2017년도 국가교통조사』의 심야시간대 주체류지 식별 기준의 일부를 준용함
- 기존 심야시간대 주체류지 식별 기준은 ‘오전 12시부터 오전 6시 사이에 3시간 이상 체류한 기록이 주 3회 이상 기록된 기지국’이나, 평풍 현상 보정 알고리즘을 보완하여 일자 구분과 관계없이 심야시간대 주체류지를 탐색할 수 있게 되어 정확성을 높이기 위해 로그 기록 시간대 기준을 기준일 전날 오후 9시부터 기준일 오전 7시까지로 확대하고, 총 로그 기록시간 기준은 그대로 3시간 이상을 준용함
- 집과 집 이외의 심야시간대 주체류지를 구분하기 위해서는 기록 횟수의 임계치를 설정할 필요가 있음
- 반복적으로 기록되는 집과 달리 집 이외의 주체류지는 반복적으로 기록되지 않는 특성을 갖고 있기 때문
- 기존 심야시간대 주체류지 식별 기준은 주 단위의 횟수를 임계치로 사용했지만, 주 단위의 기준으로는 집과 집 이외의 심야시간대 주체류지를 구분하기 어려우므로 월 단위의 기준을 설정할 필요
- 집은 특별한 이유가 없는 이상(예: 장기 출장) 매주 기록되는 반면, 집 이외의 심야시간대 주체류지(예: 호텔, 친지 집)은 매주 기록되지 않기 때문
- 원천데이터 분석 결과, ‘집’의 기록 횟수 임계치는 주 3회 이상, 월 3회 이상이 적합한 것으로 나타남
- 기록 횟수가 주 3회 이상, 월 3회 이상을 동시에 충족해야 통행자의 90% 이상이 집을 식별할 수 있는 것으로 나타남

- ‘집 이외의 심야시간대 주체류지’는 반복적으로 기록되지 않으므로 주 1회 이상을 기록 횟수 기준으로 두되, 통행 특성이 불분명한 기타(잠재체류지)와 구별할 수 있는 별도 기준이 필요
- ‘집 이외의 심야시간대 주체류지’와 ‘기타(잠재체류지)’는 ‘집’과의 거리에서 차이가 나타날 것으로 보고, 단거리 통행 거리 기준(5km)³⁾을 식별 기준으로 추가함
 - 심야시간대 기록되는 ‘기타(잠재체류지)’는 심야시간대 운영하는 상가일 가능성이 크고, 활동을 마치면 결국 집으로 귀가할 것이기 때문에 집과 가까이 있을 가능성이 높음
 - 집 이외의 심야시간대 ‘주체류지(예: 호텔, 친지 집)’는 집과 동일하게 휴식, 수면 목적이기 때문에 집으로 다시 귀가할 가능성이 없어 집과 멀리 떨어져 있을 가능성이 높기 때문
 - 주 1회 이상 기록된 기지국 중에서 집과의 거리가 5km 이상인 기지국은 모두 집 이외 심야시간대 주체류지로 두고, 집과의 거리가 5km 미만인 기지국은 하루 기록 중 가장 마지막에 기록된 기지국만 집 이외 심야시간대 주체류지로 두고, 나머지는 기타(잠재체류지)로 분류하도록 함
- <표 2-6>은 전술한 식별 기준을 종합하여 집과 집 이외 심야시간 주체류지 식별 기준을 정리한 것이며, <표 2-7>은 기준을 적용하였을 때 각 체류지의 식별율을 분석한 결과임

<표 2- 6> 집과 집 이외 심야시간대 주체류지 식별 기준

구분	집(H)	집 이외 심야시간대 주체류지(N)
기준시간 (체류시작시간 ~ 체류종료시간)	오후 9시 ~ 오전 7시	오후 9시 ~ 오전 7시
체류시간	3시간 이상	3시간 이상
반복횟수	주 3회 이상 & 월 3회 이상	주 1회
기타	-	* 집과의 거리가 5km 이상인 경우, 모두 집 이외 주체류지로 식별. * 5km 미만인 경우 심야시간대 마지막에 등장한 기지국을 집 이외 주체류지로 식별

3) 김순관·고준호·이신해(2013)에서 정의한 단거리 거리 기준 적용, 김순관·고준호·이신해(2013), 「대중교통 서비스 개선을 위한 서울시 출근통행의 질 평가」 p. 4.

<표 2- 7> 집과 집 이외 심야시간대 주체류지 식별율 검토 (2017년 기준)

구분	집	집 이외 심야시간대 주체류지
식별 인구 (전수화) ¹⁾	10,799,982 (41,538,392)	6,627,719 (25,491,227)
식별율 ²⁾	92.9%	57.0%

주 1) 2017년 3월 기준 KT 시장점유율(약 26%)을 고려하여 전수화한 값임.

2) KT 전체 가입 고객 인구 수 11,629,966명을 기준으로 식별 인구가 차지하는 비율을 산출한 값임.

2) 회사와 학교 식별 기준

- 회사와 학교 통행은 주로 낮 시간대 발생되므로, 『2017년도 국가교통조사』의 낮 시간대 주체류지 식별 기준의 로그 기록 시간대, 총 기록시간을 준용함
 - 기존 낮 시간대 주체류지 식별 기준은 ‘오전 9시부터 오후 6시 사이에 3시간 이상 체류한 기록이 주 3회 이상 기록된 기지국’이며, 여기서 로그 기록 시간대(오전 9시부터 오후 6시 사이) 기준과 총 체류 시간(3시간 이상) 기준을 준용함
- 반복 횟수 기준은 기존 기준(주 3회 이상)을 그대로 적용할 경우 회사, 학교 외에 다른 체류지가 포함될 가능성이 있어, 통계청에서 공표하고 있는 통근 통학 인구수를 비교 기준으로 두고 검증하여 반복 횟수 기준을 ‘주 2회 이상, 월 3회 이상’으로 변경함
 - 2015년 통근통학 인구(12세 이상)를 이용하여 추정한 2017년 통근 통학 인구(5세 이상 60세 미만)는 전체 인구의 약 50%였으며, 그 중 79%는 통근 인구이고, 21%가 통학 인구인 것으로 분석됨
 - 이는 5세 이상 60세 미만 휴대폰 소지자 중 50% 소지자의 해당 통행을 회사 또는 학교 통행으로 식별될 수 있으며, 회사와 학교 통행의 식별 비율이 각각 79%, 21% 여야 한다는 것을 의미함
 - 분석 결과 ‘주 2회 이상, 월 3회 이상’의 반복 횟수 기준일 때 해당 기준을 충족하는 것으로 나타남
 - 기록 횟수 기준 산출시 회사와 학교 모두 대부분 주중에 통행이 발생되기 때문에 주말 통행 기록은 분석에서 제외함

<표 2- 8> 통계청 기준 통근 통학 인구와 모바일 데이터 기준 통근 통학 인구 비교 (2017년 기준)

구분		모바일 데이터 기준		통계청 기준	
		인구 수 (전수화) ¹⁾	비율	인구 수	비율
통근 통학 인구	계	6, 173, 914 (23, 745, 823)	100.0%	25, 929, 590 ²⁾	100.0%
	통근	4, 460, 367 (17, 155, 258)	72.2%	20, 470, 134 ³⁾	78.9%
	통학	1, 713, 547 (6, 590, 565)	27.8%	5, 459, 456 ⁴⁾	21.1%
전체 인구		11, 629, 966 (44, 730, 638)		51, 422, 507	
통근 통학 인구 비율		53.1%		50.4%	

주 1) 2017년 3월 기준 KT 시장점유율(약 26%)을 고려하여 전수화한 값임.

2) 2015년 12세 이상 전체 인구에서 12세 이상 통근 통학 인구가 차지하는 비율(67%)을 2017년 5세 이상 60세 미만 인구에 적용하여 추정한 결과임.

3) 2015년 기준 통근 통학 인구에서 통근 인구가 차지하는 비율(78.9%)을 추정한 2017년 전체 통근 통학 인구에 곱하여 산출한 결과임.

4) 2015년 기준 통근 통학 인구에서 통학 인구가 차지하는 비율(21.1%)을 추정한 2017년 전체 통근 통학 인구에 곱하여 산출한 결과임.

자료: 통계청 KOSIS “성별/연령별/이용 교통수단별 통근 통학 인구(12세 이상)-시군구”, http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1PA1503 (18. 11. 05).

통계청 KOSIS “연령 및 성별 인구 - 읍면동(2015), 시군구(2016~)”, http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1IN1503&vw_cd=MT_ZTITLE&list_id=A11_2015_1_10_10&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE (18. 11. 05).

과학기술정보통신부 “2017년 12월 무선통신서비스 가입자 통계”, <https://www.msit.go.kr/>.

- 한편, 20세 이상 25세 미만은 취학시기와 취업시기에 맞물리는 시점이기 때문에 회사와 학교를 구분하기 어려우므로 방학기간의 로그 기록을 추가 검토하도록 하는 기준을 추가함
 - 방학기간(7~8월, 1~2월)에도 사전에 정의한 식별 기준대로 로그가 기록되고 있는지 검토하여, 로그 기록 패턴이 일정한 경우 회사로, 로그 기록에 변화가 있는 경우 학교로 구분하도록 함
- <표 2-9>는 전술한 식별 기준을 종합하여 회사와 학교 체류 식별 기준을 정리한 것임

<표 2- 9> 회사와 학교 체류 식별 기준

구분	회사(C)	학교(S)
기준시간 (체류시작시간 ~ 체류종료시간)	오전 9시 ~ 오후 6시	
체류시간	3시간 이상	
반복횟수	주 2회 이상 (주말 제외) & 월 3회 이상	
통행자 연령	20세 이상	25세 미만
기타	-	20세 이상 25세 미만 연령의 경우, 방학기간(7~8월, 1~2월)에도 위 기준이 충족될 경우 회사로 간주

3) 종교집회장소와 기타 식별 기준

- 종교 활동을 목적으로 한 통행은 주로 주말 내 특정 시간(예배시간 등)에 발생된다는 점을 고려하여 종교집회장소 체류지 식별을 설정함
- 교회, 성당 등 종교집회장소마다 예배시간이 다양하기 때문에 모든 시간을 반영하는데 한계가 존재함
 - 가장 빈번히 종교집회장소를 방문하는 시간인 오전 8시부터 오후 4시 사이(혹은 오후 9시부터 오후 4시 사이)를 설정하여 1시간 이상 체류한 경우를 식별기준으로 삼음
- 또한 종교 활동의 경우 여가(기타)와 달리 정기적으로 이루어지므로, 기타 잠재체류지와 달리 반복적으로 기록되는 패턴을 보일 것으로 판단됨
- 통계청 종교인구 비율을 기준으로 하여 반복횟수 기준을 추가함
 - 통계청에서 공표하는 종교 인구는 전체인구의 약 44%이며, 이와 가장 유사한 식별율을 갖는 조건은 ‘주 1회 이상의 기록이 월 2회 이상 나타난 경우’였음

<표 2- 10> 통계청 기준 종교인구와 모바일 데이터 기준 종교인구 비교 (2017년 기준)

구분	모바일 데이터 기준 인구수 (전수화) ¹⁾	통계청 기준 인구수
종교 인구	6,032,707 (23,202,719)	21,553,674 ²⁾
전체 인구	11,629,966 (44,730,638)	51,422,507
종교 인구 비율	51.8%	43.9%

주 1) 2017년 3월 기준 KT 시장점유율(약 26%)을 고려하여 전수화한 값임.

2) 2015년 종교인구를 기준으로 2017년 종교인구를 추정된 결과임.

자료: 통계청 KOSIS “성별/연령별/종교별 인구-시군구”, http://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1PM1502&checkFlag=N (18. 11. 21).

- 기타 잠재체류지는 공통된 특징이 없으므로, 전술한 모든 체류지를 제외한 나머지를 기타 잠재체류지로 볼 수 있음
- <표 2-11>은 전술한 식별 기준을 종합하여 종교집회장소와 기타 잠재체류지 식별 기준을 정리한 것임

<표 2- 11> 종교집회장소와 기타 잠재체류지 식별 기준

구분	종교집회장소(R)	기타 잠재체류지(X)
기준시간 (체류시작시간 ~ 체류종료시간)	토요일 또는 일요일 로그에 한함 오전 8시 ~ 오후 2시 / 오전 9시 ~ 오후 4시	체류 시간 25분 이상 & 다른 체류지로 식별되지 않은 체류지
체류시간	1시간 이상	
반복횟수	주 1회 이상 & 월 2회 이상	

라. 통행목적 구분

- 통행목적은 <표 2-12>와 같이 원천데이터에 체류지 식별 기준을 적용하여 체류지를 구분한 후 통행목적별 기종점 체류지 유형(<표 2-3> 참조)을 기준으로 하여 구분할 수 있음
 - 체류지는 체류시간, 빈도가 높은 집, 회사 / 학교, 집 이외 주체류지, 종교집회장소, 기타 잠재체류지 순으로 식별함

<표 2- 12> 로그별 통행목적 추정 결과 (예시)

기준일	고객식별 번호	기지국 좌표		체류 시작 시간	체류 종료 시간	체류 시간	성	연령	체류지 구분	통행 목적
		x 좌표	y 좌표							
20170126	1813080	127.132	36.475	0001	0805	484	남	40	집	출근
20170126	1813080	128.126	37.319	0850	1150	180	남	40	회사	
20170126	1813080	127.080	37.216	1210	1310	60	남	40	기타	여가
20170126	1813080	128.126	37.319	1330	1800	270	남	40	회사	기타
20170126	1813080	127.132	36.475	1845	2359	314	남	40	집	귀가

제2절 모바일 자료 기반 DB 구축

1. 모바일 자료 기반 DB 설계

- 개인별 로그 기록은 「개인정보보호법」과 「위치정보의 보호 및 이용 등에 관한 법률」 등으로 인해 공개될 수 없으므로 기반 DB는 개인의 위치와 이동궤적이 추정되지 않는 집계 형태여야 함
- 따라서 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), 통행목적 정보가 동일한 인구를 집계하여 <표 2-13>과 같은 형태로 DB를 구축하고자 함
 - 2018년에 구축한 모바일 기반 DB 구조에 통행 목적 구분 필드가 추가함
 - 개인의 위치와 이동궤적이 추정되지 않도록 하려면 기지국 위치 정보를 비식별화 처리해야 하고 동일한 기준을 갖는 값(집계된 값)이 최소 3 이상이어야 함
- 교통폴리곤은 기지국의 수신범위를 고려하여 구축한 모바일 기반 DB 분석 단위임

<표 2- 13> 모바일 기반 DB 형태 (예시)

출발				도착				통행 목적	성 별 코 드	연 령 코 드	유동 인구 수
일자	시 간	교통 폴리곤 ID	체류지 유형	일자	시 간	교통 폴리곤 ID	체류지 유형				
20171121	13	48270310	집	20171123	14	12390981	기타	여가	M	20	10
20170501	06	48270310	집	20170501	09	32789014	회사	출근	F	50	25
20170311	17	48270320	집	20170312	18	45608912	학교	하교	M	10	15

- 모바일 기반 DB 테이블 정의서는 다음 <표 2-14>와 같음

<표 2- 14> 모바일 DB 테이블 정의서

컬럼 ID	컬럼명	Type	비고
o_polygon	출발폴리곤ID	string	-
o_base_ymd	출발 일자	string	yyyymmdd
o_timezn_cd	출발시간대	string	00-23
o_trip_type	출발 체류지 유형	string	H:집 N:집 이외 심야시간대 주체류지 C: 회사 S: 학교 R: 종교집회장소 X: 기타 잠재체류지
d_polygon	도착폴리곤ID	string	-
d_base_ymd	도착 일자	string	yyyymmdd
d_timezn_cd	도착시간대	string	00-23
d_trip_type	도착 체류지 유형	string	H:집 N:집 이외 심야시간대 주체류지 C: 회사 S: 학교 R: 종교집회장소 X: 기타 잠재체류지 -: 이동정보 없음
age_itg_cd	연령대 정보	integer	0~110 예) 0:0세 이상 10세 미만 10: 10세 이상 20세 미만
sex_type_itg_cd	성별 정보	string	M, F
total	인구 수	integer	-

2. 모바일 자료 기반 DB 구축 과정

1) 데이터 전처리

- 이상치를 제거하고 왜곡데이터를 보완
 - 기지국 좌표가 Null값이거나 행정동 정보가 매칭되지 않는 데이터를 제거하고, 통행 정보를 왜곡할 수 있는 ‘필요 이상의 핸드오버 데이터(Unnecessary Handover)’, ‘핑퐁 핸드오버(Ping-pong Handover)로 인한 데이터’를 본 과업에서 개발한 전처리 기술을 통해 보정함

2) 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
 - 고객식별번호(단말기구분코드)와 데이터 생성일자를 기준으로 데이터를 개인별·일자별 데이터로 구분한 후, 구분된 데이터 내에서 체류시작시간과 체류종료시간을 기준으로 데이터를 오름차순으로 정렬하여 순번을 부여함

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

<그림 2- 2> 체류순서 부여 (개인별 이동궤적 형성)

3) 체류 정보 추출

- 이동 중에 기록된 로그 기록을 제외하고 체류 중에 기록된 로그 기록만을 추출
 - 트립 타입 필드를 추가하여 총 로그 기록시간(이하 '체류시간')이 25분 미만인 경우는 이동 중에 기록된 데이터로 '경로(Pass-by)' 데이터로 구분하고, 체류시간이 25분 이상인 로그는 체류 중에 기록된 데이터로 3시간 이상 체류한 기록이 주 3회 이상 반복되는 로그 기록은 '주상주지(Stay area)'로, 주상주 데이터가 아닌 로그 기록은 '잠재체류지(potential area)'로 구분한 후, 트립타입 필드가 '주상주지'이거나 '잠재체류지'인 로그 기록만을 추출

ID	그룹 번호	KEY 킬럼 (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	주상주
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	주상주
		⋮	⋮	⋮					
453	1	20160520-1843030	126,934	37,556	0805	1041	25	156	주상주
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	잠재
		⋮	⋮	⋮					
679	227	20160522-1843030	126,934	37,556	0003	0830	25	507	주상주
680	228	20160522-1843030	126,936	37,556	0831	0833	25	2	경로
		⋮	⋮	⋮					
711	259	20160522-1843030	126,815	37,551	0850	1100	25	130	잠재

<그림 2- 3> 이동정보와 체류정보 구분 (예시)

4) 위치정보, 연령 정보 변환

- 「개인정보보호법」과 「위치정보의 보호 및 이용 등에 관한 법률」에 따라 개인의 이동 궤적을 추적할 수 없도록 원천데이터에서 출발지와 도착지의 위치 정보는 분석 맵의 교통폴리곤 ID로 변경하고, 1세 단위의 연령은 10세 단위로 변경

ID	그룹 번호	KEY 칼럼 (기준일- 고객식별번호)	가상기지국		체류 시작 시간	체류 종료 시간	연 령 대	체류 시간	트립 타입	연령대 변환	위치 정보 변환 (교통폴리곤 ID)
			x좌표	y좌표							
1	1	20160520-17232311	126.914	37.545	1400	1750	33	240	주상주	30	210001
2	2	20160520-17232311	126.914	37.549	1751	1752	33	1	경로	30	-
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로	30	-
4	4	20160520-17232311	126.914	37.549	1810	1813	33	3	경로	30	-
5	5	20160520-17232311	126.915	37.551	1813	1814	33	1	경로	30	-
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로	30	-
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	주상주	30	230017
453	1	20160520-1843030	126.934	37.556	0805	1041	25	156	주상주	20	311014
454	2	20160520-1843030	126.936	37.556	1041	1042	25	1	경로	20	-
455	3	20160520-1843030	126.935	37.554	1043	1044	25	1	경로	20	-
456	4	20160520-1843030	126.933	37.552	1045	1050	25	5	경로	20	-
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로	20	-
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로	20	-
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로	20	-
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	잠재	20	312123
679	227	20160522-1843030	126.934	37.556	0003	0830	25	507	주상주	20	110016
680	228	20160522-1843030	126.936	37.556	0831	0833	25	2	경로	20	-
711	259	20160522-1843030	126.815	37.551	0850	1100	25	130	잠재	20	210003

<그림 2- 4> 위치정보, 연령 정보 변환 (예시)

5) 체류지 유형 구분

- 체류지 식별 기준(<표 2-15> 참조)에 따라 체류지 유형을 구분하여 필드 추가

<표 2- 15> 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 주 상 주	집	오후 9시 ~ 오전 7시	3시간 이상	주 3회 이상 기록된 주가 월 3회 이상	-	-
		집 이외	오후 9시 ~ 오전 7시	3시간 이상	주 1회	-	집과의 거리가 5km, 이상인 경우, 모두 집 이외 주체류지로 식별, 5km미만인 경우 심야시간대 마지막에 등장한 기지국을 집 이외 주체류지로 식별

지역						
낮 시간대 주상주 지역	회사	오전 9시 ~ 오후 6시	3시간 이상	주 2회 이상(주말 제외) 기록된 주가 월 3회 이상	20세 이상	20세 이상 25세 미만 연령의 경우 방학기간 (7-8월, 1-2월)에도 위 기준이 충족될 경우 회사로 간주
	학교	오전 9시 ~ 오후 6시	3시간 이상	주 2회 이상(주말 제외) 기록된 주가 월 3회 이상	25세 미만	20세 이상 25세 미만 연령의 경우 방학기간 (7-8월, 1-2월)에 위 기준이 충족되지 않는 경우 회사로 간주
잠재 활동 지역	종교 집회장소	오전 8시 ~ 오후 14시 / 오전 9시 ~ 오후 16시	1시간 이상	주 1회 이상 기록된 주가 월 2회 이상	-	‘토요일 또는 일요일’에 기록된 로그에 한 함
	기타	집, 집 이외 심야시간대 주체류지, 회사, 학교, 종교집회장소로 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	연령대 변환	위치 정보 변환 (교통플리곤 ID)	체류지 유형 구분
			X좌표	Y좌표								
1	1	20160520-17232311	126.914	37.545	1400	1750	33	240	주상주	30	210001	집
2	2	20160520-17232311	126.914	37.549	1751	1752	33	1	경로	30	-	-
3	3	20160520-17232311	126.916	37.548	1754	1809	33	15	경로	30	-	-
4	4	20160520-17232311	126.914	37.549	1810	1813	33	3	경로	30	-	-
5	5	20160520-17232311	126.915	37.551	1813	1814	33	1	경로	30	-	-
6	6	20160520-17232311	126.914	37.549	1814	1815	33	1	경로	30	-	-
18	18	20160520-17232311	126.815	37.538	1920	2359	33	279	주상주	30	230017	회사
453	1	20160520-1843030	126.934	37.556	0805	1041	25	156	주상주	20	311014	집
454	2	20160520-1843030	126.936	37.556	1041	1042	25	1	경로	20	-	-
455	3	20160520-1843030	126.935	37.554	1043	1044	25	1	경로	20	-	-
456	4	20160520-1843030	126.933	37.552	1045	1050	25	5	경로	20	-	-
457	5	20160520-1843030	126.929	37.547	1051	1052	25	1	경로	20	-	-
458	6	20160520-1843030	126.923	37.547	1053	1055	25	2	경로	20	-	-
459	7	20160520-1843030	126.918	37.547	1056	1059	25	3	경로	20	-	-
460	8	20160520-1843030	126.914	37.545	1100	1200	25	60	잠재	20	312123	기타
679	227	20160522-1843030	126.934	37.556	0003	0830	25	507	주상주	20	110016	집
680	228	20160522-1843030	126.936	37.556	0831	0833	25	2	경로	20	-	-
711	259	20160522-1843030	126.815	37.551	0850	1100	25	130	잠재	20	210003	종교집회 장소

<그림 2- 5> 체류지 유형 구분 (예시)

6) 통행목적 구분

- 앞서 부여한 체류순서에 따라 한 달 단위로 출발과 도착을 구분한 다음 출발지와 도착지의 체류지 유형 조합에 따라 통행목적은 <그림 2-6>과 같이 구분함
- 한 달 단위로 출발과 도착을 구분한다는 것은 예컨대 1월 1일부터 1월 31일까지(한 달) 발생한 로그에 체류 순서에 따라 출발, 도착을 구분하고, 2월 1일부터는 1월 31일의 마지막 체류지가 출발이든 도착이든 모조건 처음 기록된 체류지를 출발로 본다는 것을 의미함
- 마지막 일자(1월 31일)의 마지막 체류지가 출발 정보로 구분되고, 도착 정보가 없는 경우 이동하지 않은 것으로 보고 ‘무통행’으로 간주함

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

<그림 2- 6> 통행목적 구분 (예시)

7) DB 구조에 따른 인구수 집계

- 기점 정보(출발 일자, 출발 시간대, 출발 위치, 출발 체류지 유형), 종점 정보(도착 일자, 도

착 시간대, 도착 위치, 도착 체류지 유형), 통행자 정보(성, 연령), 통행목적 정보가 동일한 로그 기록을 집계함

단말기ID (고객식별번호)	기록 기지국 위치정보 (XY)	기록일자 (YYYYMMDD)	기록시작시간 (HHMM)	기록종료시간 (HHMM)	총 기록시간 (MM)	단말기 고객 정보			
						성	연령		
1843030	127.132 36.475	20160620	10:25	10:52 출발	28	남	25	1통행	
1843030	126.147 37.319	20160620	12:47	12:48	2	남	25		
1843030	127.136 36.477	20160620	도착 13:00	13:39 출발	40	남	25	1통행	
1843030	127.132 36.475	20160620	도착 15:28	17:28	121	남	25		
2752075	126.147 37.319	20160621	17:16	18:20 출발	65	여	40	1통행	
2752075	127.134 36.475	20160621	도착 19:25	23:59	275	여	40		

출발위치 (교통플리곤 ID)	도착위치 (교통플리곤 ID)	출발일자	도착일자	출발시간대	도착시간대	출발지 유형	도착지 유형	통행자 정보		통행량	통행 목적
								성	연령대		
110006	110006	20160620	20160620	10	13	기타 잠재체류지	집	남	20	8	귀가
110016	110716	20160620	20160620	13	15	학교	기타 잠재체류지	남	20	15	기타
110017	110006	20160621	20160621	18	19	회사	기타 잠재체류지	여	40	6	기타

<그림 2- 7> 인구수(통행량) 집계 (예시)

3. DB 구축 결과

- 전국에서 2017년 1월 1일부터 2017년 12월 31일까지 발생한 로그 기록을 대상으로 통행목적 이 분류된 DB를 구축하였으며, 총 약 4.3TB의 기반 DB가 생성됨(1일 기준 약 12GB)

제3장 모바일 분석 맵 고도화

제1절 분석 맵 구축 방법

제2절 분석 맵 구축 및 검증

제3장 모바일 분석 맵 고도화

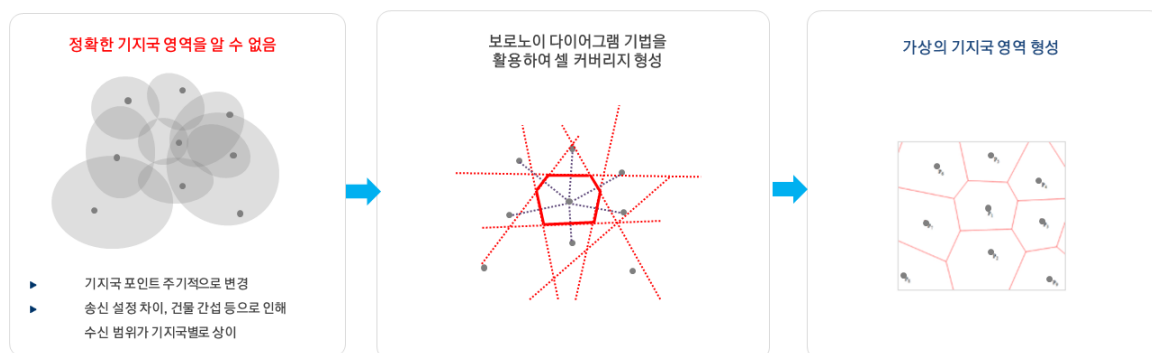
제1절 분석 맵 구축 방법

1. 기존 분석 맵 구축 방법

가. 분석 맵 구축 과정

1) 1단계: 기지국별 가상의 영역(셀 반경)을 설정

- 2016년 5월 한 달 동안 25일 이상 신호가 기록된 기지국을 주 기지국으로 두고, 이를 기준으로 보로노이 기법¹⁾을 적용하여 기지국별 영역 형성

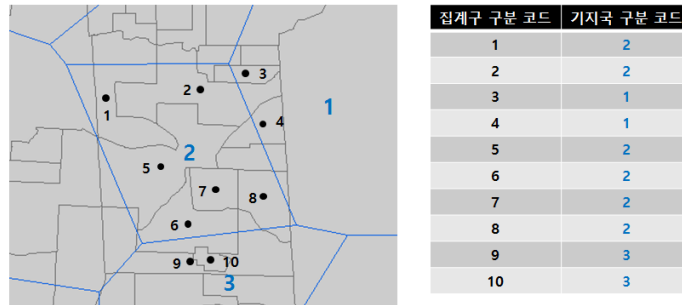


<그림 3- 1> 교통 폴리곤 구축 1단계

2) 2단계: 주기지국 수신범위를 기준으로 1차 병합

- 단일의 보로노이 셀이 여러 집계구의 중심에 중첩되는 경우 하나의 집계구로 병합하여 주 기지국과 표준 집계구가 최소 1:1(혹은 N:1)이 되도록 함
- 집계구의 중심점을 생성한 후 기지국 수신범위와 중첩시켜 동일한 수신범위에 있는 집계구를 동일 행정(읍면동) 내에서 병합

1) 임의의 한 점과 다른 점을 연결한 선분의 수직이등분선을 그린 후, 수직이등분선간의 교차점을 연결하여 공간을 분할하는 기법.

