

2014년 「국가교통조사 및 DB구축사업」

# 대용량교통정보시스템 구축 및 분석

16





# 제 출 문

국토교통부장관 귀하

본 보고서를 국가정보화사업 중 「2014년도 국가교통조사 및 DB구축사업」의 최종보고서로 제출합니다.

2014년 12월

한국교통연구원

원장 이 창 운

**본 『2014년도 국가교통조사 및 DB구축사업』은 다음  
연구진에 의해 수행되었습니다.**

**참 여 연 구 진**

<b>&lt;한국교통연구원&gt;</b>	
연구책임자	<ul style="list-style-type: none"> <li>◦ 김찬성 연구위원</li> </ul>
연 구 진	<ul style="list-style-type: none"> <li>◦ 박인기, 최정민 , 정경옥 연구위원</li> <li>◦ 조종석, 박민철, 박용일, 박상준, 이석주, 김주영, 김희경, 황순연, 홍다희, 천승훈, 연지윤, 장동익, 한진석, 김병관 부연구위원</li> <li>◦ 신영권, 성홍모, 김동호, 김진우, 김규진, 김정은, 강국수, 강재원, 고두환, 김관용, 김성민, 김은미, 김진오, 김형범, 박미란, 박준호, 박홍주, 변상진, 손강주, 서창범, 신동찬, 오연선, 이선아, 정승연, 정재훈, 정창욱, 정현진, 주진호, 최서윤, 탁지훈, 홍성표 연구원</li> <li>◦ 신지현 연구조원</li> <li>◦ 전윤미, 나선영, 소윤종, 윤황섭, 박선임</li> </ul>
<b>&lt;한국해양수산개발원&gt;</b>	
연 구 진	<ul style="list-style-type: none"> <li>◦ 이호춘 전문연구원</li> <li>◦ 이건우 전문연구원</li> <li>◦ 반영길, 김혜주 연구원</li> </ul>

# 『2014년도 국가교통조사 및 DB구축사업』

## 보고서 구성 및 담당연구진

번 호	과 제 명	연 구 진
제 1권	요약보고서	박용일, 황순연, 정경옥, 신영권 김규진, 박준호, 신동찬, 정재훈
제 2권	전국 여객 O/D 보완갱신 연구	박인기, 조종석, 김병관, 강국수, 박미란, 이선아, 탁지훈
제 3권	여객 O/D 조사방법론 개선방안 연구	조종석, 김동호, 정현진
제 4권	여객교통수요 신뢰도 개선방안 연구	천승훈, 김동호, 김성민, 강국수 이선아, 김관용, 탁지훈
제 5권	장래교통계획DB 구축 및 실행방안 연구	김희경, 서창범, 정창욱
제 6권	국가교통DB 사후평가	김주영, 박흥주
제 7권	화물 O/D 보완갱신 연구	박민철, 강재원, 김형범, 변상진
제 8권	물류거점 화물실태조사	박민철, 한진석, 김형범, 변상진
제 9권	해상화물 O/D 보완갱신 및 방법론 연구	이호춘, 이건우, 반영길, 김혜주
제10권	교통시설 인프라 조사 및 교통주제도 구축	최정민, 정승연, 김은미
제11권	교통망 성능평가 연구	홍다희, 손강주, 김진오, 최서윤
제12권	교통유발원단위 조사연구	황순연, 오연선, 고두환
제13권	교통수단 이용실태조사 연구	연지윤, 김정은, 주진호
제14권	특별교통통행실태조사 연구	성흥모, 홍성표
제15권	교통비용 및 온실가스 DB 구축 연구	연지윤, 김정은, 주진호
제16권	대용량교통정보시스템 구축 및 분석	천승훈, 이석주, 장동익, 김진우 김성민

## 『2014년도 국가교통조사 및 DB구축사업』

### 과제별 공동참여·위탁용역 사업자

#### 【공동사업 참여기관】

- 전국여객 O/D 현행화 공동사업 (부산·울산권 부문)
  - ㈜선일이엔씨, 경성대학교산학협력단
- 전국여객 O/D 현행화 공동사업 (대전광역시권 부문)
  - ㈜드림이엔지
- 전국여객 O/D 현행화 공동사업 (광주광역시권 부문)
  - ㈜유신
- 전국여객 O/D 현행화 공동사업 (수도권 부문)
  - 서울연구원, 경기개발연구원, 인천발전연구원
- 전국여객 O/D 현행화 공동사업 (대구광역시권 부문)
  - 대구경북연구원

#### 【위탁용역 사업자】

- 2014년 국가교통DB점검단 운영지원
  - (사)교통투자평가협회
- 교통수단이용실태조사
  - ㈜메트릭스코퍼레이션
- 전국 지역간 여객O/D 조사방법론 개선방안 연구
  - 홍익대학교
- 대도시권 여객O/D 조사방법론 개선방안 연구
  - 경기개발연구원
- 2014년 교통주제도 구축
  - ㈜팀지오&중앙항업(주) 컨소시엄
- 물류거점 화물실태조사
  - ㈜메트릭스코퍼레이션, 부경대
- 국가교통DB-Brief 발간대행
  - ㈜피그마리온

## 【위탁용역 사업자】

- 여객교통수요 신뢰도 개선방안 연구(시외유출입 교통량조사)
  - 동해엔지니어링, 한국교통량데이터베이스, 도시데이터시스템
- 광역시 교통네트워크 성능평가체계 구축 및 분석
  - 서울시립대 산학협력단
- KTDB 전산 인프라 유지보수
  - 아이넷시스템즈
- 첨단자료를 이용한 교통분석용 네트워크 구축방안 연구
  - 현대엠엔소프트(주)
- 특별교통통행실태조사 및 이용자 만족도 조사
  - ㈜리서치랩
- 빅데이터 기반 교통예보를 위한 핵심 요소기술 개발
  - 큐빅웨어&서울대학교
- 장래교통계획 DB 시작품 제작
  - 팀지오
- 첨단자료를 활용한 여객교통수요 신뢰도 개선방안 연구
  - 명지대학교, 큐빅웨어
- Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구
  - 서울대학교, 큐빅웨어
- 복합용도시설 교통유발원단위조사
  - 아이로드테크
- 국민생활시설 교통실태 설문조사
  - 네오알앤에스
- 교통유발원단위조사
  - 도시데이터시스템
- 복합용도시설 교통유발통행실태조사
  - 나이스알앤씨

<b>【자문용역 사업자】</b>
<ul style="list-style-type: none"><li>• 전국 장래 시군 및 읍면동 인구예측에 관한 연구<ul style="list-style-type: none"><li>- 고려대 김기환교수</li></ul></li><li>• 교통유발원단위조사 표본설계 연구<ul style="list-style-type: none"><li>- 경기대 이상은교수, 한국외대 신기일교수</li></ul></li><li>• 효율적인 차량주행거리 산정방법론의 확대방안 연구<ul style="list-style-type: none"><li>- 충북대학교 산학협력단</li></ul></li><li>• 물류시설사업 평가방법론 사례연구<ul style="list-style-type: none"><li>- 한국교통정책경제학회</li></ul></li></ul>

## 최종보고서 목차

- 제 1권 요약보고서
- 제 2권 전국 여객 O/D 보완갱신 연구
- 제 3권 여객 O/D 조사방법론 개선방안 연구
- 제 4권 여객교통수요 신뢰도 개선방안 연구
- 제 5권 장래교통계획DB 구축 및 실행방안 연구
- 제 6권 국가교통DB 사후평가
- 제 7권 화물 O/D 보완갱신 연구
- 제 8권 물류거점 화물실태조사
- 제 9권 해상화물 O/D 보완갱신 및 방법론 연구
- 제 10권 교통시설 인프라 조사 및 교통주제도 구축
- 제 11권 교통망 성능평가 연구
- 제 12권 교통유발원단위 조사연구
- 제 13권 교통수단 이용실태조사 연구
- 제 14권 특별교통통행실태조사 연구
- 제 15권 교통비용 및 온실가스 DB 구축 연구
- 제 16권 대용량교통정보시스템 구축 및 분석





# 목 차

## 요 약

### 제1장 Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구 ..... 1

- 제1절 연구개요 / 3
- 제2절 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화 / 6
- 제3절 교통혼잡비용 추정 연구 / 30
- 제4절 전국 교통혼잡지도 시스템 구축 / 53
- 제5절 전국 교통혼잡지표 및 혼잡비용 추정결과 집계 / 97
- 제6절 결론 및 향후 로드맵 / 112

### 제2장 빅데이터 기반 교통예보를 위한 핵심 요소기술 개발 ..... 115

- 제1절 연구개요 / 117
- 제2절 교통예보 관련 현황 분석 / 123
- 제3절 교통예보를 위한 기초자료 구축 / 152
- 제4절 교통예보를 위한 소통상태 예측 핵심 방법론 개발 / 164
- 제5절 교통예보를 위한 시스템 구축 방안 / 197
- 제6절 사례분석 및 평가 / 207
- 제7절 결론 및 향후 연구 / 246

### 제3장 교통정보DB 통합관리 시스템 설계 ..... 249

- 제1절 추진배경 및 경위 / 251
- 제2절 교통정보 공유현황 / 252
- 제3절 교통정보 통합관리 시스템 구축 방안 / 261
- 제4절 지자체 교통정보 현황 / 266

## 표 목 차

〈표 1-1〉 교통혼잡지표의 개선 내용 .....	12
〈표 1-2〉 Probe 대수에 따른 혼잡지표 정확도 분석결과 .....	13
〈표 1-3〉 이상치 제거 알고리즘 개선 .....	15
〈표 1-4〉 Probe건수에 따른 혼잡지표 정확도 분석결과 .....	17
〈표 1-5〉 교통량 및 속도 자료의 변화에 대한 검토내용 .....	34
〈표 1-6〉 기초 조사자료에 따른 공간적 범위 비교 .....	36
〈표 1-7〉 기존 교통혼잡비용 추정방법의 이론적 진단 및 개선방향(지역간 도로) .....	39
〈표 1-8〉 기존 교통혼잡비용 추정방법의 이론적 진단 및 개선방향(도시부 도로) .....	41
〈표 1-9〉 교통혼잡비용 추정을 위한 링크별 입력데이터 테이블 정의서 .....	46
〈표 1-10〉 도로등급 별 작업 내역 .....	55
〈표 1-11〉 표준노드링크 오류 수정 내역 .....	59
〈표 1-12〉 고속도로 오류 수정 내역 .....	60
〈표 1-13〉 분석 Map 도로등급 별 구축 현황 .....	61
〈표 1-14〉 표준노드링크의 링크 테이블 .....	66
〈표 1-15〉 표준노드링크의 노드 테이블 .....	67
〈표 1-16〉 분석맵의 링크 테이블 .....	68
〈표 1-17〉 분석맵의 링크 테이블 .....	69
〈표 1-18〉 분석맵과 표준노드링크의 맵핑테이블 .....	70
〈표 1-19〉 도로교통량 상시 테이블 .....	70
〈표 1-20〉 도로교통량 수시 테이블 .....	71
〈표 1-21〉 고속도로 영업소 테이블 .....	71
〈표 1-22〉 KOTI-코든라인 테이블 .....	72
〈표 1-23〉 Car Navigation 차량 수집 데이터 테이블 .....	72
〈표 1-24〉 Raw Probe DB 테이블(Car Navigation 데이터 + 혼잡 경계속도) .....	73

〈표 1-25〉 원단위 혼잡지표 테이블 .....	73
〈표 1-26〉 총량 혼잡지표 테이블 .....	74
〈표 1-27〉 혼잡비용 테이블 .....	74
〈표 1-28〉 Car Navigation 수집 현황 .....	75
〈표 1-29〉 수집 교통정보 가공 후 데이터 결과 비교 .....	80
〈표 1-30〉 포털지도서비스/Open Source 활용 사례 .....	95
〈표 1-31〉 전국 광역시/도 단위 혼잡강도(%) 분석 결과(CASE1) .....	102
〈표 1-32〉 전국 광역시/도 단위 혼잡강도(%) 분석 결과(CASE2) .....	103
〈표 1-33〉 전국 광역시/도 단위 혼잡강도(%) 분석 결과(CASE3) .....	104
〈표 1-34〉 전국 시/군/구 단위 혼잡강도(%) 분석 결과(CASE1) .....	105
〈표 1-35〉 전국 시/군/구 단위 혼잡강도(%) 분석 결과(CASE2) .....	106
〈표 1-36〉 전국 시/군/구 단위 혼잡강도(%) 분석 결과(CASE3) .....	107
〈표 1-37〉 전국 교통혼잡비용 추정 결과(지역간 도로) .....	109
〈표 1-38〉 전국 교통혼잡비용 추정 결과(도시부 도로, 서울특별시) .....	110
〈표 1-39〉 선행 연구의 교통혼잡비용 추정결과와의 비교(평일 기준) .....	111
〈표 1-40〉 단계별 향후 로드맵 .....	113
〈표 2-1〉 예측모형의 분류 .....	129
〈표 2-2〉 국내 교통예보 관련 기술현황 .....	137
〈표 2-3〉 국외 교통예보 관련 기술현황 .....	140
〈표 2-4〉 국내 선행연구 고찰 결과 .....	146
〈표 2-5〉 국외 선행연구 고찰 결과 .....	147
〈표 2-6〉 국내 기술의 한계 분석 .....	148
〈표 2-7〉 국외 기술의 한계 분석 .....	149
〈표 2-8〉 Car Navigation 수집자료 테이블 정의서 .....	154
〈표 2-9〉 수집자료에 대한 보정전, 후 데이터 분석 결과 비교 .....	157
〈표 2-10〉 분석 맵의 도로등급 별 구축 현황 .....	160
〈표 2-11〉 분석 맵 기준의 Car Navigation 기초자료 정의서 .....	167

〈표 2-12〉 본 연구의 공간적 범위 내 기초자료 구축 현황	168
〈표 2-13〉 본 연구의 이상치 제거 방법	174
〈표 2-14〉 과거 이력자료 구축 현황	181
〈표 2-15〉 개별 Navigation 차량 수집 데이터 테이블(기본 데이터)	199
〈표 2-16〉 표준노드링크의 노드 테이블	199
〈표 2-17〉 표준노드링크의 링크 테이블	200
〈표 2-18〉 분석 Map의 링크 테이블	201
〈표 2-19〉 분석 Map의 노드 테이블	202
〈표 2-20〉 분석 Map의 링크 맵핑 테이블	202
〈표 2-21〉 Raw Probe DB 테이블	202
〈표 2-22〉 속도 유전자 지도 DB 테이블	203
〈표 2-23〉 Historical 구간속도/Prob대수 DB 테이블	203
〈표 2-24〉 Current 구간속도/Prob대수 DB 테이블	204
〈표 2-25〉 장래 구간 예측속도 DB 테이블	204
〈표 2-26〉 분석 맵 기준의 Car Navigation 데이터 정의서 및 저장 형식	211
〈표 2-27〉 분석 시나리오 설정	218
〈표 2-28〉 종합평가 결과	220
〈표 2-29〉 전체 분석 네트워크의 평가 결과(시나리오 1)	222
〈표 2-30〉 전체 분석 네트워크의 평가 결과(시나리오 2)	227
〈표 2-31〉 전체 분석 네트워크의 평가 결과(시나리오 3)	232
〈표 2-32〉 전체 분석 네트워크의 평가 결과(시나리오 4)	237
〈표 2-33〉 전체 분석 네트워크의 평가 결과(시나리오 5)	242
〈표 3-1〉 국가교통정보센터 기본교통정보 분류 및 항목구성	252
〈표 3-2〉 도로교통량 조사 지점 수 및 구간 수 (2012년 기준)	253
〈표 3-3〉 국토관리청별 연계현황 (2012년 현재)	254
〈표 3-4〉 국가대중교통정보 데이터 현황	255
〈표 3-5〉 KTDB 신뢰도 제고 및 통계DB 구축을 위한 각 기관별 자료 수집 현황	257

〈표 3-6〉 KTDB 서버 현황 .....	261
〈표 3-7〉 보안 및 네트워크 장비 현황 .....	261
〈표 3-8〉 DB 소프트웨어 현황 .....	261
〈표 3-9〉 인구 30만 이상 지자체 교통정보 수집 및 연계현황 .....	267

## 그림목차

〈그림 1- 1〉 기 개발된 속도별 배출계수(승용차 기준) .....	8
〈그림 1- 2〉 기 개발된 속도별 연료소모량 계수(속도 40km 이하, 승용차 기준) .....	9
〈그림 1- 3〉 기 개발된 속도별 연료소모량 계수(속도 40km 초과, 승용차 기준) .....	10
〈그림 1- 4〉 교통혼잡비용 분석시스템과 연동된 교통혼잡지표 산출 흐름 .....	12
〈그림 1- 5〉 교통량과 속도와의 관계 .....	14
〈그림 1- 6〉 개별 Car Navigation 차량의 속도 분포 .....	15
〈그림 1- 7〉 속도 Profile의 이상치 제거(본 연구의 개선 알고리즘) .....	16
〈그림 1- 8〉 Probe 샘플수에 따른 혼잡 경계속도의 안정성 .....	18
〈그림 1- 9〉 서울외곽순환고속도로(서하남IC → 송파IC 구간) 혼잡경계속도 검증 결과 .....	19
〈그림 1-10〉 인천대교 고속도로(송도 → 인천공항 구간) 혼잡경계속도 검증 결과 .....	20
〈그림 1-11〉 남부순환도로(낙성대삼거리 → 서울대입구역) 혼잡경계속도 검증 결과 .....	21
〈그림 1-12〉 강남대로(교보타워 사거리 → 강남역) 혼잡경계속도 검증 결과 .....	22
〈그림 1-13〉 본 연구의 혼잡경계속도 판정 알고리즘의 구조 .....	23
〈그림 1-14〉 분석 Map 집계 예시 .....	25
〈그림 1-15〉 혼잡지표 이용자 분석방법론 .....	29
〈그림 1-16〉 지역간 도로의 교통혼잡비용 추정 프로세스 .....	31
〈그림 1-17〉 도시부 도로의 교통혼잡비용 추정 프로세스 .....	32
〈그림 1-18〉 전국 표준노드링크 데이터 설명 .....	35
〈그림 1-19〉 본 연구의 교통혼잡비용 추정 프로세스 .....	42
〈그림 1-20〉 해당 링크의 통행속도 분포도 .....	47
〈그림 1-21〉 전국 행정구역별/도로별 혼잡비용 집계 결과(예시) .....	52
〈그림 1-22〉 전체 시스템 구성 .....	53
〈그림 1-23〉 표준노드링크 및 분석 Map 비교 .....	54
〈그림 1-24〉 구축 범위 .....	55