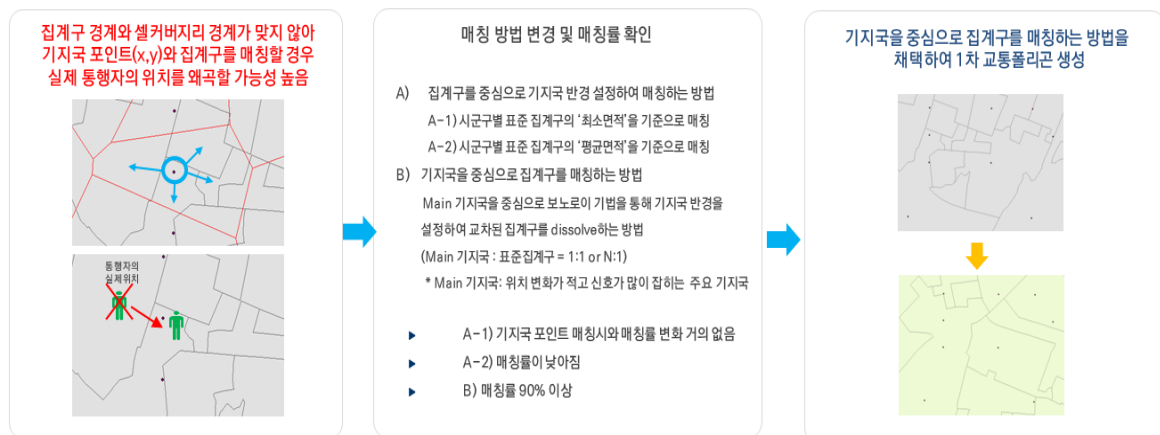


<그림 3- 2> 셀 반경 내 집계구 병합 기준 (예시)

- <표 3-1>과 같이 기지역이 매핑 되지 않은 집계구는 동일 행정동(읍면동) 내 맞닿아있는 경계선(border)이 많은 폴리곤과 병합

<표 3- 1> 집계구별 기지역 개수 (예시)

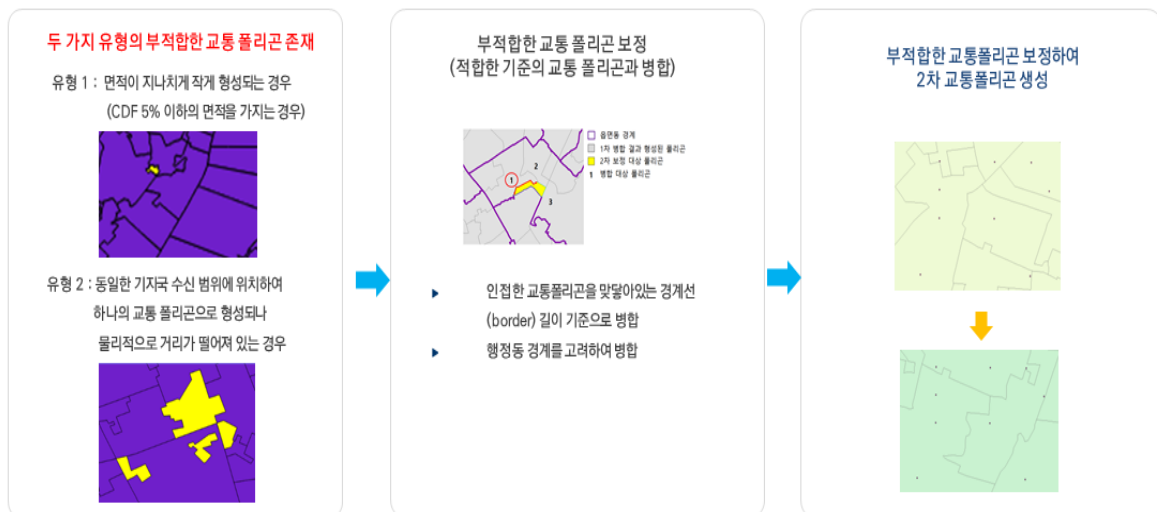
순번	읍면동 ID	집계구 ID	매핑된 기지역 개수
1	1101053	1101053010001	0
2	1101053	1101053010002	0
3	1101053	1101053010003	1
4	1101053	1101053010004	1
5	1101053	1101053010005	1



<그림 3- 3> 교통 폴리곤 구축 2단계

3) 3단계: 부적합한 교통폴리곤 보정

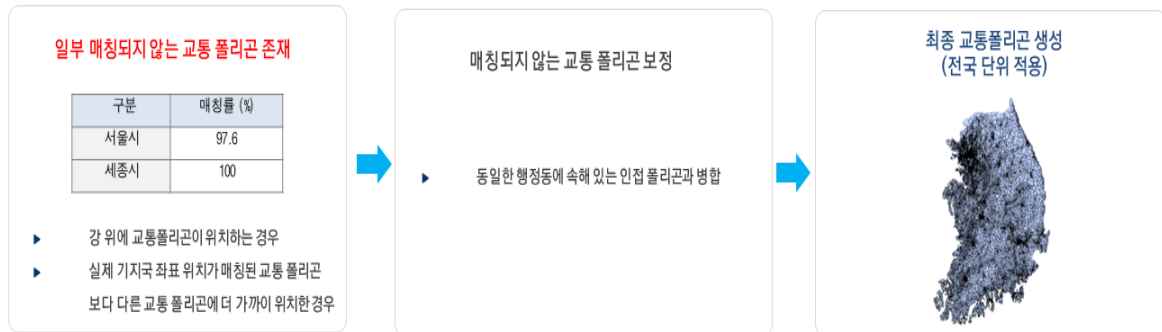
- 부적합한 교통폴리곤 생성되는 경우 인근 교통폴리곤과 병합하여 보정
 - 시도별 폴리곤 면적 누적분포함수 값(Cumulative distribution function, CDF)이 5% 이하인 폴리곤과, 1차 병합 후 기지국 셀 반경 안에 속성 값이 같으나 여러 개로 분리된 교통폴리곤(이하 'Multi-part 폴리곤')을 동일 행정동(읍면동) 내 인접한 교통폴리곤 중 가장 경계선(border)을 많이 공유하고 있는 교통폴리곤과 병합
 - Multi-part 폴리곤은 기준으로 활용한 집계구 자체가 Multi-part이고, 이들이 서로 다른 수신범위 내에 위치하는 경우 형성되거나, 집계구 내에 다른 집계구가 위치하고 이들이 서로 다른 수신범위 내에 위치하는 경우 형성됨



<그림 3- 4> 교통 폴리곤 구축 3단계

4) 4단계: 비(非)매칭 교통 폴리곤 보정

- 기지국과 매칭 되지 않은 일부 교통 폴리곤 보정
 - 강 위, 바다 위 영역(예: 섬)은 기지국이 없는 경우가 많으나 병합 과정에서 면해 있는 폴리곤(병합 대상)을 탐색하지 못해 병합되지 않을 수 있음
 - 이러한 비매칭 교통폴리곤은 인근 폴리곤의 중심점간의 거리 기준으로 동일 읍면동 내에서 가까운 폴리곤과 병합하여 보정



<그림 3- 5> 교통 폴리곤 구축 4단계

나. 기존 방식의 문제점 및 개선방향

1) 문제점

○ 주기지국 안정성 미흡

- 주기지국을 한 달 기준으로 선별하였기 때문에 특정 기간에만 기록된 기지국이 주기지국 정보에 포함되어 있을 수 있음
- 특정 기간에만 기록된 기지국이 포함된 폴리곤은 실제보다 통행량이 적게 집계될 수 있음

○ 일부 지역에는 부적합한 최소면적 기준

- ‘시도별 폴리곤 면적 누적분포함수 값 5% 이하를 최소면적으로 적용할 경우, 세종시 등 일부 지역은 큰 규모의 교통폴리곤이 형성될 수밖에 없음
- 전국으로 보면 교통폴리곤은 기존 분석구의 약 1/5 수준의 규모로 구축되었으나, 일부 지역은 그 보다 크게 구축됨
- 시도별 기지국 셀 평균 반경을 기준으로 형성할 수 있는 최대 교통폴리곤 개수와 실제 구축된 교통폴리곤 개수를 비교해보면 일부 지역이 과다하게 병합된 것을 알 수 있음 (<표 3-2> 참조)

<표 3- 2> 시도별 최대 교통폴리곤 형성 개수와 실제 구축된 교통폴리곤 개수 비교

시도	면적(㎡)	셀 반경 평균 (㎡)	시도별 최대 교통폴리곤 형성 개수 (1)	‘17년 기준 교통폴리곤 개수 (2)	(1) / (2)
서울	606,430,426	254	2,387,521	1,617	1,477
부산	793,540,061	353	2,247,989	789	2,849
대구	879,775,266	268	3,282,744	555	5,915
인천	1,167,435,302	512	2,280,147	440	5,182
광주	498,134,459	297	1,677,220	422	3,974
대전	539,199,225	309	1,744,981	437	3,993
울산	1,069,415,828	628	1,702,891	205	8,307
세종	465,430,058	896	519,453	70	7,421
경기	10,401,806,219	679	15,319,302	3,204	4,781
강원	16,933,074,307	1,140	14,853,574	776	19,141
충북	7,409,073,778	1,069	6,930,845	726	9,547
충남	8,323,924,705	1,063	7,830,597	1,218	6,429
전북	8,035,471,571	1,143	7,030,159	1,055	6,664
전남	12,657,093,637	1,316	9,617,852	1,389	6,924
경북	19,040,801,161	1,222	15,581,670	1,433	10,873
경남	10,616,566,548	1,152	9,215,770	1,272	7,245
제주	1,888,752,859	846	2,232,568	329	6,786

2) 개선방향

○ 주기지국 기준 변경

- 1년 기준으로 365일 이상 신호가 기록된 기지국을 주기지국으로 선별하여 특정 기간에만 기록되는 기지국 정보는 제외함

<표 3- 3> 주기지국 기준 변화

‘17년 주기지국	‘18년 주기지국
2016년 5월 한 달 동안 25일 이상 신호가 기록된 기지국	1년 기준으로 365일 이상 신호가 기록된 기지국

○ 최소면적 기준 강화

- 더욱 작은 규모의 최소면적 기준을 적용할 수 있도록 최소면적 기준을 시군구 단위로 변경

<표 3- 4> 최소면적 기준 변화

'17년 최소면적 기준	'18년 최소면적 기준
시도별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하	시군구별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하

3. 모바일 분석 맵 구축 개선안

나. 모바일 분석 맵 구축 개선안

1) 1단계: 기지국별 가상의 영역(셀 반경)을 설정

- 365일 이상 기록된 기지국을 (이하 '주 기지국') 기준으로 보로노이 기법을 적용하여 기지국별 영역 형성

2) 2단계: 주기지국 수신범위를 기준으로 1차 병합

- 집계구의 중심점을 생성한 후 기지국 수신범위와 중첩시켜 동일한 수신범위에 있는 집계구를 동일 행정(읍면동) 내에서 병합
- 기지국이 매핑 되지 않은 집계구는 동일 행정동(읍면동) 내 맞닿아있는 경계선(border)이 많은 폴리곤과 병합

3) 3단계: 부적합한 교통폴리곤 보정

- 지나치게 작게 형성된 폴리곤, Multi-part 폴리곤 보정
- 시군구별 집계구 면적 누적분포함수 값(Cumulative distribution function, CDF) 5%를 폴리곤 면적 최소 기준으로 두고(<표 3-5>, <표 3-6> 참조), 이보다 작게 형성된 폴리곤, 하나의 폴리곤으로 인식되나, 실제로는 떨어져 위치하여 하나의 폴리곤으로 보기 어려운 폴

리곤을 동일 행정동(읍면동) 내 맞닿아있는 경계선(border)이 많은 폴리곤과 병합

<표 3- 5> 시군구별 최소면적 기준

시도	시군구ID	기준면적 (㎡)	시도	시군구ID	기준면적 (㎡)	시도	시군구ID	기준면적 (㎡)	
서울	11010	10,536	대구	22010	2,806	경기	31011	15,785	
	11020	2,950		22020	87,965		31012	21,206	
	11030	10,226		22030	7,411		31013	6,223	
	11040	6,397		22040	6,327		31014	12,278	
	11050	7,800		22050	45,013		31021	19,543	
	11060	6,601		22060	36,420		31022	11,708	
	11070	8,065		22070	30,718		31023	33,712	
	11080	9,849		22310	207,300		31030	38,458	
	11090	9,719	인천	23010	71,497		31041	17,412	
	11100	10,300		23020	788		31042	10,930	
	11110	16,493		23030	9,903		31050	25,174	
	11120	12,245		23040	27,883		31060	19,229	
	11130	8,045		23050	25,888		31070	225,716	
	11140	9,983		23060	14,149		31080	47,493	
	11150	6,518		23070	21,919		31091	27,382	
	11160	19,961		23080	63,751		31092	56,402	
	11170	8,721		23310	184,957		31101	81,452	
	11180	5,836		23320	80,540		31103	25,139	
	11190	10,965	광주	24010	20,780		31104	18,194	
	11200	7,179		24020	22,537		31110	14,448	
	11210	14,417		24030	27,454		31120	14,094	
	11220	21,400		24040	56,484		31130	229,202	
	11230	17,727		24050	106,633		31140	17,861	
	11240	16,612	대전	25010	65,288		31150	69,931	
	11250	11,126		25020	29,645		31160	16,314	
				25030	46,069		31170	25,409	
부산	21010	1,487	울산	25040	86,588		31180	42,218	
	21020	6,987		25050	33,005		31191	230,655	
	21030	4,915		26010	17,055		31192	40,417	
	21040	5,869		26020	36,213		31193	18,018	
	21050	13,135	세종	26030	16,215		31200	339,210	
	21060	7,872		26040	75,516		31210	230,126	
	21070	12,011		26310	373,356		31220	273,958	
	21080	17,781		29010	226,391		31230	135,823	
	21090	23,475					31240	352,275	
	21100	23,021					31250	211,012	
	21110	31,562					31260	155,129	
	21120	87,218					31270	408,794	
	21130	5,286					31280	296,757	
	21140	5,163					31350	317,100	
	21150	16,939					31370	403,574	
	21310	106,940					31380	398,416	

<표 3- 6> 시군구별 최소면적 기준 (계속)

시도	시군구ID	기준면적 (㎡)	시도	시군구ID	기준면적 (㎡)	시도	시군구ID	기준면적 (㎡)
강원	32010	550, 097	전북	35011	42, 327	경북	37310	191, 871
	32020	431, 997		35012	51, 731		37320	535, 637
	32030	516, 118		35020	200, 275		37330	418, 431
	32040	84, 256		35030	249, 549		37340	350, 178
	32050	148, 912		35040	344, 807		37350	308, 864
	32060	48, 220		35050	368, 600		37360	302, 471
	32070	574, 164		35060	266, 564		37370	174, 168
	32310	848, 635		35310	404, 726		37380	266, 863
	32320	457, 099		35320	67, 982		37390	222, 176
	32330	531, 523		35330	281, 985		37400	304, 339
	32340	669, 488		35340	219, 015		37410	598, 126
	32350	501, 795		35350	235, 013		37420	469, 035
	32360	411, 951		35360	192, 682		37430	15, 396
	32370	453, 152		35370	281, 522		38030	352, 103
	32380	278, 533		35380	237, 439		38050	123, 076
	32390	688, 584	전남	36010	26, 807	경남	38060	195, 096
	32400	282, 286		36020	256, 294		38070	230, 675
	32410	295, 624		36030	454, 700		38080	397, 050
충북	33020	482, 959		36040	292, 025		38090	204, 464
	33030	433, 774		36060	228, 367		38100	241, 639
	33041	197, 960		36310	155, 142		38111	100, 565
	33042	57, 788		36320	206, 103		38112	40, 450
	33043	96, 078		36330	142, 259		38113	115, 906
	33044	104, 367		36350	396, 983		38114	42, 806
	33320	237, 172		36360	226, 526		38115	58, 394
	33330	238, 200		36370	384, 040		38310	176, 888
	33340	384, 648		36380	238, 479		38320	198, 347
	33350	195, 966		36390	201, 828		38330	230, 785
	33360	415, 366		36400	498, 966		38340	250, 049
	33370	252, 173		36410	288, 503		38350	173, 229
	33380	383, 269		36420	221, 153		38360	279, 429
	33390	39, 728		36430	193, 601		38370	350, 822
충남	34011	218, 328	제주	36440	221, 479		38380	359, 607
	34012	95, 141		36450	204, 072		38390	396, 177
	34020	421, 491		36460	195, 326		38400	475, 575
	34030	286, 959		36470	211, 601		39010	482, 328
	34040	267, 083		36480	339, 043		39020	437, 457
	34050	364, 158	경북	37011	197, 039			
	34060	266, 129		37012	365, 414			
	34070	28, 500		37020	650, 737			
	34080	353, 881		37030	496, 532			
	34310	286, 334		37040	758, 343			
	34330	268, 817		37050	302, 170			
	34340	170, 057		37060	330, 601			
	34350	179, 969		37070	449, 261			
	34360	219, 687		37080	601, 929			
	34370	268, 640		37090	440, 035			
	34380	266, 623		37100	204, 239			

4) 4단계: 비(非)매칭 교통 폴리곤 보정

- 기지국이 '0'개 매핑된 폴리곤 중에서 2단계에서 병합되지 않은 폴리곤을 동일 읍면동 내에서 가까운 폴리곤과 병합하여 보정
 - － 인근 폴리곤과 중심점간의 거리 기준으로 병합 대상 탐색

제2절 분석 맵 구축 결과 및 검증

1. 분석 맵 구축 결과

- 전국 총 16,335개의 교통폴리곤이 형성됨
 - － 교통폴리곤이 가장 작게 형성된 지역은 서울특별시이고, 가장 크게 형성된 지역은 강원도임
 - 서울특별시는 약 605km²의 면적이 1,664개로 분할되면서 평균 약 0.36km²의 교통폴리곤이 형성되었으며, 강원도는 약 1681km²의 면적이 941개로 분할되면서 평균 약 17.9km²의 교통폴리곤이 형성됨

<표 3- 7> 교통 폴리곤 구축 결과

구분	개수	평균 면적 (m ²)	최소 면적 (m ²)	최대 면적 (m ²)
서울특별시	1,664	363,731	12,062	9,855,240
부산광역시	764	1,026,162	24,521	21,945,462
대구광역시	582	1,511,316	58,253	37,441,954
인천광역시	441	2,492,872	44,783	51,352,949
광주광역시	449	1,109,139	58,826	27,426,655
대전광역시	469	1,150,260	61,139	28,315,975
울산광역시	253	4,195,228	170,947	30,651,525
세종특별자치시	89	5,222,970	236,411	23,661,753
경기도	2,762	3,711,108	6,554	83,428,425
강원도	941	17,863,176	49,696	181,537,669
충청북도	821	9,023,623	165,854	69,705,478
충청남도	1,289	6,410,660	139,520	63,102,745
전라북도	1,190	6,798,629	55,292	82,033,354
전라남도	1,469	8,455,669	55,895	63,545,867
경상북도	1,496	12,720,036	217,495	154,212,850
경상남도	1,328	7,946,745	125,034	71,988,821
제주특별자치도	328	5,685,216	292,897	95,926,333
계	16,335	-		



<그림 3- 6> 전국 교통폴리곤 형성 결과

가. '17년도 교통폴리곤과 '18년도 교통폴리곤 비교

- '17년도 교통폴리곤 대비 '18년도 교통폴리곤은 약 2.5% 증가함
 - 시도별로는 평균 약 7.2% 증가하였으며, 면적은 약 6.8%로 감소함
 - 부산광역시, 경기도, 제주특별자치시를 제외하고 나머지 지역은 모두 '17년도 보다 세분되었으며, 특히 세종특별자치시, 울산광역시, 강원도, 충청북도, 전라북도 지역이 많이 개선된 것으로 나타남
 - 경기도는 교통폴리곤 구축 기준을 변경하면서 불안정한 기지국을 제외한 영향을 많이 받은 것으로 판단됨

<표 3- 8> 교통 폴리곤 구축 결과

구분	'17년도 교통폴리곤		'18년도 교통폴리곤		증감율(%)	
	개수	평균 면적 (㎡)	개수	평균 면적 (㎡)	개수	면적
서울특별시	1,617	375,320	1,664	363,731	2.9	-3.1
부산광역시	789	1,006,100	764	1,026,162	-3.2	2.0
대구광역시	555	1,585,891	582	1,511,316	4.9	-4.7
인천광역시	440	2,655,597	441	2,492,872	0.2	-6.1
광주광역시	422	1,181,355	449	1,109,139	6.4	-6.1
대전광역시	437	1,234,709	469	1,150,260	7.3	-6.8
울산광역시	205	5,218,281	253	4,195,228	23.4	-19.6
세종특별자치시	70	6,653,727	89	5,222,970	27.1	-21.5
경기도	3,204	3,248,864	2,762	3,711,108	-13.8	14.2
강원도	776	21,832,096	941	17,863,176	21.3	-18.2
충청북도	726	10,211,441	821	9,023,623	13.1	-11.6
충청남도	1,218	6,839,509	1,289	6,410,660	5.8	-6.3
전라북도	1,055	7,622,155	1,190	6,798,629	12.8	-10.8
전라남도	1,389	9,119,564	1,469	8,455,669	5.8	-7.3
경상북도	1,433	13,292,880	1,496	12,720,036	4.4	-4.3
경상남도	1,272	8,350,587	1,328	7,946,745	4.4	-4.8
제주특별자치도	329	5,745,843	328	5,685,216	-0.3	-1.1
계	15,937	-	16,335	-		

나. 기존 분석구와 비교

- 교통폴리곤은 기존 시군구 대비 약 1/65, 읍면동 대비 약 1/5 수준의 규모로 구축되어 기존보다 섬세한 수요 분석이 가능해질 것으로 사료됨
- 특히 제주도는 기존 시군구 대비 약 1/164, 읍면동 대비 약 1/8 수준으로 구축되어 그 효과가 가장 클 것으로 기대

<표 3- 9> 교통폴리곤과 기존 분석구 비교

구분	기존 분석구				교통폴리곤		시군구 대비 ¹⁾ (배)	읍면동 대비 ²⁾ (배)
	시군구		읍면동					
	개수	평균 면적 (㎡)	개수	평균 면적 (㎡)	개수	평균 면적 (㎡)		
서울	25	24, 228, 390	424	1, 428, 561	1, 664	363, 731	66. 6	3. 9
부산	16	49, 016, 111	205	3, 825, 648	764	1, 026, 162	47. 8	3. 7
대구	8	109, 997, 567	139	6, 330, 795	582	1, 511, 316	72. 8	4. 2
인천	10	110, 032, 326	150	7, 335, 488	441	2, 492, 872	44. 1	2. 9
광주	5	99, 680, 245	95	5, 246, 329	449	1, 109, 139	89. 8	4. 7
대전	5	107, 968, 256	79	6, 833, 434	469	1, 150, 260	93. 8	5. 9
울산	5	212, 344, 456	56	18, 959, 326	253	4, 195, 228	50. 6	4. 5
세종	1	465, 174, 895	14	33, 226, 778	89	5, 222, 970	89. 0	6. 4
경기도	42	244, 226, 677	561	18, 284, 350	2, 762	3, 711, 108	65. 8	4. 9
강원도	18	934, 323, 824	193	87, 139, 009	941	17, 863, 176	52. 3	4. 9
충청북도	14	529, 487, 831	153	48, 449, 867	821	9, 023, 623	58. 6	5. 4
충청남도	16	516, 852, 302	207	39, 949, 936	1, 289	6, 410, 660	80. 6	6. 2
전라북도	15	539, 754, 611	241	33, 594, 685	1, 190	6, 798, 629	79. 3	4. 9
전라남도	22	565, 052, 276	297	41, 855, 724	1, 469	8, 455, 669	66. 8	4. 9
경상북도	24	793, 211, 576	332	57, 340, 596	1, 496	12, 720, 036	62. 3	4. 5
경상남도	22	479, 937, 799	314	33, 626, 215	1, 328	7, 946, 745	60. 4	4. 2
제주도	2	933, 180, 426	43	43, 403, 741	328	5, 685, 216	164. 0	7. 6
총계	250	-	3, 503	-	16, 335	-	65. 3	4. 7

1) 교통폴리곤 개수 / 시군구 개수

2) 교통폴리곤 개수 / 읍면동 개수

2. 분석 맵 구축 검증

- 기존 분석구와 교통폴리곤 전국 분포 비율을 비교한 결과 대체로 기존 분석구 분포 비율과 교통폴리곤 분포 비율이 유사하게 나타남

<표 3- 10> 기존 교통존과 교통 폴리곤 비교

구분	기존 분석구				교통폴리곤	
	시군구		읍면동			
	개수	비율	개수	비율	개수	비율
서울특별시	25	10.0%	424	12.1%	1,664	10.2%
부산광역시	16	6.4%	205	5.9%	764	4.7%
대구광역시	8	3.2%	139	4.0%	582	3.6%
인천광역시	10	4.0%	150	4.3%	441	2.7%
광주광역시	5	2.0%	95	2.7%	449	2.7%
대전광역시	5	2.0%	79	2.3%	469	2.9%
울산광역시	5	2.0%	56	1.6%	253	1.5%
세종특별자치시	1	0.4%	14	0.4%	89	0.5%
경기도	42	16.8%	561	16.0%	2,762	16.9%
강원도	18	7.2%	193	5.5%	941	5.8%
충청북도	14	5.6%	153	4.4%	821	5.0%
충청남도	16	6.4%	207	5.9%	1,289	7.9%
전라북도	15	6.0%	241	6.9%	1,190	7.3%
전라남도	22	8.8%	297	8.5%	1,469	9.0%
경상북도	24	9.6%	332	9.5%	1,496	9.2%
경상남도	22	8.8%	314	9.0%	1,328	8.1%
제주특별자치도	2	0.8%	43	1.2%	328	2.0%
계	250	-	3,503	-	16,335	-

제4장 모바일 자료 기반 View T 2.0 서비스 기능 개발

제1절 기능 정의 및 UI 설계

제2절 기반 DB 구축

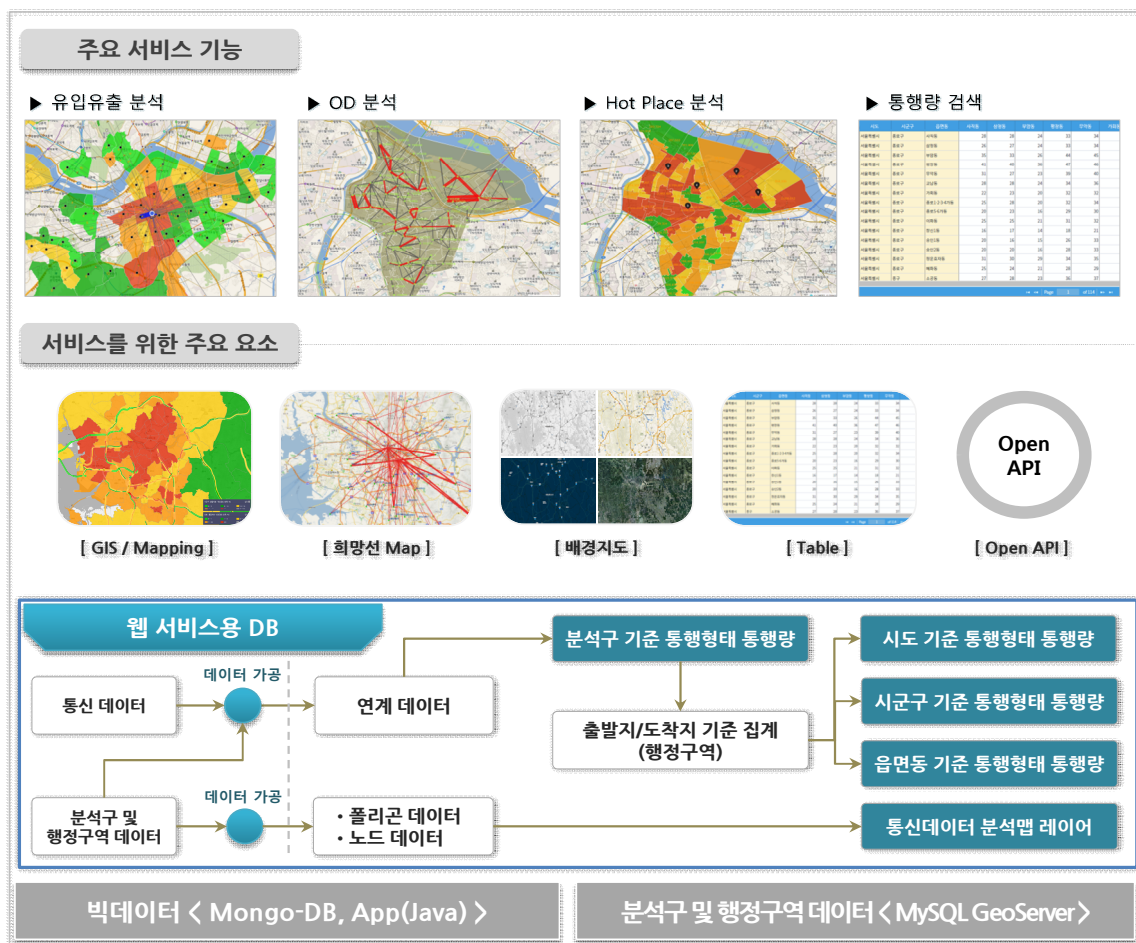
제3절 시스템 운영 및 유지보수

제4장 모바일 자료 기반 View-T 2.0 서비스 기능 개발

제1절 기능 정의 및 UI 설계

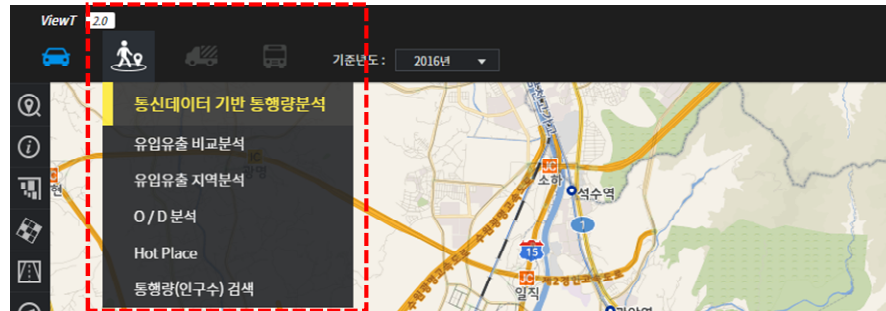
1. 개요

- GIS 기반으로 유입유출 지역분석, 유입유출 비교분석, OD 분석, Hot Place 분석을 할수 있는 서비스와 사용자가 조건을 설정하여 해당되는 통행량(이동인구수)를 검색하고 이를 파일형태로 출력할 수 있는 서비스를 개발함



<그림 4- 1> 서비스의 구성

- 개발된 서비스는 2017년에 개발된 교통 모니터링·데이터 제공·분석 플랫폼(View-T)의 통행량 분석 메뉴를 통해 이용 가능함



<그림 4-2> 통신데이터 기반의 통행량분석의 메뉴

<표 4-1> 메뉴의 구성

페이지메뉴	상세 메뉴	기능	설명
통신데이터 기반 통행량 분석	유입유출 비교분석	유입량/유출량 선택	<ul style="list-style-type: none"> 2개 지역에 대한 비교분석 통행량 선택
		출발지, 도착지 선택	<ul style="list-style-type: none"> 비교지역에 대한 출발지, 도착지 선택 도착지(유입), 출발지(유출) 2개 지역 선택 지도를 통한 선택, ID 검색을 통한 선택 통행량의 집계단위 설정 출발지(유입), 도착지(유출)의 집계단위 선택
		상세설정	<ul style="list-style-type: none"> 기간, 요일, 시간대, 트립타임, 성별, 연령대 설정
		표출지도 설정	<ul style="list-style-type: none"> 비교지역의 분석결과에 대한 희망선도, 분포도의 범례설정
		출력	<ul style="list-style-type: none"> 분석결과에 대한 통신데이터 파일출력(csv)
	유입유출 지역분석	유입량/유출량 설정	<ul style="list-style-type: none"> 특정지역의 분석 통행량 선택
		지도 Scale 설정	<ul style="list-style-type: none"> 시도, 시군구, 읍면동, 분석구 단위의 표출지도에 대한 ZOOM 범위 설정

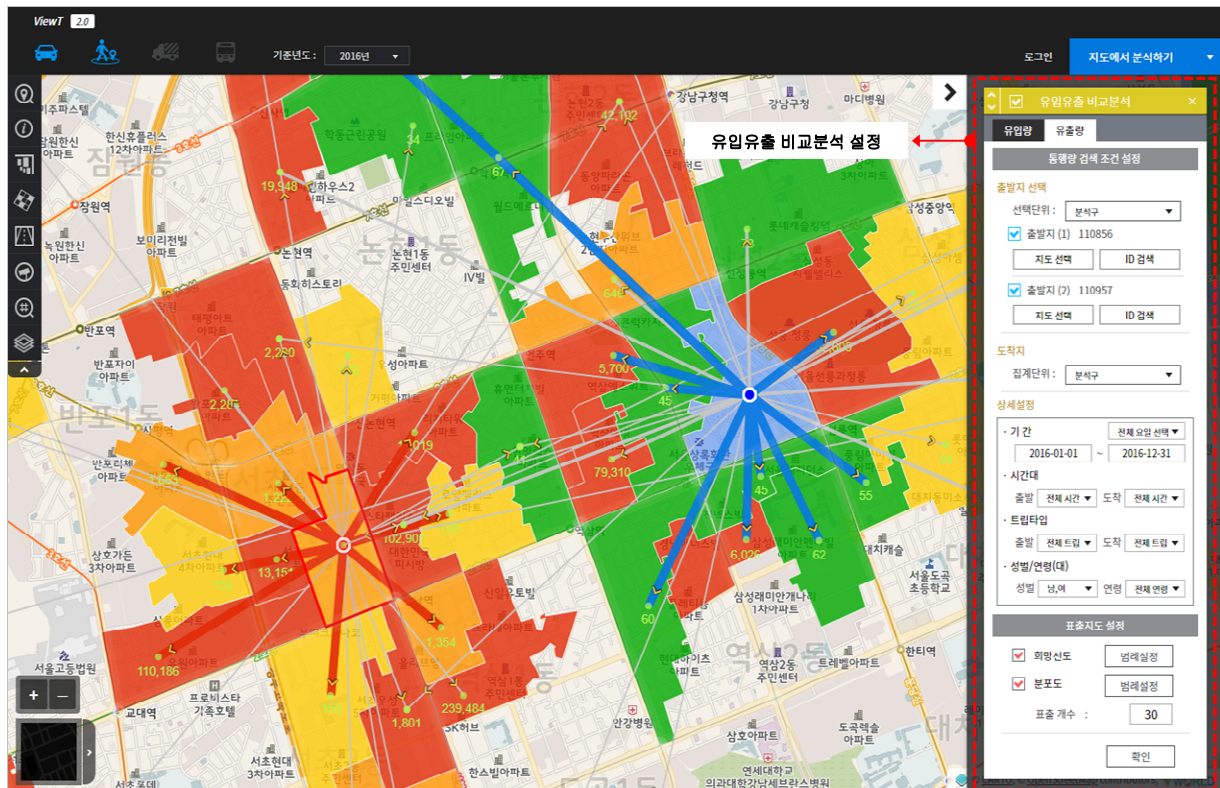
페이지메뉴	상세 메뉴	기능	설명
통신데이터 기반 통행량 분석	유입유출 지역분석	출발지, 도착지 선택	<ul style="list-style-type: none"> • 유입 : 도착지 선택 • 유출 : 출발지 선택 • 지도를 통한 선택, ID 검색을 통한 선택
		상세설정	<ul style="list-style-type: none"> • 기간, 요일, 시간대, 트립타입, 성별, 연령대 설정
		표출지도 설정	<ul style="list-style-type: none"> • 분석결과에 대한 희망선도 범례설정 • 분석결과에 대한 분포도 범례설정
		출력	<ul style="list-style-type: none"> • 분석결과에 대한 통신데이터 파일출력 (csv)
	OD 분석	분석지역 선택	<ul style="list-style-type: none"> • 시도, 시군구, 읍면동 단위로 선택
		통행량 산출방식 선택	<ul style="list-style-type: none"> • 단방향 통행량, 양방향 통행량 설정
		상세설정	<ul style="list-style-type: none"> • 기간, 요일, 시간대, 트립타입, 성별, 연령대 설정
		표출지도 설정	<ul style="list-style-type: none"> • 분석결과에 대한 희망선도 범례설정
		출력	<ul style="list-style-type: none"> • 분석결과에 대한 통신데이터 파일출력 (csv)
	HotPalce 분석	분석지역 설정	<ul style="list-style-type: none"> • 시도, 시군구, 읍면동 단위로 선택
		상세설정	<ul style="list-style-type: none"> • 기간, 요일, 시간대, 성별, 연령대 설정
		표출지도 설정	<ul style="list-style-type: none"> • 분석결과에 대한 Heat Map 범례설정 • 분석결과에 대한 분포도 범례설정
		출력	<ul style="list-style-type: none"> • 분석결과에 대한 통신데이터 파일출력 (csv)
	통행량 (인구수) 검색	검색조건 설정	<ul style="list-style-type: none"> • 기간, 요일, 시간대, 트립타입, 성별, 연령대 설정
		출력	<ul style="list-style-type: none"> • 검색결과에 대한 통신데이터 파일출력 (csv)

2. 주요 기능

가. 유입유출 비교분석

- 시도, 시군구, 읍면동, 분석구 단위로 복수개의 지역을 선택하여 각 지역으로 유입되는 통행량 또는 각 지역에서 유출되는 통행량을 비교분석할 수 있도록 개발함
 - 최대 5개의 지역 비교 분석 가능
- 유입량 분석은 출발지의 집계단위, 유출량 분석은 도착지의 집계단위를 시도, 시군구, 읍면동, 분석구 단위로 설정할 수 있도록 개발함

(※ 집계단위 : 통행량 산출의 최소 공간범위)
- 기간, 요일, 출발 및 도착시간대, 출발 및 도착 트립타입, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 분석결과를 희망선도, 분포도로 지도에 표출하고 표출지도의 범례를 설정할 수 있도록 개발함
- 유입유출 비교분석에 활용된 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 4-3> 유입유출 비교분석의 실행 화면

1) 기능설명

① 유입유출 비교분석의 설정 UI

- 선택한 지역의 유입량, 유출량을 비교분석할 수 있도록 UI를 구성함
- 통행량 검색 조건 설정을 통해서 분석할 지역을 선택하고 통신데이터의 검색조건을 설정할 수 있도록 개발함
- 분석결과를 희망선도, 분포도로 지도에 표출하고 표출지도의 범례를 설정할 수 있도록 개발함

■ 유입량 분석 시 설정 UI



The UI for Inflow Analysis Settings is divided into several sections. At the top, there are tabs for '유입량' (Inflow) and '유출량' (Outflow). Below this is a '통행량 검색 조건 설정' (Travel Volume Search Condition Setting) section. It includes a '도착지 선택' (Destination Selection) section with a '선택단위' (Selection Unit) dropdown set to '분석구' (Analysis District). There are two checked checkboxes for '도착지 (1) 110856' and '도착지 (2) 110957', each with a '지도 선택' (Map Selection) and 'ID 검색' (ID Search) button. Below this is a '출발지' (Origin) section with a '집계단위' (Aggregation Unit) dropdown set to '분석구'. The '상세설정' (Detailed Settings) section includes '기간' (Period) with a date range from 2016-01-01 to 2016-12-31, '시간대' (Time Period) with '출발' (Departure) and '도착' (Arrival) time selection, '트립타입' (Trip Type) with '출발' (Departure) and '도착' (Arrival) trip type selection, and '성별/연령(대)' (Gender/Age Group) with '성별' (Gender) set to '남,여' (Male, Female) and '연령' (Age) set to '전체 연령' (All Ages). The '표출지도 설정' (Map Display Setting) section has three checked checkboxes for '희망선도' (Desired Line Map), '분포도' (Distribution Map), and '표출 개수' (Display Count) set to 30. There are '범례설정' (Legend Setting) buttons for each map type and '출력' (Output) and '확인' (Confirm) buttons at the bottom.

① 출발지 및 도착지 설정

- 2개의 비교지역 설정
 - 유입 : 도착지
 - 유출 : 출발지
 - 레이어 ON/OFF 설정

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

③ 표출지도 설정

- 희망선도 범례 설정
- 분포도 범례 설정
- 표출지도 ON/OFF
- 표출개수 설정

■ 유출량 분석 시 설정 UI



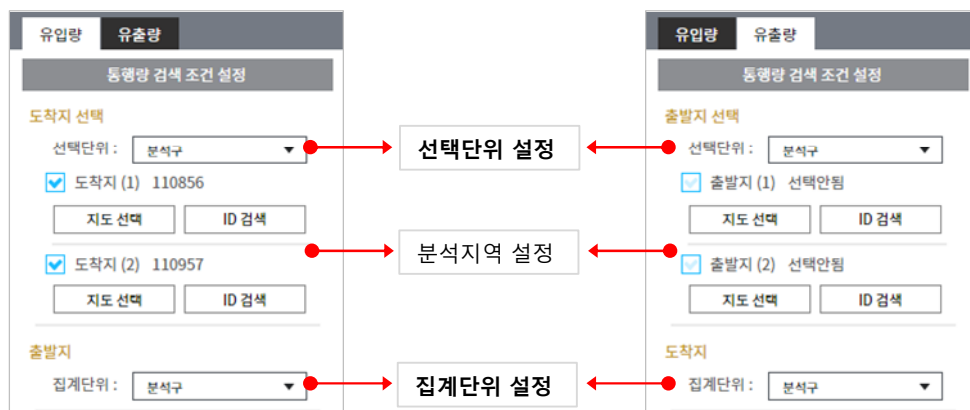
The UI for Outflow Analysis Settings is similar to the Inflow UI but with some differences. It has tabs for '유입량' (Inflow) and '유출량' (Outflow). The '통행량 검색 조건 설정' (Travel Volume Search Condition Setting) section includes a '출발지 선택' (Origin Selection) section with a '선택단위' (Selection Unit) dropdown set to '분석구' (Analysis District). There are two checked checkboxes for '출발지 (1) 선택안됨' (Origin (1) Not Selected) and '출발지 (2) 선택안됨' (Origin (2) Not Selected), each with a '지도 선택' (Map Selection) and 'ID 검색' (ID Search) button. Below this is a '도착지' (Destination) section with a '집계단위' (Aggregation Unit) dropdown set to '분석구'. The '상세설정' (Detailed Settings) section is identical to the Inflow UI. The '표출지도 설정' (Map Display Setting) section is also identical to the Inflow UI, with checked checkboxes for '희망선도' (Desired Line Map), '분포도' (Distribution Map), and '표출 개수' (Display Count) set to 30. There are '범례설정' (Legend Setting) buttons for each map type and '출력' (Output) and '확인' (Confirm) buttons at the bottom.

<그림 4-4> 유입유출 비교분석 설정 UI

- 사용자가 분석결과를 쉽게 확인할 수 있도록 다양한 ON/OFF 기능을 제공함
 - 선택 지역에 대한 레이어 ON/OFF 기능을 개발함
 - 희망선도 및 분포도에 대한 ON/OFF 기능을 개발함
- 지도에 표출되는 분포도, 희망선의 개수를 설정할 수 있도록 개발함

② 통행량 검색조건 설정 및 분석

- 출발지 및 도착지 설정
 - 선택단위 설정을 통해서 분석지역을 시도, 시군구, 읍면동, 분석구 단위로 선택할 수 있도록 개발함
 - 집계단위 설정을 통해서 분석지역으로 유입되는 통행량, 분석지역에서 유출되는 통행량을 시도, 시군구, 읍면동, 분석구 단위로 지도에 표출할 수 있도록 개발함
 - 유입량 분석의 출발지 및 도착지 설정에 관한 주요내용은 다음과 같음
 - 도착지는 시도, 시군구, 읍면동, 분석구 단위로 최대 5개의 지역을 선택할 수 있도록 개발함
 - 출발지는 설정한 도착지로 유입되는 통행량을 시도, 시군구, 읍면동, 분석구 단위로 집계할 수 있도록 개발함
 - 유출량 분석의 출발지 및 도착지 설정에 관한 주요내용은 다음과 같음
 - 출발지는 시도, 시군구, 읍면동, 분석구 단위로 최대 5개 지역을 선택할 수 있도록 개발함
 - 도착지는 설정한 출발지에서 유출되는 통행량을 시도, 시군구, 읍면동, 분석구 단위로 집계할 수 있도록 개발함

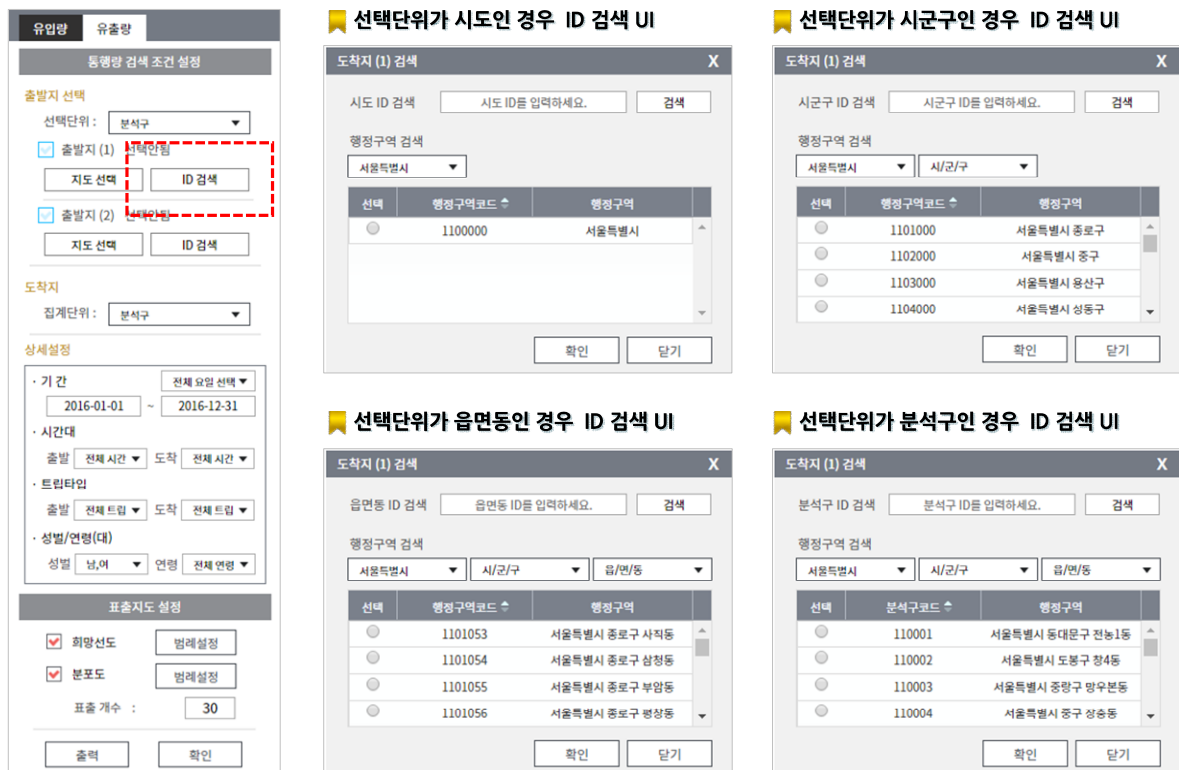


<그림 4- 5> 유입유출 비교분석의 출발지 및 도착지 설정

- 지도 선택과 ID 검색을 통해서 사용자가 출발지, 도착지를 쉽게 선택할 수 있도록 개발함



<그림 4- 6> 지도 선택을 통한 분석지역 선택



<그림 4- 7> ID 검색을 통한 분석지역 선택

○ 상세조건 설정

- 사용자가 기간, 요일, 출발시간대, 도착시간대, 출발트립타입, 도착트립타입, 성별, 연령대에 대한 조건을 설정하여 유입량, 유출량을 분석할 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 트립타입 : 주간상주, 야간상주, 잠재상주
- 성별 : 남, 여
- 연령대 : 00대, 10대, 20대, 30대, 40대, 50대, 60대, 70대, 80대, 90대, 100대, 110대

○ 분석

- 출발지 및 도착지 설정 정보, 상세설정 정보에 맞는 통신데이터를 모바일 통행형태 통행량 DB에서 검색함
- 유입량 분석은 검색된 통신데이터를 기반으로 다음과 같이 분석을 실시함
 - 출발지-도착지를 데이터의 분류기준으로 설정하고 분류기준에 따라 데이터를 분류함
 - 출발지-도착지별 이동인구수 합계를 산출함
- 유출량 분석은 검색된 통신데이터를 기반으로 다음과 같이 분석을 실시함
 - 도착지-출발지를 데이터의 분류기준으로 설정하고 분류기준에 따라 데이터를 분류함
 - 도착지-출발지별 이동인구수 합계를 산출함

③ 표출지도의 범례설정

- 유입유출 비교분석의 결과는 행정구역(시도, 시군구, 읍면동), 분석구 단위로 산출됨
- 행정구역 및 분석구 공간정보와 분석결과를 연계하여 폴리곤으로 분석결과를 표출할 수 있도록 개발함
- 또한, 행정구역 및 분석구 공간정보에 대한 노드(점) 정보와 분석결과와 출발지 및 도착지 정보를 연계하여 분석결과를 희망선으로 지도에 표출할 수 있도록 개발함
- 선택 지역에 대한 유입량과 유출량을 지도에서 명확히 구분할 수 있도록 희망선도의 범례설정 UI를 구성함
- 선택 지역에 대한 유입량과 유출량의 합계를 지도에서 명확히 구분할 수 있도록 분포도의 범례설정 UI를 구성함

- 희망선도 범례설정의 주요내용은 다음과 같음
 - 2개 지역에 대한 희망선도의 범례설정
 - 표출등급 구분 및 범위 설정
 - 등급별 희망선의 지도표출, 색상, 두께 설정
 - 표출결과 라벨표출 설정 및 희망선도의 투명도 설정
 - 사용자의 편의성을 고려하여 분석결과의 최소값, 최대값 제공

① 희망선도 ON/OFF 설정

② 희망선도 적용 지역 선택

③ 부가정보

- 전체결과값 : 분석결과의 최소값, 최대값

④ 희망선의 등급 설정

- 최소값, 최대값 설정
- 지도표출 설정
- 색상 설정
- 희망선의 두께 설정

⑤ 외곽선, 라벨, 투명도 설정

- 분석결과의 라벨표출 설정
- 희망선의 투명도 설정

<그림 4- 8> 유입유출 비교분석의 희망선도 범례설정 UI

- 분포도 범례설정의 주요내용은 다음과 같음

- 표출등급 구분 및 범위 설정
- 등급별 지도표출, 색상 설정
- 정보 없음(NULL)의 지도표출 및 색상 설정
- 분포도 외곽선의 지도표출 및 색상 설정
- 분포도의 투명도 설정
- 분석결과의 라벨표출 설정(라벨의 크기 및 색상 설정)
- 사용자의 편의성을 고려하여 분석결과의 최소값, 최대값 제공

① 분포도 ON/OFF 설정

② 부가정보

- 전체결과값 : 분석결과의 최소값, 최대값

③ 분포도의 등급 설정

- 최소값, 최대값 설정
- 지도표출 설정
- 색상 설정

④ 외곽선, 라벨, 투명도 설정

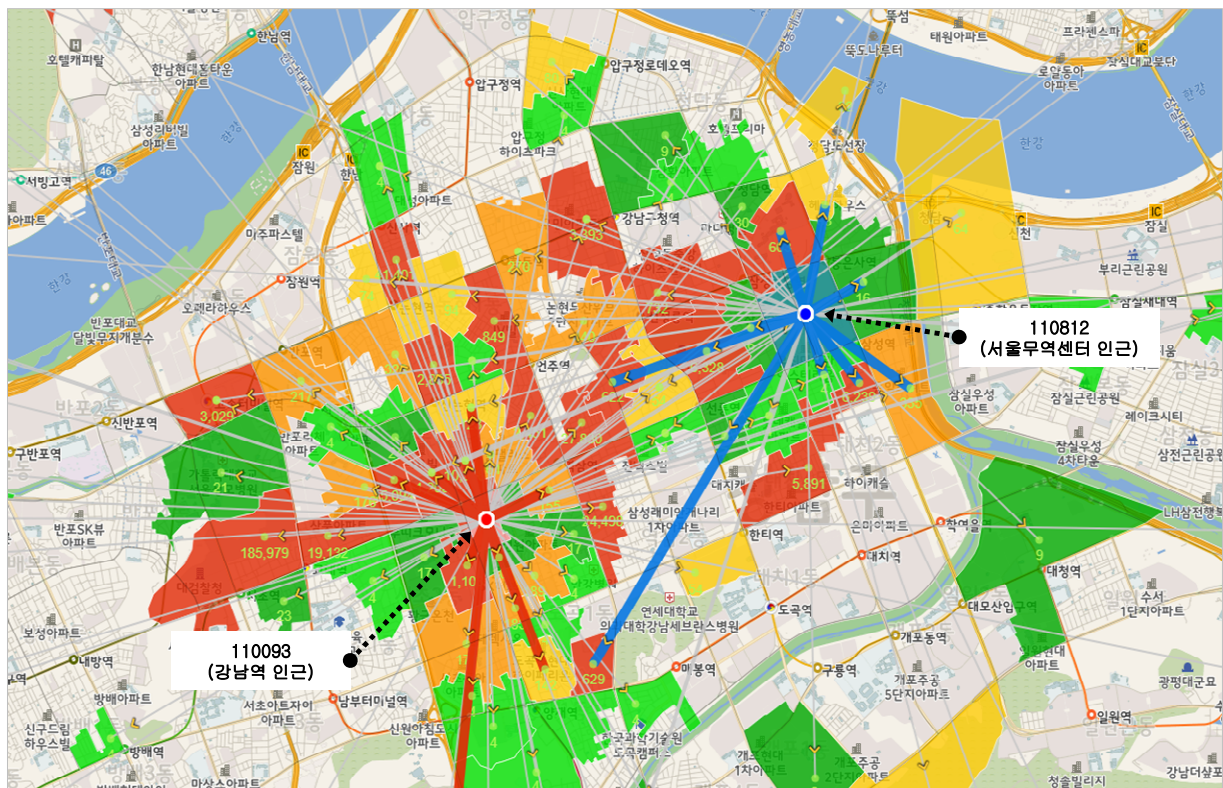
- 분포도 외곽선의 지도표출 및 색상 설정
- 분석결과의 라벨표출 설정
- 분포도의 투명도 설정

<그림 4- 9> 유입유출 비교분석의 분포도 범례설정 UI

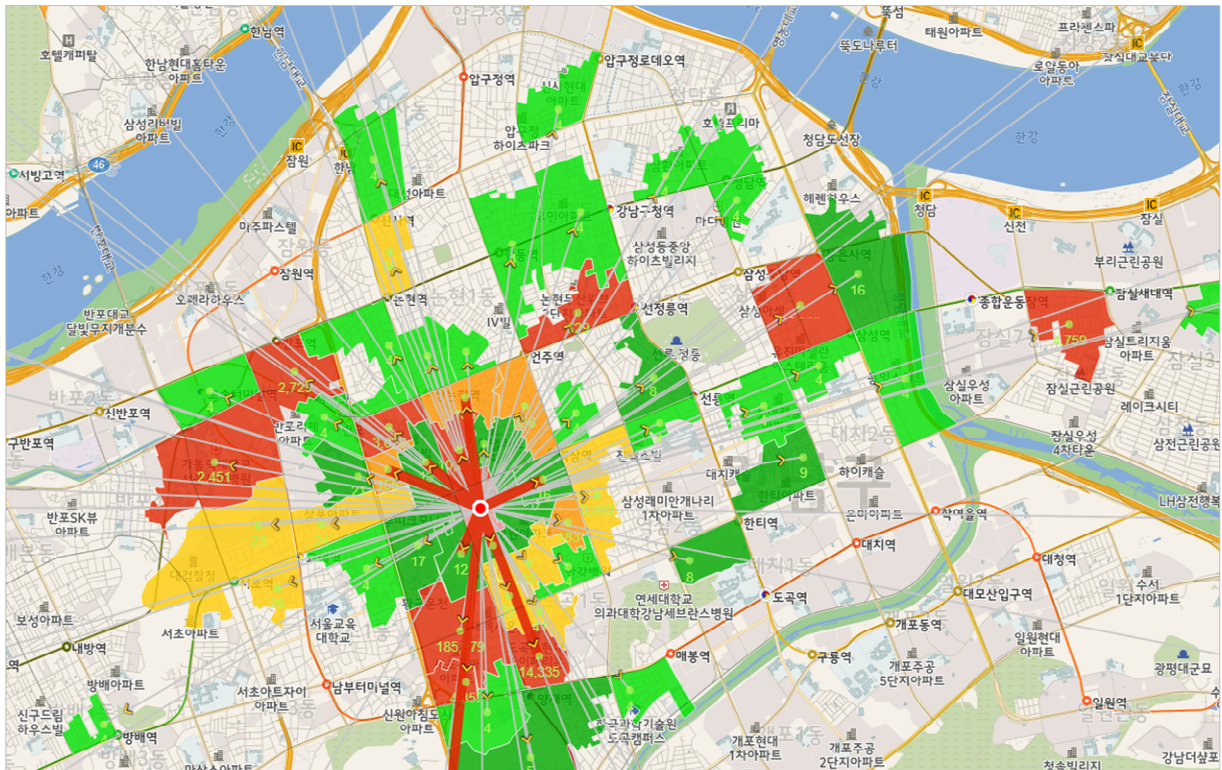
④ 유입유출 비교분석의 결과 예시

○ 분석 조건

- 통행량 : 유출량
- 분석지역
 - 출발지(1) : 110093(강남역 인근)
 - 출발지(2) : 110812(서울무역센터 인근)
- 도착지 집계단위 : 분석구
- 기간 : 2016.06.01. ~ 2016.06.30.
- 요일 : 전체 요일
- 시간대 : 출발 및 도착 전체시간대
- 성별 : 남, 여
- 연령대 : 전체 연령대