〈그림 1-25〉 분석 Map 관리시스템 구성	56
〈그림 1-26〉 분석Map 편집 UI	57
〈그림 1-27〉 표준노드링크 오류체크 UI	58
〈그림 1-28〉 분석맵 오류체크 UI	58
〈그림 1-29〉 맵핑테이블 생성 예시	59
〈그림 1-30〉 고속도로 형상 수정 결과 화면	60
〈그림 1-31〉 서울시의 표준노드링크 및 분석 Map 비교 화면	61
〈그림 1-32〉 부산시의 표준노드링크 및 분석 Map 비교 화면	62
〈그림 1-33〉 시스템 화면 구성	63
〈그림 1-34〉 전국교통혼잡지도 시스템 ERD	64
〈그림 1-35〉 Car Navigation 수집 유형(특이사항)	76
〈그림 1-36〉 표준노드링크 오류 예시	77
〈그림 1-37〉 통행 그룹 생성	78
〈그림 1-38〉 통행 그룹 경로 보정	78
〈그림 1-39〉 ‘A’ 단말기의 경로 보정 결과 화면	79
〈그림 1-40〉 ‘B’ 단말기의 경로 보정 결과 화면	79
〈그림 1-41〉 수집교통정보 가공 및 Raw Profile 생성	80
〈그림 1-42〉 원단위 혼잡지표 생성 UI	81
〈그림 1-43〉 관측지점 및 표준노드링크 맵핑테이블 생성	82
〈그림 1-44〉 수도권 및 전국 관측지점 분포도	83
〈그림 1-45〉 교통량 전수화 UI	83
〈그림 1-46〉 스케일 설정에 따른 크기 변화	84
〈그림 1-47〉 주제도 UI	85
〈그림 1-48〉 도로 및 교차로 UI	86
〈그림 1-49〉 행정구역 주제도 UI 및 시군구 주제도 예시 화면	87
〈그림 1-50〉 읍면동 - 충청도 지역만 선택 시 디스플레이 결과	87
〈그림 1-51〉 Grid 주제도 UI	88

〈그림 1-52〉 그리드 주제도 예시 화면 .....	88
〈그림 1-53〉 관측교통량 UI 및 예시 화면 .....	89
〈그림 1-54〉 레이어 정의 UI .....	90
〈그림 1-55〉 2013년 12월 네비게이션 통계 결과 .....	90
〈그림 1-56〉 WebBook 화면구성 .....	91
〈그림 1-57〉 메뉴 별 화면 구성 .....	92
〈그림 1-58〉 행정구역별 혼잡도 .....	93
〈그림 1-59〉 도로별 혼잡도/평균속도 .....	93
〈그림 1-60〉 GIS 시스템 구성 비교 .....	94
〈그림 1-61〉 혼잡지도 Web 시스템 구성도 .....	95
〈그림 1-62〉 웹 서비스 예시 화면 .....	96
〈그림 1-63〉 전국 고속도로 전수화 교통량 평가결과(전체, 평일기준) .....	100
〈그림 1-64〉 전국 일반국도 전수화 교통량 평가결과(전체, 평일기준) .....	100
〈그림 1-65〉 전국 지방도 전수화 교통량 평가결과(전체, 평일기준) .....	101
〈그림 1-66〉 서울특별시 전수화 교통량 평가결과(전체, 평일기준) .....	101
〈그림 2- 1〉 연구의 수행체계 .....	120
〈그림 2- 2〉 본 연구의 교통예보 개념(1) .....	121
〈그림 2- 3〉 본 연구의 교통예보 개념(2) .....	122
〈그림 2- 4〉 장래 교통예보시스템 정보표출 예시 .....	122
〈그림 2- 5〉 거시적 교통량과 밀도의 관계(장현호 등., 2004) .....	124
〈그림 2- 6〉 교통류 상태의 시/공간적 진화 .....	125
〈그림 2- 7〉 교통류 상태의 시간적 진화 .....	127
〈그림 2- 8〉 k-최근린 이웃의 전략 .....	136
〈그림 2- 9〉 한국도로공사 교통정보(ROAD PLUS) .....	138
〈그림 2-10〉 서울시 교통정보센터(TOPIS) .....	138
〈그림 2-11〉 엔나비/SK T-Map 교통정보 .....	139
〈그림 2-12〉 독일 OLSIM 교통정보 .....	141



〈그림 2-13〉 영국 Help2Travel 교통정보 .....	142
〈그림 2-14〉 영국 Traffic England 교통정보 .....	143
〈그림 2-15〉 미국 INRIX Traffic 교통정보 .....	144
〈그림 2-16〉 미국 Beat the Traffic .....	145
〈그림 2-17〉 링크별 Probe건수와 속도 정보생성 개념 .....	153
〈그림 2-18〉 ITS 표준노드링크 자료 설명 .....	154
〈그림 2-19〉 수집자료의 텍스트 저장 형태 .....	155
〈그림 2-20〉 본 연구의 Car Navigation 자료구축 현황 .....	157
〈그림 2-21〉 분석 맵의 구축 개념도 .....	158
〈그림 2-22〉 분석 맵 단위의 Car Navigation 자료 구축 프로세스 .....	159
〈그림 2-23〉 분석 맵 단위의 자료 집계 기준(CASE 1) .....	159
〈그림 2-24〉 분석 맵 단위의 자료 집계 기준(CASE 2) .....	160
〈그림 2-25〉 분석 맵 단위의 자료 집계 기준(CASE 3) .....	160
〈그림 2-26〉 연속류 구간의 기초자료 특성 분석 .....	162
〈그림 2-27〉 단속류 구간의 기초자료 특성 분석 .....	163
〈그림 2-28〉 장래 예측 방법론의 개발 구조 .....	165
〈그림 2-29〉 과거 이력자료 구축을 위한 개발 구조 .....	166
〈그림 2-30〉 분석 맵 단위의 텍스트 저장 형태 .....	168
〈그림 2-31〉 Probe건수의 DB구축 형태 .....	169
〈그림 2-32〉 속도 유전자 지도 개발 형태 .....	170
〈그림 2-33〉 속도 유전자 지도 개발 개념도 .....	171
〈그림 2-34〉 초기 속도 profile의 형태 .....	172
〈그림 2-35〉 개별 Car Navigation 차량의 속도 분포 .....	173
〈그림 2-36〉 이상치 제거 전과 후의 속도 분포 .....	174
〈그림 2-37〉 유전자 지도 DB 구축 형태 .....	175
〈그림 2-38〉 경부고속도로(양재IC → 서초IC 구간) 속도 유전자 지도 .....	176
〈그림 2-39〉 올림픽대로(한남대교→동호대교 구간) 속도 유전자 지도 .....	177

〈그림 2-40〉 인천대교 고속도로(송도→인천공항 구간) 속도 유전자 지도 .....	177
〈그림 2-41〉 강남대로(신논현역 → 강남역 구간) 속도 유전자 지도 .....	178
〈그림 2-42〉 종로(종로2가→종각역 구간) 속도 유전자 지도 .....	178
〈그림 2-43〉 Probe 샘플이 최소샘플수 이상인 구간의 속도 산출방법 .....	179
〈그림 2-44〉 Probe 샘플이 낮은 구간의 구간 통행속도 산출방법 .....	180
〈그림 2-45〉 과거 이력자료 DB 구축 저장 형태 .....	181
〈그림 2-46〉 교통예보를 위한 장래 통행속도 예측 개념도 .....	183
〈그림 2-47〉 모수기반 모형과 비모수 기반의 KNN모형 구조 .....	184
〈그림 2-48〉 개발모형의 예측 영역 .....	186
〈그림 2-49〉 본 연구의 예측모형 구조 .....	187
〈그림 2-50〉 KNN 구축을 위한 상태벡터의 개념 .....	188
〈그림 2-51〉 KNN 군집구축 알고리즘의 DB구조 .....	191
〈그림 2-52〉 KNN 군집 구축 알고리즘 .....	192
〈그림 2-53〉 두 변수의 상태간 거리에 따른 의사결정 그룹 군집화 .....	193
〈그림 2-54〉 예측 단계의 입력값 .....	194
〈그림 2-55〉 예측 단계의 출력값 .....	195
〈그림 2-56〉 최종 의사결정 군집 선택 .....	196
〈그림 2-57〉 DB구조 설계 개념도 .....	197
〈그림 2-58〉 교통예보 시스템 ERD .....	198
〈그림 2-59〉 사례분석 대상구간 위치도(연속류) .....	208
〈그림 2-60〉 사례분석 대상구간 위치도(단속류) .....	209
〈그림 2-61〉 사례분석 대상구간의 분석 네트워크 구축 .....	210
〈그림 2-62〉 Probe 이력자료 특성분석(경부 고속도로, 서초IC→양재IC) .....	214
〈그림 2-63〉 속도 이력자료 특성분석(경부 고속도로, 서초IC→양재IC) .....	214
〈그림 2-64〉 Probe 이력자료 특성분석(올림픽대로, 반포대교→한남대교) .....	215
〈그림 2-65〉 속도 이력자료 특성분석(경부 고속도로, 반포대교→한남대교) .....	215
〈그림 2-66〉 Probe 이력자료 특성분석(강남대로, 신논현역→강남역) .....	216

〈그림 2-67〉 속도 이력자료 특성분석(강남대로, 신논현역→강남역) .....	216
〈그림 2-68〉 본 연구의 적중률(%) .....	217
〈그림 2-69〉 기존 방법과의 MAPE(%)와 적중률(%) 비교 .....	220
〈그림 2-70〉 시나리오 1의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC) .....	223
〈그림 2-71〉 시나리오 1의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교) .....	224
〈그림 2-72〉 시나리오 1의 단속류 구간 평가결과(강남대로, 신논현역~강남역) .....	225
〈그림 2-73〉 시나리오 2의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC) .....	228
〈그림 2-74〉 시나리오 2의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교) .....	229
〈그림 2-75〉 시나리오 2의 단속류 구간 평가결과(강남대로, 신논현역~강남역) .....	230
〈그림 2-76〉 시나리오 3의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC) .....	233
〈그림 2-77〉 시나리오 3의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교) .....	234
〈그림 2-78〉 시나리오 3의 단속류 구간 평가결과(강남대로, 신논현역~강남역) .....	235
〈그림 2-79〉 시나리오 4의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC) .....	238
〈그림 2-80〉 시나리오 4의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교) .....	239
〈그림 2-81〉 시나리오 4의 단속류 구간 평가결과(강남대로, 신논현역~강남역) .....	240
〈그림 2-82〉 시나리오 5의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC) .....	243
〈그림 2-83〉 시나리오 5의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교) .....	244
〈그림 2-84〉 시나리오 5의 단속류 구간 평가결과(강남대로, 신논현역~강남역) .....	245
〈그림 3- 1〉 국가교통정보센터 기본교통정보 중 ‘교통소통정보’ 데이터 예시 .....	253
〈그림 3- 2〉 DGT 샘플 자료 .....	254
〈그림 3- 3〉 차량번호 No.5011(일반화물)의 11월 8일 운행궤적 및 통행구분 예시 .....	260
〈그림 3- 4〉 시스템 구성 현황 .....	262
〈그림 3- 5〉 시스템 구성 현황 .....	263
〈그림 3- 6〉 시스템 구성 현황 .....	265
〈그림 3- 7〉 시스템 구성 현황 .....	265



요약





## 1. Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구

### 가. 연구개요

#### 1) 연구의 배경 및 목적

- 교통시설의 공급은 한정된 자원과 토지로 인하여 지속적으로 늘어나는 교통수요에 미치지 못하므로 이는 교통 혼잡의 주된 원인이 되며, 교통 혼잡은 국민의 경제활동 및 실생활에 큰 불편을 끼치고 있음
- 전국에서 교통 혼잡으로 발생하는 비용은 약 29조원(2010년 기준)으로 국가예산 291조(2010년 기준)의 10%에 해당되며, 이 중 절반이 승용차 보급의 확대에 인하여 발생하는 혼잡비용임
- 전국 자동차등록대수의 경우도 2011년 기준 1,844만대에서 장래 2015년경에는 2,000만대에 도달할 것으로 예상되어 도로의 교통 혼잡은 갈수록 심해질 것으로 판단됨
- 이러한 교통 혼잡에 대한 객관적이고 정확한 진단을 위하여 과거 「2013년 국가교통수요조사 및 DB구축사업」에서는 2013년 9월 한 달 분량의 내비게이션 자료를 활용하여 전국 교통 혼잡지도를 구축하는 연구가 기 수행되었음
- 본 연구에서는 Car Navigation 자료의 범위를 기존 연구보다 확장(2013년 10월~2014년 1월, 4개월치)하여 보다 더 세밀하고 정확한 전국 교통혼잡지도를 구축하고자 함
- 또한 기 수행되었던 연구를 기반으로 알고리즘 검증 및 고도화, 애플리케이션 고도화, 분석용 Map 구축, WebBook 작성, 교통 혼잡비용 추정, Web service를 위한 GIS 웹 엔진 구축 등을 통하여 전국 교통 혼잡지도의 활용성을 증진하고자 함
- 더불어 도시 전반에 걸친 교통 혼잡상태를 보다 더 효율적으로 나타낼 수 있도록 관련 혼잡 지표를 검증 및 보완하며, 이용자 중심의 혼잡지도로서 다양한 분석을 통하여 해당 지역의 교통상태 파악 및 도로망 성능평가 등에 활용될 수 있도록 전국 교통 혼잡지도의 고도화 연구를 수행하고자 함

## 2) 연구의 범위

### ① 공간적 범위

- 대상범위 : 전국권 대상

### ② 시간적 범위

- 분석연도 : 과업기간 내 수집된 Car Navigation 자료
  - 데이터 수집기간(4개월) : 2013년 10월 1일 ~ 2014년 1월 31일

### ③ 내용적 범위

- 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화
  - 교통혼잡지표 검증 및 개선
  - 기 개발된 교통혼잡지도의 구축 알고리즘의 정확도 개선
  - 기 개발된 혼잡판정 알고리즘의 도로특성별 검증
  - 혼잡판정 알고리즘 수행속도 개선
  - 혼잡지표의 시/공간적 집계 방법론 고도화
  - 혼잡지표의 이용자 분석 방법론 고도화
- 전국 교통 혼잡지도 혼잡비용 추정 연구
  - 기존 교통혼잡비용 산정방법론에 대한 이론적 진단
  - GPS프로브 기반의 교통량 전수화 방법론 및 알고리즘 개발
  - GPS프로브 기반의 교통량 전수화 알고리즘의 파라미터 최적화
  - 전국 도로구간별 혼잡비용 항목별 산정 방법론
  - 전국 도로구간/도로등급/행정구역별 혼잡비용 집계 방법론 개발
- 전국 교통 혼잡지도 시스템 구축
  - 전국 교통 혼잡지도 애플리케이션 고도화
  - 전국 교통 혼잡지도 분석용 Map 구축
  - 전국 교통 혼잡지도 WebBook 작성
  - 전국 교통 혼잡지도의 Web service를 위한 GIS 웹 엔진 구축



- 전국 교통혼잡지표 및 혼잡비용 집계 산출
  - 혼잡지표의 시/공간적 집계 산출 결과
  - 전국 행정구역별/ 도로별 혼잡비용 추정 결과
  - 기존 교통혼잡비용과의 비교 및 원인 분석
- 향후 로드맵 구축
  - 향후 로드맵 구축

#### 나. 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화

- 지난 2013년 한국교통연구원에서 수행한 “2013년 국가교통수요조사 및 DB구축사업 중 Car Navigation 자료를 이용한 교통혼잡지도 연구”의 성과물로 전국 교통혼잡지도 시스템을 개발하였음. 해당 시스템에서 제공하고 있는 교통혼잡지도는 특정한 공간(도로, 지역)이나 특정한 시간대의 혼잡의 정도를 한눈에 파악할 수 있도록 전국 도로의 혼잡상태를 GIS Map상으로 표출한 21세기형 도로망 정보지도로서 활용되고 있음
- 하지만, 지난 과업에서는 3개월이라는 짧은 사업 기간으로 인하여 불가피하게 전국 교통혼잡지도 시스템 개발에만 집중하게 되어 시스템 내 교통혼잡지표를 산출하기 위한 관련 알고리즘의 검증과정이 미미하였음. 또한, 교통혼잡지표를 산출하기 위하여 이용되고 있는 Car Navigation 자료도 2013년 9월 한 달 분량의 데이터만을 이용함에 따라 데이터 수급의 한계도 발생하였음
- 본 연구에서는 먼저 Car Navigation 자료 수급의 범위가 지난 과업보다 확장되어 보다 더 세밀하고 정확한 전국 교통혼잡지도를 구축하기 위한 제반 여건이 갖춰지게 되었음
  - 기존 연구 : 1개월치 자료 수급(2013년 9월 1일 ~ 30일)
  - 본 연구 : 4개월치 자료 수급(2013년 10월 1일 ~ 2014년 1월 31일)
- 확장된 Car Navigation 자료를 이용하여 전국 도로의 혼잡 상태를 보다 더 효율적으로 나타낼 수 있도록 관련 교통혼잡지표 산출 알고리즘을 검증하고, 교통 혼잡에 대한 객관적이고 정확한 진단을 위하여 이용자 중심의 다양한 분석을 통하여 전국 교통 혼잡지도의 고도화 연구를 수행하고자 함
- 본 연구의 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화에 대한 세부 내용은 다음과 같음
  - 교통혼잡지표 검증 및 개선
  - 기 개발된 교통혼잡지도의 구축 알고리즘의 정확도 개선

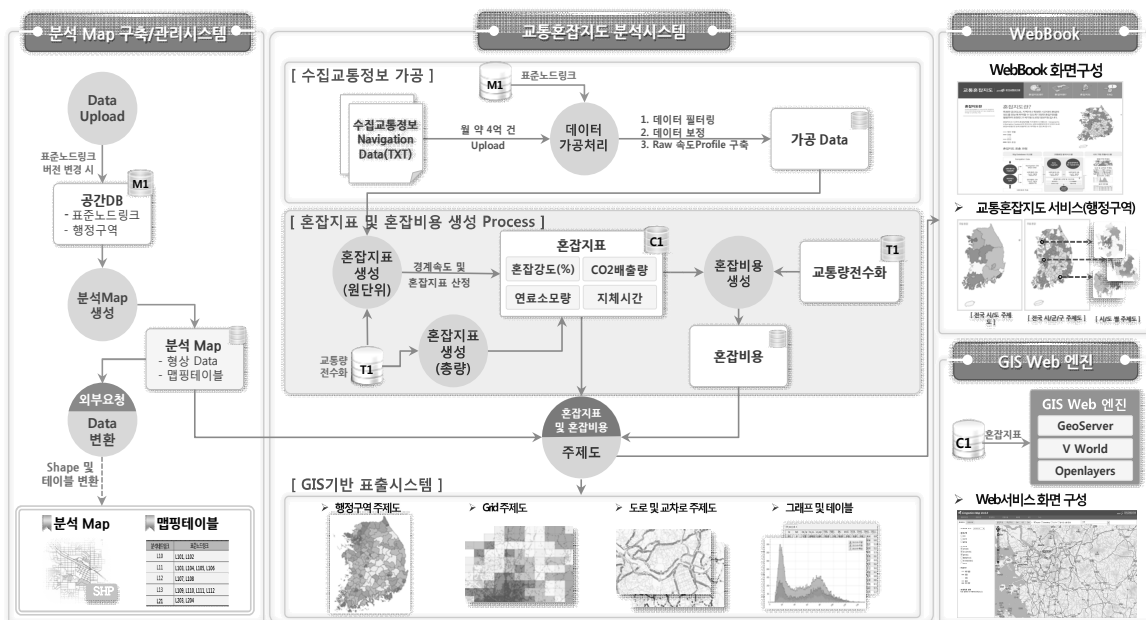
- 기 개발된 혼잡판정 알고리즘의 도로특성별 검증
- 혼잡판정 알고리즘 수행속도 개선
- 혼잡지표의 시/공간적 집계 방법론 고도화
- 혼잡지표의 이용자 분석 방법론 고도화