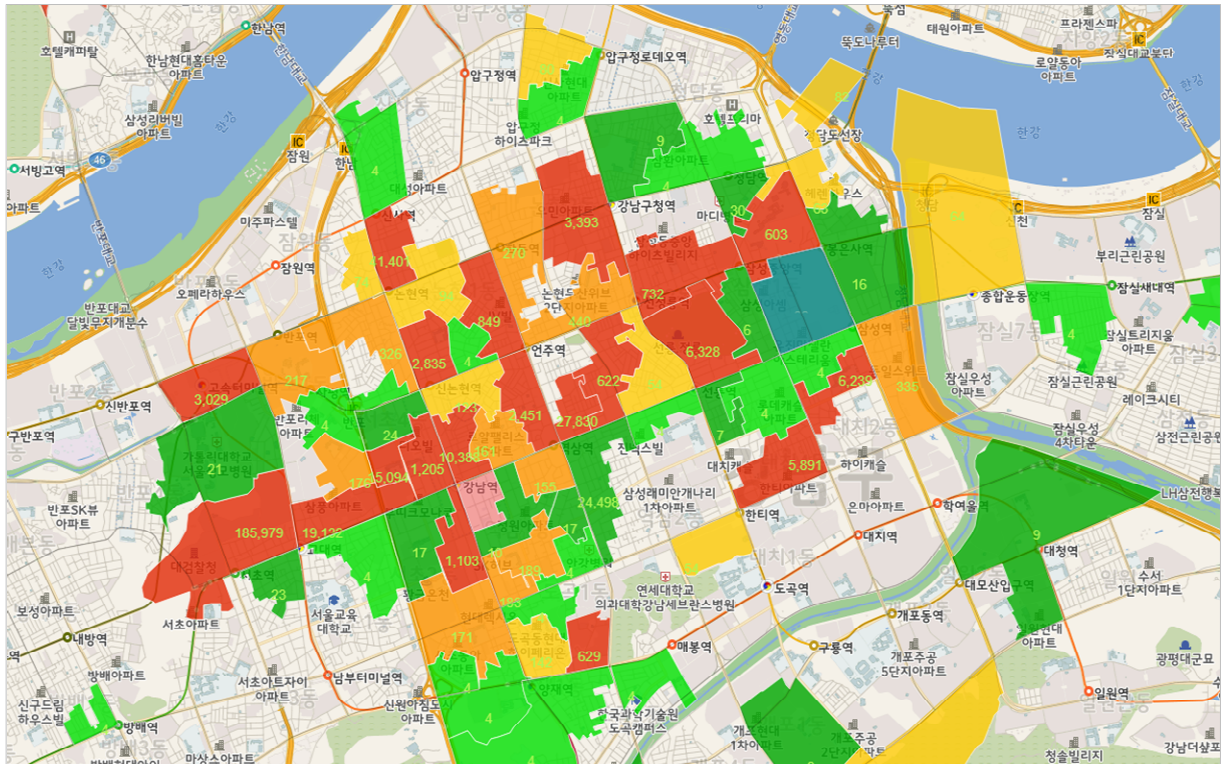
<그림 4- 10> 유출량 비교분석의 결과화면 예시 (출발지(1) ON, 출발지(2) ON)



<그림 4- 11> 유출량 비교분석의 결과화면 예시 (출발지(1) ON, 출발지(2) OFF)



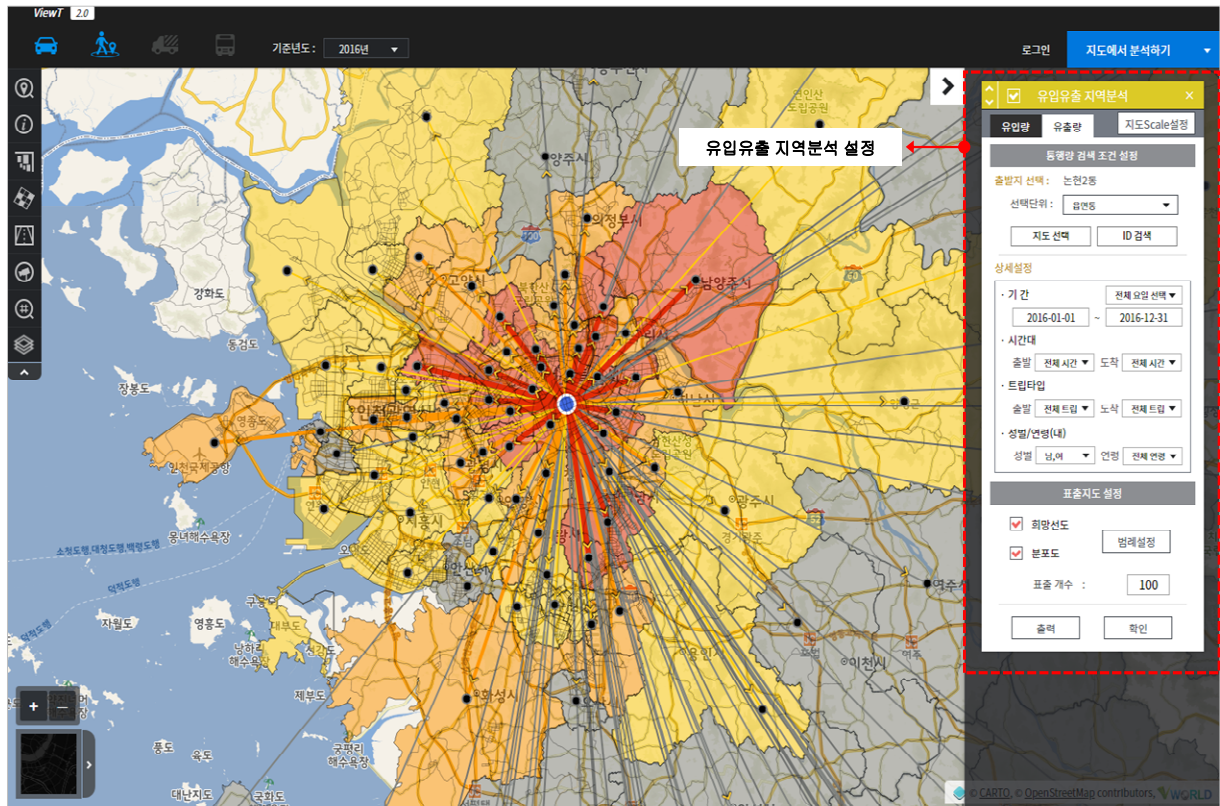
<그림 4- 12> 유출량 비교분석의 결과화면 예시 (희망선도 ON, 분포도 OFF)



<그림 4- 13> 유출량 비교분석의 결과화면 예시 (희망선도 OFF, 분포도 ON)

나. 유입유출 지역분석

- 전국의 시도, 시군구, 읍면동, 분석구에서 특정지역으로 유입되는 통행량과 특정지역에서 전국의 시도, 시군구, 읍면동, 분석구로 유출되는 통행량을 분석할 수 있도록 개발함
- 시도, 시군구, 읍면동, 분석구 단위로 집계된 유입 통행량, 유출 통행량을 지도의 ZOOM 단계에 따라 지도에 표출되도록 개발함
- 유입 통행량, 유출 통행량의 분석결과에 대한 시도, 시군구, 읍면동, 분석구 단위의 표출지도가 표현되는 ZOOM 단계를 사용자가 설정할 수 있도록 개발함
- 기간, 요일, 출발 및 도착시간대, 출발 및 도착 트립타입, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 분석결과를 희망선도, 분포도로 표출할 수 있도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 유입유출 지역분석에 활용된 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 4- 14> 유입유출 지역분석의 실행 화면

1) 기능설명

① 유입유출 지역분석의 설정 UI

- 선택단위를 통해서 선택된 시도, 시군구, 읍면동, 분석구에 대하여 유입량, 유출량 분석이 가능하도록 UI를 구성함
- 분석결과가 반영된 시도, 시군구, 읍면동, 분석구 단위의 지도가 표출되는 ZOOM 단계를 사용자가 설정할 수 있도록 개발함
- 검색조건 설정을 통해서 분석대상 통신데이터를 설정할 수 있도록 개발함
- 분포도와 희망선도로 표출되는 지도의 범례를 설정할 수 있도록 개발함
- 지도에 표출되는 분포도, 희망선의 개수를 설정할 수 있도록 개발함

유입량 분석 시 설정 UI



The UI for Inflow Analysis Settings includes a top navigation bar with '유입량' (Inflow), '유출량' (Outflow), and '지도Scale설정' (Map Scale Setting). The main section is titled '통행량 검색 조건 설정' (Travel Volume Search Condition Setting). It features a '도착지 선택' (Destination Selection) dropdown set to '선택안됨' (Not Selected), a '선택단위' (Selection Unit) dropdown set to '분석구' (Analysis District), and buttons for '지도 선택' (Map Selection) and 'ID 검색' (ID Search). Below this is the '상세설정' (Detailed Setting) section, which includes a date range from '2016-01-01' to '2016-12-31', a '시간대' (Time Period) dropdown, and checkboxes for '출발' (Departure) and '도착' (Arrival). The '트립타입' (Trip Type) section has checkboxes for '출발' (Departure) and '도착' (Arrival). The '성별/연령(대)' (Gender/Age Group) section has dropdowns for '성별' (Gender) and '연령' (Age Group). At the bottom, there is a '표출지도 설정' (Map Display Setting) section with checkboxes for '희망선도' (Desired Map) and '분포도' (Distribution Map), a '범례설정' (Legend Setting) button, and a '표출 개수' (Display Count) input field set to '100'. Finally, there are '출력' (Output) and '확인' (Confirm) buttons.

① 표출지도 ZOOM 단계 설정

- 시도, 시군구, 읍면동, 분석구 단위의 표출지도

② 출발지 및 도착지 설정

- 유입 : 도착지
- 유출 : 출발지

③ 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

④ 표출지도 설정

- 희망선도 범례 설정
- 분포도 범례 설정

유출량 분석 시 설정 UI



The UI for Outflow Analysis Settings is similar to the Inflow version but includes a '출발지 선택' (Departure Location Selection) dropdown set to '110580'. The '상세설정' (Detailed Setting) section is identical to the Inflow version. The '표출지도 설정' (Map Display Setting) section has the '분포도' (Distribution Map) checkbox checked and the '표출 개수' (Display Count) input field set to '50'. The '출력' (Output) and '확인' (Confirm) buttons are at the bottom.

<그림 4- 15> 유입유출 지역분석의 설정 UI

② 표출지도의 ZOOM 단계 설정

- 특정지역으로 유입되는 통행량과 특정지역에서 유출되는 통행량을 지도의 ZOOM 단계에 따라 시도, 시군구, 읍면동, 분석구 단위로 표출할 수 있도록 개발함
- 지도 Scale 설정을 통해서 지도가 표출되는 ZOOM 단계를 설정할 수 있도록 개발함

지도Scale설정

· 적용가능한 지도의 줌 레벨은 7 ~ 18 입니다.

지도 표시단위	줌 레벨
시도	8
시군구	10
읍면동	12
분석구	14

적용 **닫기**

<그림 4- 16> 유입유출 지역분석의 표출지도 ZOOM 단계 설정 UI

③ 통행량 검색조건 설정 및 분석

○ 출발지 및 도착지 설정

- 선택단위 설정을 통해서 시도, 시군구, 읍면동, 분석구 단위로 분석지역을 선택할 수 있도록 개발함
- 유입량 분석은 도착지, 유출량 분석은 출발지 선택을 통해서 특정지역의 유입 통행량과 유출 통행량을 분석할 수 있도록 개발함
- 지도 선택과 ID 검색을 통해서 출발지, 도착지를 선택할 수 있도록 개발함

○ 상세조건 설정

- 사용자가 기간, 요일, 출발시간대, 도착시간대, 출발트립타입, 도착트립타입, 성별, 연령대에 대한 조건을 설정하여 유입 통행량, 유출 통행량을 분석할 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 트립타입 : 주간상주, 야간상주, 잠재상주
 - 성별 : 남, 여
 - 연령대 : 00대, 10대, 20대, 30대, 40대, 50대, 60대, 70대, 80대, 90대, 100대, 110대

○ 분석

- 모바일 통행형태 통행량 DB에서 출발지 및 도착지 설정 정보, 상세설정 정보에 맞는 통신 데이터를 검색함
- 유입량 분석은 출발지를 시도, 시군구, 읍면동, 분석구로 구분하여 데이터를 분류하고 다음과 같이 통행량을 산출함
 - 시도 출발지-도착지별 이동인구수 합계를 산출함
 - 시군구 출발지-도착지별 이동인구수 합계를 산출함
 - 읍면동 출발지-도착지별 이동인구수 합계를 산출함
 - 분석구 출발지-도착지별 이동인구수 합계를 산출함
- 유출량 분석은 도착지를 시도, 시군구, 읍면동, 분석구로 구분하여 데이터를 분류하고 다음과 같이 통행량을 산출함
 - 출발지-시도 도착지별 이동인구수 합계를 산출함
 - 출발지-시군구 도착지별 이동인구수 합계를 산출함
 - 출발지-읍면동 도착지별 이동인구수 합계를 산출함
 - 출발지-분석구 도착지별 이동인구수 합계를 산출함

④ 표출지도의 범례설정

- 유입유출 지역분석의 결과는 행정구역(시도, 시군구, 읍면동), 분석구 단위로 산출됨
- 행정구역 및 분석구 공간정보와 분석결과를 연계하여 폴리곤으로 분석결과를 표출할 수 있도록 개발함
- 또한, 행정구역 및 분석구 공간정보에 대한 노드(점) 정보와 분석결과와 출발지 및 도착지 정보를 연계하여 분석결과를 희망선으로 지도에 표출할 수 있도록 개발함
- 사용자가 분포도와 희망선도에 대한 범례를 설정할 수 있도록 개발함
 - 희망선도 및 분포도 ON/OFF
 - 지도에 표출되는 분포도 및 희망선의 표출개수 설정
 - ZOOM 단계에 따라 시도, 시군구, 읍면동, 분석구로 표출되는 지도의 형태 설정
 - 분포도 및 희망선의 표출 등급 구분 및 등급별 표출 설정
 - 등급별 분포도의 색상, 희망선의 두께 설정
 - 정보없음의 색상 및 지도표출 설정
 - 분포도의 외곽선 색상 설정

- 표출결과 라벨표출 설정 및 표출지도의 투명도 설정
- 해당지도의 ZOOM 범위 정보 제공
- 분석결과 최소값, 최대값 제공
- 사용자 편의성을 고려하여 부가정보를 제공하는 기능을 개발함
- 해당지도의 ZOOM 범위 정보 제공
- 분석결과 최소값, 최대값 제공

표출지도 설정

① ☒ 희망선도

② ☒ 분포도

표출 개수 : 50

범례설정

③ 시도 시군구 읍면동 분석구

④ 표출범위(zoom): 14 ~ 18
전체 결과값: 4 ~ 284,449

선택	심볼	최소	최대	희망선 두께
<input checked="" type="checkbox"/>	■	0	10	2 ▼
⑤ <input checked="" type="checkbox"/>	■	10	24	2 ▼
<input checked="" type="checkbox"/>	■	24	84	4 ▼
<input checked="" type="checkbox"/>	■	84	441	4 ▼
<input checked="" type="checkbox"/>	■	441	284,449	10 ▼
<input type="checkbox"/>	■	0	0	0 ▼
<input type="checkbox"/>	■	0	0	0 ▼
<input type="checkbox"/>	■	정보없음		

⑥ ☒ 외곽선 투명도 30 ▼

· 설정

① 희망선도 ON/OFF 설정

② 분포도 및 희망선의 표출개수 설정

③ ZOOM 단계별 지도범례 설정

- 시도, 시군구, 읍면동, 분석구

④ 부가정보

- 표출범위 : 해당지도의 ZOOM 범위
- 전체결과값 : 분석결과 최소값, 최대값

⑤ 분포도 및 희망선 등급 설정

- 최소값, 최대값 설정
- 지도표출 설정
- 색상 설정
- 희망선의 두께 설정
- 정보없음의 색상 및 지도표출 설정

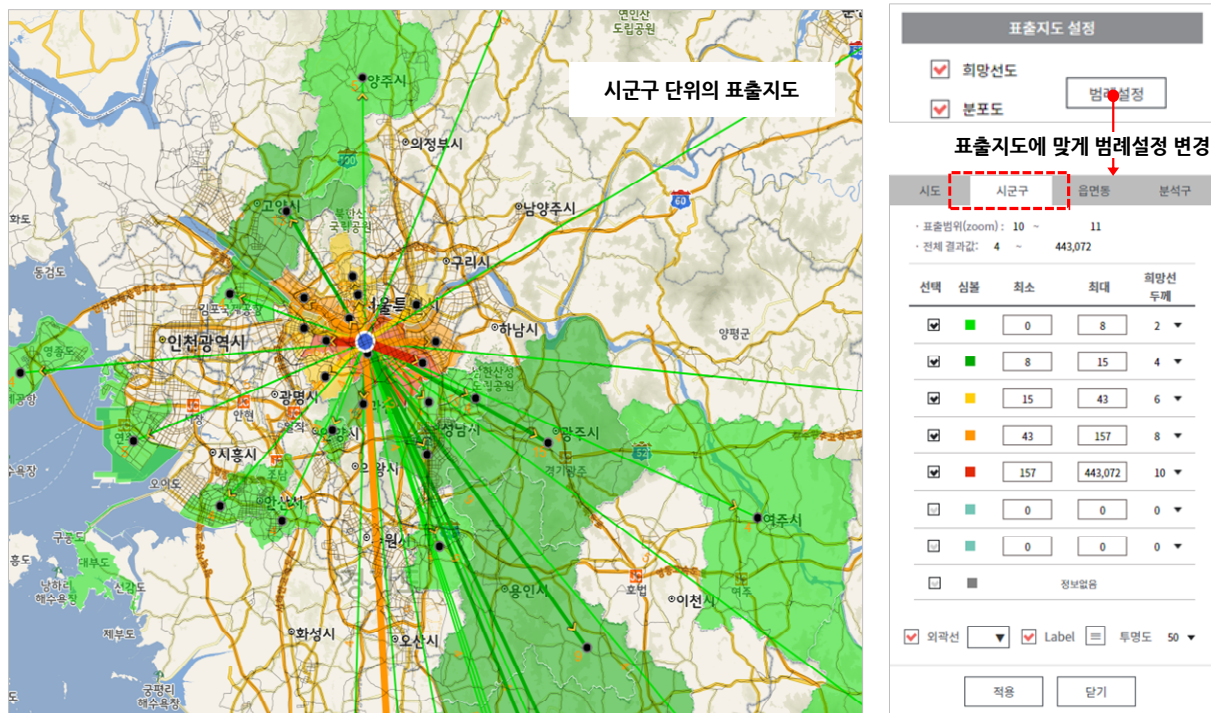
⑥ 외곽선, 라벨, 투명도 설정

- 외곽선 : 분포도의 외곽선 색상 설정
- 분석결과 라벨표출 설정
- 분포도, 희망선의 투명도 설정

<그림 4- 17> 유입유출 지역분석의 표출지도 범례설정 UI

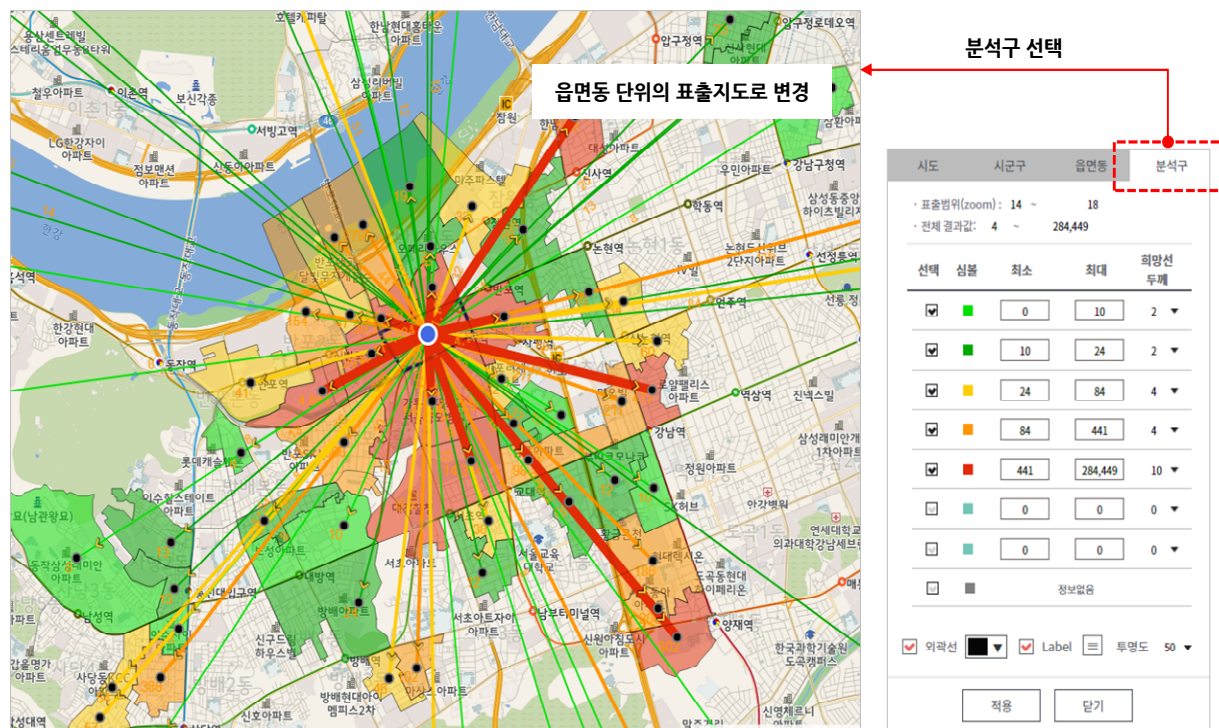
⑤ 범례설정의 연계 기능

- 유입유출 지역분석은 ZOOM 단계에 따라 분석결과가 시도, 시군구, 읍면동, 분석구 단위로 다르게 표출됨
- 본 연구에서는 표출되는 지도와 범례설정의 연계기능을 개발하여 사용자 쉽게 표출지도의 범례를 설정할 수 있도록 개발함
- 현재 ZOOM 단계에서 범례설정 버튼을 선택할 경우 해당 ZOOM 단계의 범례설정이 활성화 되도록 개발함



<그림 4- 18> ZOOM 단계에 따른 범례설정의 변경 예시

- 범례설정에서 시도, 시군구, 읍면동, 분석구 탭을 선택할 경우 지도가 설정된 ZOOM 단계로 자동 변경되도록 개발함

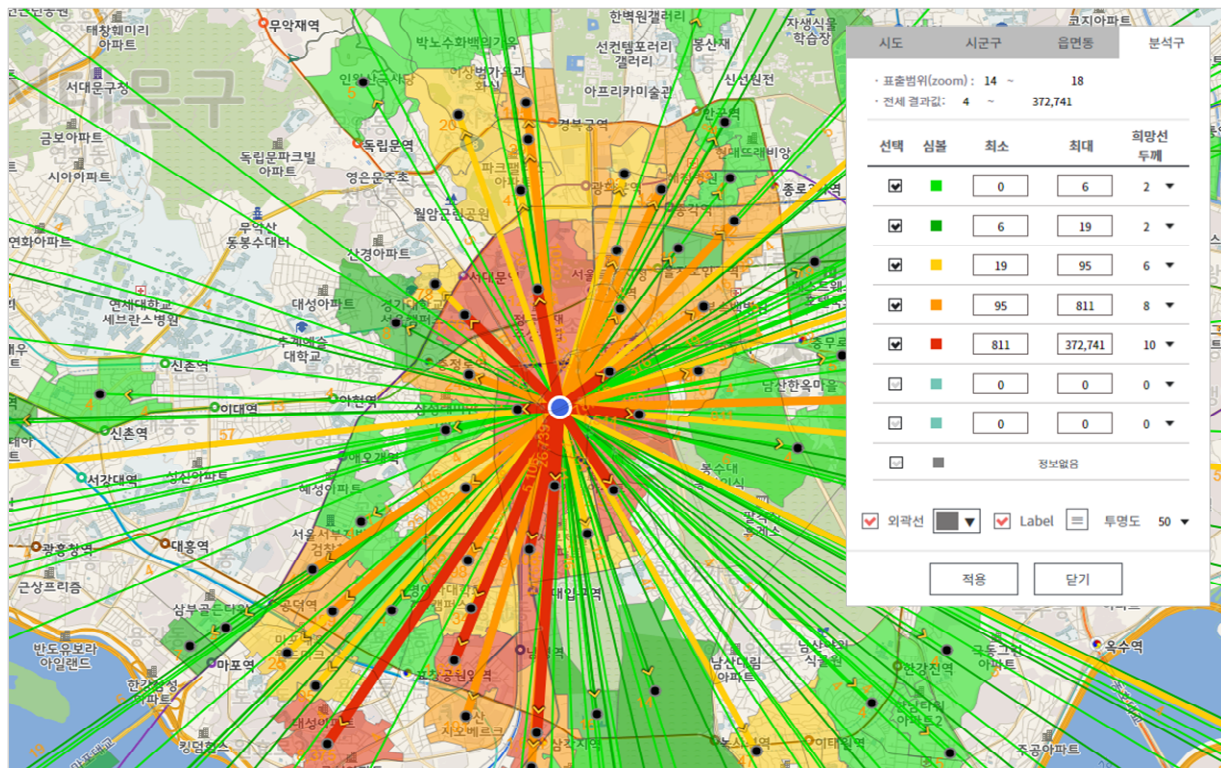


<그림 4- 19> 범례설정의 표출단위 선택에 따른 표출지도 변경 예시

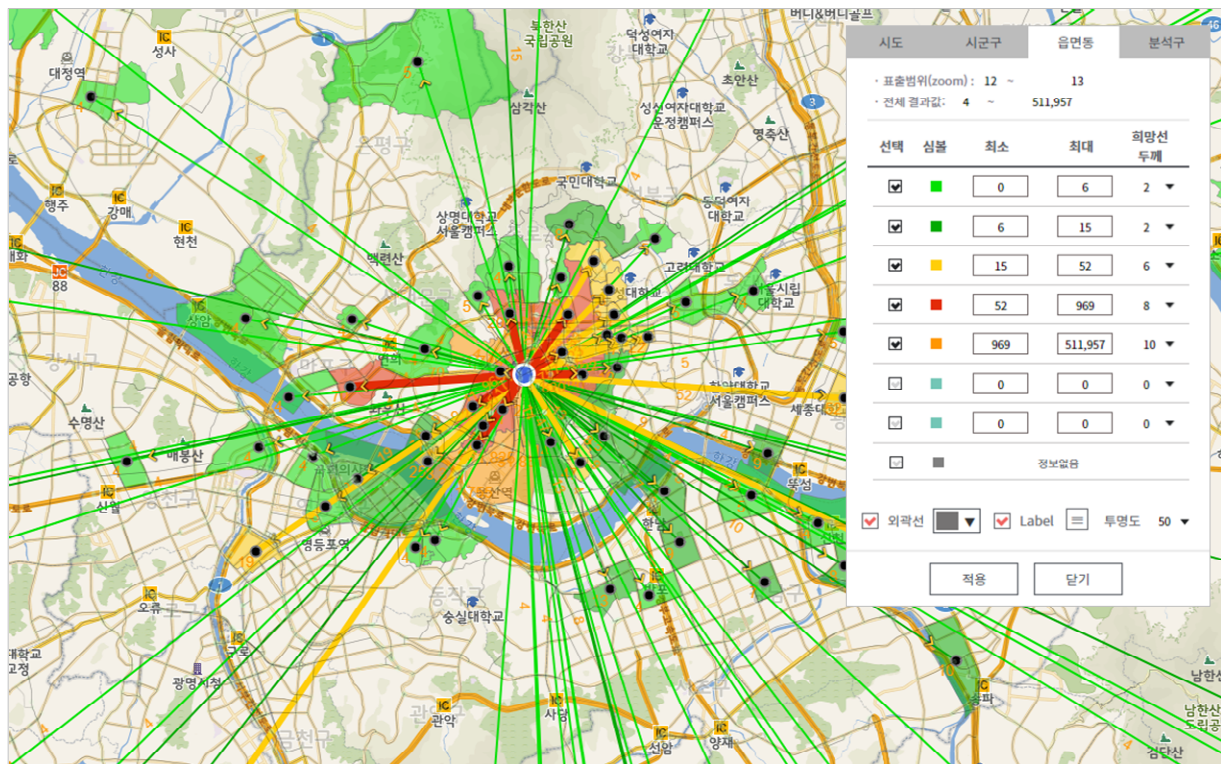
⑥ 유입유출 비교분석의 결과 예시

○ 통행량 검색 조건

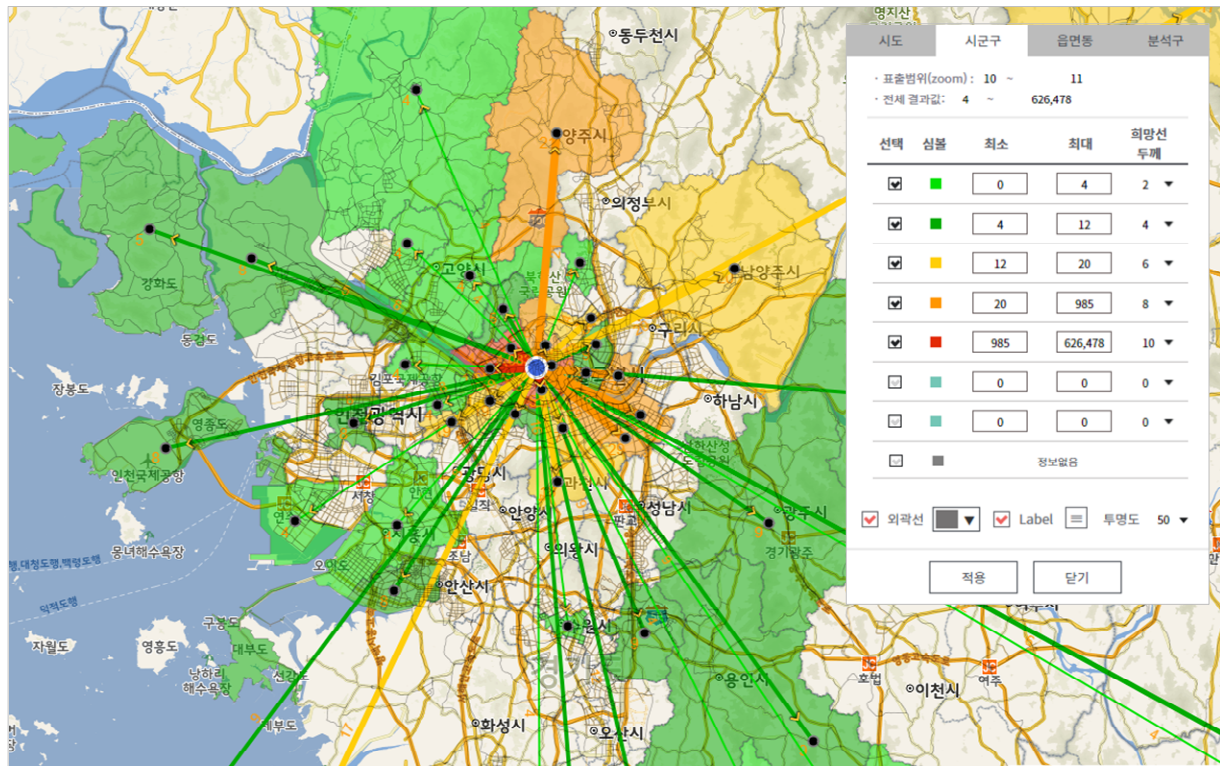
- 통행량 : 유출 통행량
- 선택단위 : 분석구 110105(서울역 인근)
- 기간 : 2016.06.01. ~ 2016.06.30.
- 요일 : 전체 요일
- 시간대 : 출발 및 도착 전체 시간대
- 트립타입 : 출발 및 도착 전체 트립타입(주간상주, 야간상주, 잠재상주)
- 성별 : 남, 여
- 연령대 : 전체 연령대