#### 다. 교통혼잡비용 추정 연구

- 전국 교통혼잡비용은 2010년 기준 29조원으로 2010년 기준 국가 예산이었던 291조의 9.9%에 해당하며, 승용차 보급이 늘어남에 따라 발생한 비용이 교통혼잡비용의 절반을 차지하고 있음
- 자동차 등록대수는 2004년 1,493만대에서 2014년 1,950만대(1월 기준)로 지속적으로 증가하는 추세를 보이고 있으며, 이로 인해 도로의 교통 혼잡이 갈수록 심해질 것으로 판단됨
- 교통혼잡비용은 교통의 외부불경제 규모를 제시하는 정책지표로, 각종 교통정책(국가기간교통망계획, 도로망계획, 국가철도망계획, 물류기본계획, 대중교통기본계획 등)에서 가장 많이 인용되고 중요한 정책지표로 활용되고 있음
- 따라서 지속적으로 교통혼잡비용에 대한 모니터링/관리/개선이 이루어져야 하나, 교통혼잡과 관련한 기초 데이터가 부족한 이유로 관련 연구 및 대책이 적절히 수립되지 않고 있음
- 과거의 교통혼잡비용 추정 관련 연구는 교통조사의 어려움으로 인해 풍부한 자료를 취득하는 것이 어려워 연구의 범위가 제한적이었음
- 교통혼잡비용을 추정하기 위한 분석자료의 계절적/시간적 특성이 반영되지 못하는 문제가 존재하였으며, 2014년 현재 2012년을 기준으로 한 교통혼잡비용 추정 결과가 제시되고 있어 교통혼잡비용 추정 년도에 대한 시의성 확보가 이루어지지 못함
- 하지만 최근 네비게이션 수요의 증가 및 스마트폰의 보급에 따라 교통분야 빅데이터 이용이 폭발적으로 증가하고 있으며, 차량의 네비게이션에서 제공되는 개별 차량의 이동 궤적 정보를 활용하는 것이 가능해짐
- 즉, 기존 교통데이터의 단편적, 편향적 한계를 극복하여 네트워크 커버리지, 이력/실시간 정보 및 객관성과 일관성을 확보할 수 있게 됨
- 본 연구에서는 네비게이션에서 제공되는 개별 차량 이동궤적 정보를 활용하여 교통혼잡비용을 추정하는 방법론을 개발하고자 하며, 개발된 교통혼잡비용 추정방법론을 기존에 적용해 온 교통혼잡비용 추정방법과 비교/분석 수행하도록 함

## 라. 전국 교통혼잡지도 시스템 구축

- 전국교통혼잡지도 시스템은 분석맵 관리시스템, 교통혼잡지도 분석시스템, Web service부분으로 구성
- 분석맵 관리시스템은 분석맵을 구축하는 편집기능, 검증기능, 관리기능으로 구성
- 교통혼잡지도 분석시스템은 데이터 가공 프로세스, 혼잡지표 및 혼잡비용 생성 프로세스, GIS기반 주제도 표출시스템으로 구성
  - 데이터 가공 프로세스는 수집교통정보의 단말기 별 1차 필터링, 단말기 별 데이터 보정, Raw 속도 Profile 생성 단계이며, 표준노드링크 단위로 Output이 생성되도록 시스템 개발
  - 혼잡지표 및 혼잡비용 생성 프로세스는 수집 데이터의 Probe 및 관측교통량을 기반으로 교통량 전수화 후 혼잡지표(총량) 및 혼잡비용을 산정하며, 분석맵 단위로 Output 생성
  - 주제도 표출시스템은 표준노드링크, 분석맵, 행정구역, Grid 단위로 표출하도록 개발
- 전국교통혼잡지도 분석시스템에 의하여 생성된 다양한 콘텐츠는 GIS Web 엔진을 기반으로 Web 서비스함



<그림 1> 전체 시스템 구성

## 마. 전국 교통혼잡지표 및 혼잡비용 추정결과 집계

### 1) 교통량 전수화 평가 결과

#### ① 전국 지역간 도로

##### a) 고속도로

- 전국 고속도로의 관측교통량 조사지점 총 509개소(1,018개 링크 매칭)에 대하여 실제 관측교통량과 전수화 교통량에 대한 평가결과는 전체 관측교통량 조사지점을 기준으로 MAPE 11.5%,  $R^2=0.97$ 로 전체적으로 모형의 평가가 우수한 것으로 분석되었음

##### b) 일반국도

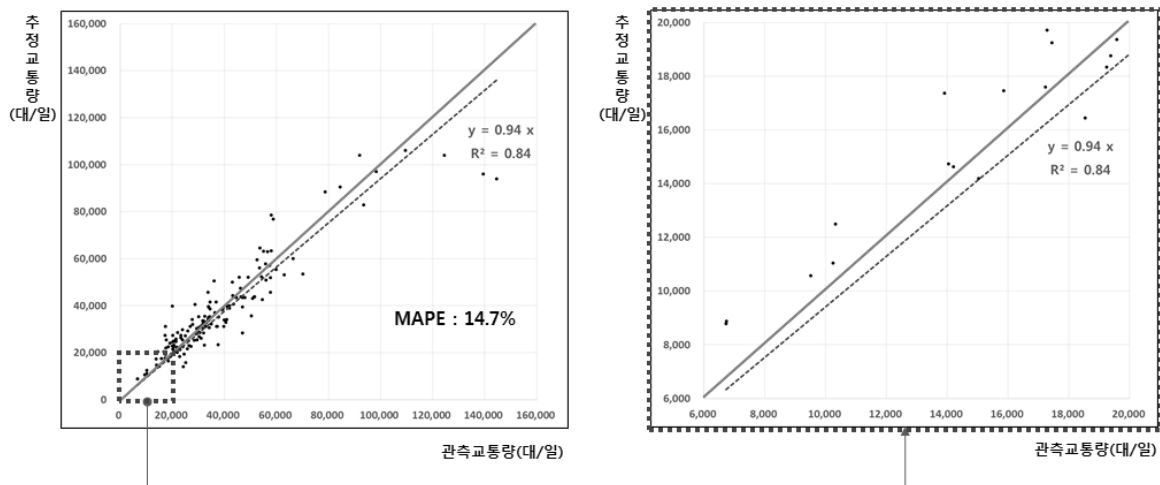
- 전국 고속도로의 관측교통량 조사지점 총 1,198개소(2,396개 링크 매칭)에 대하여 실제 관측교통량과 전수화 교통량에 대한 평가결과는 전체 관측교통량 조사지점을 기준으로 MAPE 29.0%,  $R^2=0.95$ 로 앞서 고속도로의 경우 보다 추정교통량의 다소 오차가 높은 것으로 분석되었음

##### c) 지방도

- 전국 고속도로의 관측교통량 조사지점 총 1,494개소(2,988개 링크 매칭)에 대한 평가결과는 전체 관측교통량 조사지점을 기준으로 MAPE 38.7%,  $R^2=0.93$ 로 앞서 고속도로와 일반국도의 경우보다 추정교통량의 오차가 다소 높은 것으로 분석되었음

#### ② 도시부 도로 : 서울특별시

- 서울특별시의 관측교통량 조사지점 총 96개소(192개 링크 매칭)에 대한 평가결과는 전체 관측교통량 조사지점을 기준으로 MAPE 14.7%,  $R^2=0.84$ 로 전체적으로 모형의 평가가 우수한 것으로 분석되었음



<그림 2> 서울특별시 전수화 교통량 평가결과(전체, 평일기준)

## 2) 교통혼잡지표 분석 결과

- 전국 고속도로를 대상으로 광역시/도 단위의 집계를 수행한 결과, 주중 기준 전국 평균 11.7% 보다 혼잡한 광역시/도는 총 6개 지역으로 서울특별시, 인천광역시, 부산광역시 등의 순으로 분석되었음
- 주말의 혼잡강도(%)는 고속도로의 특성 상 도시지역 보다는 지방의 관광지역을 많이 포함하는 강원도, 경상북도, 충청북도를 제외하고 지역별로 주중 보다 낮은 것으로 분석되었음
- 전국 고속도로를 제외한 나머지 도로를 대상으로 광역시/도 단위의 집계를 수행한 결과, 앞서 주중 기준 고속도로의 경우보다 전체적으로 혼잡강도가 높게 분석되었음. 이는 지역간 고속도로를 제외한 지역별 도시 내부도로들에 대한 교통혼잡이 포함되어 있기 때문임
- 또한, 지역별로 혼잡강도(%)를 살펴보면 주중 기준 전국 평균 29.0% 보다 혼잡한 광역시/도는 총 7개 지역으로 인천광역시, 대구광역시, 부산광역시 등의 순으로 분석되었음. 이는 지역간 고속도로를 제외한 나머지 도로의 혼잡강도(%)를 집계하므로 지역 간 통행 보다는 해당 지역의 내부통행량이 높은 지역일수록 다른 지역보다 상대적으로 혼잡강도가 높은 것으로 판단됨
- 주중 대비 주말의 혼잡강도(%)는 지역별 모두 낮게 분석되었으며, 이는 지역별로 주중보다 주말 도시 내부도로의 전체 통행량이 감소하기 때문인 것으로 판단됨
- 전국 전체 도로를 대상으로 광역시/도 단위의 집계를 수행한 결과, 전체적으로 CASE1과 CASE2의 중간 정도의 혼잡강도로 분석되었음. 이는 지역간 고속도로와 지역별 내부도로의

교통혼잡이 모두 포함되어 있기 때문임

- 또한, 지역별로 혼잡강도(%)를 살펴보면 주중 기준 전국 평균 25.7% 보다 혼잡한 광역시/도는 총 7개 지역으로 서울특별시, 부산광역시, 대구광역시 등 인구밀도와 자동차 등록대수가 높은 대도시 순으로 분석되었음
- 주중 대비 주말의 혼잡강도(%)는 지역별 모두 낮게 분석되었으며, 이는 지역별로 주중보다 주말 도시 내부도로의 전체 통행량이 감소하기 때문인 것으로 판단됨

### 3) 교통혼잡비용 추정 결과

#### ① 전국 교통혼잡비용 추정 결과

##### a) 지역간 도로

<표 1> 전국 교통혼잡비용 추정 결과(지역간 도로)

(단위: 억원)

구분		평일			주말			전체		
		기준	본연구	증감	기준	본연구	증감	기준	본연구	증감
고정비 포함	고속도로	32,035	11,020	-21,015	10,357	3,961	-6,396	42,392	14,981	-27,411
	일반국도	98,083	41,196	-56,887	37,543	15,018	-22,525	135,626	56,214	-79,412
	국지도	13,662	5,517	-8,145	4,742	1,733	-3,009	18,404	7,250	-11,154
	지방도	23,180	8,452	-14,728	8,734	2,818	-5,916	31,913	11,270	-20,643
	합계	166,960	66,185	-100,775	61,376	23,530	-37,846	228,335	89,715	-138,620
고정비 제외	고속도로	24,431	8,590	-15,841	8,932	3,448	-5,484	33,363	12,038	-21,325
	일반국도	77,496	33,059	-44,437	30,836	12,481	-18,355	108,332	45,540	-62,792
	국지도	10,687	4,421	-6,266	3,884	1,435	-2,449	14,571	5,856	-8,715
	지방도	17,691	6,606	-11,085	7,124	2,322	-4,802	24,815	8,928	-15,887
	합계	130,305	52,676	-77,629	50,776	19,686	-31,090	181,081	72,362	-108,719

주1 : 기준은 선행연구의 혼잡기준속도를 적용한 경우이며, 본 연구는 본 연구의 혼잡경계속도를 적용한 경우임.

주2 : 본 연구는 4개월 분량의 내비게이션 자료를 사용하여 추정된 값으로 참값이 아니며, 본 연구의 방법론을 적용한 추정값임. 향후 내비게이션 자료의 확대 구축 및 방법론의 고도화를 통해 보다 정밀한 교통혼잡비용의 추정이 필요함.

## b) 도시부 도로(서울특별시)

&lt;표 2&gt; 전국 교통혼잡비용 추정 결과(도시부 도로, 서울특별시)

(단위: 억원)

구분	평일			주말			전체		
	기존	본연구	증감	기존	본연구	증감	기존	본연구	증감
고정비 포함	32,616	27,253	-5,363	11,304	9,381	-1,923	43,920	36,634	-7,286
고정비 제외	27,684	23,179	-4,505	9,405	7,825	-1,580	37,089	31,004	-6,085

주1 : 기존은 선행연구의 혼잡기준속도를 적용한 경우이며, 본 연구는 본 연구의 혼잡경계속도를 적용한 경우임.

주2 : 본 연구는 4개월 분량의 내비게이션 자료를 사용하여 추정한 값으로 참값이 아니며, 본 연구의 방법론을 적용한 추정값임. 향후 내비게이션 자료의 확대 구축 및 방법론의 고도화를 통해 보다 정밀한 교통혼잡비용의 추정이 필요함.

## 바. 결론 및 향후 로드맵

## 1) 결론

- IT기술의 발달과 더불어 전 세계적으로 대용량 자료(Big data)에 대한 많은 연구가 수행되고 있으며, 교통분야의 ITS분야 또한 기존의 루프 검지기와 같은 고정식 지검검지체계에서 GPS 및 DSRC(단거리 이동통신)을 기반으로 하는 이동식 개별검지로 교통정보의 수집영역이 확대됨으로 인하여 방대한 양의 자료를 수집하고 있음
- 이러한 ITS분야의 정형 빅 데이터를 활용한 사례로 지난 2013년 한국교통연구원에서 수행한 “2013년 국가교통수요조사 및 DB구축사업 중 Car Navigation 자료를 이용한 교통혼잡지도 연구”에서는 방대한 양의 차량이동궤적 정보를 활용한 전국 교통혼잡지도 시스템을 개발하였음
- 하지만, 지난 과업에서는 짧은 사업 기간으로 인하여 불가피하게 시스템 개발에만 집중하게 되어 교통혼잡지표를 산출하기 위한 관련 알고리즘에 대한 검증과정이 미미하였음. 또한, 교통혼잡지표 산출을 위한 Car Navigation 자료도 2013년 9월 한 달 분량의 데이터만을 이용함에 따라 데이터 수급의 한계도 발생하였음
- 본 연구에서는 Car Navigation 자료를 기존 연구보다 확장(2013년 10월 ~ 2014년 1월, 4개월치)하고 이를 이용하여 전국 교통혼잡지도 구축을 위한 교통혼잡지표 산출 알고리즘의 검증 및 고도화 연구를 수행하였음. 또한, Car Navigation 자료의 활용으로 기존 교통혼잡비

용 추정 방법의 한계를 극복할 수 있는 교통량 전수화 모형, 링크별 교통혼잡비용 추정방법, 지역 단위 교통혼잡비용 집계방법 등의 방법론을 제시할 수 있었음

- 향후 개발모형의 정확도를 더욱 높이기 위해서는 4개월이 아닌 1년 이상의 축적된 Car Navigation 자료가 필요할 것으로 판단되며, 이에 따른 본 연구의 활용 방안은 아래와 같음
  - 도로의 운영현황을 효율적으로 파악 과학적인 도로 운영체계의 수립 지원
  - 다양한 도로망 성능평가 : 지역, 도시, 축, 도로별 위계 등
  - 비 혼잡 교통류 상태에서 속도분포 최고/최저 제한속도 등 운영규제
  - 교통수요분석 도구의 통행비용함수에 활용 가능
  - 신호교차로 운영상태(제어지체, 연동효과, 이동류별 구성비)를 지속적으로 모니터링 가능
  - 속도의 밀도 Profile 이력 DB 혼잡예보 및 다중시간대 동적 통행시간 예측에 활용 가능



## 2) 향후 로드맵 구축

&lt;표 3&gt; 단계별 향후 로드맵

구분	단기(2014년)		중기(2015~2016년)	장기(2017~2018년)
	거시적		-----▶	미시적
혼잡지도	• 교통혼잡지도시스템 구축 (이력)		• 교통혼잡지도시스템 구축 및 운영(이력, 실시간)	• 시스템 운영 및 유지 보수
	• 교통혼잡지도시스템 DB 제공 서비스(이력)		• 교통혼잡지도시스템 DB 제공 서비스(이력, 실시간) • 인터넷 포털 사이트 콘텐츠 제공 • 비즈니스 인텔리전스(BI) 애플리케이션 활용	
	• 교통혼잡지도 제공 서비스(이력) - 혼잡 및 추가 지표 - 평일/주말 일단위 - 전국 광역시/도, 시/군/구, 도로등급별, 링크구간 단위		• 교통혼잡지도 제공 서비스(이력, 실시간) - 혼잡 및 추가 지표 - 년/분기/월/일/시 - 전국 광역시/도, 시/군/구, 읍/면/동, 도로(등급)별, 교차로, 링크구간 단위	
SOC 투자사업 종합 진단 시스템	진단 시스템	-	• SOC투자사업 종합 진단을 위한 지표 개발 • SOC투자사업 종합 진단 분석방법론 개발 • SOC투자사업 종합 진단시스템 구축	• SOC투자사업 종합 진단시스템 운영
	진단 구성 요소 기술 개발	-	• 공간 클러스터링 방법론 개발 • 사회경제지표 분석 클러스터링 • 도로이용 노출도 지표 개발 • 통행기반 수형망 지표 분석 연계방법론 개발 • 시각화 분석 지원 도구 개발	-
	도로 안전 진단	-	• 도로안전진단 방법론 개발 • 도로안전진단을 위한 기초자료 제공	• 도로안전진단 시스템 구축 - 기존도로 교통사고 안전진단 모니터링 • 도로안전진단을 위한 DB 제공

## 2. 빅데이터 기반 교통예보를 위한 핵심 요소기술 개발

### 가. 연구 개요

#### 1) 연구의 배경 및 목적

- 현재 우리나라는 지속적인 경제성장으로 인하여 삶의 수준이 향상되면서 도시의 출퇴근 통행의 증가와 더불어 여가통행의 비중도 꾸준히 증가하고 있음
- 특히 여가를 목적으로 하는 통행의 경우는 그만큼 운전자가 이동하는 날짜와 시간대에 이용하고자 하는 도로들에 대한 교통상황 정보(원활, 서행, 정체)에 민감할 수밖에 없으며, 대부분의 운전자들은 출발 전에는 스마트폰과 각종 인터넷 포털사이트 등에서 제공하는 교통상황 정보에 의존하고, 출발 후에는 차량 내비게이션에 의해 제공되는 이동경로 및 교통상황 정보에 의존하게 됨
- 따라서, 운전자들에게 제공되고 있는 교통상황 정보는 21세기 정보화 시대에 중요한 정보임과 동시에 운전자들이 신뢰할 수 있도록 정확한 교통상황 정보를 제공하는 것이 무엇보다도 중요하다고 볼 수 있음
- 이러한 측면에서 20세기 말 국내에 정보수집 → 정보가공 → 정보제공이라는 일련의 지능형 교통체계(Intelligent Transport System, 이하 ITS)가 도입된 후, 지속적인 기술 발전으로 최근에는 첨단 IT기술을 기반으로 하는 정보수집 매체를 통하여 다양한 정보제공 서비스를 실현하고 있음
- 하지만, 현재 ITS는 대부분 실시간 정보를 이용한 교통상황 정보만을 제공하고 있으며, 장래 교통상태 예측정보(교통예보)의 경우는 일부의 민간 교통정보 제공서비스 업체와 한국도로공사를 제외하면 미미한 수준임. 즉, 21세기 IT기술의 급속한 발달과 더불어 시스템 분야의 하드웨어와 다양한 정보제공 콘텐츠 개발에 따른 기술적인 발전에도 불구하고 정보제공 서비스의 핵심요소인 장래 교통상태 예측분야는 이러한 기술적 발전과 병행하지 못하고 있는 실정임
- 또한, 일부 ITS의 장래 교통상태 예측방법은 대부분 현재시각 기준 5~15분 이후의 장래 교통상태를 예측하는 단기에측(Short-term forecasting)기법에 의존하고 있으며, 단기에측 모형의 입력 값으로 이용되는 자료의 시간적인 차원도 현행(Current)의 실시간 수집정보로 한정시켜왔음. 즉, 기존의 단기에측 방법은 방대한 양의 이력자료 관리 및 활용이라는 측면을 고려할 필요가 없었음

- 이러한 단기예측 기법은 익일 이상의 중장기 장래 교통소통 정보를 예측하고 평가를 수행할 수 없는 모형의 구조적 단점을 가지고 있으며, 기존 모형으로 개발할 경우 시스템의 지원능력도 방대한 이력자료의 활용이라는 요구를 충족하지 못하였음
- 이러한 현실 속에서 ITS에서의 장래 교통정보 예측분야가 한걸음 더 도약하기 위해서는 현재의 단기예측 기술이 넘지 못하고 있는 예측의 시·공간적 범위(Forecasting horizon)를 극복해야 함
- 그 동안 선행 연구들에서 수많은 단기예측 모형이 개발이 되었으나, van Hinsbergen et al.(2007)이 지적한 바와 같이 타 모형에 비하여 매우 우수한 성능을 가진 모형은 없다고 지적하였음. Chang et al.(2012)가 지적한 바와 같이 결국 21세기의 ITS 예측분야의 개척 영역은 수용 가능한 수준 이내에서 예측영역의 시/공간적 확장이라 판단됨
- 또한 최근 내비게이션과 스마트폰의 보급 확대 등과 같이 Big Data로 표현할 수 있는 21세기 첨단 IT기술과 정보화는 기존 ITS 예측기법의 변화를 요구하고 있음. 즉, 스마트폰을 선두로 하는 첨단 개인단말은 다양한 교통정보 Needs의 증대와 동시에 실시간 교통정보를 포함하여 과거의 방대한 이력자료를 활용한 익일 이상의 중장기 장래 교통상태 예측정보를 요구하고 있음
- 더불어 최근의 발전된 Database관리 능력과 검색 엔진도 교통 Big Data를 활용한 장래 교통예보정보 생성에 있어 새로운 가능성을 제시해주고 있음
- 따라서 본 연구의 목적은 개별 차량의 내비게이션 단말에서 수집되는 교통 빅데이터(Probe 건수, 속도)를 이용하여 아래와 같이 익일의 장래 통행속도 예측이 가능한 빅데이터 기반의 통행속도 예측 방법론을 개발하는데 있음
  - 빅데이터 기반으로 하는 장래 통행속도 예측을 위한 핵심 요소기술 개발
  - 개발된 요소기술의 검증을 통한 실용화 가능성 검토
  - 개발 기술의 개선방향 및 실용화를 위한 시스템 요구사항 제시

## 2) 연구의 범위 및 방법

### ① 시간적 범위

- 분석연도 : 2014년(→ 과업기간 내 수집된 Car Navigation 자료 이용)
- 데이터 수집기간(4개월) : 2013년 10월 1일 ~ 2014년 1월 31일

### ② 공간적 범위

- 대상범위 : 서울특별시 강남 일대(→ 전국 모든 도로에 적용이 가능한 방법론 개발)

### ③ 연구의 방법

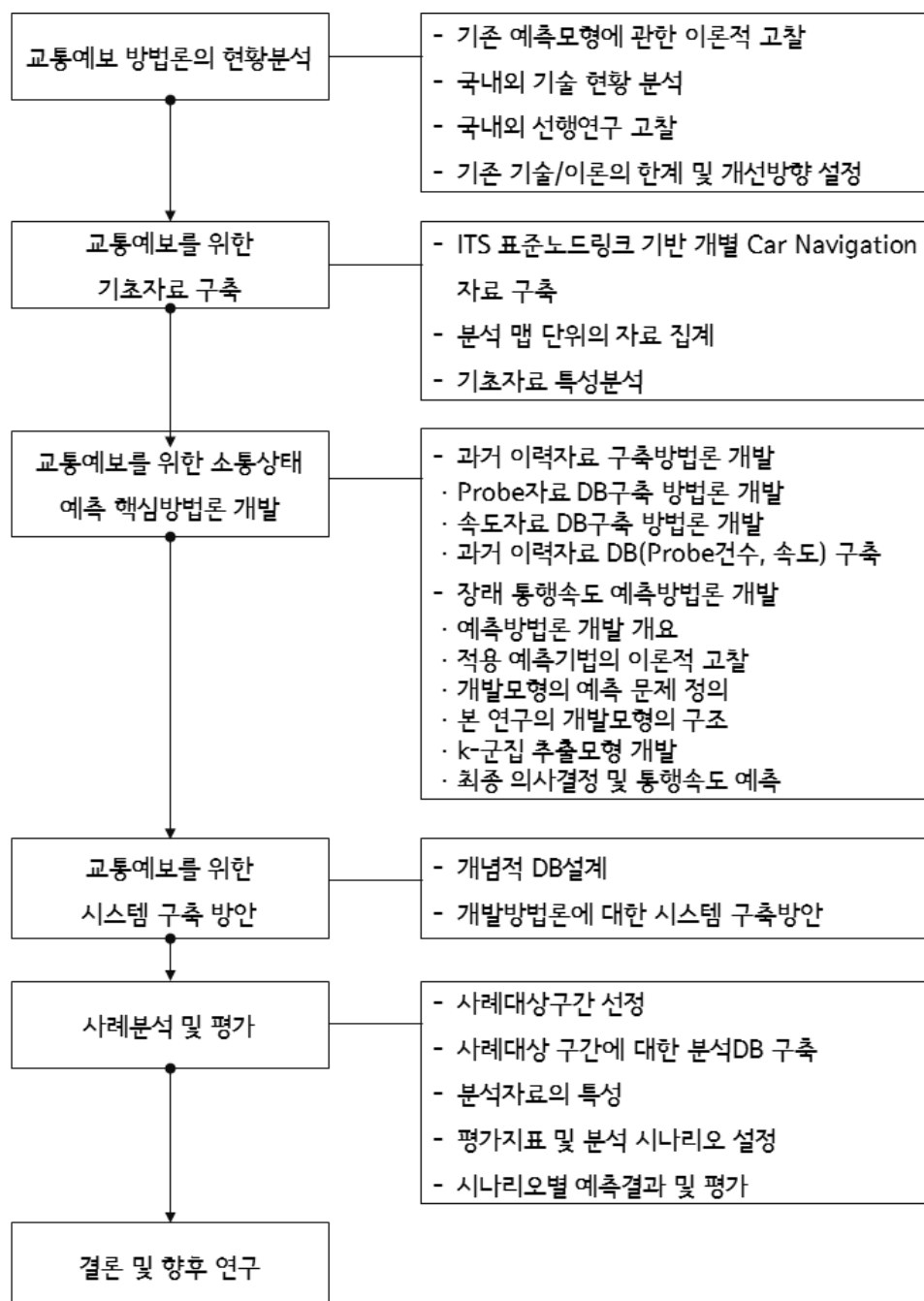
- 본 연구의 목표는 향후 교통예보 시스템에 탑재되기 위한 빅데이터 기반의 장래 통행속도를 예측하는 방법론을 개발하는데 있음. 즉, ITS 표준노드링크 기반의 Car Navigation 자료 (Probe건수, 속도)를 이용하여 익일의 링크별/시간대별 통행속도를 예측하기 위한 핵심방법론을 개발하는 것으로 전체 연구의 방법은 아래와 같이 크게 4단계로 구성됨
- Step 1 : ITS 표준노드링크 기반의 개별 Car Navigation 기초자료 구축
- Step 2 : 과거 이력자료 구축 방법론 개발
  - Probe건수 DB구축 방법론 개발
  - 속도자료 DB구축 방법론 개발(속도 유전자 지도, 장래 구간 통행속도 산출)
  - 과거 이력자료 DB(probe건수, 속도) 구축
- Step 3 : 구축된 이력자료 기반의 장래 통행속도 예측 방법론 개발
  - KNN<sup>1)</sup> 단계 : k-군집 추출모형 개발
  - 예측 단계 : 최종 의사결정 및 통행속도 예측모형 개발
- Step 4 : 방법론 적용 및 평가
  - 평가지표 및 분석 시나리오 설정
  - 시나리오별 예측결과 및 평가

---

<sup>1)</sup> k개의 최근린 이웃(k-nearest neighbor, 이하 KNN)

### 3) 연구의 수행 체계

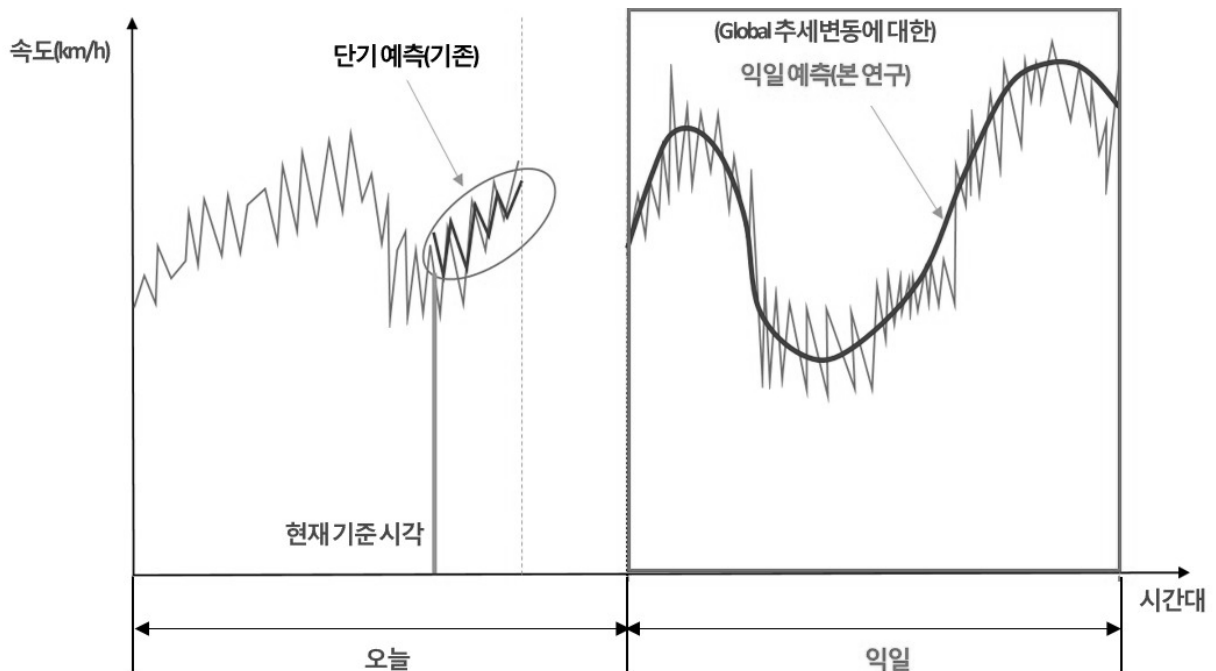
- 본 연구는 교통예보 관련 현황분석, 교통예보를 위한 기초자료 구축, 교통예보를 위한 소통 상태 예측 핵심방법론 개발, 교통예보를 위한 시스템 구축 방안, 사례분석 및 평가, 결론 및 향후 연구로 수행되며, 연구의 수행체계는 아래의 그림과 같음



<그림 3> 연구의 수행체계

#### 4) 본 연구의 교통예보 개념

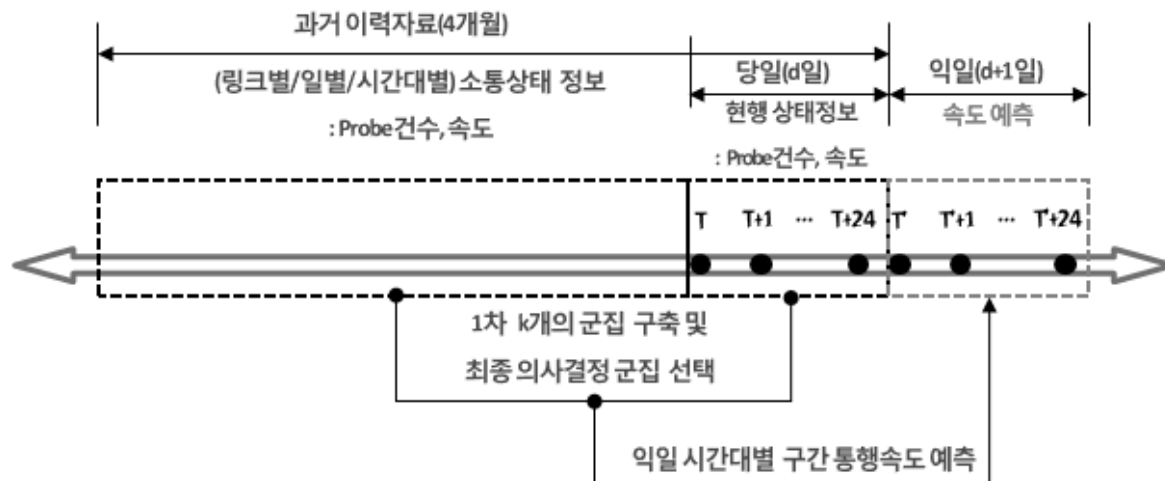
- 본 연구의 교통예보는 기존의 ITS의 예측분야 관점에서 단기예측(Short-term forecasting)이 아닌 중/장기 예측에 해당하는 익일 이후의 장래 소통상태(통행속도)를 예측하는 것임
  - 단기 예측 : 해당 시간대 간격(1~15분)으로 과거 5~15분 이전의 실시간 소통상태 정보를 이용하여 현재시각 기준 5~15분 이후의 장래 소통 상태를 예측
  - 중/장기 예측(본 연구) : 해당 시간대 간격(15분~1시간)으로 당일의 실시간 소통상태 정보를 이용하여 익일 이후(+n일)의 장래 소통 상태를 예측함 → 즉, 본 연구에서는 익일의 소통 상태를 예측



<그림 4> 본 연구의 교통예보 개념(1)

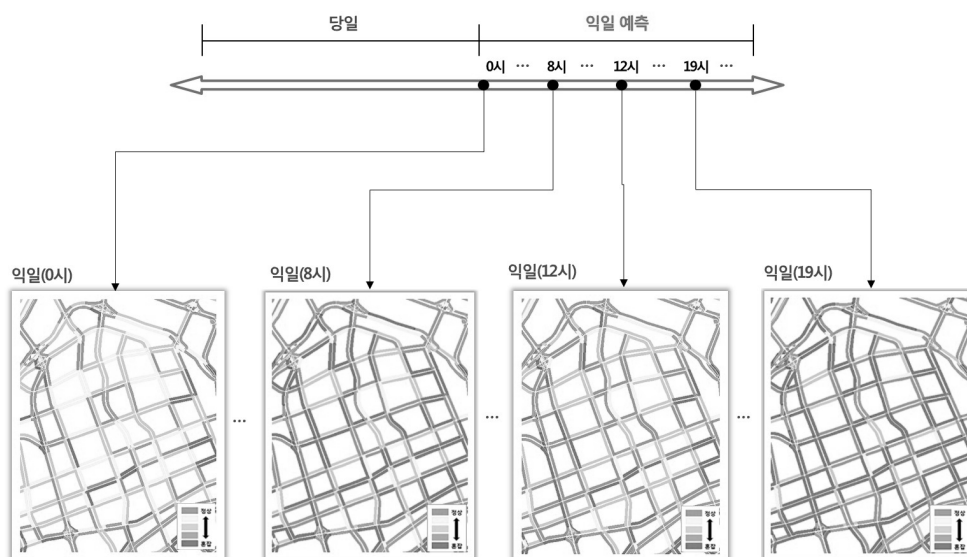
- 본 연구에서는 장래 교통예보를 익일의 링크별/시간대별 통행속도 예측으로 정의하며, 이를 위한 분석 자료는 본 연구의 시/공간적 범위에서 수집된 ITS 표준노드링크 기반의 개별 Car Navigation Probe 데이터를 시간 단위로 집계하여 구축한 링크별/일별/시간대별 Probe 건수와 속도 자료를 이용함
- 익일의 링크별/시간대별 통행속도를 예측하기 위해서는 현재시각 기준으로 당일의 링크별/시간대별 실시간 소통상태 정보(Probe건수, 속도)가 필요하나, 현재는 교통예보시스템을 구축하기 이전이므로 실시간 소통상태 정보를 과거 어느 시점에서의 이력자료로 대체함. 즉, 과

거의 이력자료 중 임의로 선택된 일자(d일)를 당일로 지정하고, 해당 일자의 링크별/시간대별 소통상태 정보를 이용하여 익일(d+1일)의 링크별/시간대별 통행속도를 예측하는 것임



<그림 5> 본 연구의 교통예보 개념(2)

- 향후 장래 교통예보 시스템이 구축될 경우, 익일의 교통예보는 링크별/시간대별 통행속도 예측결과를 교통 혼잡 상태(정상 ↔ 혼잡)로 구분하여 아래의 그림과 같이 표출하고자 함



<그림 6> 장래 교통예보시스템 정보표출 예시

## 나. 교통예보 관련 현황 분석

### 1) 기존 예측모형에 관한 이론적 고찰

<표 4> 예측모형의 분류

구분	예측모형
정적(Naive)기법	동시(Instantaneous)기법
	과거평균(Historical average)기법
	군집(Clustering) 모형
모수(Parametric)기법	교통류 모의실험 모형
	시계열(Time Series)모형
비모수(Nonparametric)기법	퍼지 이론(Fuzzy Logic)
	신경망(Neural Network)
	k-최근린 이웃(K-Nearest Neighbor)

#### ○ 정적기법

- 정적기법(Naive)은 모수 및 비모수 모형에 비하여 낮은 예측력을 보이지만, 실제 적용에 있어 현재로는 가장 현실적이며 보편적인 방안 중 하나임
- 현재 중/장거리 소통상태의 정보제공에 있어 유력한 대안이라 할 수 있으나, 중/장거리 소통상태 정보 예측에 적용할 수 있을 정도로 단기 예측모형의 예측의 시간적 범위가 수용할 만한 수준에서 확장되지 않았기 때문에 중/장거리 예측에 적용하기에는 무리가 있음