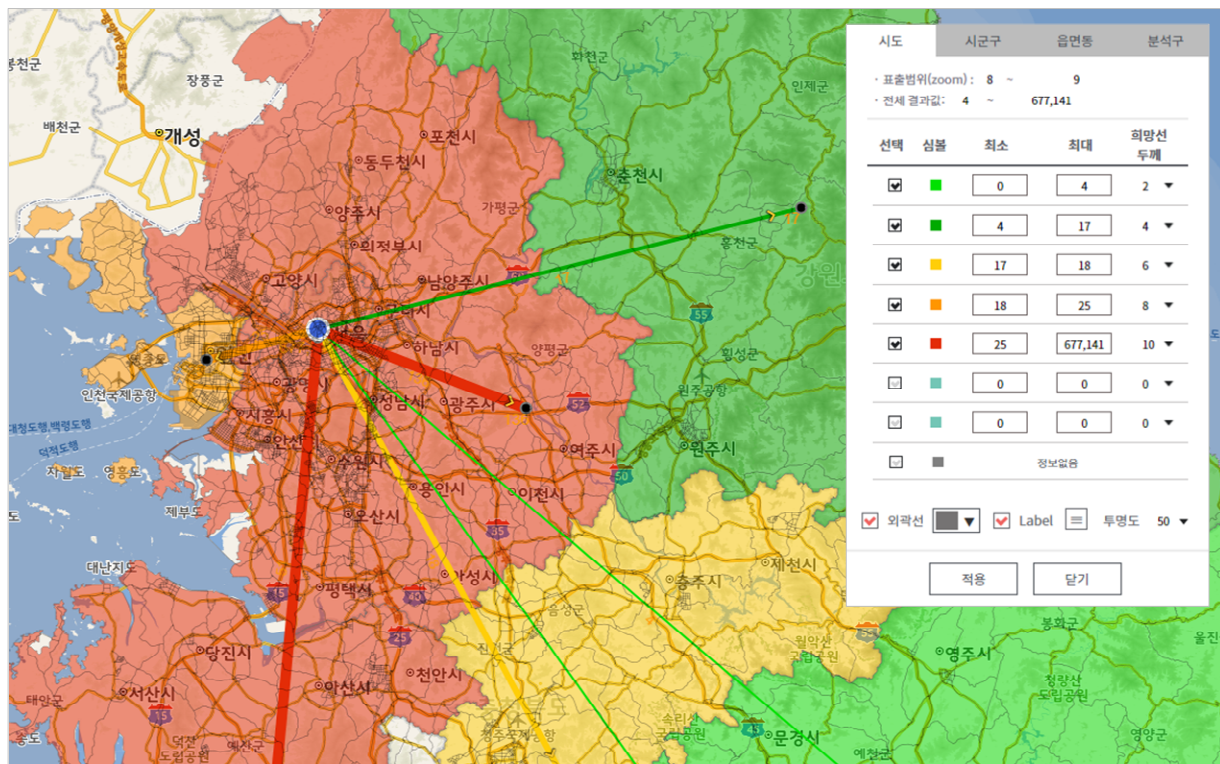
<그림 4-20> 유입유출 지역분석의 분석구단위 유출량 분석의 화면 예시



<그림 4-21> 유입유출 지역분석의 읍면동단위 유출량 분석의 화면 예시



<그림 4-22> 유입유출 지역분석의 시군구단위 유출량 분석의 화면 예시



<그림 4-23> 유입유출 지역분석의 시도단위 유출량 분석의 화면 예시

다. O/D 분석

- 사용자가 분석지역을 선택하여 분석구간의 통행량을 분석할 수 있도록 개발함
- 분석지역은 행정구역 단위로 선택할 수 있도록 개발함
- 분석구 간의 단방향, 양방향 통행량을 분석할 수 있도록 개발함
- 표출지도는 분석구 노드를 직선으로 연결한 희망선으로 표출되도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 검색조건에 해당되는 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함

1) 기능설명

① O/D 분석의 설정 UI

- 분석지역을 시도, 시군구, 읍면동 단위로 선택하여 O/D 분석을 할 수 있도록 개발함
- 분석지역에 포함된 분석구 간의 O/D 통행량을 분석할 수 있도록 개발함
- 단방향 통행량, 양방향 통행량을 사용자 선택해서 분석할 수 있도록 개발함
- 기간, 요일, 출발 및 도착시간대, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 출발 분석구 노드와 도착 분석구 노드를 직선으로 연결한 희망선으로 분석결과를 지도에 표출할 수 있도록 개발함
- 희망선도의 범례를 설정할 수 있도록 개발함

① 분석지역 선택 및 통행량 산출방법 설정

- 분석지역 : 시도, 시군구, 읍면동 선택
- O/D 통행량 : 단방향 통행량 산출
- O/D 합계 통행량 : 양방향 통행량 산출

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

④ 표출지도 설정

- 희망선도 범례 설정
- 희망선도 표출 개수 설정

<그림 4- 24> O/D 분석의 설정 UI

② 통행량 검색조건 설정 및 분석

○ 분석지역 설정 통행량 산출방법 설정

- 시도, 시군구, 읍면동, 분석구 단위로 분석지역을 선택할 수 있도록 개발함
- 지도선택과 ID 검색을 통해서 분석지역을 선택할 수 있도록 개발함
- 단방향 통행량, 양방향 통행량을 사용자가 선택할 수 있도록 개발함
- O/D 통행량 : 단방향 통행량 산출
- O/D 합계 통행량 : 양방향 통행량 산출

○ 상세조건 설정

- 사용자가 기간, 요일, 출발시간대, 도착시간대, 성별, 연령대 조건을 설정하여 O/D 분석을

할 수 있도록 개발함

- 기간 : 일단위 날짜 설정
- 요일 : 월, 화, 수, 목, 금, 토, 일
- 출발 및 도착 시간대 : 00시 ~ 23시
- 출발 및 도착 트립타입 : 주간상주, 야간상주, 잠재상주
- 성별 : 남, 여
- 연령대 : 00대, 10대, 20대, 30대, 40대, 50대, 60대, 70대, 80대, 90대, 100대, 110대

○ 분석

- 모바일 통행형태 통행량 DB에서 출발분석구와 도착분석구가 분석지역에 포함되고 상세설정 정보에 맞는 통신데이터를 검색함
- 검색된 통신데이터를 기반으로 다음과 같이 단방향 O/D 통행량을 분석함
 - 출발분석구-도착분석구를 데이터의 분류기준으로 설정하고 분류기준에 따라 데이터를 분류함
 - 출발분석구-도착분석구별로 이동인구수 합계를 산출함
- 검색된 통신데이터를 기반으로 다음과 같이 양방향 O/D 통행량을 분석함
 - 단방향 O/D 통행량을 분석함
 - 단방향 O/D 통행량 분석결과에서 출발분석구-도착분석구의 통행량을 추출하고 도착분석구-출발분석구로 다시 통행량을 추출하여 추출된 2 데이터의 이동인구수 합계를 산출함

③ 표출지도의 범례설정

- O/D 분석결과는 출발분석구와 도착분석구 간의 통행량으로 산출됨
- 분석구 공간정보에 대한 노드(점) 정보와 분석결과의 출발분석구 및 도착분석구 정보를 연계하여 분석결과를 희망선으로 지도에 표출할 수 있도록 개발함
- 사용자가 희망선도에 대한 범례를 설정할 수 있도록 개발하였으며, 범례설정의 주요내용은 다음과 같음
 - 지도에 표출되는 희망선의 개수 설정
 - 표출되는 희망선 중 강조범위 설정
 - 상위 비율(%)로 설정
 - 강조 색상 설정

- 사용자 편의성을 고려하여 강조구간의 최소값, 최대값 정보 제공
- 두께 범위 설정
 - 표시되는 희망선의 최소값, 최대값을 기준으로 두께의 변화 범위를 설정할 수 있도록 개발함
 - 20단계까지 설정
- 표시되는 희망선의 색상을 설정할 수 있도록 개발함
- 분석결과의 라벨표출 및 표출지도의 투명도를 설정할 수 있도록 개발함

① 희망선의 표출개수 설정

② 부가정보

- 전체결과값
 - 검색범위의 최소값, 최대값
- 표출범위 결과값
 - 표시되는 희망선의 최소값, 최대값

③ 희망선 강조 범위 설정

- 상위 비율(%) 설정
- 강조색상 설정
- 강조 구간의 최소값, 최대값 제공

④ 희망선 두께 설정

- 결과값에 비례하여 설정

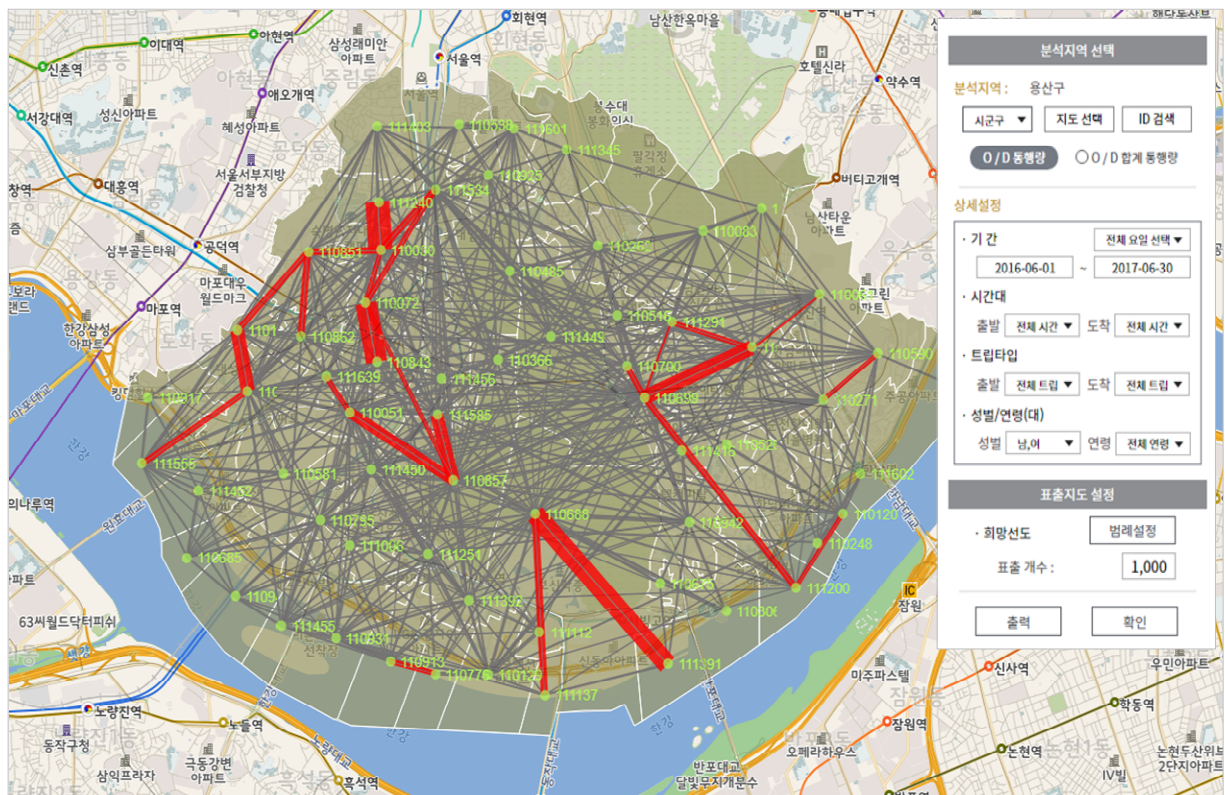
⑤ 희망선의 색상, 라벨표출, 투명도 설정

<그림 4- 25> O/D 분석의 희망선도 범례설정 UI

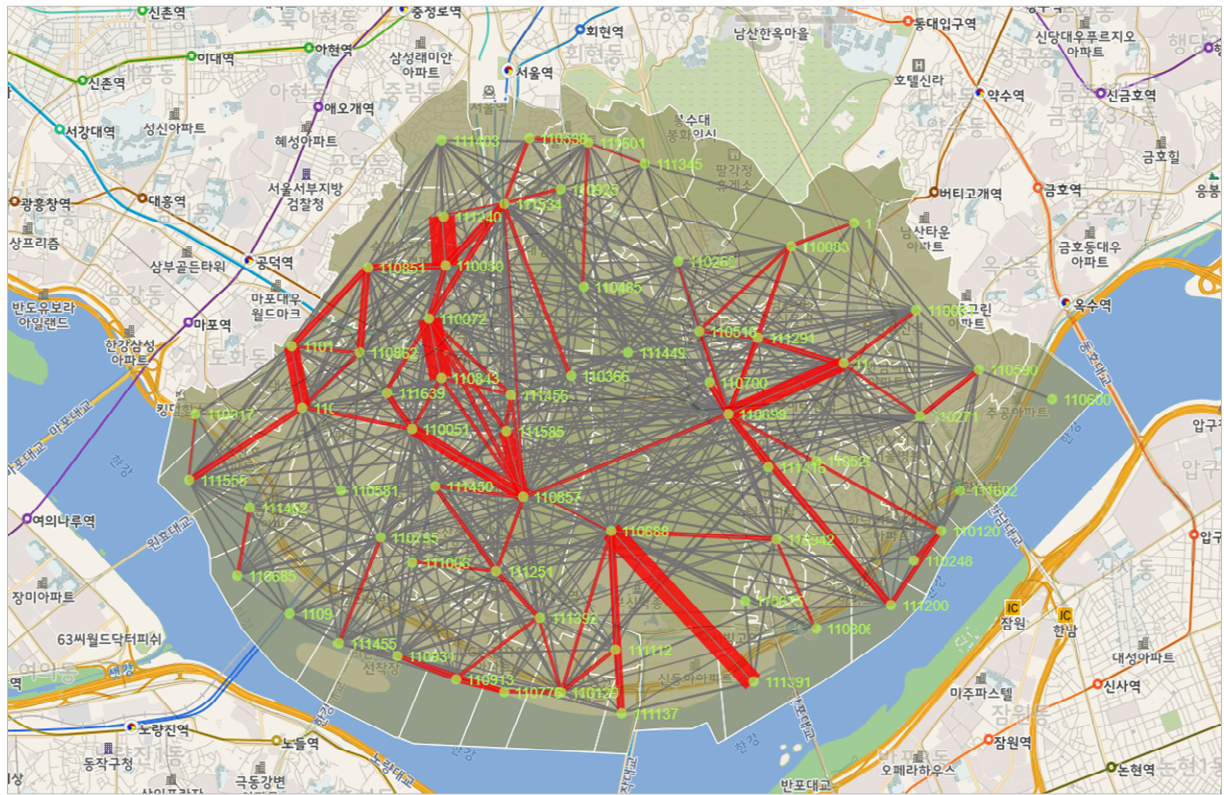
④ O/D 분석의 결과 예시

○ 분석 조건

- 분석지역 : 서울시 용산구
- 통행량 산출방법 : 단방향 O/D 통행량
- 기간 : 2016.06.01. ~ 2016.06.30.
- 요일 : 전체 요일
- 시간대 : 출발 및 도착 전체시간대
- 트립타입 : 출발 및 도착 전체 트립타입(주간상주, 야간상주, 잠재상주)
- 성별 : 남, 여
- 연령대 : 전체 연령대



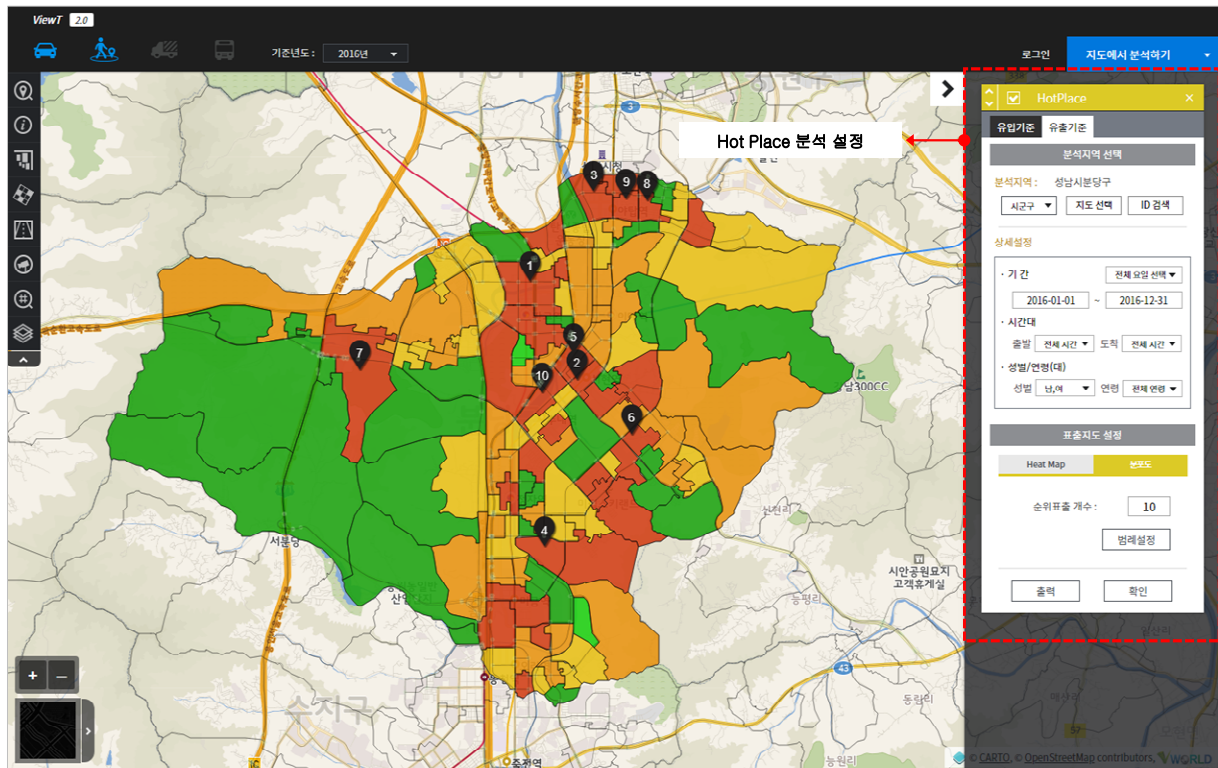
<그림 4-26> O/D 분석결과 화면 예시(상위 10% 강조)



<그림 4- 27> O/D 분석결과 화면 예시(상위 30% 강조)

라. Hot Place 분석

- 사용자가 선택한 특정지역의 범위 내에서 사람이 가장 많이 몰리는 곳, 외부로 가장 많이 나가는 곳을 분석할 수 있도록 개발함
- 분석구 단위로 분석결과를 지도에 표출할 수 있도록 개발함
- 표출지도에 대한 범례를 설정할 수 있도록 개발함
- 검색조건에 해당되는 통신데이터를 파일형태(csv)로 출력할 수 있도록 개발함



<그림 4- 28> Hot Place 분석의 실행 화면

1) 기능설명

① Hot Place 분석의 설정 UI

- 분석지역을 시도, 시군구, 읍면동, 분석구로 선택하여 Hot Place 분석을 할 수 있도록 개발함
- 유입기준 및 유출기준의 Hot Place 분석이 가능하도록 개발함
- 기간, 요일, 출발 및 도착시간대, 성별, 연령대에 대하여 사용자가 검색조건을 설정할 수 있도록 개발함
- 분석결과를 Heat Map, 분포도로 지도에 표출하고 표출되는 지도의 형태를 사용자가 설정할 수 있도록 개발함

■ 유입기준 분석 시 설정 UI



The UI for Inflow Analysis Settings is divided into three main sections. The top section, '분석지역 선택' (Select Analysis Area), includes a dropdown for '분석지역' (Analysis Area) set to '선택안됨' (Not Selected), a circled 1 next to a '시군구' (City/Gun/Gu) dropdown, and buttons for '지도 선택' (Select Map) and 'ID 검색' (Search by ID). The middle section, '상세설정' (Detailed Settings), contains a date range '2016-01-01 ~ 2016-12-31' with a circled 2, a '시간대' (Time Period) section with '출발' (Departure) and '도착' (Arrival) times set to '전체 시간' (All Time), and a '성별/연령(대)' (Gender/Age Group) section with '성별' (Gender) set to '남,여' (Male/Female) and '연령' (Age) set to '전체 연령' (All Ages). The bottom section, '표출지도 설정' (Map Display Settings), has 'Heat Map' selected over '분포도' (Distribution Map), with sliders for '반경' (Radius) at 25 and '흐림' (Blur) at 30, and buttons for '출력' (Output) and '확인' (Confirm).

① 분석지역 선택

- 시도, 시군구, 읍면동

② 통신데이터 검색조건 설정

- 기간, 요일, 시간대
- 트립타입, 성별, 연령대

③ 표출지도 설정

- HeatMap 범례 설정
- 분포도 범례 설정

■ 유출기준 분석 시 설정 UI



The UI for Outflow Analysis Settings is similar to the Inflow version but with some differences. The '분석지역 선택' (Select Analysis Area) section has '분석지역' (Analysis Area) set to '서울특별시' (Seoul Special City). The '상세설정' (Detailed Settings) section is identical to the Inflow version. The '표출지도 설정' (Map Display Settings) section has '분포도' (Distribution Map) selected over 'Heat Map', and includes a '순위표출 개수' (Number of Ranked Outputs) set to 10 and a '범례설정' (Legend Settings) button. The '출력' (Output) and '확인' (Confirm) buttons are at the bottom.

<그림 4- 29> Hot Place 분석의 설정 UI

② 통행량 검색조건 설정 및 분석

○ 분석지역 설정

- 시도, 시군구, 읍면동, 분석구 단위로 분석지역을 선택할 수 있도록 개발함
- 지도 선택과 ID 검색을 통해서 분석지역을 선택할 수 있도록 개발함

○ 상세조건 설정

- 사용자가 기간, 요일, 출발시간대, 도착시간대, 성별, 연령대 조건을 설정하여 HotPlace 분석을 할 수 있도록 개발함

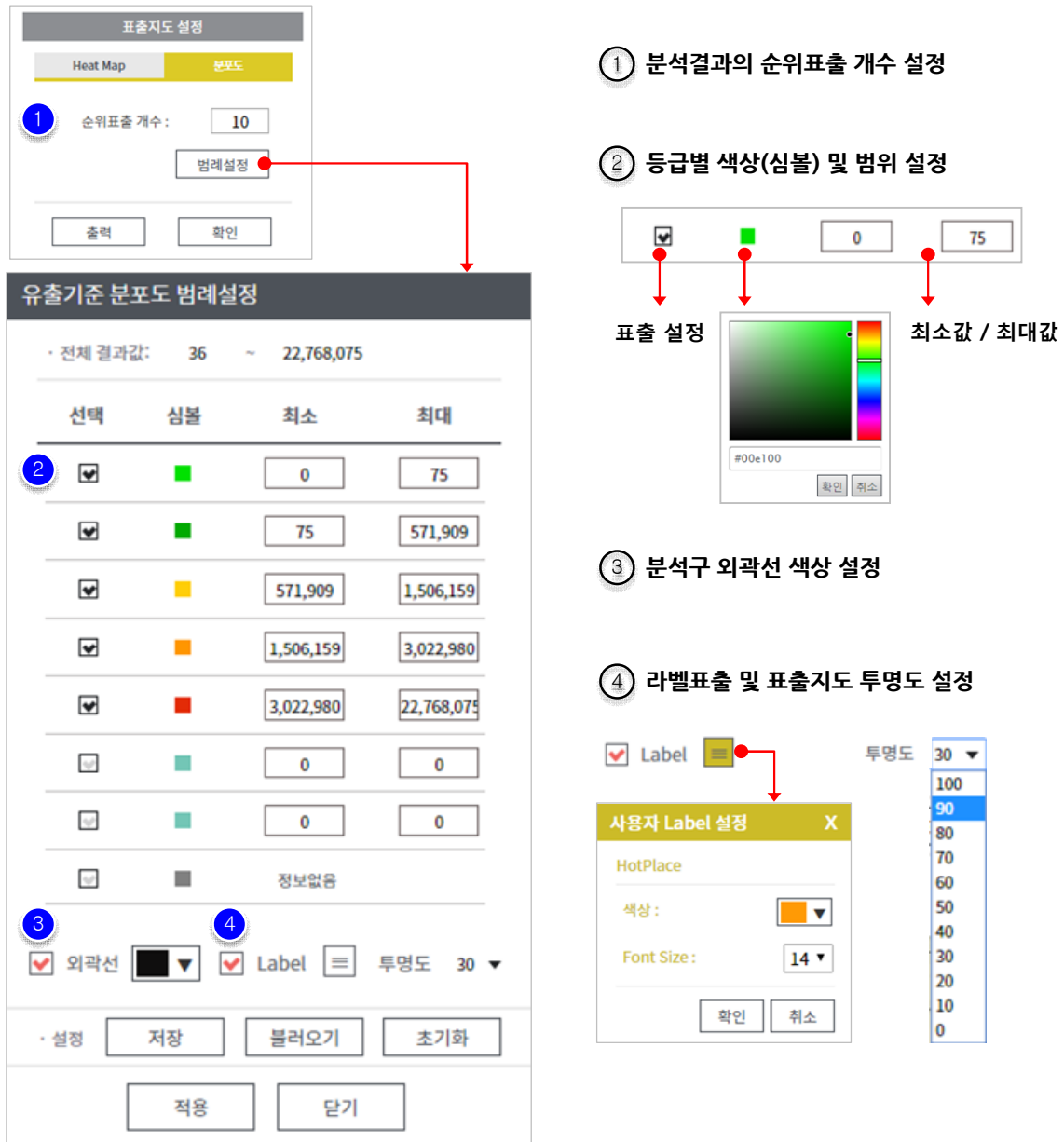
- 기간 : 일단위 날짜 설정
- 요일 : 월, 화, 수, 목, 금, 토, 일
- 출발 및 도착 시간대 : 00시 ~ 23시
- 성별 : 남, 여
- 연령대 : 00대, 10대, 20대, 30대, 40대, 50대, 60대, 70대, 80대, 90대, 100대, 110대

○ 분석

- 모바일 통행형태 통행량 DB에서 출발지 또는 도착지의 행정구역 정보가 분석지역과 동일하고 상세설정 정보에 맞는 통신데이터를 검색함
- 유입기준 분석은 검색된 통신데이터를 기반으로 다음과 같이 분석을 실시함
 - 도착지의 행정구역 정보가 분석지역과 동일한 데이터를 추출함
 - 분석지역 범위에 포함되는 분석구의 정보(ID)를 기준으로 데이터를 재분류함
 - 분석구별 이동인구수 합계를 산출함
- 유출기준 분석은 검색된 통신데이터를 기반으로 다음과 같이 분석을 실시함
 - 출발지의 행정구역 정보가 분석지역과 동일한 데이터를 추출함
 - 분석지역 범위에 포함되는 분석구의 정보(ID)를 기준으로 데이터를 재분류함
 - 분석구별 이동인구수 합계를 산출함

③ 표출지도의 범례설정

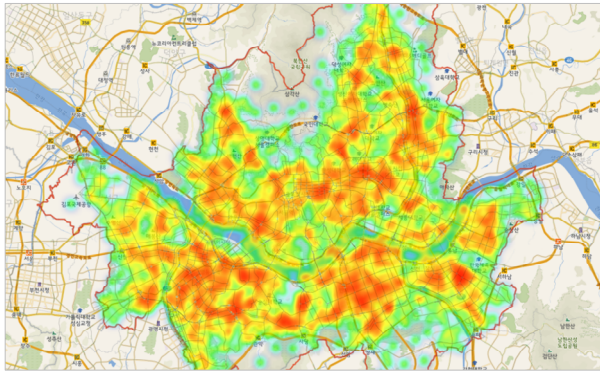
- Hot Place 분석의 결과는 분석구 단위로 산출됨
- 분석구 공간정보와 분석결과를 연계하여 폴리곤으로 분석결과를 표출할 수 있도록 개발함
- 분석구 공간정보에 대한 노드(점) 정보와 분석결과를 연계하여 분석결과를 Heat Map으로 지도에 표출할 수 있도록 개발함
- 본 연구에서는 사용자가 분포도와 HeatMap에 대한 범례를 설정할 수 있도록 개발함
- 분포도 범례설정의 주요내용은 다음과 같음
 - 지도에 분석결과의 순위를 표출하는 개수 설정
 - 등급구분 및 등급별 색상, 범위, 지도표출 설정
 - 분석구의 외곽선 색상 설정
 - 분석결과 값의 라벨 색상 및 크기, 지도표출 설정
 - 분포도의 투명도 설정 등