#### ○ 모수(Parametric) 모형

- 먼저 모의실험 모형들은 교통사고 등의 비 반복적 상태와 다양한 교통제어 전략을 포함하는 동적 교통류 관리에 적합하나, 교통류 및 차량 모형의 파라미터 정산이 쉽지 않고 장래 예측이 아닌 동적 O/D통행량 추정에 더 적합함. 현재까지 동적 O/D통행량 추정은 지속적인 연구 분야이며, 모형의 연산 수행 속도가 타 차량모형에 비하여 상대적으로 느리다는 단점이 있음
- 시계열 분석 기법들은 확률적 상태이론을 기반으로 시계열적으로 정적(Stationary) 상태를 가정하지만 실제 시스템은 동적 시스템임. 일반적으로 시계열 모형은 계절성(Seasonality)이 필요하며, 교통류 이론보다는 통계적 함수에 의존함. 특히 ARIMA모형은 평균값에 의존하는 모형의 본질적 특성으로 인하여 상태변화가 안정정일 때 예측력은 우수한 반면, 교통상태가 크게 변화하는 전환점에서 다음 상태를 예측하지 못 하는 한계가 있음(Vlahogianni



et al., 2005). 특히 교통상태가 급격하게 변하는 단속류의 행태를 정확히 설명하지는 못함 (Chang et al., 2012b; Yoon and Chang, 2014)

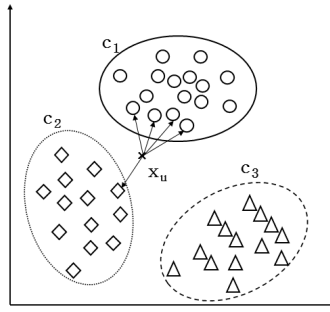
○ 비모수(Non-parametric) 모형

- 신경망 모형은 그 동안 많은 연구에도 불구하고, 한편의 논문(Park and Rilett, 1998)을 제외하면 네트워크 차원의 적용이 이루어지지 않았음. 이는 대다수의 연구가 단일 지점 또는 단일 구간의 자료를 이용하여 수행되었기 때문이며, ITS의 시스템에서 볼 때 신경망 모형의 연산 수행속도는 동적 교통류 관리 및 정보 제공에 불리하기 때문임
- 현재 ITS의 자료수집 기술이 발달하고, 대용량 자료를 지원하게 되면서 비모수 기반의  $k$ -최근린 이웃 기법에 대한 연구가 증가하고 있음(Chang et al., 2012a). Chang et al.(2010)과 Kim et al.(2011)은 해당 모형이 타 모형에 비해 통행시간의 예측에 우수한 예측 정확도를 보였으며, Chang et al.(2012a, 2012b)는 교통량의 예측과 추정에 있어 다중 시간대 예측에도 불구하고 단일 시간대 예측에 주로 이용되는 Seasonal ARIMA, 칼만필터링 모형의 정확도 보다 우수한 것으로 분석됨
- 앞서 살펴본 단기 예측의 모형들은 수용할 만한 예측오차의 수준에서 예측영역을 확장하는데 실패한 반면, 비모수 기반의  $k$ -최근린 이웃 기법은 예측영역의 확장에 성공한 사례를 보고하고 있음(Smith and Demetsky, 1996; Chang et al., 2010, 2012b). 또한 과거 시계열 자료의 변동이 극심한 단속류의 상태 예측에 있어서도 Seasonal ARIMA와 칼만필터링 모형 보다 훨씬 우수한 성능을 보이는 것으로 분석됨(Yoon and Chang, 2014)

2) 본 연구의 적용 기법에 대한 이론적 고찰

- $k$ -최근린 이웃(K-Nearest Neighbor)기법<sup>2)</sup>은 앞서 설명한 비모수(Non-parametric) 모형 중에서 가장 많이 이용되고 있는 방법이며,  $k$ 개의 최근접하는 이웃을 이용한다는 의미에서 붙은 명칭으로 학습 데이터 집합에 있는 표본 간의 유사도에 따라 라벨이 붙지 않은 표본들을 분류하는 매우 직관적인 방법이라고 할 수 있음

<sup>2)</sup> 한빛아카데미(2014), “패턴인식 개론”, p.233



<그림 7> k-최근린 이웃의 전략

- 즉, 라벨이 없는 표본  $x_u \in R^D$ 이 주어질 경우, 학습 데이터 집합에서 가장 가까운  $k$ 개의 라벨 표본을 찾고,  $k$ 개의 부분집합 내에 가장 빈도가 많이 나타나는 클래스에  $x_u$ 를 할당하는 방법으로 단지 상수  $k$ , 라벨이 있는 학습데이터 집합의 표본과 거리 척도만 필요함
- 여기서, k-최근린 이웃기법의 거리 척도는 주로 유클리드 거리(Euclidean distance)를 이용하며, 유클리드 거리를 구하는 식은 아래와 같음
  - 데이터 간의 유클리드 거리( $ED$ ) 측정

$$ED(x, y) = \left( \sum_{k=1}^d |x_k - y_k|^m \right)^{1/m} \quad m=2, \quad k \leq d$$

- <그림 7>과 같이 3개의 클래스( $c_1, c_2, c_3$ )가 있을 경우, 미지의 표본  $X_u$ 에 대한 라벨을 결정하여 이 표본이 속하는 클래스를 찾고자 함. 앞서 설명한 유클리드 거리를 척도로 이용할 경우  $k$ 를 5로 하고 인접한 표본에 대한 거리를 각각 계산하면, 이들 중 4개가 클래스  $c_1$ 에 속하고 나머지 1개가 클래스  $c_2$ 에 속하므로  $x_u$ 는 클래스  $c_1$ 으로 할당됨

### 3) 국내외 기술 현황 분석

#### ① 국내 기술현황

- 국내 교통예보 관련 기술현황은 한국도로공사, 서울시 토피스(TOPIS) 등 주요 공공기관에서 수집된 과거 교통 이력자료를 바탕으로 제한적인 교통예측 정보를 제공 중에 있음
- 또한, SK플래닛, KT 등 이동통신사를 중심으로 한 일부 민간기업에서 휴대폰 어플리케이션을 이용하여 실시간 빠른 길 찾기, 예측 통행시간 정보 등을 제공하고 있음

&lt;표 5&gt; 국내 교통예보 관련 기술현황

기관	대상도로	예측수준	비고
한국도로공사	고속도로	주간예보, 월간예보	과거 2~4주의 평균값을 이용 ※ 정확도 낮은 실정이며, 타 도로에 적용이 불가능함
서울시 교통정보센터 (TOPIS)	서울시 주요도로, 도시고속도로	-	실시간 교통소통정보 위주 (예측시스템은 현재 개발 중)
엔나비/ SK T Map	전국도로	일중	일중 예측정보는 제공하고 있으나, 익일 이상의 교통예보는 제공하지 않음

## ② 국외 기술현황

- 미국의 Beat the Traffic, IBM, INRIX, 영국의 Highway Agency는 대표적인 해외 교통 정보 서비스 제공기관임
- 패턴화된 교통정보를 가공하여 주요 간선도로 및 고속도로를 대상으로 교통정보를 제공함

&lt;표 6&gt; 국외 교통예보 관련 기술현황

기관	대상도로	예측수준	비고
OLSIM(독일)	고속도로	30분~1시간	Autobahn 고속도로 대상
Help2Travel(영국)	Midland 지역	1주후, 2주후 주기	행사/여행정보 제공 ※ 도로별 교통 예측정보 제공하지 않음
Traffic England(영국)	주요 도시별 고속도로 및 국도	10시간	현재 개발 중
INRIX Traffic(미국)	고속도로, 간선도로, 보조간선도로	12시간	-
Beat the Traffic(미국)	고속도로, 간선도로, 보조간선도로	2시간	-

## 4) 국내외 선행연구 고찰

- 장래 교통소통상태(통행시간 등) 예측과 관련한 국내 선행연구들은 현재시각 기준 1~15분 후의 단기 예측의 수준에 머물러 있으며, 본 연구와 같이 익일의 시간대별 구간 통행속도를 예측하는 중/장기 예측에 관한 연구는 미미한 실정임

&lt;표 7&gt; 국내 선행연구 고찰 결과

구분	대상구간	교통변수	예측주기	데이터 수집	적용 방법론
강정규외 1명(2002년)	고속도로	통행시간	10분/15분	TCS <sup>1)</sup>	신경망모형
이의은외 1명(2002년)	고속도로	통행시간	5분	TCS	신경망모형
오세창외 2명(2003년)	고속도로	통행시간	1분	루프검지기 영상검지기 자석식검지기	칼만필터링
이승재외 1명(2003년)	네트워크	교통량, 통행시간	5분	영상검지기	CA모형 칼만필터링
남궁성(2005년)	고속도로	경로통행시간	Estimation	TCS	전진반복 전후방탐색법
이영인외 2명(2005년)	일반국도	통행시간	5분	영상검지기 AVI <sup>2)</sup>	칼만필터링
김재진외 2명(2006년)	고속도로	통행시간	5분	TCS	베이지안 추론
심상우외 1명(2006년)	단속류	통행시간	-	GPS	대기행렬 이론
여태동외 2명(2009년)	단속류	통행시간	-	VDS <sup>3)</sup> , AVI	히스토리컬 프로파일 모형
심상우외 3명(2013년)	고속도로	통행시간	5분	하이패스 DSRC	PBM/LBM

주1 : 고속도로 통행료수납시스템(TCS, Toll Collection System)

주2 : 주행차량자동인식(AVI, Automatic Vehicle Identification)

주3 : 교통정보수집장치(VDS, Vehicle Detection System)

- 장래 교통소통상태(통행시간 등) 예측과 관련한 국외 선행연구도 국내 선행연구와 유사하게 시간대의 길이는 1~15분이며, 장래 예측 시간대의 개수는 1~4개로 분석되어 단기 예측의 수준에 머물러 있음을 알 수 있음
- 이는 시간대의 길이를 증가하거나 장래 예측 시간대의 길이를 연장할 경우, 장래에 대한 불확실성이 증가하기 때문에 예측의 정확도는 낮아지기 때문임

&lt;표 8&gt; 국외 선행연구 고찰 결과

연번	저자	도로 유형	시간대 길이(분)	예측 시간대개수	자료수집 방법	모형
1	Innamaa (2009)	M	5	4	D	신경망
2	Jintanakul et al. (2009)	M	5	1	S	베이지안
3	van Hinsbergen et al. (2009)	M	5	3	D	신경망
4	Zou et al. (2009)	M	5	1	D	신경망
5	Innamaa (2009)	M	1	1	D	신경망
6	Yang et al. (2010)	M	15	1	A	통계기반
7	Abu-Lebdeh and Singh (2011)	A	5	1	S	Hybrid
8	Bustillos and Chiu (2011)	M	15	1	S	통계기반
9	Fei et al. (2011)	M	1	1	D	베이지안
10	Khosravi et al. (2011)	M	5	1	D	Hybrid
11	Li and Rose (2011)	M	10	6	A	신경망
12	Myung et al. (2011)	M	5	1	D/A	통계기반
13	Oh and Park (2011)	M	1	1	A	신경망
14	Simroth and Zehle (2011)	M	1	1	GPS	통계기반
15	Soriguera and Robusté (2011)	M	5	1	D/A	통계기반
16	Xia et al. (2011)	M	15	1	D	통계기반
17	Wang and Shi (2012)	M	1	1	D	베이지안
18	Mu et al. (2012)	M	1	1	S	통계기반
19	Du et al. (2012)	M/A	2	1	S	통계기반
20	Haworth and Cheng (2012)	M/A	5	1	A	통계기반
21	Khan (2012)	M	5	1	G	베이지안
22	Lu (2012)	M	5	1	D	베이지안
23	Ma et al. (2012)	M	1	1	S	통계기반
24	Qiao et al. (2012)	M	5	1	B	통계기반
25	Zheng and Van Zuylen (2012)	A	1	1	G	신경망

## 5) 개선방향 설정

- 국내외 기술현황과 이론의 고찰과 한계, 그리고 시사점 분석결과, 기존의 단기 예측 (Short-term prediction)방법을 통해서는 본 연구의 최종목표인 익일 교통소통상태 예보를 위한 장래 상태를 수용할 만한 예측력을 가지기는 어려울 것으로 판단됨
- 이는 기존의 단기예측 방법은 시스템의 낮은 성능으로 이하여 단기적인 예측 전략에 적합하게 개발되었으며, 장기 교통예보를 위한 장래의 교통상태 예측을 수행하지 않았기 때문임. 즉, 개발모형의 목표와 용도가 다름
- 그러나 최근의 급속하게 발전된 Database처리기술과 연산능력, 그리고 검색엔진의 성능을

고려할 때, 이론적인 접근법은 Data기반의 접근법이 가장 합리적이라고 판단됨. 그 이유는 예보수준에서 장래 교통상태에 대한 불확실성을 극복하기 위한 적절한 방법론이 제시되고 있지 않기 때문임

- 따라서 기존의 단기에측모형은 본 연구에 적합하지 않으며, 앞서 설명한 한계 및 시사점에 근거하여 본 연구의 모형 개발방향을 설정하면 다음과 같음
  - 첫째, 중·장기 상태의 변화를 잘 설명하기 위해서는 Link기반의 공간 정의를 Network기반으로 정의해야 함
  - 둘째, Network의 소통상태 변화는 Link의 국부적인 변동보다는 순환성(Recurrency)과 경향성(Pattern), 그리고 내재된 알려지지 않은 Network의 상태 진화에 대한 교통 현상(Traffic Behavior)인 속도와 수요(Probe 건수)의 변화 추세에 의하여 설명되어야 함
  - 셋째, Network의 소통 상태를 설명하기 위해서는 Network라는 유기체를 구성하는 개별 Link의 시간대별 변화 속성을 기반으로 Network 상태를 고려해야 함
  - 넷째, 인위적인 통계적 모델링 보다는 데이터 기반의 지식 발견(Knowledge discovery) 접근법이 본 과제의 현실적인 목표달성에 타당할 것으로 판단됨

## 다. 교통예보를 위한 기초자료 구축

### 1) 구축 개요

- 최근 IT기술의 급속한 발전에 따라 Car Navigation, 스마트폰 등 첨단기기의 보급이 보편화 되었으며, 이를 통하여 과거 수집하지 못했던 다양한 정보들을 수집할 수 있게 되었음. 즉, 개인 또는 개별 차량의 이동궤적을 확인할 수 있는 시/공간상의 연속적 동적 자료<sup>3)</sup>를 수집할 수 있음
- 이러한 자료는 과거 1세대와 2세대의 정적인 지점자료의 형태가 아닌 개인이 이동하면서 수집되는 동적인 3세대 자료라고 볼 수 있음
- 과거 2013년 한국교통연구원에서 수행한 “빅데이터를 이용한 교통정책 개발 및 활용성 증대 방안”에서는 3세대 자료는 차량의 연속적 이동을 통해 자동적으로 자료가 수집되기 때문에 링크와 링크 간의 시/공간적인 연결성을 통해서 링크별 통행속도, 이동경로 등 다양한 정보를 활용할 수 있는 측면에서 그 가치가 매우 높다고 설명하고 있음
- 따라서, 본 연구에서는 개별 차량의 이동궤적을 확인할 수 있는 Car Navigation 자료를 장래 교통예보를 위한 기초자료로서 구축하였으며, 이 자료는 국토교통부에서 제공하고 있는 ITS 표준노드링크 체계의 링크 단위로 개별 차량의 링크별 통행속도 정보를 수집하고 있음
- 여기서, 먼저 ITS 표준노드링크 기반의 개별 Car Navigation 자료 생성과정과 그 구축 결과를 제시하였으며, 기존 ITS 표준노드링크의 문제점을 보완하기 위하여 기 구축한 분석 맵의 링크 단위 기초자료 집계 방법과 그 결과를 제시하였음
- 마지막으로 구축된 Car Navigation 자료에 대한 기초자료 특성을 분석하여 해당 자료의 우수성을 제시하였으며, 본 장의 주요 내용은 아래와 같음

(1) ITS 표준노드링크 기반의 개별 Car Navigation 자료 구축

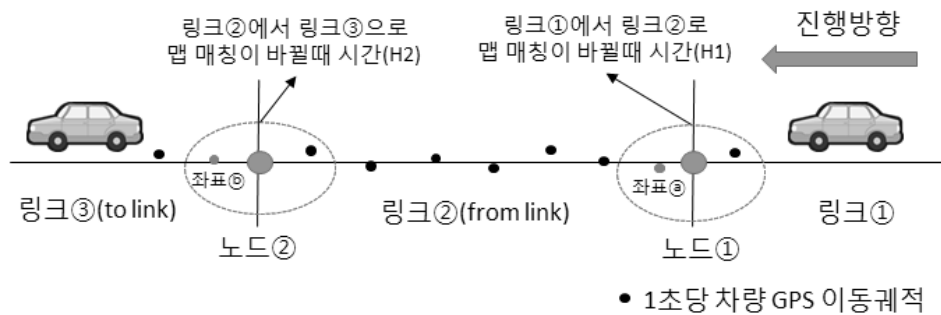
(2) 분석 맵 단위의 자료 집계

(3) 기초자료 특성분석

<sup>3)</sup> Dixon, M. P. and L. R. Rillet(2002), “Real-time OD estimation using automatic vehicle identification and traffic count data”, Journal of Computer-Aided Civil and Infrastructure Engineering, pp. 7-21.

## 2) ITS 표준노드링크 기반의 개별 Car Navigation 자료 구축

- 본 연구에서 이용하고 있는 Car Navigation 자료는 개별 차량의 Navigation 단말기에서 수집되는 차량의 이동궤적 정보로서 국토교통부에서 제공하고 있는 ITS 표준노드링크 체계의 링크 단위로 개별 Probe 차량의 이동궤적을 맵 매칭하여 각 차량에 대한 링크별 Probe 건수와 통행속도 정보를 수집하고 있음
- 여기서, 개별 차량의 이동궤적 정보로부터 링크별 Probe건수와 속도 정보로 데이터를 생성하는 과정은 <그림 8>과 같음



- 링크 ②의 Probe건수(건) = 링크 ②에 맵 매칭이 이루어지는 Probe 차량수
- 링크 ②의 속도(km/h) = (좌표㉓와 좌표㉔의 거리) / (H2 - H1)

<그림 8> 링크별 Probe건수와 속도 정보생성 개념

- 본 연구의 Car Navigation 자료는 앞서 설명한 ITS 표준링크를 기반으로 일 단위의 수집 시간 순으로 수집되며, 수집 자료의 저장 형태는 시/공간적 범위에 따라 아래와 같이 총 5 가지 Column 형태로 수집일자별 텍스트로 저장됨
- 수집 Column(총 5개) : ①auth\_key/②link\_time/③from\_link/④to\_link/⑤speed
- 공간적 범위 : 전국
- 데이터 수집기간 : 2013년 10월 1일 ~ 2014년 1월 31일, 4개월치 이력자료

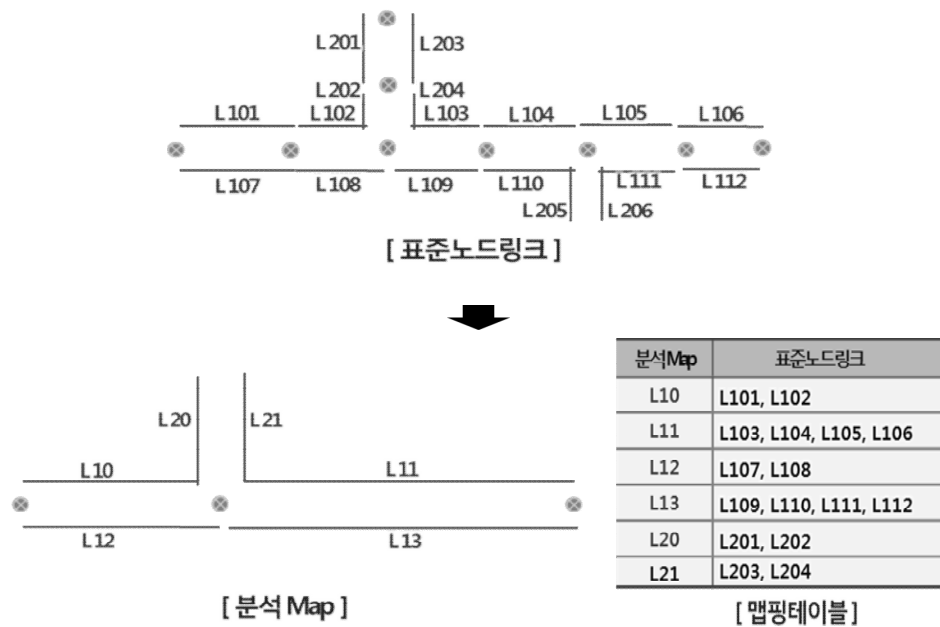
<표 9> Car Navigation 수집자료 테이블 정의서

Column	설명	Data Type	PK	FK
auth_key	단말기 ID	Datetime	-	-
link_time	수집 시간 (년/월/일/시/분/초)	Integer	-	-
from_link	진입 링크번호	Integer	-	-
to_link	진출 링크번호	Integer	-	-
speed	속도(km/h)	Integer	-	-



### 3) 분석맵 단위의 자료 집계

- ITS 표준노드링크에서는 동일한 구간이 교차로, 속성변환지점(속도, 차선수 등), 행정경계 등으로 인하여 상세하게 링크로 분할되어 있어 향후 예측정보 표출 시 동일 도로구간의 소통정보(정상 ↔ 혼잡)를 이용자들에게 효율적으로 제공하기 어려움
- 따라서, 동일 도로구간에 대한 소통정보를 이용자들이 모두 인지하기 쉽도록 동일 도로구간에서 여러 개로 분할되어 있는 링크를 병합하여 장래 소통정보를 한눈에 파악 할 수 있도록 하는 분석 맵의 구축이 필요함
- 이러한 분석 맵은 2014년 한국교통연구원이 수행한 “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”에서 기 구축이 되어 있으며, 본 연구의 분석 맵은 기 구축되어 있는 전국 교통혼잡지도의 분석 맵을 이용하였음



<그림 9> 분석 맵의 구축 개념도

자료 : “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”, 한국교통연구원, 2014

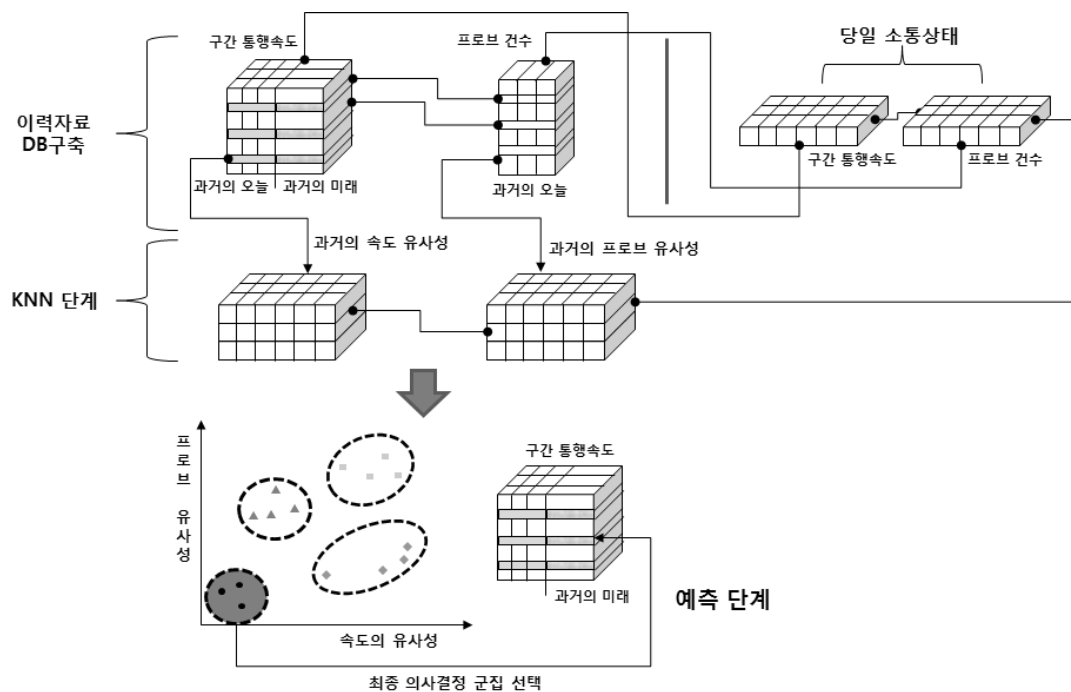
## 라. 교통예보를 위한 소통상태 예측 핵심 방법론 개발

### 1) 개발개요

- Smart phone 내비게이션 App과 차량 내비게이션 단말기를 통하여 제공되는 교통소통정보는 IT의 기술 발달과 더불어 높은 정확도를 보유하고 있으며, 수집되는 정보의 양과 공간적 규모도 국내에서 가장 큰 교통 Big data라고 표현할 수밖에 없음. 이와 같이 가용자료의 양과 정확도가 급변하는 시점에서 장래 교통상태의 예측방법론도 이에 부합하게 변화해야 함
- 따라서, 본 연구에서는 기존 단기 예측기법의 한계를 극복함으로써 익일 이상의 보다 먼 장래 통행속도를 수용 가능한 오차의 범위 내에서 예측하고, 향후 교통예보 시스템에 탑재가 용이하고 교통류의 행태와 예측 모델링에 대한 전문지식이 없이도 운영이 가능한 데이터 기반의 예측방법론을 개발하고자 함
- 본 연구의 장래 소통상태 예측 방법론은 아래와 같이 크게 2가지의 핵심 요소기술로 구분하고 각 요소기술별로 세부적인 개발모형들을 포함하고 있음
  - 핵심 요소기술① : Car Navigation 자료를 이용한 과거 이력자료 구축방법론 개발
    - 링크별/일별/시간대별 Probe건수와 속도 이력자료 DB 구축
    - ▶ Probe건수 산출 방법론 개발
    - ▶ 속도 유전자 지도 개발 및 구간 통행속도 산출 방법론 개발
  - 핵심 요소기술② : 구축된 이력자료를 이용한 장래 통행속도 예측방법론 개발
    - KNN 단계 : k-군집 추출모형 개발
    - 예측단계 : 최종 의사결정 및 통행속도예측 모형 개발
- 장래 교통예보를 위한 전체 개발 구조는 <그림 10>에서 보는 바와 같이 크게 4단계 걸쳐 개발을 수행하고자 함
  - Step 1 : Car Navigation 자료를 이용한 과거 이력자료 DB 구축(Probe건수, 속도)
    - 데이터 수집기간 : 2013년 10월 1일 ~ 2014년 1월 31일(총 4개월)
    - 대상범위 : 서울특별시 강남 일대
  - Step 2 : 장래 통행속도 예측을 위한 시·공간적 범위 설정
    - 시간적 범위 : 과거 이력자료 DB 중 예측하고자 하는 예측일(익일) 선택
    - 공간적 범위 : 과거 이력자료 DB 중 예측하고자 하는 분석 네트워크(링크번호) 선택
    - ▶ 당일의 소통 상태 : 선택된 예측일 하루 전 당일의 링크별/시간대별 소통상태 집합
    - ▶ 과거의 소통 상태 : 당일 기준보다 이전의 링크별/일별/시간대별 소통상태 집합

여기서 말하는 소통 상태는 Probe건수와 속도를 의미함

- Step 3 : k-군집 추출모형 개발(KNN 단계)
  - 과거 소통상태의 집합 중 당일 소통 상태와 유사한 k개의 군집 추출
- Step 4 : 최종 의사결정 및 통행속도예측 모형 개발(예측 단계)
  - 이전 단계의 k개의 군집 중 최종의사결정 군집 선택(유사성이 가장 높은 군집)
  - 장래 구간 통행속도 예측 : 각 구성요소들에 대한 익일의 링크별/시간대별 통행속도의 가중평균값



<그림 10> 장래 예측 방법론의 개발 구조

## 2) 과거 이력자료 구축방법론 개발

- 본 연구는 기존의 확률적인(Stochastic) 상태를 가정하는 대부분의 단기 예측을 위한 통계기반 모형이 아닌 데이터 기반의 비모수(Non-parametric) 모형을 적용하며, 비모수 모형에서 최근린 이웃은 독립변수의 상태간 거리가 가장 가까운, 즉 유사성이 가장 높은 k개의 최근린 이웃(k-nearest neighbor)의 집합을 의미함
- 이러한 모형의 예측력은 모형에 적용되는 분석 데이터의 질과 양에 의해 크게 좌우되므로 무엇보다도 정확한 과거 이력자료 구축이 필요함
- 본 연구의 과거 이력자료는 분석 맵 단위의 Car Navigation 자료(건별 이벤트 자료)를 이

용하여 Probe건수와 속도라는 두 가지 변수로 링크별/일별/시간대별로 이력자료 DB를 구축해야 함. 이를 위하여 장래 교통예보를 위한 Probe건수와 속도 이력자료 구축 방법론을 개발하고자 함

- Probe건수(건/시) : 건별 이벤트 자료 → 링크별/일별/시간대별로 누적 합계한 Probe건수
- 속도(km/h) : 건별 이벤트 자료 → 속도 유전자 지도를 이용한 링크별/일별/시간대별 가중 평균속도
- 과거 이력자료 DB의 규모는 시/공간적 범위와 Car Navigation 데이터 수집기간에 따라 각각 752,760개( $255 \times 123 \times 24 = 752,760$ )씩 Probe건수와 속도 자료가 구축되었음
- 대상 링크개수 : 총 255개(서울특별시 강남 일대)
- 구축일 개수 : 총 123개(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
- 시간대 개수 : 총 24개(01~24시)
- 구축된 링크별/일별/시간대별 Probe건수와 속도의 DB구축 현황 및 저장 형태는 아래와 같음

<표 10> 과거 이력자료 구축 현황

구분		Probe건수(건/시)			속도(km/h)		
		데이터 개수	평균	표준편차	데이터개수	평균	표준편차
연속류	경부고속도로	88,560개	60.2	25.4	88,560개	64.2	24.0
	올림픽대로	138,744개	56.8	28.7	138,744개	65.7	19.2
단속류	기타 내부도로	525,456개	31.9	11.4	525,456개	41.8	11.8
전체		752,760개	49.6	21.8	752,760개	57.2	18.3

- Probe건수(건/시) DB

구분	링크번호(총 255개)				
	174400	174500	...	3251400	3251600
2013년10월1일 00~01시	24	21	...	22	25
2013년10월1일 01~02시	18	20	...	45	17
⋮	⋮	⋮	⋮	⋮	⋮
2014년1월30일 22~23시	24	45	...	34	21
2014년1월31일 23~24시	20	30	...	45	33

- 속도(km/h) DB

구분	링크번호(총 255개)				
	174400	174500	...	3251400	3251600
2013년10월1일 00~01시	120	109	...	99	94
2013년10월1일 01~02시	110	95	...	100	92
⋮	⋮	⋮	⋮	⋮	⋮
2014년1월30일 22~23시	100	110	...	108	98
2014년1월31일 23~24시	95	90	...	114	100

<그림 11> 과거 이력자료 DB 구축 저장 형태

### 3) 장래 통행속도 예측방법론 개발

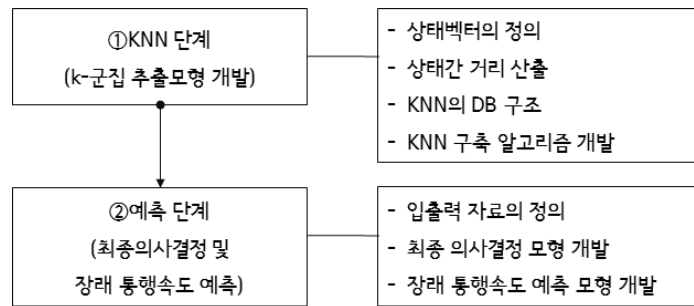
#### ① 예측방법론 개발 개요

- 본 연구는 기존의 단기 예측기법에 대한 한계를 극복함으로써 보다 먼 장래의 통행속도를 수용 가능한 오차의 범위 내에서 예측하고, 교통류의 행태와 예측 모델링에 대한 전문적인 지식이 없이도 운영이 가능한 데이터 기반의 예측방법론을 개발하고자 함
- 이러한 예측모형은 통계기반의 모형에 데이터를 맞추던 기존의 방식을 벗어나 개발모형이 데이터베이스(DB)를 탐색함으로써 선별적으로 데이터를 추출하며, 추출된 데이터를 지식발견(Knowledge discovery, 이하 KD)과 의사결정의 과정을 통하여 장래 통행속도를 예측하는 지식 탐색형 모형을 의미함
- 즉, 본 연구의 개발모형은 데이터 탐색을 기반으로 하는 예측 관련 선행 연구의 고찰을 통하여 연산 수행속도가 빠르면서 모형의 예측력이 우수한 비모수(non-parametric) 기반의 k-최근린 이웃 기법(k-nearest neighbor)을 선정하였음
- 이 기법의 파라미터인 k값에 대한 전역(Global) 최적화의 단점을 극복하기 위하여 모형의 수행 과정에서 국부(Local) 파라미터 최적화 과정을 통하여 기존의 전역 최적화의 문제점을 극복하고자 함. 또한, 국부(Local) k값의 최적화는 연산 속도가 빠르면서 효과적으로 군집을 구성할 수 있는 안정적인 군집화 통하여 달성하고자 함
- 먼저, 본 연구의 장래 통행속도 예측모형을 개발하기 위한 전제조건은 예측하고자 하는 분석 네트워크의 링크별 통행속도는 해당 네트워크의 교통량(즉, Probe 건수)과 통행속도의 관계에 의하여 정의됨
  - 장래 링크별 통행속도 =  $F_n\{\text{링크별 Probe 건수, 링크별 속도}\}$
- 위의 전제조건에 대한 상태벡터의 차원은 링크로 구성되는 네트워크로 정의되며, 각 상태벡터는 네트워크 구간별 프로브 대수와 네트워크 구간별 속도로 정의됨. 그리고 현행 상태벡터와 과거 이력 상태벡터 간의 유사성은 유클리디언 거리(Euclidean Distance)로 정의함
- 즉, Probe건수(네트워크 수요)의 유사성과 속도의 유사성을 동시에 고려하여 장래 네트워크의 링크별 통행속도를 예측하며, Probe건수와 속도의 유사성 높다는 것은 두 변수에 대한 현행 상태벡터와 과거 이력 상태벡터 간의 유클리디언 거리(Euclidean Distance)가 가깝다는 것을 의미함
  - 입력상태 벡터(Input State Vector) : 네트워크 상태벡터

- 속도 상태벡터 : 시간대 개수 (24개) \* 링크별 속도
- Probe 상태벡터 : 시간대 개수 (24개) \* 링크별 Probe건수
- 출력상태 벡터(Output State Vector)
  - 속도 상태벡터 유사성
  - Probe 상태벡터 유사성
  - 익일의 링크별/시간대별 통행속도
- 또한, 네트워크 상태벡터는 연결로를 제외한 본선만을 이용하며, 이는 연결로에서 발생하는 국부적 Variation에 의한 영향을 배제함과 동시에 전체 네트워크의 전역(Global) 변화를 탐색하고자 함

## ② 본 연구의 예측 모형 구조

- 본 연구의 예측모형은 과거 이력자료 DB(Probe건수, 속도)를 이용하여 경험적 지식을 기반으로 하는 상향식(Bottom-up)체계 의사결정모형으로서 아래와 같이 ①KNN 단계, ②예측 단계로 구성됨
  - KNN 단계는 과거 이력자료 DB(Probe건수, 속도)에서 현행 교통상태와 유사한 과거 교통상태를 탐색·추출하여 장래 교통상태를 설명하기 위한 의사결정집단인  $k$ 개의 최근린 이웃 ( $k$ -nearest neighbor, 이하 KNN)을 구성하는 일종의 정보추출 및 구축 모형으로 이용됨. 이상의 과정을 달성하기 위하여 ①입·출력 상태벡터(State vector), ②현행 상태와 과거 상태의 유사성(Similarity)을 설명하기 위한 상태거리(State metric), ③탐색된 유사성이 높은 과거 상태를 추출하고 KNN을 구축하기 위한 자료구조(Data structure)가 설정됨. 설정된 3개 구성요소는 KNN 정보추출 및 구축 모형을 구성하게 되며, 개발된 KNN구축 알고리즘으로 구현함
  - 예측 단계는 이전 단계에서 구축된  $k$ 개의 군집 집단들 중 최종 의사결정 집단을 결정하고, 최종 의사결정 집단을 구성하는  $m$ 개 구성원( $m < k, m > 1$ )의 장래 링크별 소통 상태를 결정하고 구간 통행속도를 예측하는 단계임. 해당 단계는 최종 의사결정 집단 선택모형과 장래 구간통행속도 예측모형으로 구성되며, 최종의사결정 집단 선택모형은 예측 단계에서 선택된 군집들 중 현행 상태와 가장 유사한 군집부터 군집순위를 정하여 현행 상태와 가장 유사한 군집 집단을 선택함. 장래 구간통행속도 예측모형은 최종적으로 선택된 군집 집단의  $m$ 개의 구성원에 해당하는 과거 상태와 KNN단계에서 산정한 개별 구성원의 상태 거리를 기반으로 하는 유사성을 이용하여 장래 통행속도를 예측함