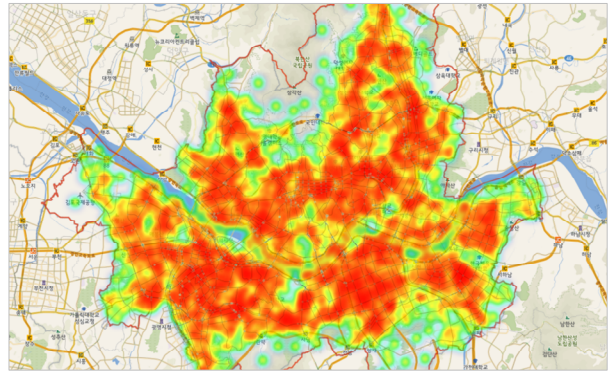
<그림 4- 30> Hot Place 분석의 분포도 범례설정 UI

- Heat Map 범례설정의 주요내용은 다음과 같음
 - － 반경 설정 : 각 포인트 지점에서 다른 포인트를 찾는 반경 설정(범위 : 1~50)
 - － 흐림 설정 : 색상의 흐림 정도를 설정(범위 : 1~50)

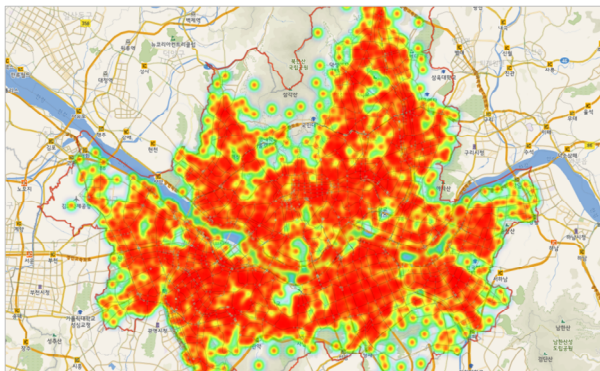
반경 : 11, 흐림 : 37



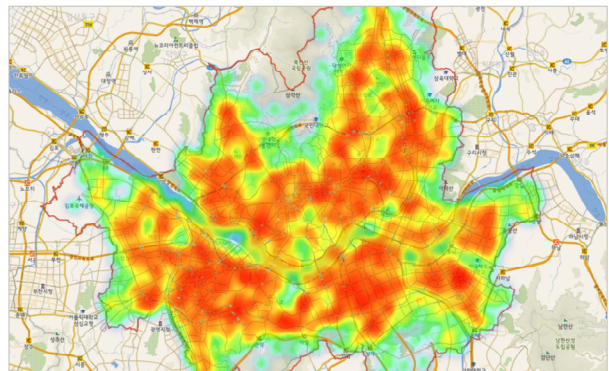
반경 : 13, 흐림 : 37



반경 : 13, 흐림 : 20



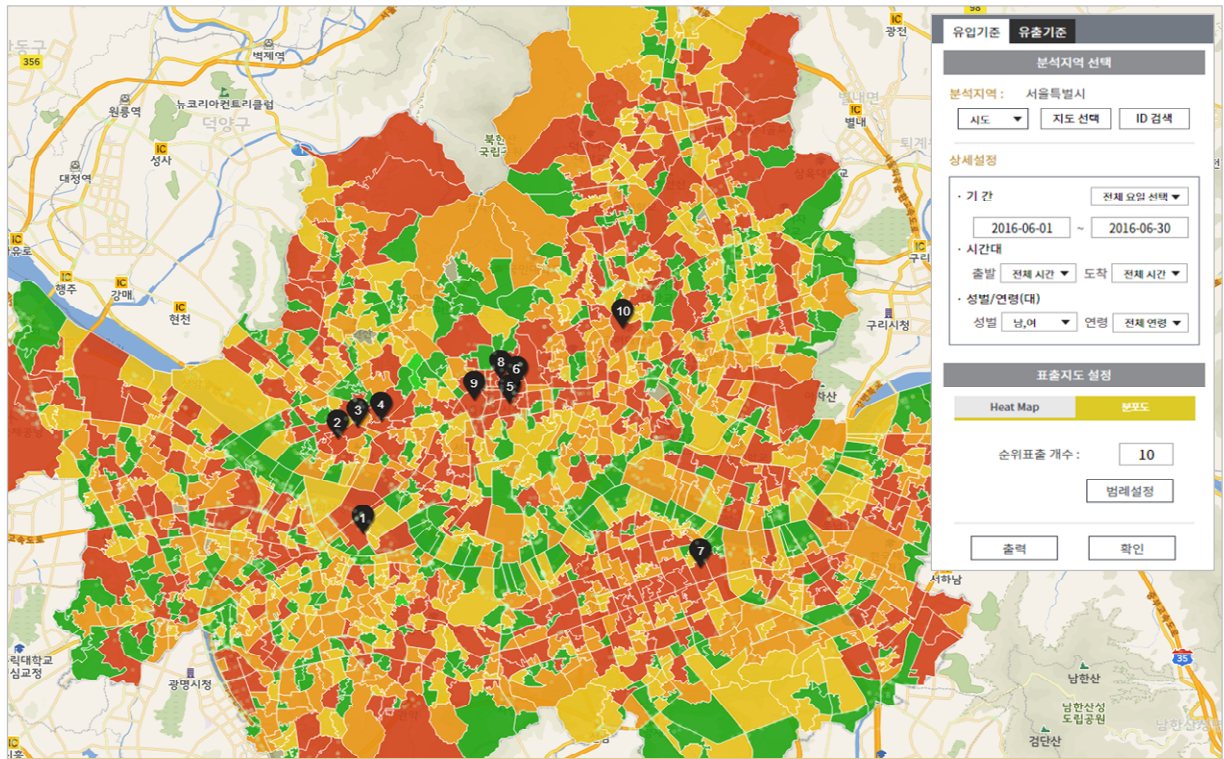
반경 : 13, 흐림 : 40



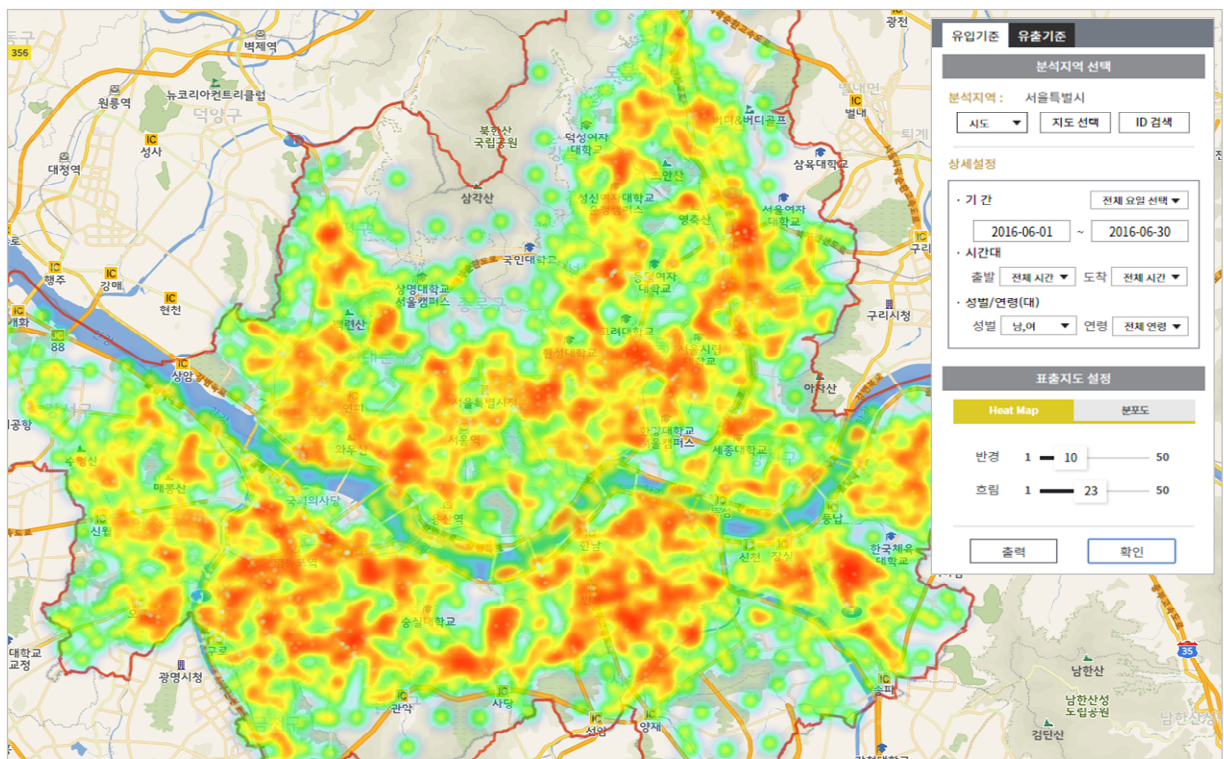
<그림 4- 31> HeatMap의 반경과 흐림

④ Hot Place 분석의 결과 예시

- 통행량 검색 조건
 - － 통행량 : 유입기준 통행량
 - － 분석지역 : 서울시
 - － 기간 : 2016.06.01. ~ 2016.06.30.
 - － 요일 : 전체 요일
 - － 시간대 : 출발 및 도착 전체시간대
 - － 성별 : 남, 여 / 연령대 : 전체 연령대



<그림 4-32> 유입기준 Hot Place 분석결과와 분포도 화면 예시(순위표출개수 : 10)



<그림 4-33> 유입기준 Hot Place 분석결과와 HeatMap 화면 예시

마. 통행량(인구수) 검색

- 본 연구에서는 사용자가 사용목적에 따라 통신데이터를 직접 가공할 수 있도록 통신데이터를 검색할 수 있는 기능을 개발함
- 검색된 결과를 파일형태(csv)로 출력할 수 있도록 개발함
- 사용자가 기간, 요일, 출발시간대, 도착시간대, 성별, 연령대 조건을 설정하여 통신데이터를 검색하고 출력할 수 있도록 개발함
- 요일, 출발시간대, 도착시간대, 성별, 연령대의 설정은 다중선택이 가능하도록 개발함

통행량(인구수) 검색

· 기간 설정
2016-06-01 ~ 2016-06-10

· 요일 설정
월 화 수 목 금 토 일

· 시간대 설정
출발 전체 시간 ~ 도착 전체 시간

· 트립타입 설정
출발 ☒ 주간상주 ☒ 야간상주 ☒ 잠재상주
도착 ☒ 주간상주 ☒ 야간상주 ☒ 잠재상주

· 성별 설정
☒ 남자 ☒ 여자

· 연령대 설정
0 10 20 30 40 50
60 70 80 90 100 110

출력

<그림 4-34> 통행량(인구수) 검색의 설정 UI

- 출력파일은 설정된 기간에 해당하는 데이터를 날짜별로 제공되도록 개발함

- 수집된 통신데이터에 출발분석구와 도착분석구에 대한 행정구역 정보를 추가하여 사용자가 통신데이터를 행정구역 단위로 분석할 수 있도록 개발함
- 출력파일의 구성은 다음과 같음

<표 4-2> 통행량(인구수) 검색의 출력파일 구성

NO	필드명	타입	비고
1	출발폴리곤	STRING	-
2	출발 일자	STRING	-
3	출발시간대	STRING	-
4	출발지 OD트립타입	STRING	-
5	도착폴리곤	STRING	-
6	도착 일자	STRING	-
7	도착시간대	STRING	-
8	도착지 OD트립타입	STRING	-
9	연령대 정보	STRING	-
10	성별 정보	STRING	-
11	이동 인구수	STRING	-
12	기준일	STRING	-
13	출발행정구역	STRING	읍면동 행정구역 코드
14	도착행정구역	STRING	읍면동 행정구역 코드

바. 기타 기능

1) 표출지도의 범례정보 저장 및 불러오기

- 본 연구에서는 기능별로 사용자가 설정한 범례의 정보를 저장하거나 불러오는 기능을 개발하여 사용자의 편의성을 향상시킴

· 전체 결과값: 36 ~ 22,768,075

선택	심볼	최소	최대
<input checked="" type="checkbox"/>	■	0	75
<input checked="" type="checkbox"/>	■	75	571,909
<input checked="" type="checkbox"/>	■	571,909	1,506,159
<input checked="" type="checkbox"/>	■	1,506,159	3,022,980
<input checked="" type="checkbox"/>	■	3,022,980	22,768,075
<input type="checkbox"/>	■	0	0
<input type="checkbox"/>	■	0	0
<input type="checkbox"/>	■	정보없음	

☒ 외곽선 ☒ Label ☐ 투명도 30

· 설정
저장
블러오기
초기화

적용
닫기

[범례정보 파일 예시]

```

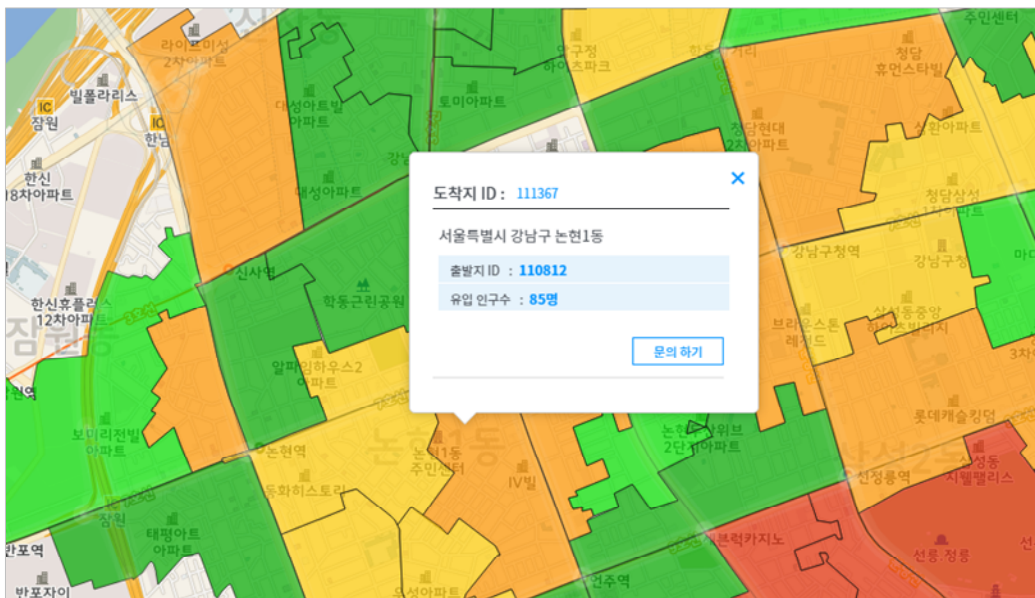
{
  "MinWidth": 1,
  "MaxWidth": 20,
  "Color": "#47FF7E",
  "Opacity": 30
},
{
  "Classification": {
    "Opacity": 30,
    "Grade": {
      "Select": true,
      "Color": "#00E100",
      "Width": "4.0",
      "Min": "0",
      "Max": "600"
    },
    "Select": true,
    "Color": "#00AA00",
    "Width": "4.0",
    "Min": "600",
    "Max": "3000"
  },
  "Select": true,
  "Color": "#FFCE00",
  "Width": "4.0",
  "Min": "3000",
  "Max": "6000"
},
{
  "Select": true,
  "Color": "#FFCE00",
  "Width": "4.0",
  "Min": "6000",
  "Max": "15000"
},
{
  "Select": true,
  "Color": "#E32B07",
  "Width": "4.0",
  "Min": "15000",
  "Max": "200000"
},
{
  "Select": false,
  "Color": "#73CEB6",
  "Width": "4.0",
  "Min": "",
  "Max": ""
}

```

<그림 4- 35> 범례정보 저장 및 블러오기

2) 팝업

- 표출지도에서 마우스 클릭 시 해당지역의 정보와 분석결과를 제공할 수 있도록 개발함



<그림 4- 36> 팝업 기능을 통한 분석결과의 정보제공

3) 분석결과에 대한 통신데이터의 파일출력

- 유입유출 비교분석, 유입유출 지역분석, O/D 분석, Hot Place 분석 시 활용된 통신데이터를 설정된 기간 동안의 날짜별로 출력할 수 있도록 개발함
- 출력파일(csv) 구성은 통행량(인구수) 검색의 출력파일 구성과 동일하게 구성함

분석지역 선택

분석지역: 서초구

시군구 ▼
지도선택
ID 검색

☒ O/D 통행량
☐ O/D 함께 통행량

상세설정

· 기간

2016-06-01

~

2016-06-30

전체 요일 선택 ▼

· 시간대

출발

전체 시간 ▼

도착

전체 시간 ▼

· 트립타입

출발

전체 트립 ▼

도착

전체 트립 ▼

· 성별/연령(대)

성별

남,여 ▼

연령

전체 연령 ▼

표준지도 설정

· 희망선도

범례설정

표출 개수:

1,000

출력

확인

[출력파일의 예시]

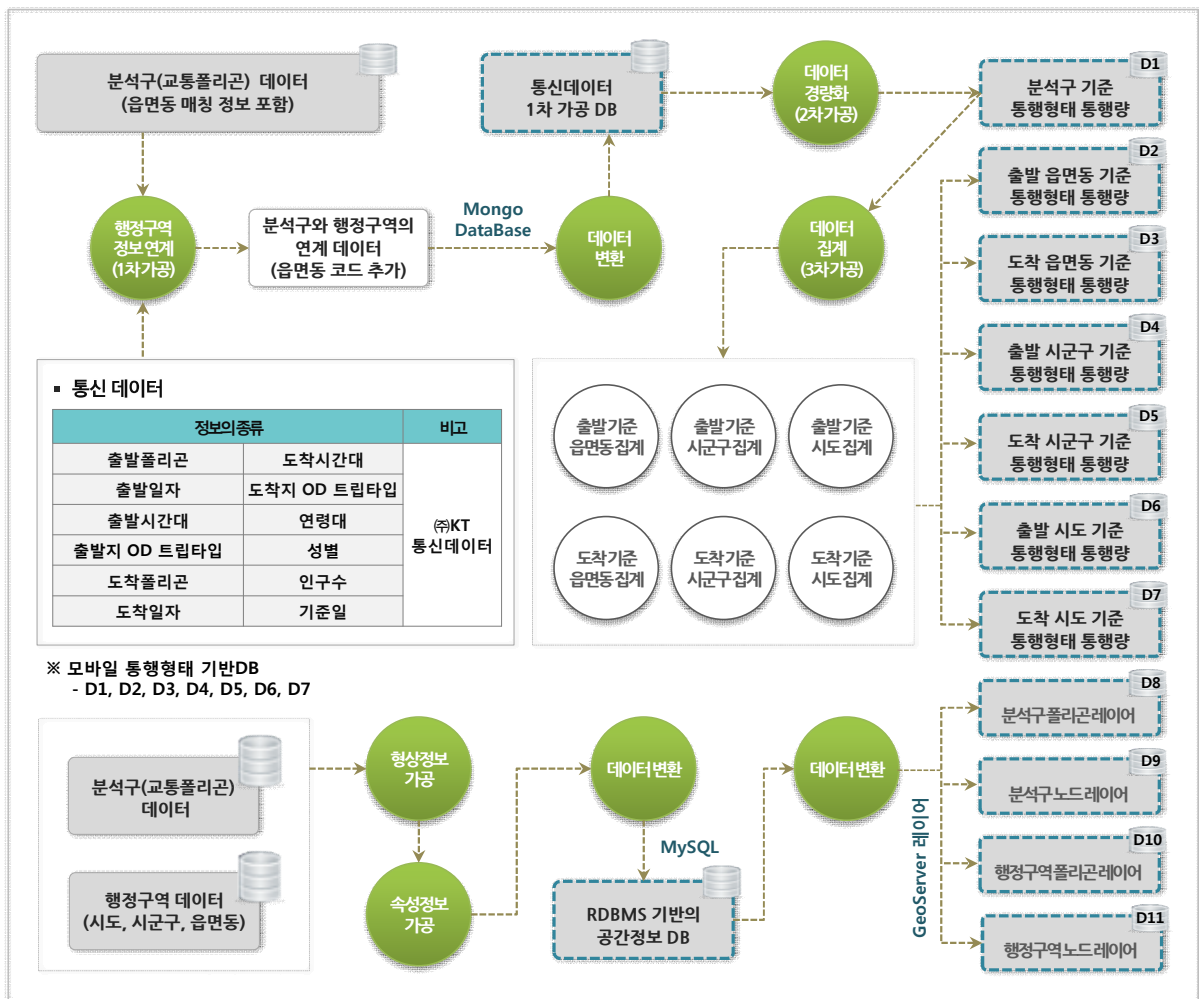
출발영리군	출발영리	출발시각대	출발지OO도합타입	도착영리군	도착영리	도착시각대	도착지	OO도합타입	연령대	성별	이동인구수	기준일	출발영리구역	도착영리구역
330188	20160601	0	D	330188	20160601	0	X	30	M	3	20160601	3304151	3304151	
330192	20160601	0	X	330194	20160601	0	D	50	F	3	20160601	3304153	3304155	
330201	20160601	0	N	330201	20160601	0	D	40	M	3	20160601	3304157	3304157	
330239	20160601	0	X	330239	20160601	2	X	20	M	3	20160601	3304255	3304255	
330250	20160601	0	X	330230	20160601	0	N	10	F	3	20160601	3304259	3304253	
330250	20160601	0	N	330251	20160601	0	X	50	M	3	20160601	3304299	3304299	
330250	20160601	0	D	330250	20160601	10	X	20	F	3	20160601	3304299	3304299	
330261	20160601	0	N	330262	20160601	0	X	50	M	3	20160601	3304311	3304311	
330267	20160601	0	X	330267	20160601	0	N	40	M	3	20160601	3304321	3304321	
330295	20160601	0	D	330294	20160601	0	X	20	F	3	20160601	3304354	3304354	
330295	20160601	0	N	330292	20160601	1	X	20	M	3	20160601	3304354	3304354	
330310	20160601	0	N	330393	20160601	0	D	30	F	3	20160601	3304357	3304354	
330310	20160601	0	X	330310	20160601	0	N	20	M	3	20160601	3304357	3304357	
330329	20160601	0	N	330329	20160601	0	X	50	F	3	20160601	3304411	3304411	
330340	20160601	0	N	330340	20160601	0	D	30	F	3	20160601	3304412	3304412	
330340	20160601	0	X	330340	20160601	0	N	30	F	3	20160601	3304412	3304412	
330367	20160601	0	N	330367	20160601	0	D	60	M	13	20160601	3304452	3304452	
330376	20160601	0	D	330367	20160601	0	X	60	M	3	20160601	3304455	3304452	
330376	20160601	0	X	330376	20160601	0	N	20	M	3	20160601	3304455	3304455	
330388	20160601	1	X	330388	20160601	1	N	50	F	3	20160601	3304151	3304151	
330393	20160601	1	N	330227	20160601	7	X	40	M	3	20160601	3304154	3304251	
330394	20160601	1	N	330199	20160601	1	X	50	M	3	20160601	3304155	3304157	
330234	20160601	1	X	330183	20160601	3	D	40	M	3	20160601	3304312	3304135	
330230	20160601	1	N	330230	20160601	1	D	10	F	3	20160601	3304253	3304253	
330230	20160601	1	X	330298	20160601	1	N	30	F	3	20160601	3304253	3304253	
330230	20160601	1	N	330310	20160601	8	X	20	M	3	20160601	3304253	3304257	
330230	20160601	1	N	330250	20160601	14	X	20	M	3	20160601	3304253	3304259	
330231	20160601	1	X	330227	20160601	1	N	30	M	3	20160601	3304253	3304251	
330246	20160601	1	N	330246	20160601	1	D	40	M	3	20160601	3304258	3304258	
330246	20160601	1	X	330246	20160601	1	N	40	M	3	20160601	3304258	3304258	
330250	20160601	1	D	330227	20160601	1	X	20	M	3	20160601	3304251	3304251	
330250	20160601	1	X	330230	20160601	1	D	20	M	4	20160601	3304259	3304253	
330255	20160601	1	X	330269	20160601	1	N	30	M	3	20160601	3304311	3304311	
330284	20160601	1	N	330285	20160601	1	X	50	M	3	20160601	3304352	3304352	
330286	20160601	1	X	330298	20160601	3	N	20	F	3	20160601	3304352	3304355	
330294	20160601	1	X	330290	20160601	1	D	30	F	3	20160601	3304354	3304353	
330294	20160601	1	X	330294	20160601	2	D	60	F	3	20160601	3304354	3304354	
330298	20160601	1	X	330230	20160601	2	N	20	F	3	20160601	3304355	3304253	
330301	20160601	1	N	330200	20160601	7	D	30	F	3	20160601	3304356	3304157	
330311	20160601	1	X	330311	20160601	1	N	20	M	3	20160601	3304357	3304357	
330311	20160601	1	N	330285	20160601	2	X	0	F	3	20160601	3304357	3304352	
330327	20160601	1	N	330370	20160601	8	X	50	M	3	20160601	3304411	3304413	
330329	20160601	1	N	330329	20160601	2	X	20	F	3	20160601	3304411	3304411	
330340	20160601	1	D	330341	20160601	1	X	40	F	3	20160601	3304412	3304412	
330342	20160601	1	X	330342	20160601	2	N	40	F	3	20160601	3304412	3304412	
330343	20160601	1	X	330337	20160601	1	X	20	F	3	20160601	3304412	3304412	
330343	20160601	1	X	330337	20160601	2	N	40	M	3	20160601	3304412	3304412	
330344	20160601	1	X	330344	20160601	1	N	20	M	3	20160601	3304412	3304412	
330364	20160601	1	X	330376	20160601	1	N	60	F	3	20160601	3304451	3304455	
330368	20160601	1	X	330188	20160601	2	N	10	M	3	20160601	3304452	3304151	
330370	20160601	1	N	330373	20160601	1	X	30	M	3	20160601	3304453	3304455	
330373	20160601	1	N	330281	20160601	1	X	20	M	3	20160601	3304455	3304332	
330373	20160601	1	N	330370	20160601	1	X	30	M	3	20160601	3304455	3304453	
311880	20160601	2	N	330221	20160601	7	X	60	F	3	20160601	3118803	3304232	
330159	20160601	2	D	330160	20160601	2	N	60	M	3	20160601	3304132	3304132	
330177	20160601	2	D	330176	20160601	2	X	40	F	3	20160601	3304134	3304134	
330193	20160601	2	X	330200	20160601	2	D	10	M	3	20160601	3304154	3304157	
330193	20160601	2	N	330245	20160601	2	X	40	M	3	20160601	3304154	3304257	
330196	20160601	2	D	330195	20160601	2	N	10	M	3	20160601	3304156	3304155	
330201	20160601	2	X	330194	20160601	4	D	10	M	3	20160601	3304157	3304155	
330203	20160601	2	X	330203	20160601	2	N	40	M	3	20160601	3304158	3304158	
330227	20160601	2	X	330302	20160601	3	X	20	F	3	20160601	3304251	3304356	
330232	20160601	2	N	330251	20160601	2	X	50	M	3	20160601	3304254	3304259	

<그림 4- 37> 분석결과에 대한 통신데이터의 파일출력 예시

제2절 기반 DB 구축

1. 데이터 구축 프로세스

- 데이터 구축은 통신데이터 기반의 모바일 통행형태 기반 DB 구축과 GeoServer에 탑재되는 통신데이터 분석 맵 레이어 구축으로 크게 구분 함
- 본 연구의 모바일 통행형태 기반 DB는 발주처에서 제공하는 통신 데이터, 분석구 데이터, 행정구역 데이터를 기반으로 구축 함
- 통신데이터 분석맵 레이어는 분석구 데이터, 행정구역 데이터를 기반으로 GeoServer 레이어로 구축 함



<그림 4- 38> 데이터 구축 프로세스

2. 통신데이터 분석 맵 레이어 구축

가. 분석구 레이어 구축

1) 자료수집

- 지역 : 전국
- 형상 : 폴리곤
- 개체 수 : 15,937
- 좌표계 : WGS84 경위도 좌표계
- 용량 : 약 108 MB
- 데이터형식 : shp

<표 4- 3> 수집된 분석구 데이터의 속성정보

NO	필드명	설명	데이터 타입	비고
1	FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	KEY	교통폴리곤 ID	DOUBLE	-
4	행정동 ID	읍면동 코드	STRING	-

2) 구축절차



- 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- 39> 분석구 레이어의 구축 절차

- GeoServer, Openlayers 기반의 웹 GIS 형태로 서비스를 개발하였음

- GeoServer : GIS 서버 개발에 적용
- Openlayer : 클라이언트 서비스 개발에 적용
- 수집된 분석구 데이터를 활용하여 GeoServer 형식의 분석구 폴리곤 레이어와 분석구 노드 레이어를 구축함
- 분석구 폴리곤 레이어는 유입유출 지역분석 및 비교분석, OD 분석 등에서 분석결과에 대한 분포도 생성의 기반 공간정보로 활용됨
- 분석구 노드 레이어는 회망선 지도표출, Hot Place 분석결과와 지표표출 등에서 기반 공간 정보로 활용됨
- 분석구 노드는 분석구 폴리곤의 대표 포인트(노드)를 의미함