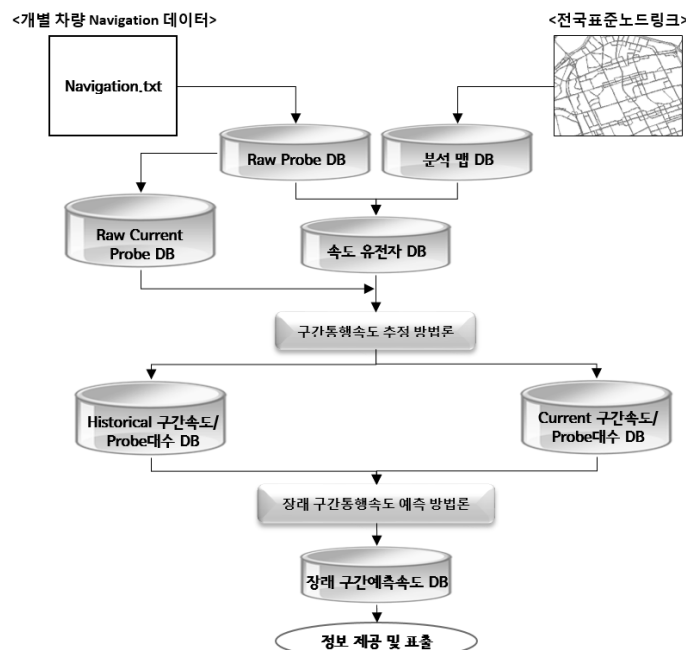


<그림 12> 본 연구의 예측모형 구조

## 마. 교통예보를 위한 시스템 구축 방안

### 1) 시스템 DB 구조 설계

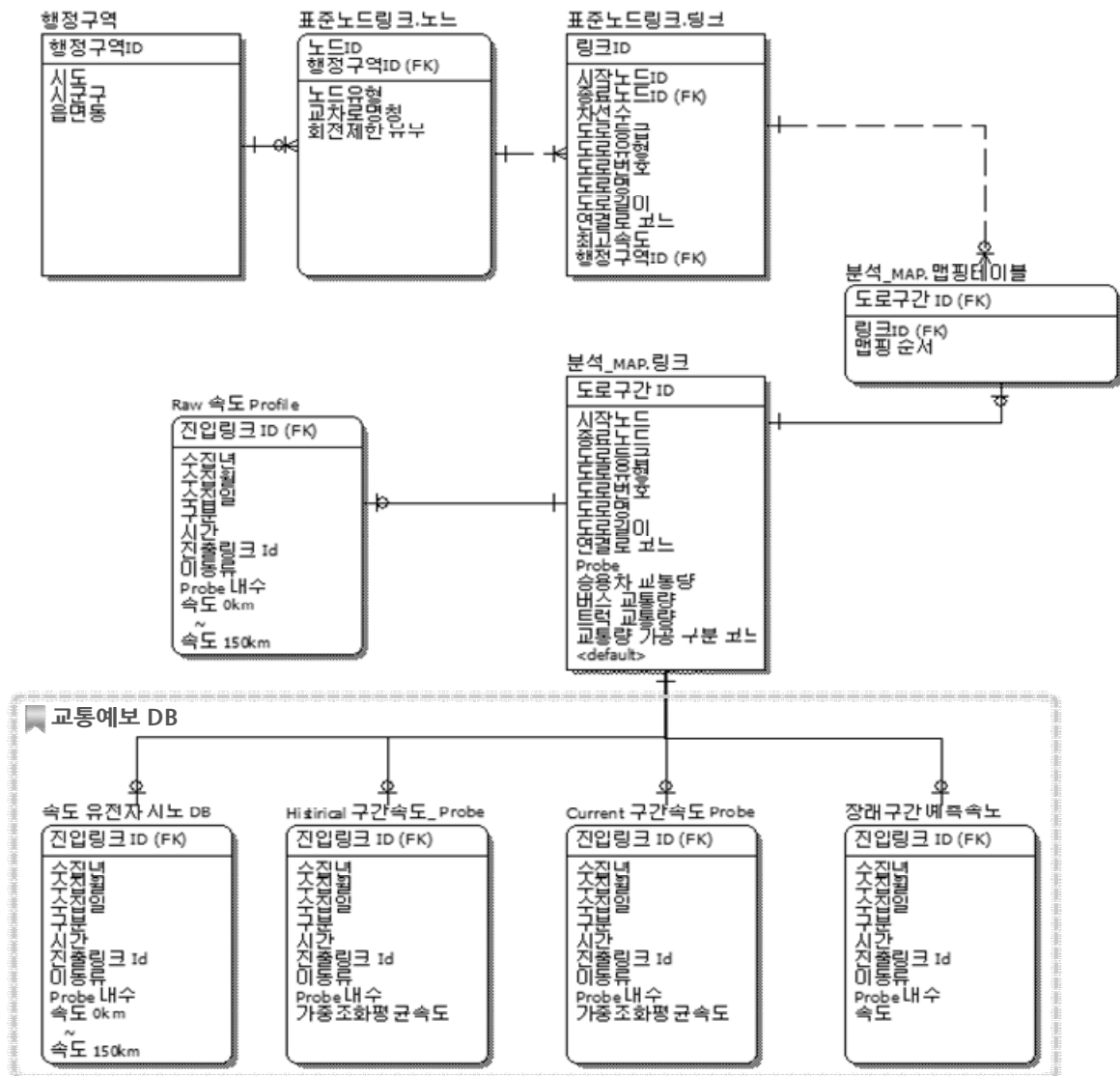
- 본 연구의 DB 구조는 먼저 개별 차량의 Navigation에서 추출되는 이동계적 데이터(속도, Probe대수)를 Raw Probe DB로 구축하고, 입력 및 결과 정보를 표출하기 위한 분석 Map DB를 구축함
- 위의 Raw Probe DB와 분석 Map DB를 이용하여 구축하는 속도 유전자 DB를 기반으로 Historical DB(이력자료 DB)와 Current DB(현재자료 DB)를 구축하고, 이 두 DB를 이용하여 장래 구간별 통행속도를 예측하고 그 결과를 DB로 구축함



<그림 13> DB구조 설계 개념도

## 2) 시스템 ERD 설계

- 교통예보의 주요 DB구성은 Navigation 차량 수집 데이터, 네트워크(표준노드링크, 분석맵), 속도 유전자 지도 DB 등으로 구성되며, 교통예보 시스템 ERD(Entity Relation Diagram)는 아래와 같음



&lt;그림 14&gt; 교통예보 시스템 ERD



### 3) 개발 방법론을 적용하기 위한 시스템 요구사항

- 본 연구에서 제안한 기술은 기존의 ITS에서 적용해 오면 현행자료기반의 단기 예측기법에서 벗어나 방대한 양의 Raw data와 1/2차 가공된 DB를 관리/저장해야 함. 따라서 고성능의 검색 및 연산능력을 가진 시스템의 구축이 필요할 것을 판단됨
- 따라서 본 연구에서 개발되는 기술을 적용하기 시스템 요구사항으로서 ①본 연구의 장래 예측 기술에 요구되는 방대한 양의 DB 크기, ②정보제공 시스템의 정보제공 요구시간에 적합하게 예측을 수행할 수 있는 예측시스템의 성능, ③다중 접속자에 의한 정보제공시스템의 처리능력 등 다양한 시스템 요구사항을 분석하고 제시하도록 함
- 시스템 요구사항을 정의하고 세부 요구사항을 제시하는 다음의 5개 사항을 고려하여 수행하도록 함
  - 첫째, 개념적 시스템 설계를 통하여 실제 구축되는 시스템에서 요구되는 요구사항을 정의하고, HW/SW 그리고 운영인력에 대한 요구사항을 정의하도록 함
  - 둘째, 개발 모형의 실제 운영이 가능한 자료의 수집범위와 크기에 대한 정의를 통하여 지속적인 자료의 수집이 가능하도록 요구사항을 제시하도록 함
  - 셋째, Big data 기반의 ITS 시스템 전문가 집단의 검토를 통하여 실제 시스템의 구축, 개발 모형의 탑재 시 DB의 구조, DB검색과 추출을 위한 데이터 추론기술 등 전반적인 시스템의 HW와 SW의 최적화에 대한 요구사항을 제시하도록 함
  - 넷째, ITS 표준노드링크의 링크개수는 약 250,000로서 본 연구에서 개발된 방법론은 국내에서 가장 큰 Network 사이즈와 개수 그리고 데이터의 크기를 처리해야 함. 이는 용이하게 Big data를 처리할 수 있는 기존의 슈퍼컴퓨터보다 매우 뛰어난 성능을 보유한 실시간 Hadoop을 도입하고 이를 지원할 수 있는 시스템을 설계해야 함을 의미함
  - 다섯째, 전국을 대상으로 장래 혼잡예보 정보를 다양한 요구수준을 가진 국민에게 제공하기 위해서는 정보제공을 위한 소통정보 제공용 맵과 효율적인 정보제공 콘텐츠가 필요함. 따라서 생성된 장래정보를 효과적으로 제공하기 위한 서비스 방안을 도출함과 더불어 서비스 콘텐츠 개발에 필요한 요구사항을 제시하도록 함

## 바. 사례분석 및 평가

### 1) 평가 개요

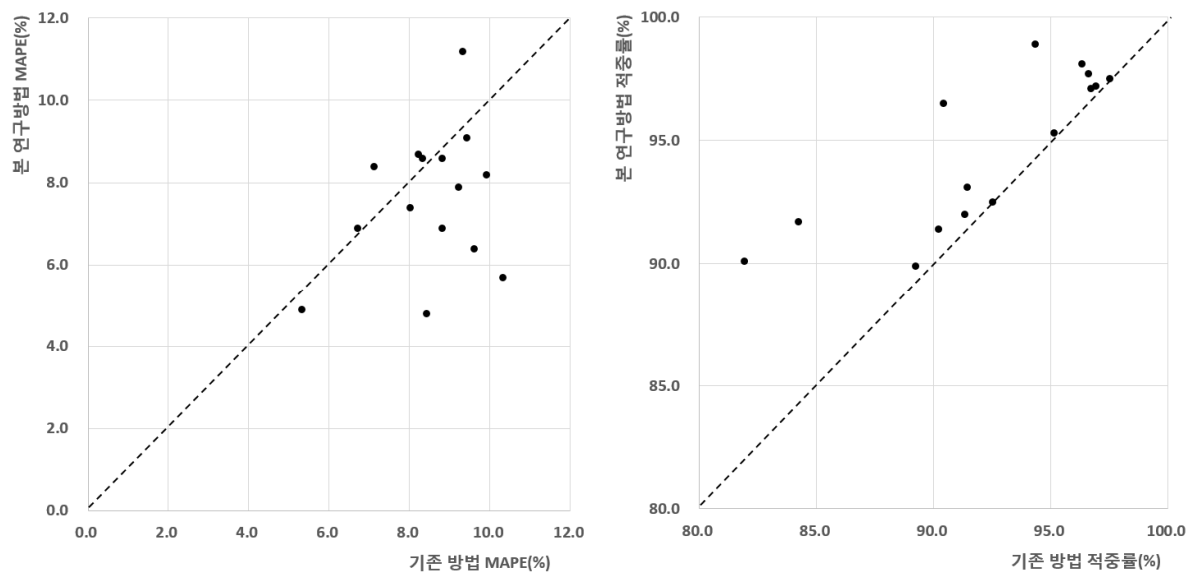
- 본 장에서는 앞서 구축한 링크별/일별/시간대별 과거 이력자료 DB(Probe건수, 속도)를 이용하여 개발된 장래 통행속도 예측모형의 성능 평가를 수행하고자 함
- 성능 평가의 수행과정은 ①사례대상 구간 선정, ②사례대상 구간의 분석DB 구축, ③분석 자료의 특성, ④평가지표 및 분석시나리오 설정, ⑤방법론 적용 및 평가 그리고 소결로 구성됨
- 먼저, 사례대상 구간 선정은 연속류와 단속류가 혼재되어 있는 서울시 강남구 일대의 전체 도로축을 대상구간으로 선정하였음
  - 연속류 구간 : 경부 고속도로(한남대교 ~ 판교JC), 올림픽 대로(원효대교 ~ 올림픽대교)
  - 단속류 구간 : 강남 일대 내부 14개 도로축
- 분석DB 구축은 본 연구의 사례대상 구간에 해당하는 분석 맵의 링크 기준으로 링크별/일별/시간대별 Probe건수와 속도 자료를 구축하였음
  - 대상링크 개수 : 총 255개
  - 구축일 개수 : 총 123개(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
  - 시간대 개수 : 총 24개(01~24시)
- 분석 자료의 특성에서는 링크별/일별/시간대별로 구축되어 있는 Probe건수와 속도 자료에 대한 Historical 데이터 분석을 수행하여 각 자료에 대한 시계열적 Variation을 검토하였음
- 사례분석 및 평가를 위한 지표는 백분율 절대평균오차(MAPE, %: Mean absolute percentage error)와 적중률(%)을 적용하였으며, 분석 시나리오는 예측하고자 하는 일자에 따라 경부고속도로, 올림픽대로, 기타 내부도로로 분석 네트워크로 구분하여 설정하였음. 여기서, 본 예측모형의 평가를 위한 비교모형으로는 현재 ITS시스템에서 많이 이용되고 있는 이력자료 평균(Historical average) 기법을 적용하였음
- 방법론 적용 및 평가에서는 설정된 평가지표를 이용하여 분석 시나리오별로 관측치와의 오차 분석을 수행하여 이를 종합적으로 평가하였음. 마지막으로 소결에서 본 개발모형의 목표 달성과 적용의 측면을 앞서 평가한 결과들을 토대로 설명하였음

## 2) 종합평가 결과

- 본 연구에서는 앞서 설명한 장래 예측일과 분석 네트워크 설정에 따라 총 5가지의 분석 시나리오를 설정하였으며, 각 시나리오별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 그 결과, 전체 분석 시나리오와 모든 네트워크에서 본 연구의 예측모형이 기존 과거평균기법보다 MAPE는 0.6~1.4%, 적중률은 0.7~4.0%의 예측력을 향상시키는 것으로 분석되었음. 이는 앞서 4절에서 설명한 바와 같이 과거 속도자료의 패턴이 대체로 일정한 비첨두시간대(00~06시)와 혼잡 시간대(07~09시, 18~20시)를 제외한 나머지 속도의 변화가 큰 시간대에서 본 모형의 예측력이 높은 것으로 판단됨
  - MAPE 1.4%의 향상은 기존 방법 보다 본 연구의 예측모형 오차가 평균 1.4% 감소했다는 의미하며, 적중률 4.0%의 향상은 기존 방법 보다 본 연구의 예측모형이 허용오차 범위( $\pm 10\text{km/h}$ )에 속하는 표본이 86개(24시간 $\times$ 255개 링크 $\times$ 4%)가 더 증가된 것을 의미함
- 시나리오 3~5의 경우, 일부 분석 네트워크에서 기존 과거평균기법의 예측력이 본 연구의 예측모형과 비슷하거나 오히려 다소 향상된 결과가 산출된 것은 해당 분석 네트워크의 과거 일별/시간대별 속도자료 분포의 Variation이 크지 않고 전체 시간대에서 일정한 패턴을 유지하는 데에서 기인된 것으로 판단됨
- 위의 종합 평가결과에 대하여 먼저, 과거 속도자료의 일별/시간대별 패턴변화가 크지 않고 일정한 경우(일별/시간대별 패턴의 주기성이 있는 경우)에는 기존 방법과 본 연구방법의 예측 결과에 대한 차이는 크지 않으며, 그 반대로 과거 속도자료의 일별/시간대별 패턴변화가 크고 일정하지 않은 경우(일별/시간대별 패턴의 주기성이 없는 경우)에는 기존 방법보다 본 연구방법의 예측력이 더 우수한 것으로 분석되었음
- 이는 기존의 과거평균방법이 본 연구방법에 비해 시계열적으로 변화하는 통행속도의 차이에 민감하게 반응하지 못하며, 또한, 기존의 개별 링크가 아닌 네트워크의 차원에서 장래 상태를 설명하기 때문에 장래 교통상태의 불확실성을 어느 정도 감소시킬 수 있었기 때문으로 판단됨

&lt;표 11&gt; 종합평가 결과

구분	평가지표	연속류 네트워크						단속류 네트워크		
		경부고속도로			울림픽대로			기타 내부도로		
		기존	본연구	증감	기존	본연구	증감	기존	본연구	증감
시나리오1	MAPE(%)	10.3	5.7	-4.6	5.3	4.9	-0.4	8.8	8.6	-0.2
	적중률(%)	90.4	96.5	6.1	95.1	95.3	0.2	96.9	97.2	0.3
시나리오2	MAPE(%)	8.8	6.9	-1.9	9.6	6.4	-3.2	9.9	8.2	-1.7
	적중률(%)	84.2	91.7	7.5	81.9	90.1	8.2	96.3	98.1	1.8
시나리오3	MAPE(%)	8.4	4.8	-3.6	8.0	7.4	-0.6	8.2	8.7	0.5
	적중률(%)	94.3	98.9	4.6	91.3	92.0	0.7	97.5	97.5	-
시나리오4	MAPE(%)	9.3	11.2	1.9	8.3	8.6	0.3	9.4	9.1	-0.3
	적중률(%)	92.5	92.5	-	90.2	91.4	1.2	96.7	97.1	0.4
시나리오5	MAPE(%)	7.1	8.4	1.3	6.7	6.9	0.2	9.2	7.9	-1.3
	적중률(%)	91.4	93.1	1.7	89.2	89.9	0.7	96.6	97.7	1.1
전체평균	MAPE(%)	8.8	7.4	-1.4	7.6	6.8	-0.7	9.1	8.5	-0.6
	적중률(%)	90.6	94.5	4.0	89.5	91.7	2.2	96.8	97.5	0.7



&lt;그림 15&gt; 기존 방법과의 MAPE(%)와 적중률(%) 비교

## 사. 결론 및 향후 연구

### 1) 결론

- 본 연구는 개별 Car Navigation의 차량이동계적 데이터를 이용하여 장래 교통예보를 위한 핵심 요소기술(①과거 이력자료 DB구축 방법론 개발, ②장래 통행속도 예측방법론 개발)을 개발하였으며, 개발모형의 평가 결과는 보편적으로 이용되고 있는 기존의 이력자료 평균(Historical average)방법 보다 전체적으로 우수한 예측력을 보였음
- 또한, 본 연구의 핵심 요소기술 개선방향 및 실용화를 위하여 향후 개념적인 시스템 구축방안도 제시하였음
- 본 연구의 빅데이터(Car Navigation의 차량이동계적 자료) 기반 교통예보(익일 링크별/시간대별 통행속도 예측)를 위한 핵심 요소기술 개발에 대한 결론을 요약하면 아래와 같음
  - 첫째, 기존의 단기예측(일중 소통상태 예측)의 예측영역을 중장기 예측영역(익일 소통상태 예측)으로 확장하였으며, 이를 위하여 통계기반의 인위적 모델링에 의존하지 않고 과거 데이터를 탐색하여 내재된 정보를 도출하는 데이터 탐색 기반의 예측모형을 적용하였음
  - 둘째, Network의 소통상태 변화는 Link의 국부적인 변동보다는 네트워크 차원에서의 순환성(Recurrency)과 경향성(Pattern), 그리고 내재된 알려지지 않은 Network의 상태 진화에 대한 교통 현상(Traffic Behavior)인 속도와 수요(Probe 건수)의 변화 추세에 의하여 영향을 받는다는 것을 알 수가 있었음
  - 셋째, 모형의 정확도가 요구수준을 만족하면 모형의 연산 수행속도와 더불어 운영상의 용이성을 고려해야 함. 개발 모형의 예측력이 뛰어나더라도 시스템 측면에서 연산 수행시간이 추정정보를 활용할 수 있는 시간적 범위를 초과하거나, 운영측면에서 지속적으로 모형의 파라미터를 조정해야 한다면 실제 모형의 적용성은 현저히 떨어지게 됨. 따라서 본 연구에서는 파라미터의 조정이 시스템적으로 간편하면서 모형의 연산속도가 빠른 모형으로 개발하였음
  - 넷째, 현재 보다 방대한 양의 Car Navigation 자료를 이용할 수 있을 경우, 익일 이상의 주간 단위(1주일)의 교통소통 상태까지도 예측할 수 있는 가능성을 제시하였음
  - 다섯째, 본 연구에서 개발한 예측모형은 향후 교통예보 시스템 탑재를 목표로 하기 때문에 입력 자료의 요구조건, 파라미터의 정산, 예측모형 적용 등에 대하여 관련 전문가가 아닌 일반 현장 요원도 시스템을 운영함에 있어 큰 어려움이 없도록 개발하였음