① 분석구 폴리곤 레이어



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

<표 4- 4> MySQL 기반의 분석구 폴리곤 공간정보의 속성

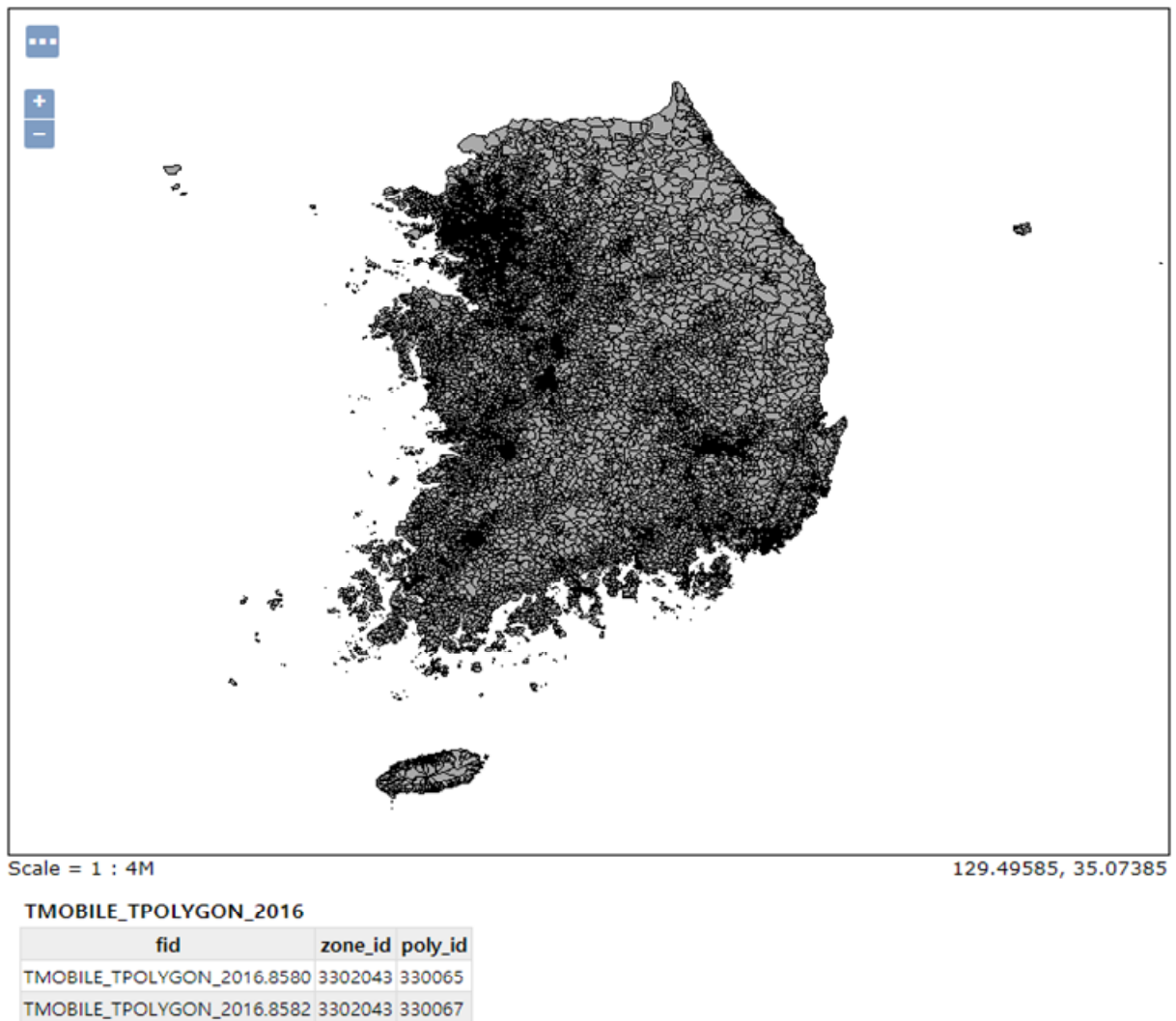
NO	필드명	설명	데이터 타입	비고
1	FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	zone_id	읍면동 코드	INT	-
4	poly_id	분석구 ID	INT	

- MySQL 분석구 폴리곤 공간정보를 활용하여 GeoServer 기반의 분석구 폴리곤 레이어를 구축함
- 전국은 15,937개의 분석구로 구성되어 있으며 분석결과를 분석구 단위의 지도로 표출할 경우 사용자의 컴퓨터 환경에서 지도 표출이 원활하여야 함
- GeoServer의 VectorTile, WMS 서비스는 지도의 뷰 레벨에 따라 최적화된 형상의 이미지를 제공 함(지도 일반화)

<표 4- 5> VectorTile, WMS의 지도 일반화

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

- VectorTile, WMS 서비스가 가능한 형태로 교통폴리곤 레이어를 구축함



<그림 4- 40> GeoServer 기반의 분석구 폴리곤 레이어

<표 4- 6> GeoServer 분석구 레이어의 속성정보

NO	필드명	설명	데이터 타입	비고
1	fid	ID	INT	
2	zone_id	읍면동 코드	INT	
3	poly_id	교통폴리곤 ID	INT	
4	geom	형상정보	MULTIGOLYGON	

① 분석구 노드 레이어

- 희망선 지도표출, 유입유출 분석, Hot Place 분석 등 분석결과의 라벨표출 시 사용되는 분석구 폴리곤의 대표 포인트(노드)를 구축 함
- 상용 GIS툴(ARCGIS)을 기반으로 분석구 폴리곤 데이터를 활용하여 노드 레이어를 구축함
- 분석구 폴리곤 레이어에 좌표 정보를 추가하여 속성정보를 구축함
- MySQL 기반의 공간정보를 구축하고 이를 활용하여 GeoServer 기반의 분석구 폴리곤 노드 레이어를 구축함
- MySQL 기반의 분석구 폴리곤 노드 공간정보는 OD 분석의 희망선 표출 시 좌표정보를 활용할 수 있도록 DB를 구축함
- VectorTile, WMS 서비스가 가능한 형태로 분석구 폴리곤 노드 레이어를 구축함

<표 4- 7> GeoServer 분석구 폴리곤 노드 레이어의 속성정보

NO	필드명	설명	데이터 타입	비고
1	fid	ID	INT	
2	zone_id	읍면동 코드	INT	
3	poly_id	교통폴리곤 ID	INT	
4	lon	X 좌표	DOUBLE	
5	lat	Y 좌표	DOUBLE	
6	geom	형상정보	POINT	



TMOBILE_TPOLYGON_P_2016

fid	zone_id	poly_id	lon	lat
TMOBILE_TPOLYGON_P_2016.11256	3532040	350802	127.426519	36.006297
TMOBILE_TPOLYGON_P_2016.11257	3532040	350803	127.437504	35.971132

<그림 4- 41> GeoServer 기반의 분석구 노드 레이어

나. 행정구역 레이어 구축

1) 자료수집

- 지역 : 전국
- 형상 : 폴리곤
- 개체수 : 3,771
- 좌표계 : Korean 1985 Katch 좌표계
- 용량 : 약 84.3 MB

- 데이터형식 : shp
- 속성정보

<표 4- 8> 수집된 행정구역 데이터의 속성정보

NO	필드명	설명	데이터 타입	비고
1	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	

2) 구축절차

- 행정구역 레이어는 분석구 레이어 구축과 동일한 절차에 따라 구축하였음
- 수집된 행정구역 데이터를 활용하여 GeoServer 형식의 폴리곤 레이어와 노드 레이어를 구축함
- 시도, 시군구, 읍면동에 대한 각각의 폴리곤 레이어와 노드 레이어를 구축함
- 행정구역 폴리곤 레이어는 분석구 폴리곤 레이어와 같이 유입유출 지역분석 및 비교분석, OD 분석 등에서 분석결과에 대한 분포도 생성의 기반 공간정보로 활용됨
- 행정구역 노드 레이어 또한, 희망선 지도표출, Hot Place 분석결과와 지표표출 등에서 기반 공간정보로 활용됨

① 행정구역 폴리곤 레이어

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

<표 4- 9> MySQL 기반의 행정구역 폴리곤 공간정보의 속성

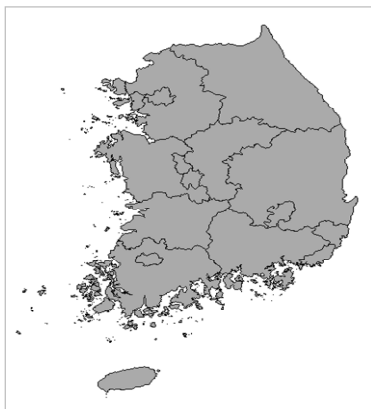
NO	필드명	설명	데이터 타입	비고
1	FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	zone_id	해당 행정구역 코드	INT	-
4	name	해당 행정구역 명칭	STRING	
5	full_name	상위행정구역 명칭 + 해당 행정구역 명칭	STRING	

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

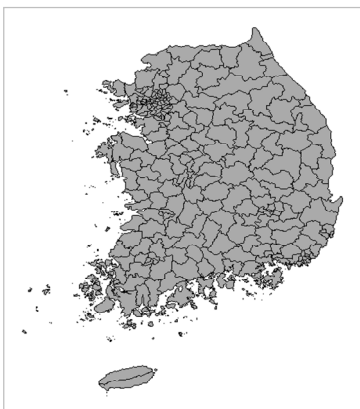
<표 4- 10> GeoServer 행정구역 폴리곤 레이어의 속성정보

NO	필드명	설명	데이터 타입	비고
1	fid	ID	INT	
2	zone_id	해당 행정구역 코드	INT	-
3	name	해당 행정구역 명칭	STRING	
4	full_name	상위행정구역 명칭 + 해당 행정구역 명칭	STRING	
5	geom	형상정보	MULTIGOLYGON	

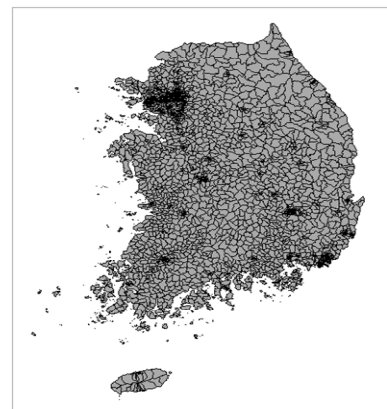
■ 시도 행정구역 레이어



■ 시군구 행정구역 레이어



■ 읍면동 행정구역 레이어



<그림 4- 42> GeoServer 기반의 행정구역 폴리곤 레이어

② 행정구역 노드 레이어

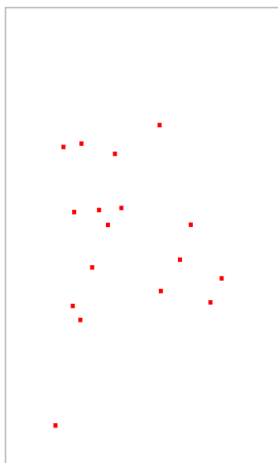
- 분석구 폴리곤 데이터에 좌표 정보를 추가하여 MySQL 기반의 공간정보를 구축함
(시도, 시군구, 읍면동 동일함)

<표 4- 11> MySQL 기반의 행정구역 노드 공간정보의 속성

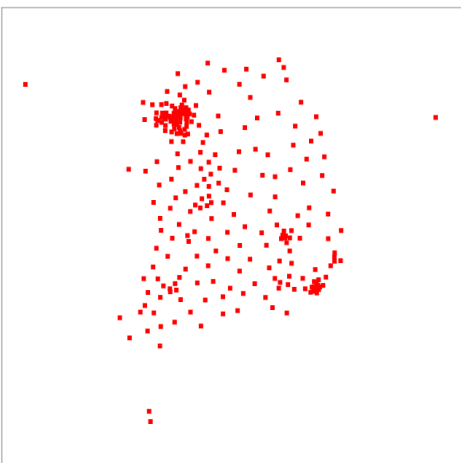
NO	필드명	설명	데이터 타입	비고
1	FID	ID	Object ID	-
2	Shape	형상정보	Geometry	-
3	zone_id	해당 행정구역 코드	INT	-
4	name	해당 행정구역 명칭	STRING	
5	full_name	상위행정구역 명칭 + 해당 행정구역 명칭	STRING	
6	lon	X 좌표	DOUBLE	
7	lat	Y 좌표	DOUBLE	

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

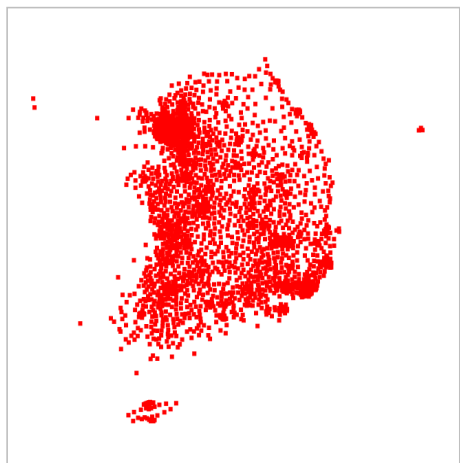
■ 시도 행정구역 노드 레이어



■ 시군구 행정구역 노드 레이어



■ 읍면동 행정구역 노드 레이어



<그림 4- 43> GeoServer 기반의 행정구역 노드 레이어

3. 모바일 통행형태 기반 DB 구축

가. 자료 수집

- 지역 : 전국
- 기간 : 2016.03.07. ~ 2016.12.31.
- 용량 : 약 555 GB
- 일별로 구분된 통신 빅데이터를 수집하였으며 데이터 형태는 csv로 되어 있음
- 수집된 통신 빅데이터는 csv 형태이며 출발폴리곤, 출발일자, 출발시간대, 출발지 OD 트립 타입, 도착폴리곤, 도착일자, 도착시간대, 도착지 OD 트립타입, 연령대, 성별, 인구수, 기준일의 정보를 수록하고 있으며 컬럼명 및 데이터 타입은 다음과 같음

<표 4- 12> 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

290068,20160701,00,N,330259,20160701,10,D,30,F,3.0,20160701
 330189,20160701,00,N,20160701,11,X,50,F,3.0,20160701,3304151
 330194,20160701,00,X,330194,20160701,01,N,30,M,3.0,20160701

<그림 4- 44> 수집된 통신 빅데이터 형태

- 연령별, 성별, 출발 및 도착 시간대, 출발지 및 도착지 OD트립타입 정보는 코드로 기록되어 있으면 코드정보는 다음과 같음

<표 4- 13> 통신 빅데이터의 연령별 코드

구분	코드	내용	비고
연령별	0	0세 - 9세	-
	1	10세 - 19세	-
	2	20세 - 29세	-
	3	30세 - 39세	-
	4	40세 - 49세	-
	5	50세 - 59세	-
	6	60세 - 69세	-
	7	70세 - 79세	-
	8	80세 - 89세	-
	9	90세 - 99세	-
	10	100세 - 109세	-
	11	110세 - 119세	-

<표 4- 14> 통신 빅데이터의 성별 코드

구분	코드	내용	비고
연령별	M	남성	-
	F	여성	-

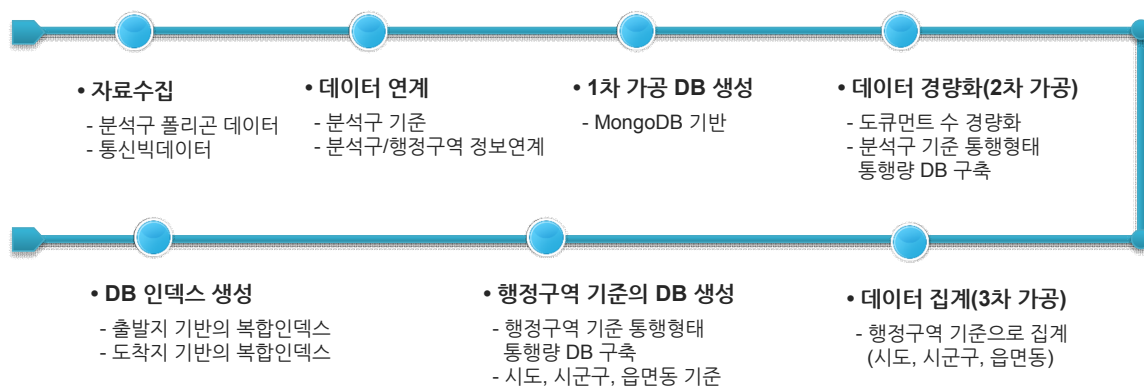
<표 4- 15> 통신 빅데이터의 시간대 코드

구분	코드	내용	비고
시간대	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- 16> 통신 빅데이터의 O/D 트립타입

구분	코드	내용	비고
O/D 트립타입	D	주간상주	-
	N	야간상주	-
	X	잠재상주	-

나. DB 구축절차



<그림 4- 45> 모바일 통행형태 기반 DB의 구축 절차

- 수집되는 통신 빅데이터는 많은 양의 정보가 주기적으로 발생하는 빅데이터 임
- 본 연구에서는 데이터의 검색 및 처리속도, 빅데이터 관리의 원활함을 위해서 수집된 데이터를 가공하여 MongoDB 기반의 빅데이터 DB를 구축함
- MongoDB에서 도큐먼트 수(테이블의 row 수)는 데이터의 검색 및 처리속도와 직접적인 관계를 갖고 있음
- 본 연구의 경우 수집된 데이터의 건수(도큐먼트 수)가 매우 커서 가공 없이 DB를 구축할 경우 데이터 검색 및 처리 속도의 저하로 정상적인 서비스를 제공하는데 한계가 있음
 - 1년 데이터 건수 : 약 130억건(용량 : 약 730GB) 예상
 - (1일 데이터 건수 : 약 3,600만건(용량 : 약 2G) 예상)
- 본 연구에서는 수집된 데이터의 도큐먼트 수를 경량화 시킨 분석구 기준의 모바일 통행형태 기반 DB를 구축하여 정상적인 서비스를 제공할 수 있도록 함
- 이동인구수의 분석을 행정구역 단위로 하는 경우 행정구역에 포함되는 분석구의 이동인구수를 검색하고 처리하는데 많은 시간이 소요됨

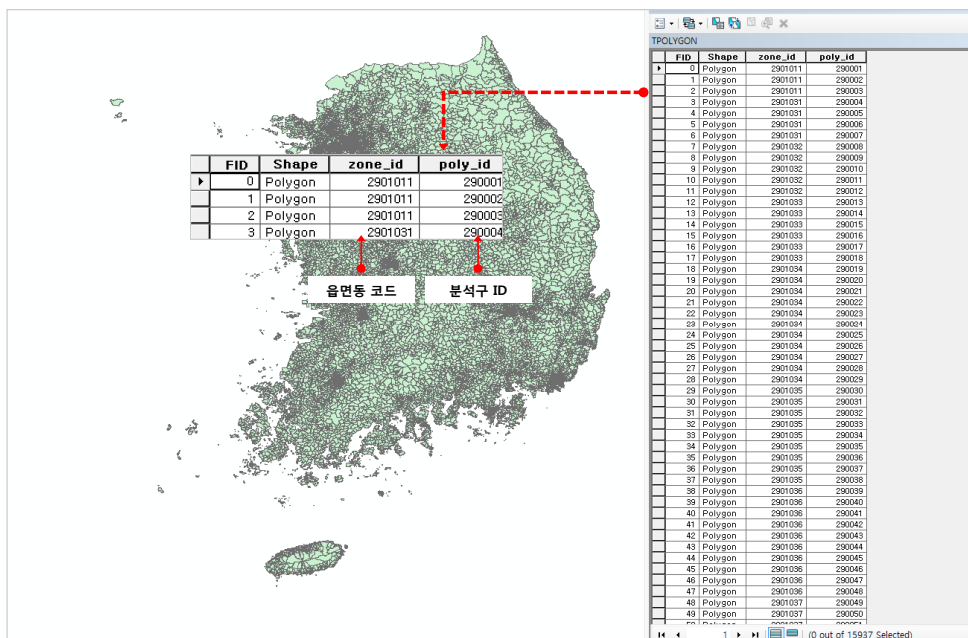
- 출발지 또는 도착지가 행정구역(시도, 시군구, 읍면동)인 경우 서비스 속도향상을 위하여 행정구역 기준의 모바일 통행형태 기반 DB를 구축함
- 입력된 DB에 대한 인덱스를 생성하여 검색속도를 향상시킴
 - 출발지 및 도착지 정보를 활용한 복합 인덱스를 생성함(시도, 시군구, 읍면동, 분석구 ID)

1) 분석구 기준의 모바일 통행형태 통행량 DB 구축

- 분석구 기준의 모바일 통행형태 통행량 DB는 서비스의 속도를 결정하는 DB이며 행정구역 기준의 모바일 통행형태 통행량 DB를 구축하는데 활용되는 기반 DB 임
- 본 연구에서는 수집된 통신 빅데이터를 1차 가공하여 MongoDB에 입력하고 이를 기반으로 경량화된 분석구 기준의 모바일 통행형태 통행량 DB를 구축함

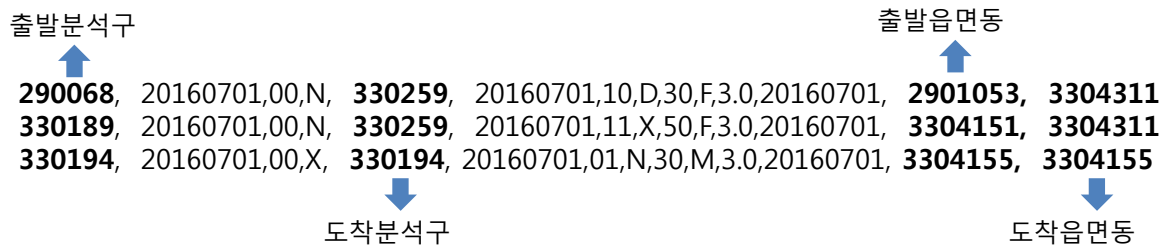
① 통신 빅데이터 1차 가공

- 수집된 통신 빅데이터에 행정구역 코드를 추가하여 MongoDB 입력파일 형태로 데이터를 1차 가공함
- 출발지, 도착지 분석구에 대한 읍면동 행정구역 코드를 추가함
- 분석구와 행정구역 매칭정보는 수집된 분석구 공간정보의 속성정보를 활용함



<그림 4- 46> 분석구와 읍면동 행정구역 코드의 매칭정보

- 가공된 MongoDB 입력파일 형태는 다음과 같음



<그림 4- 47> 1차 가공데이터의 MongoDB 입력데이터 예시

② 데이터 경량화

- 본 연구에서는 검색속도 및 데이터 처리속도 향상을 위하여 MongoDB에 입력된 1차 가공데이터를 기반으로 데이터를 경량화 하였음
 - object 타입으로 ID를 정의하고 ID를 기준으로 1차 가공 DB를 재분류함
 - ID : 출발분석구, 도착분석구, 성별, 연령대, 출발트립타입, 도착트립타입의 복합정보
 - ID별로 출발 및 도착시간대 인구수를 날짜별로 구분하여 DB를 구축함
- 행정구역 기준의 모바일 통행형태 통행량 DB 구축 시 활용되는 행정구역정보 등을 추가하여 분석구 기준의 모바일 통행형태 통행량 DB를 구축함
- 컬렉션(MongoDB의 테이블명)의 구성은 다음과 같음

<표 4- 17> 분석구 기준의 모바일 통행형태 통행량 컬렉션의 구성

컬렉션명	TMOBILE_2016			
컬렉션 설명	분석구 기준의 모바일 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고

식별자	id	Object	"_id" : { "A" : 출발분석구코드, "E" : 도착분석구코드, "J" : 성별, "I" : 연령대, "D" : 출발트립타입, "H" : 도착트립타입 }	A, E, D, H, J, I : integer 트립타입 코드 · D : 1, N : 2, X : 3 성별 코드 · F : 1, M : 2
출발도착시간대 인구수	CG	Object 배열	CG" : { { "C" : 출발시간대, "G" : 도착시간대, "K" : { { "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 } } } }	C, G, L, V : integer K : object 배열 V : _id, C, G, D에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	"A1" : 출발읍면동 코드	
도착읍면동	E1	INT	"E1" : 도착읍면동 코드	
출발시군구	A2	INT	"A2" : 출발시군구 코드	
도착시군구	E2	INT	"E2" : 도착시군구 코드	
출발시도	A3	INT	"A3" : 출발시도 코드	
도착시도	E3	INT	"E3" : 도착시도 코드	
날짜별 인구수 합계	KS	Object 배열	"KS" : { { "LS" : 날짜, "VS" : 인구수 } }	LS, VS : integer VS : 날짜별 CG.K의 V 합계

- 경량화의 결과는 다음과 같음

<표 4- 18> 통신 빅데이터의 문서 수의 경량화 결과

구분	수집데이터	경량화 (분석구 기준의 통행형태 통행량)	감소율
문서 수	9,160,803,237 건	480,983,056 건	약 95% 감소

2) 행정구역 기준의 모바일 통행형태 통행량 DB 구축

- 분석구 기준의 모바일 통행형태 통행량 DB를 기반으로 행정구역 ↔ 행정구역, 행정구역 ↔ 분석구의 통행량을 분석할 경우 행정구역에 포함되는 분석구의 이동인구수를 검색하고 처리하는데 많은 시간이 소요됨
- 행정구역 ↔ 행정구역, 행정구역 ↔ 분석구의 통행량 분석을 위해서는 분석구 기준의 모바일 통행형태 통행량 DB의 문서 수를 경량화 시킨 별도의 통행형태 통행량 DB가 필요함
- 본 연구에서는 다음과 같이 행정구역 기준의 모바일 통행형태 통행량 DB를 구축함
 - 시도, 시군구, 읍면동에서 출발하는 통행량을 분석하기 위한 DB를 구축함
 - 시도, 시군구, 읍면동으로 도착하는 통행량을 분석하기 위한 DB를 구축함
- 출발지 및 도착지의 행정구역 단위로 인구수를 각각 집계하여 MongoDB 기반의 행정구역 기준 통행형태 통행량 DB를 구축함
- 행정구역 기준의 통행형태 통행량 컬렉션의 구성

<표 4- 19> 읍면동 출발 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_UMD_ALL_2016			
컬렉션 설명	읍면동 출발 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고

식별자	id	Object	"_id" : { "A1" : 출발읍면동코드, "E" : 도착분석구코드, "J" : 성별, "I" : 연령대, "D" : 출발트립타입, "H" : 도착트립타입 }	A1, E, D, H, J, I : integer 트립타입 코드 · D : 1, N : 2, X : 3 성별 코드 · F : 1, M : 2
출발도착시간대 인구수	CG	Object 배열	CG" : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 }]]	C, G, L, V : integer K : object 배열 V : _id, C, G, D에 해당하는 원시데이터의 인구수
도착읍면동	E1	INT	"E1" : 도착읍면동 코드	
출발시군구	A2	INT	"A2" : 출발시군구 코드	
도착시군구	E2	INT	"E2" : 도착시군구 코드	
출발시도	A3	INT	"A3" : 출발시도 코드	
도착시도	E3	INT	"E3" : 도착시도 코드	
날짜별 인구수 합계	KS	Object 배열	"KS" : [{ "LS" : 날짜, "VS" : 인구수 }]	LS, VS : integer VS : 날짜별 CG.K의 V 합계

<표 4- 20> 읍면동 도착 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_ALL_UMD_2016			
컬렉션 설명	읍면동 도착 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	id	Object	"_id" : { "A" : 출발분석구코드, "E1" : 도착읍면동코드, "J" : 성별, "I" : 연령대, "D" : 출발트립타입, "H" : 도착트립타입 }	A, E1, D, H, J, I : integer 트립타입 코드 · D : 1, N : 2, X : 3 성별 코드 · F : 1, M : 2
출발도착시간대 인구수	CG	Object 배열	"CG" : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 }]]	C, G, L, V : integer K : object 배열 V : _id, C, G, D에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	"A1" : 출발읍면동 코드	
출발시군구	A2	INT	"A2" : 출발시군구 코드	
도착시군구	E2	INT	"E2" : 도착시군구 코드	

출발시도	A3	INT	“A3” : 출발시도 코드	
도착시도	E3	INT	“E3” : 도착시도 코드	
날짜별 인구수 합계	KS	Object 배열	“KS” : { { “LS” : 날짜, “VS” : 인구수 } } }	LS, VS : integer VS : 날짜별 CG.K의 V 합계

<표 4- 21> 시군구 출발 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_SGG_ALL_2016			
컬렉션 설명	시군구 출발 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	id	Object	“_id” : { “A2” : 출발시군구코드, “E” : 도착분석구코드, “J” : 성별, “I” : 연령대, “D” : 출발트립타입, “H” : 도착트립타입 }	<ul style="list-style-type: none"> ▪ A2, E, D, H, J, I : integer ▪ 트립타입 코드 <ul style="list-style-type: none"> · D : 1, N : 2, X : 3 ▪ 성별 코드 <ul style="list-style-type: none"> · F : 1, M : 2

출발도착시간대 인구수	CG	Object 배열	CG" : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 } }]	<ul style="list-style-type: none"> ▪ C, G, L, V : integer ▪ K : object 배열 ▪ V : _id, C, G, D에 해당하는 원시데이터의 인구수
도착읍면동	E1	INT	"E1" : 도착읍면동 코드	
도착시군구	E2	INT	"E2" : 도착시군구 코드	
출발시도	A3	INT	"A3" : 출발시도 코드	
도착시도	E3	INT	"E3" : 도착시도 코드	
날짜별 인구수 합계	KS	Object 배열	"KS" : [{ "LS" : 날짜, "VS" : 인구수 }]	<ul style="list-style-type: none"> ▪ LS, VS : integer ▪ VS : 날짜별 CG.K의 V 합계

<표 4- 22> 시군구 도착 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_ALL_SGG_2016			
컬렉션 설명	시군구 도착 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	id	Object	<pre> "_id" : { "A" : 출발분석구코드, "E2" : 도착시군구코드, "J" : 성별, "I" : 연령대, "D" : 출발트립타입, "H" : 도착트립타입 } </pre>	A, E2, D, H, J, I : integer 트립타입 코드 · D : 1, N : 2, X : 3 성별 코드 · F : 1, M : 2
출발도착시간대 인구수	CG	Object 배열	<pre> "CG" : [{ "C" : 출발시간대, "G" : 도착시간대, "K" : [{ "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 }] }] </pre>	C, G, L, V : integer K : object 배열 V : _id, C, G, D에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	"A1" : 출발읍면동 코드	
출발시군구	A2	INT	"A2" : 출발시군구 코드	
출발시도	A3	INT	"A3" : 출발시도 코드	
도착시도	E3	INT	"E3" : 도착시도 코드	

날짜별 인구수 합계	KS	Object 배열	<pre> "KS" : { { "LS" : 날짜, "VS" : 인구수 } } </pre>	LS, VS : integer VS : 날짜별 CG.K의 V 합계
---------------	----	--------------	---	---

<표 4- 23> 시도 출발 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_SIDO_ALL_2016			
컬렉션 설명	시도 출발 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	id	Object	<pre> "_id" : { "A3" : 출발시도코드, "E" : 도착분석구코드, "J" : 성별, "I" : 연령대, "D" : 출발트립타입, "H" : 도착트립타입 } </pre>	A3, E, D, H, J, I : integer 트립타입 코드 · D : 1, N : 2, X : 3 성별 코드 · F : 1, M : 2
출발도착시간대 인구수	CG	Object 배열	<pre> "CG" : { { "C" : 출발시간대, "G" : 도착시간대, "K" : { { "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 } } } } </pre>	C, G, L, V : integer K : object 배열 V : _id, C, G, D에 해당하는 원시데이터의 인구수
도착읍면동	E1	INT	"E1" : 도착읍면동 코드	
도착시군구	E2	INT	"E2" : 도착시군구 코드	

도착시도	E3	INT	"E3" : 도착시도 코드	
날짜별 인구수 합계	KS	Object 배열	<pre> "KS" : { { "LS" : 날짜, "VS" : 인구수 } } </pre>	LS, VS : integer VS : 날짜별 CG.K의 V 합계

<표 4- 24> 시도 도착 기준의 통행형태 통행량의 컬렉션의 구성

컬렉션명	TMOBILE_ALL_SIDO_2016			
컬렉션 설명	시도 도착 기준의 통행형태 통행량			
KEY Name	KEY	VALUE TYPE	데이터 구성	비고
식별자	id	Object	<pre> "_id" : { "A" : 출발분석구코드, "E3" : 도착시도코드, "J" : 성별, "I" : 연령대, "D" : 출발트립타입, "H" : 도착트립타입 } </pre>	A, E3, D, H, J, I : integer 트립타입 코드 · D : 1, N : 2, X : 3 성별 코드 · F : 1, M : 2
출발도착시간대 인구수	CG	Object 배열	<pre> "CG" : { { "C" : 출발시간대, "G" : 도착시간대, "K" : { { "L" : 날짜, "V" : 인구수 }, { "L" : 날짜, "V" : 인구수 } } } } </pre>	C, G, L, V : integer K : object 배열 V : _id, C, G, D에 해당하는 원시데이터의 인구수
출발읍면동	A1	INT	"A1" : 출발읍면동 코드	

출발시군구	A2	INT	“A2” : 출발시군구 코드	
출발시도	A3	INT	“A3” : 출발시도 코드	
날짜별 인구수 합계	KS	Object 배열	“KS” : { { “LS” : 날짜, “VS” : 인구수 } }	LS, VS : integer VS : 날짜별 CG.K의 V 합계

- 행정구역 기준의 통행형태 통행량의 경량화 결과

<표 4- 25> 읍면동 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

분석구 기준의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
480, 983, 056 건	278, 093, 705 건	약 43% 감소

<표 4- 26> 읍면동 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

분석구 기준의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
480, 983, 056 건	277, 898, 381 건	약 43% 감소

<표 4- 27> 시군구 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

분석구 기준의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
480, 983, 056 건	77, 754, 473 건	약 83% 감소

<표 4- 28> 시군구 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

분석구 기준의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
480, 983, 056 건	76, 036, 667 건	약 84% 감소

<표 4- 29> 시도 출발 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

분석구 기준의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
480,983,056 건	13,529,772 건	약 97% 감소

<표 4- 30> 시도 도착 기준의 통행형태 통행량의 도큐먼트수의 경량화 결과

분석구 기준의 통행형태 통행량	읍면동 출발 기준의 통행형태 통행량	감소율
480,983,056 건	13,114,843 건	약 97% 감소

4. Sharding 기반의 DB 구축

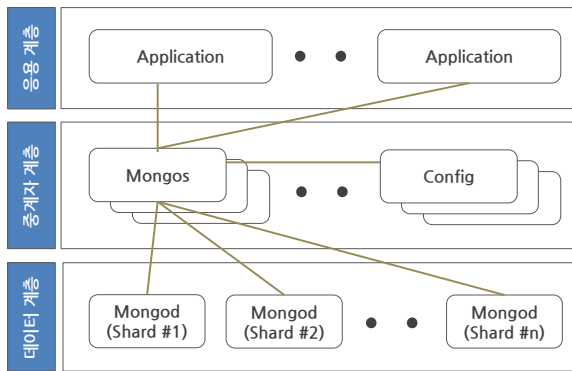
가. Sharding

- 대규모 어플리케이션에서 발생할 수 있는 문제중 하나는 데이터베이스에 저장해야 할 데이터의 수가 방대한 것임
- Sharding이란 데이터셋이 단일 데이터베이스에 저장하기에 매우 큰 경우, 데이터셋을 다수의 데이터베이스에 분산 저장하는 것을 말함
- 본 사업에서는 MongoDB의 Sharding 기법을 적용하여 모바일 통행형태 기반 DB를 구성함

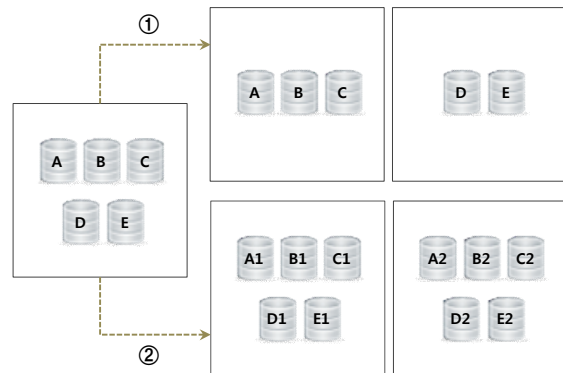
나. Sharding의 구성

- Config 서버, Mongos 서버, Mongod(Shard) 서버로 Sharding을 구성함
 - Config 서버 : 중계자 계층, Shard Meta 정보 저장
 - Mongos 서버 : 중계자 역할 수행, 응용계층의 질의와 데이터 계층의 Shard 서버의 응답을 중계하는 역할을 수행함
 - Mongod 서버 : 데이터 서버
- 데이터셋의 분산저장 방식은 테이블 단위로 분리하는 방법(①)과 테이블 자체(②)를 분할하는 방법이 있음
- 본 연구에서는 테이블 자체(②)를 분할하는 방법을 적용함
 - 데이터가 한쪽으로 몰리는 단점이 있는 분리방법(①)을 보완한 방법
 - Shard Key의 해쉬(hash)를 통해서 데이터를 균등하게 배분
 - 다음과 같이 복합정보로 구성된 _id를 Shard Key로 설정함
 - 출발지 정보(시도, 시군구, 읍면동, 분석구)
 - 도착지 정보(시도, 시군구, 읍면동, 분석구)
 - 성별 정보
 - 연령대 정보
 - 출발트립타입 정보
 - 도착트립타입 정보

■ MongoDB의 Sharding 서버 구성



■ MongoDB의 분산저장 방식



<그림 4- 48> MongoDB Sharding의 구성

다. 모바일 통행형태 기반 DB의 분사저장 결과

■ 분석구 기준의 통행형태 통행량 DB의 분산저장 결과

```
data : 442.67GiB docs : 430983056 chunks : 868
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 : 6035
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- 49> 모바일 통행형태 기반 DB의 Sharding 구성 결과

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

제3절 시스템 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- 31> 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 추가되었음

② 실행/중지

- Apache HTTP Server는 윈도우즈 서비스로 등록되어 시스템 부팅시 자동으로 실행되나 프로그램의 오류 등으로 중지되어 부득이하게 수동으로 실행 시에는 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 서비스 명을 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- 32> 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/중지

- Apache Tomcat는 윈도우즈 서비스로 등록되어 시스템 부팅 시 자동으로 실행되나 프로그램의 오류 등으로 중지되어 부득이하게 수동으로 실행 시에는 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 서비스 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 어플리케이션은 다음과 같음
 - ViewT 1.0 서비스 : [Tomcat 설치경로]\webapps\cong
 - GoeServer : [Tomcat 설치경로]\webapps\geoserver

나. 백업 및 복구

- 시스템 운영 중 발생할 수 있는 장애에 대응하기 위해서 평시 백업과 복구에 대한 절차가 필요함

1) 백업 대상 및 방법

- 데이터베이스 : DB 데이터 파일, 로그파일
 - 대상 파일을 별도의 드라이브로 복사하여 백업하거나 DB관리 툴(ex: Microsoft SQL Server management Studio 이용)의 백업 기능을 이용하여 백업을 실행 함

<표 4- 33> 데이터베이스 백업 대상 파일

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

- 응용프로그램 파일 : 프로그램 실행파일, 환경설정 파일

- 시스템의 응용프로그램 백업 대상은 Apache HTTP Server와 Apache Tomcat Server를 대상으로 함

<표 4- 34> 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\shard 아래의 파일 중 데이터를 제외한 파일

2) 백업 및 복구 절차

<표 4- 35> 백업 및 복구 절차

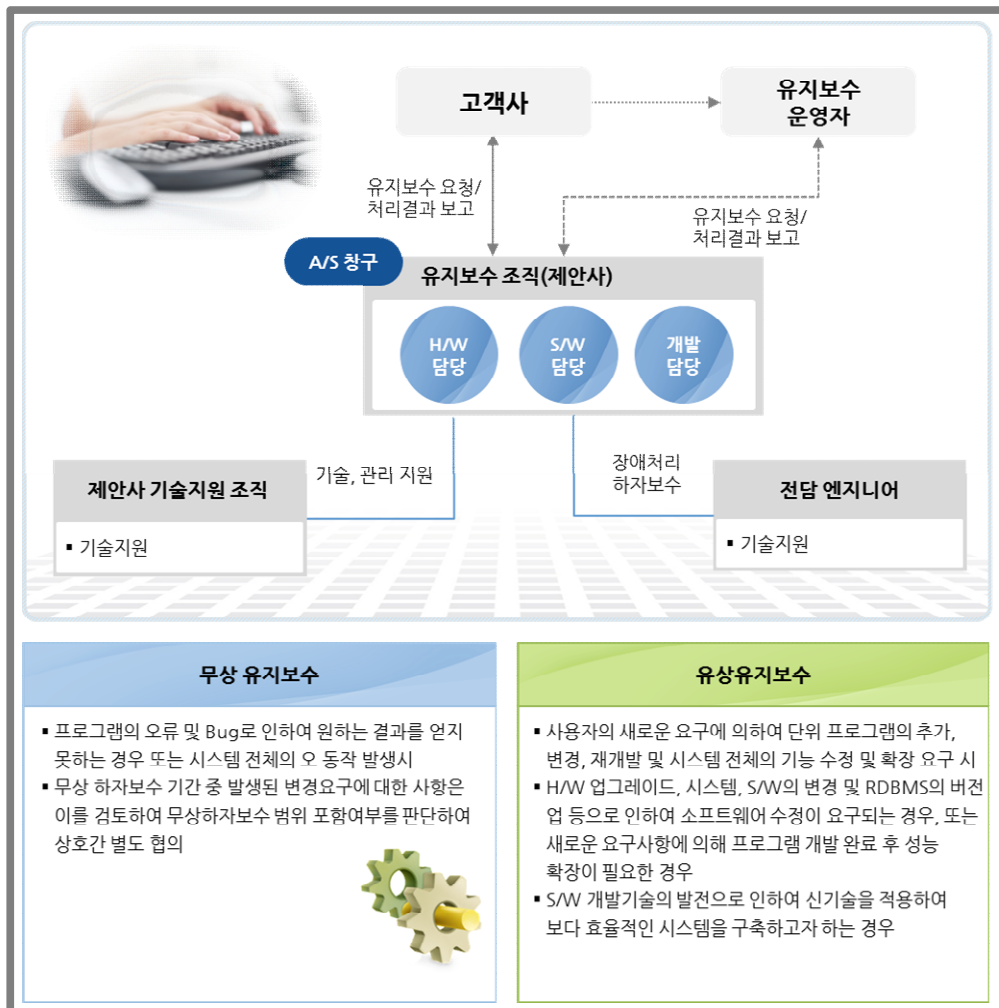
범위	수행단계	설명
백업	백업 수행 및 모니터링	<ul style="list-style-type: none"> 백업 담당자는 계획된 백업에 대해 직접 또는 스케줄링 된 백업에 대한 모니터링 실시
	백업실패 조치	<ul style="list-style-type: none"> 백업 담당자는 백업 중 발생한 에러 유형을 확인하고 조치
	백업 재수행	<ul style="list-style-type: none"> 백업 중 발생한 에러를 해결 한 후 백업을 재수행
	백업관리대장 반영	<ul style="list-style-type: none"> 백업 담당자는 백업 결과를 관리대장에 반영하고, 백업 중 에러가 발생한 경우 해당 내용에 대한 조치경과를 기록
백업	백업 종료	<ul style="list-style-type: none"> 백업 종료 후 백업 결과를 확인하고 현황을 관련자에게 통보

복구	복구요청 접수	<ul style="list-style-type: none"> 데이터의 복구가 필요한 경우 백업담당자에게 복구요청을 신청
	복구계획 수립	<ul style="list-style-type: none"> 백업 담당자는 복구 요청을 분석하고 복구계획을 수립
	복구계획 승인	<ul style="list-style-type: none"> 백업 관리자는 백업 담당자가 수립한 복구계획을 확인하고 이상이 없으면 복구를 승인
	복구 실행	<ul style="list-style-type: none"> 백업 담당자는 복구계획에 따라 백업 데이터를 복구함
	복구 검증	<ul style="list-style-type: none"> 백업 데이터에 대한 검증과 응용프로그램 실행의 이상유무를 확인하고 결과를 관련자에게 통보
	후속조치	<ul style="list-style-type: none"> 백업담당자 및 관리자는 검증결과에 이상이 있을 경우 복구나 백업장비 교체 등의 후속조치를 수행

2. 유지보수

가. 유지보수 범위 및 조직

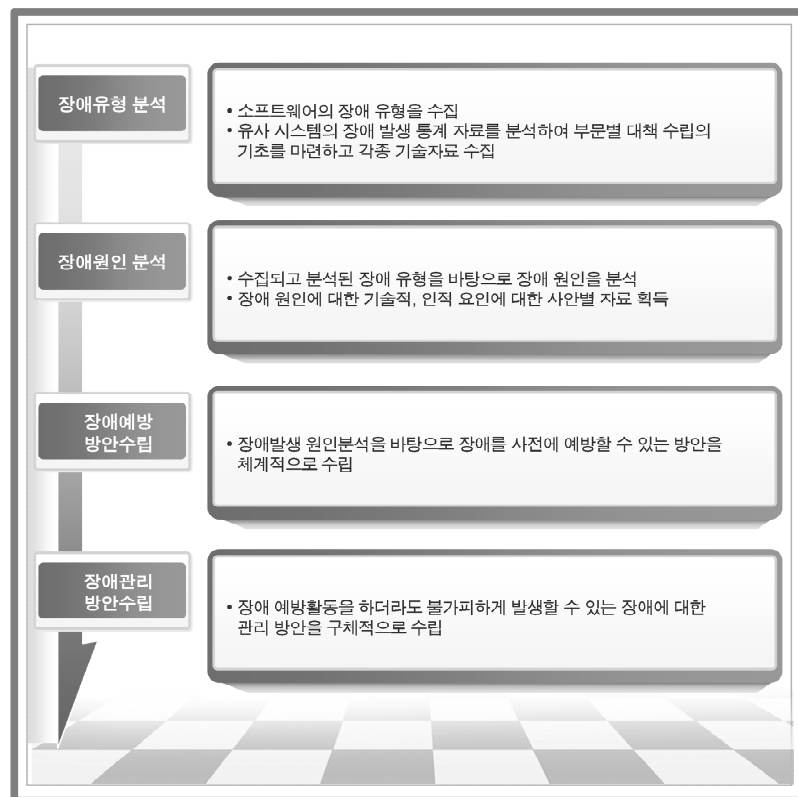
- 본 사업 완료 후 유지보수 기간 동안 도입 시스템의 안정적이고 효율적인 시스템 운영, 신속한 장애처리 서비스를 위하여 개발 참여인력들로 운영지원팀을 구성하여 안정적인 유지보수 서비스를 제공함
- 장애발생에 대한 신속한 대응과 문제에 대한 근본적인 원인을 종합적으로 분석하여 장애 재발을 최소화할 수 있도록 고객사와 대화창구 단일화를 위한 유지보수총괄조직을 구성하고, 각 분야별 유지보수 팀을 운영하여 문제발생 시 신속히 대처함



<그림 4- 50> 유지보수 범위 및 조직

나. 장애관리

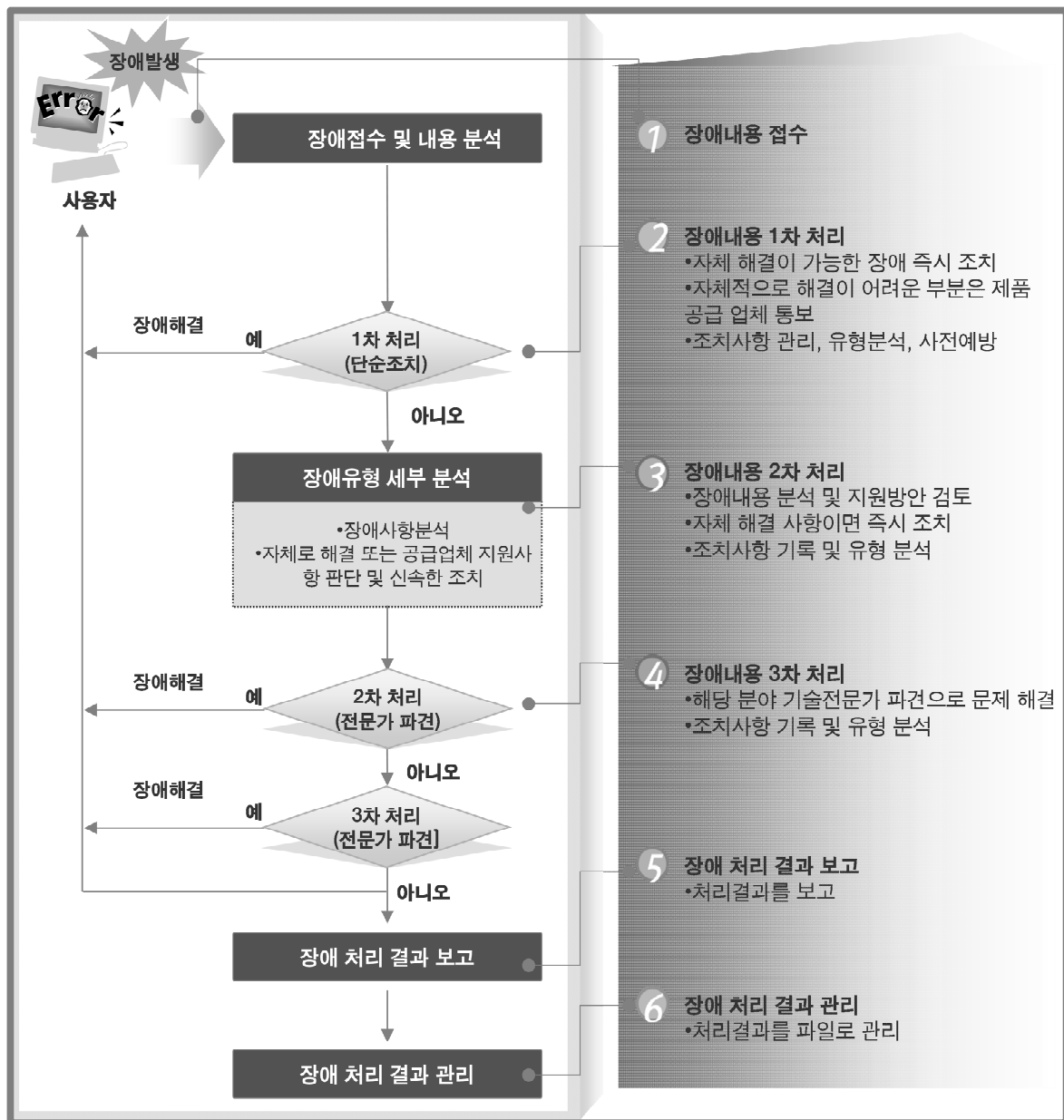
- 시스템의 원활한 서비스와 효율적인 운영을 위해서는 시스템의 안정성, 신뢰성, 유연성 및 장애대책이 확보되어야 함. 따라서 각 시스템 자원에 대한 원활한 운영 및 장애 발생에 대한 조속한 대응이 가능하도록 중단 없는 서비스, 신속한 시스템 복구, 시스템 장애의 Zero화를 장애 대책의 최우선 목표로 함
- 예상되는 데이터 오류 등의 장애요인을 파악하고 대처 방안을 마련하여 즉각적인 조치가 가능하도록 함
- 사전에 장애 및 오류 사항에 대해 원인을 분석하여 해당 장애 발생 시 신속하고 효율적인 장애처리가 가능하도록 함
- 시스템을 원활하고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



<그림 4-51> 장애관리 절차

다. 장애 대응

- 장애 발생 시 신속 정확한 장애원인 분석 및 조치로 장애시간을 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 장애시간을 최소화 함



<그림 4- 52> 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 ‘17년 개발한 모바일 가공 알고리즘과 모바일 분석 맵 구축 알고리즘을 보완하였으며, 이를 기반으로 17년 기준 모바일 DB를 구축하였음
- 핑퐁 핸드오버 데이터의 주요 발생원인(야간시간대 전압 조절로 인한 신호 탐색)을 고려하여 하루를 넘어가면서 발생하는 핑퐁 데이터가 최대한 보정될 수 있도록 탐색 기준을 일 단 위에서 일주일 단위로 변경함
- 사람의 이동특성을 파악하는데 필수 요소인 ‘통행목적’정보를 ‘17년에 개발한 DB에 추가하여 모바일 DB를 구축함
 - 모바일 DB를 통해 통행목적을 분류하려면 기점과 종점의 기준이 되는 체류지 유형을 기존보다 구체화 시킬 필요가 있어, 심야시간대 주체류지, 낮 시간대 주체류지, 잠재체류지 세 가지로 구분되었던 유형을 각각 나누어 낮 시간대 주체류지는 집과 집 이외 심야시간대 주체류지로, 낮 시간대 주체류지는 회사와 학교로 잠재체류지는 종교집회 장소와 기타로 세분류함
 - 통행 로그에서 각 체류지를 구분하기 위해 공간적 특성이 개인의 통행과 활동에 영향을 줄 수 있다는 가정 하에 주로 방문하는 시간대, 평균 체류 시간, 방문 빈도, 주요 통행자(연령 특성 반영), 관심지점정보 위치 여부를 기준으로 체류지를 구분하는 기준을 개발함
 - 이를 통해 추출할 수 있는 통행목적 유형은 총 5가지(출근 / 등교 / 여가 / 귀가 / 기타)이며, 도착지에 따라 귀가 목적 통행은 ‘퇴근’, ‘하교’, ‘귀가’로 기타 목적 통행은 ‘종교 활동’, ‘친지 방문’ ‘기타’로 세분 가능함
 - 통행 목적이 구분된 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 2017년 1월 1일부터 2017년 12월 31일까지 발생한 로그 기록을 가공하여 DB를 구축함

<표 5- 1> 모바일 기반 DB 형태 (예시)

출발				도착				통행 목적	성 별 코 드	연 령 코 드	유동 인구 수
일자	시 간	교통 폴리곤 ID	체류지 유형	일자	시 간	교통 폴리곤 ID	체류지 유형				
20171121	13	48270310	집	20171123	14	12390981	기타	여가	M	20	10
20170501	06	48270310	집	20170501	09	32789014	회사	출근	F	50	25
20170311	17	48270320	집	20170312	18	45608912	학교	하교	M	10	15

- 전국 약 15,937개로 나뉘었던 교통폴리곤('17년 기준)은 주 기지국 기준을 변경하고, 교통폴리곤 최소면적 기준을 강화하여 16,335개로 세분화 되었으며, 이로 인해 시도별 교통폴리곤 면적은 평균 약 6.8% 감소하였음
- 주 기지국 기준을 한 달 동안 25일 이상 신호가 기록된 기지국에서 1년 365일 이상 신호가 기록된 기지국으로 변경하고, 교통폴리곤 최소면적을 시도별 폴리곤 면적 누적 분포함수 값(CDF) 5% 이하에서 시군구별 폴리곤 면적 누적분포함수 값(CDF) 5% 이하로 변경함
- 부산광역시, 경기도, 제주특별자치시를 제외하고 나머지 지역은 모두 '17년도 보다 세분되었으며, 특히 세종특별자치시, 울산광역시, 강원도, 충청북도, 전라북도 지역이 많이 개선된 것으로 나타남
- 또한 '17년 개발된 교통 모니터링·데이터 제공·분석 플랫폼(View-T)를 통해 모바일 자료 기반의 분석 서비스가 제공될 수 있도록 기반 정보 DB를 설계하고, 모바일 DB 기반 분석 기능(통행자 기반의 O/D 분석, 유입유출 분석, Hot Place 분석 등)을 개발하였음

<표 5- 2> View-T 2.0 모바일 자료 기반 분석 서비스 개발 주요 내용

구분		내용
데이터구축	통신데이터 분석 맵 레이어	<ul style="list-style-type: none"> • 분석구(교통폴리곤) 구축 • 분석구 단위의 노드 구축 • 행정구역 단위의 폴리곤/노드 구축
	모바일 통행형태 기반 DB 구축	<ul style="list-style-type: none"> • 서비스 기능 구현 및 확장성을 고려한 DB 설계 • 검색속도 향상을 위한 MongoDB 기반의 DB 설계 • 통신 빅데이터 기반의 통행량 DB 구축

통신데이터 기반의 서비스 기능 개발	O/D 분석	<ul style="list-style-type: none"> • 분석지역에 대한 OD 분석 • 행정구역(시도, 시군구, 읍면동) 단위로 분석지역 설정 • 분석구 단위로 OD 분석 • 단방향, 양방향 통행량 분석 • OD 희망선 표출지도 생성 및 표출지도 범례 설정 등
	유입유출 지역분석	<ul style="list-style-type: none"> • 특정지역에 대한 유입유출 통행량 분석 • 시도, 시군구, 읍면동, 분석구 단위로 특정지역 설정 • 전국 시도, 시군구, 읍면동, 분석구에서 특정지역으로 유입되는 통행량을 분석 • 특정지역에서 전국의 시도, 시군구, 읍면동, 분석구로 유출되는 통행량 분석 • 분석결과에 대한 분포도(폴리곤) 생성 • 도착지-출발지(유입), 출발지-도착지(유출) 희망선 표출지도 생성 • 분포도, 희망선 표출지도 범례 설정 등
	유입유출 비교분석	<ul style="list-style-type: none"> • 2지역의 유입유출 통행량 분석 • 시도, 시군구, 읍면동, 분석구 단위로 2지역으로 유입되는 통행량을 비교분석 • 2지역에서 시도, 시군구, 읍면동, 분석구 단위로 유출되는 통행량을 비교분석 • 분포도, 희망선 표출지도 생성 및 표출지도 범례 설정 등
	Hot Place 분석	<ul style="list-style-type: none"> • 시도, 시군구, 읍면동을 대상으로 특정조건에서 유입유출되는 통행량의 분포(도) 분석 • 기간, 연령별, 성별, 시간대 등에 따른 유입유출 통행량을 분석구 단위로 분석 • 분석결과에 대한 분포 및 순위 표출지도 생성 • 표출지도에 대한 범례설정 등
통행량(인구수) 데이터 검색		<ul style="list-style-type: none"> • 검색조건에 따른 일별 통행량 데이터 검색 • 검색조건 : 기간, 연령별, 성별, 시간대, 트립타입 • 검색결과에 대한 파일 출력 등

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	기지국 X좌표	기지국 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 야간시간대 발생하는 핑퐁 핸드오버 데이터를 추가적으로 보완할 방안을 모색해야 할 것이며, 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, 금년도에 개발된 모바일 DB를 기준으로 교통 모니터링·데이터 제공·분석 플랫폼(View-T)에 제공할 분석 서비스가 추가적으로 개발될 필요가 있으며, 서비스 만족도를 높이기 위해 검색, 표출 속도를 개선할 필요가 있음
 - 새로 추가된 통행목적 정보를 기준으로 통행목적별 유출/유입 분석, 통행목적별 주요 통행구간(O/D) 분석 등 새로운 분석 기능을 개발할 수 있음