## 2) 향후 연구

- 본 연구에서는 지난 4개월 간(2013년 10월 1일 ~ 2014년 1월 31일)의 Car Navigation 자료만을 이용하여 장래 구간 통행속도 예측모형을 개발하였음. 즉, 구축된 과거 입력 자료의 양이 많지 않으므로 예측 모형의 k군집 추출 단계에서 충분한 표본 수를 확보하지 못하였음
- 향후 개발 모형의 예측력을 더욱 높이기 위해서는 다 년간 축적된 Car Navigation 자료의 이용이 필요하며, 이를 위하여 실시간 Car Navigation 이동궤적 정보를 수집 및 저장할 수 있는 교통예보시스템 개발이 요구되어짐
- 즉, 개발모형에 대한 장래 소통상태 예측기법 및 현행 상태를 재구성하기 위한 다양한 형태의 접근방법을 적용하여 모형을 고도화 할 수 있을 것으로 판단됨
- 본 연구의 교통예보는 익일의 교통 소통상태만을 예측하지만, 향후 입력 자료가 충분히 확보된다면 익일 이상의 주간 교통예보까지도 예측이 가능할 것으로 판단됨
- 마지막으로 익일 이상의 중장기 교통 소통 상태를 예측할 때 발생하는 오차는 특히 교통사고, 공사 등과 같은 이벤트 상황 즉, 비 반복적 행태의 소통상태가 발생 시 이를 고려하여 장래 소통 상태를 예측하는 연구가 지속적으로 이루어져야 된다고 봄

### 3. 교통정보DB 통합관리 시스템 설계

#### 가. 추진배경 및 경위

##### 1) 추진배경

- 국회 등에서 KTDB를 활용한 예측 교통량과 관측 교통량간의 오차가 발생함에 따라 KTDB의 신뢰성 제고 필요성 지적
- 국가교통통계 생성을 위한 교통정보DB의 연계·공유가 필요
- 이에 공공기관·지자체 등에서 수집하는 교통량·속도 정보 등을 연계·공유하여 교통정보 활용을 극대화하는 교통정보 통합관리 방안 마련('14.9.30), 단계적 추진
  - 1단계 국토부·공공기관 교통정보, 2단계 지자체 교통정보, 3단계 민간 교통정보 통합관리
- 공공기관 교통정보 연계를 현행 한국교통연구원·한국도로공사에서 한국건설기술연구원·교통안전공단까지 추가 확대·운영
  - 국토부·한국교통연구원(KTDB센터), 한국도로공사(국가교통정보센터) 교통정보DB 협의회 구성·운영 중('14.7)

##### 2) 추진경위

- 2014. 6. 국가교통DB센터와 국가교통정보센터 연계 협력 회의
  - 양 센터간 인터넷 전용선 설치 등 수집정보 연계·공유 방안 논의
- 2014. 7. 교통정보 연계·공유를 위한 교통정보DB 협의회 구성
  - 국토부(교통정책조정과, 첨단도로환경과), 한국교통연구원(국가교통DB센터), 한국도로공사(국가교통정보센터)로 구성
- 2014. 8. 제1차 교통정보DB 협의회 회의 개최
- 2014. 9. 국가교통DB 신뢰도 제고를 위한 교통정보 통합관리 방안 수립
- 2014.11. 교통정보DB 협의회 확대를 위한 실무협의회 회의 개최
  - 국토부(도로운영과·신교통개발과·교통안전복지과), 한국건설기술연구원(도로교통량정보회사업단), 교통안전공단(정보운영처, 교통안전처) 참석
- 2014.12. 제2차 교통정보DB 협의회 회의 개최

- 공공기관 교통정보 연계·공유 추진현황 및 부문별 향후계획 제시

## 나. 교통정보 공유현황

### 1) 주요기관 DB 종류 및 형태

#### ① 국토교통부(국가교통정보센터)

- 고속국도·국도·지방도·시/군도 등 교통시설에 대한 지능형교통체계 구축·운영시 필요한 기본적인 교통정보를 교환·연계중에 있음
  - 기본교통정보는 아래와 같이 9개로 구성되어 있으며, 이들 중 KTDB 신뢰도 제고를 위한 관측교통량 분석 및 교통분석 효율성 제고 등을 위해서는 ‘교통소통 정보’의 연계 필요
  - 실제로 ‘교통소통정보’를 중심으로 제공되고 있으며, 이는 KTDB와 기 연계 중에 있음

#### ② 한국건설기술연구원(교통량정보시스템)

- 도로법 제87조 (도로에 관한 조사)에 따라 일반국도의 통과 교통량 자료 등을 정확하게 파악하기 위해 매년 상시조사와 수시조사를 수행하고, 조사된 자료를 통계 처리하고 도로교통량 통계연보를 작성하고 있음

#### ③ 교통안전공단 DTG(운행기록분석시스템)

- 교통안전법에 근거하여 전체 사업용 차량에 대하여 디지털운행기록계 장착이 의무화 되었으며, 교통안전공단에서 통합 시스템을 구축하여 주행궤적 자료를 수집중에 있음(매10초 간격)

#### ④ 교통안전공단 TAGO(국가대중교통정보센터)

- 버스, 철도, 항공, 해운에 대하여 노선 및 운행정보 등에 관한 DB를 구축

#### ⑤ 기타기관

- 한국철도공사, 지방도시철도공사, 전국고속버스운송사업조합, 한국공항공사, 인천국제공항공사, 한국해운조합, 해양수산부 등 교통관련 기관의 운행현황 및 수송실적 통계



## 2) KTDB 자료 수집 현황 및 공유·연계 방안

## ① KTDB 자료 수집 현황

&lt;표 12&gt; KTDB 신뢰도 제고 및 통계DB 구축을 위한 각 기관별 자료 수집 현황

수집목적	자료명	기관	요청 주기	자료 주요내용
O/D 구축 및 검증	일반철도 출발·도착지 간 여객·화물 수송실적	한국철도공사	매분기	- 여객: 주운행선, 열차종류, 출발역, 도착역, 인원 - 화물: 주운행선, 품목분류, 출발역, 도착역, 수송톤수
	지하철 출발·도착지 간 수송실적	한국철도공사, 지방도시철도공사	매분기	- 출발역코드, 출발역명, 도착역코드, 도착역명, 인원, 거리
	고속버스 터미널 간 여객 수송실적	고속버스조합	매분기	- 노선, 거리, 버스등급, 총 운행횟수, 일평균 회수, 총 인원, 일 평균 인원, 공급 좌석수, 이용률, 인-km
	공항 간 항공 여객·화물 수송실적	한국공항공사, 인천국제공항공사	매분기	- 국내선: 연월, 출발공항코드, 도착공항코드, 운행횟수, 여객수, 여객 인-km, 화물톤수, 우편물톤수, 화물 톤-km, 우편물 톤-km - 국제선: 연월, 노선, 국가, 도시, 운행횟수, 여객수
	해운 여객·화물 수송실적	한국해운조합 (국내)	매분기	- 계획회수, 결항회수, 여객수송능력, 수송인원, 인-km, 차량대수, 화물톤수, 화물 톤-km, 여객선수입
		해양수산부(국제)	매분기	- 노선, 여객수(입항, 출항), 여객 인-km, 차량대수, 일반화물 톤수, 컨테이너 수송실적
코드/스크린 라인 검증	고속도로 TCS 자료	한국도로공사	매년	- 평일/주말 요금소간 통행량
	교통량 자료	한국건설 기술연구원	매년	- 도로위계별 지점별 교통량, 교통량 조사지점
교통네트워 크 구축	도로/철도 시설 현황	지자체 한국철도공사	매년	- 도로 : 준공도로 현황, 차선수, 연장 등 - 철도 : 운행시각표, 역간거리, 운영현황 등
	대중교통 노선 현황	고속버스조합, 지자체 등	매년	- 버스노선번호, 노선유형, 시·종점, 운행횟수, 순번별 정차정류장 등
장래 O/D, 네트워크 구축	토지이용계획	지자체, LH공사, 지역별 도시공사	매년	- 택지개발사업, 주택건설사업, 도시개발사업, 산업단지개발사업, 도시재정비촉진사업, 주거환경개선사업, 주택재개발사업, 주택재건축사업
	교통시설계획	지자체, 한국철도시설공단, 지방도시철도공사	매년	- 개발계획 연장, 준공연도, 노선도, 역간거리, 운영계획

## ② 자료 공유·연계 방안

- 노선 및 운행정보 등 동적자료: 교통량·속도 및 노선별 등 자료용량이 방대하고 실시간성이 강한 자료의 경우 KTDB와 실시간 연계를 통한 통합관리 필요
- 수송 실적 등 정적자료: 자료용량이 비교적 작으며, 일년에 몇 번 변하지 않는 자료에 대해서는 현재와 같은 수집체계를 유지하며, 자료공유를 통한 통합관리 필요

## 3) 자료 공유·연계에 따른 기대효과

### ① 국토교통부(국가교통정보센터) 및 교통안전공단 TAGO(국가대중교통정보센터)

- 교통량, 속도, 수송실적 등의 교통정보를 활용하여 O/D 총량 검증 및 네트워크 성능 검증 등 KTDB 신뢰도 제고
- 교통정보 연계확보를 통해 조사비용 절감 및 세밀화된 대중교통 정보를 이용한 대중교통 수요분석의 효율성 및 정확성 증대
- 내비게이션 자료를 이용한 교통망에 차량궤적정보를 통합하여 보다 현실적인 통행패턴 분석 가능
- 대국민 교통정보 서비스 및 정책 의사결정 지원 콘텐츠 개발

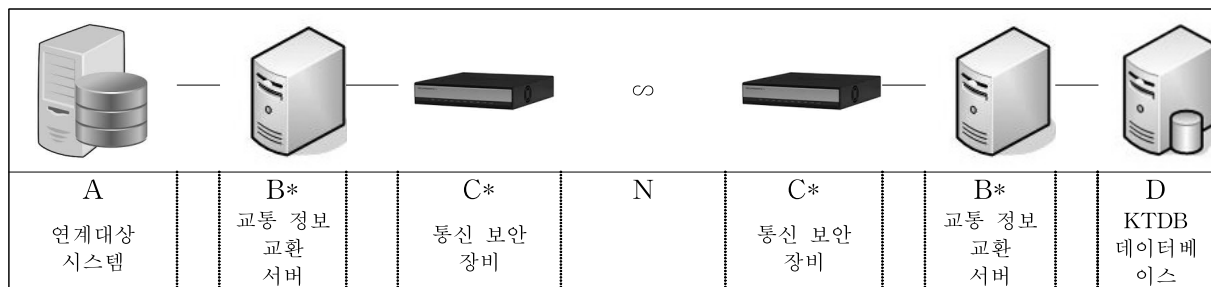
### ② 교통안전공단 DTG(운행기록분석시스템)

- 여객부분 수송실적 산출 (택시 및 전세버스)
- 화물부분 O/D 정확성 검증 및 물류지도 작성
- 문제점 : 낮은 자료 수집율

## 다. 교통정보 통합관리 시스템 구축 방안

### 1) 교통정보 통합관리 시스템 구축

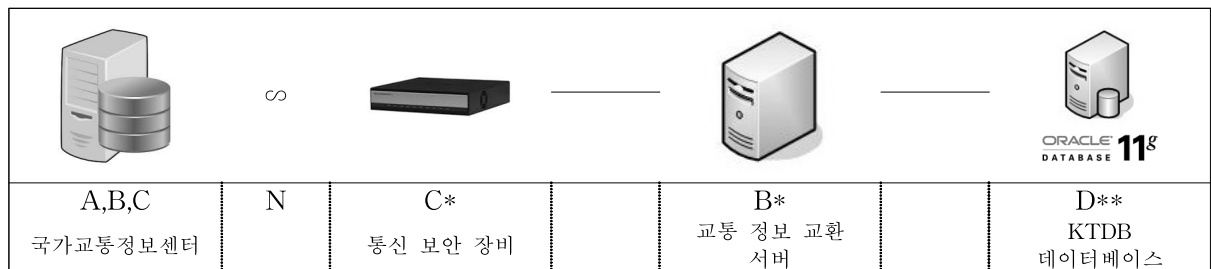
#### ① 일반적인 데이터 연계 시스템



\*: 신규 도입 장비

<그림 16> 시스템 구성 현황

#### ② 국가교통정보센터 시스템 연계 및 구축

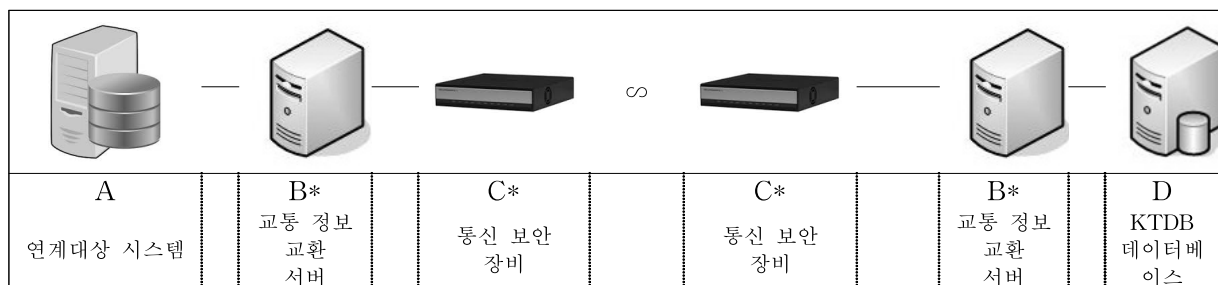


\*: 2014년 도입 완료

\*\* : 하드웨어는 기존 장비 재활용. DB SW는 2014년 도입 완료

<그림 17> 시스템 구성 현황

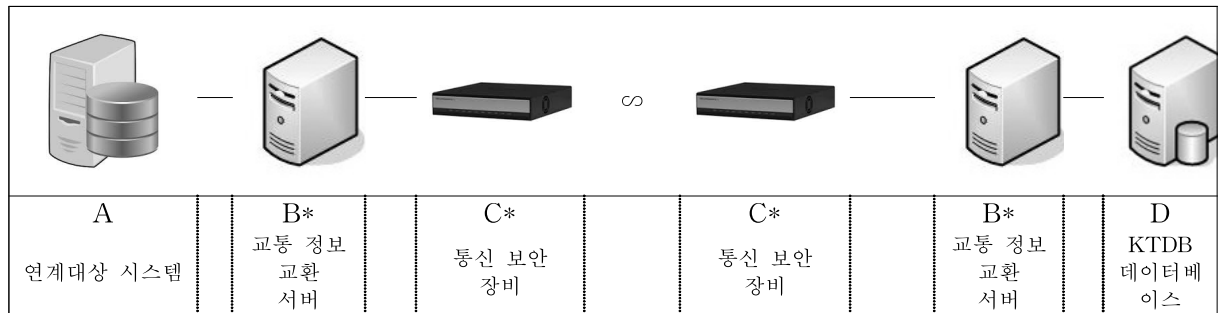
#### ③ 교통안전공단 TAGO 시스템 연계 및 구축



\*: 신규 도입 장비

<그림 18> 시스템 구성 현황

④交通安全공단 DTG



\*: 신규 도입 장비

<그림 19> 시스템 구성 현황

라. 지자체 교통정보 현황

1) 지자체 교통정보 수집 및 관리현황

- 지자체 교통정보 수집은 크게 ITS를 통한 상시조사와 지자체 자체 예산으로 실시되는 수시조사로 구분됨
- 상시조사로 도심 내 도로의 교통정보가 실시간으로 수집되어 지자체 교통정보정보센터 등에서 관리하고 현재 국가교통정보센터에 연계 및 저장되고 있음
- 수시조사는 도심 내 도로와 지방도 및 국지도를 대상으로 지자체가 매년 일년에 한번 조사하고 있으며 각각의 연계 및 관리체계는 상이함

&lt;표 13&gt; 인구 30만 이상 지자체 교통정보 수집 및 연계현황

인구	도시		교통량				속도			
			ITS		현장조사 <sup>4)</sup>		ITS		현장조사	
			수집 여부	관리 및 연계	수집 여부	관리 및 연계	수집 여부	관리 및 연계	수집 여부	관리 및 연계
1,000만 이상	서울특별시		○		x		○		x	
100만 이상	울산광역시		○		○		○		○	
	대전광역시		○		○		○		○	
	부산광역시		x		○		○		○	
	대구광역시		x		○		○		○	
	광주광역시		○		○		○		○	
	인천광역시		○		○		○		x	
	경기도	수원시	○		x		x		x	
	경남	창원시	x		x	해당 지자체 관리,	x		x	
	경기도	고양시	○		○		○		x	
50~ 100만	경기도	성남시	○	지자체 교통정보 센터 및 국가교통 정보센터	x	지방도 및 국지도: 한국건 설기술 연구원 연계  , 도심내 도로 : 미연계	○	지자체 교통정 보센터 및 국가교 통정보 센터	x	해당 지자체 관리,  모두 미연계
		안양시	○		x		○		x	
		용인시	x		○		○		x	
		안산시	○		x		○		x	
		남양주시	x		○		○		x	
		화성시	○		○		○		x	
		부천시	○		○		○		x	
	충북	청주시	x		○		x		x	
	충남	천안시	○		○		○		x	
	전북	전주시	x		○		x		x	
	경북	포항시	x		x		x		x	
	경남	김해시	x		○		x		x	
30~ 50만	경기도	의정부시	x		○		x		x	
		광명시	○		x		○		x	
		평택시	x		x		x		x	
		시흥시	○		x		○		x	
		파주시	x		x		x		x	
		김포시	x		x		○		x	
	강원도	원주시	x		x		x		x	
	전북	익산시	x		x		x		x	
	경북	구미시	x		x		x		x	
	경남	진주시	x		○		x		x	
	제주도	제주시	○		○		○		x	

4) 현장조사 : 지자체 자체적으로 조사하는 도심내 도로 또는 지방도 및 국지도에서 실시하는 수시조사를 의미

## 2) 지자체 교통정보 공유방안

- 지자체 ITS로 실시간으로 수집되는 도심 내 도로의 교통정보는 현재 국가교통정보센터와의 연계되고 있으므로 연계공유 가능(교통정보통합관리방안 1단계에서 완료)
- 수시조사를 통하여 수집하는 지방도 및 국지도의 교통정보는 한국건설기술연구원에 연계되어 관리되고 있으므로 이와 연계하여 공유가능
- 반면 도심 내 도로의 교통정보는 수시조사를 통하여 지자체가 수집 및 관리하지만 연계·제공은 이루어지지 않음. 이에 이를 공유할 수 있는 별도의 시스템이 필요함
- 수시조사를 통하여 수집되는 교통정보의 통합관리함으로서 지자체 교통정보 수집지점 관리, 중복조사 방지, 도심 내 도로의 다양한 교통정책분석 가능 등의 기대효과를 제시할 수 있음

## 3) 지자체 교통정보 활용방안

- 지자체 교통정보를 연계·공유시, 이를 활용하여 교통분석 및 정책 등에 다양한 활용이 가능

# 제1장 Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구

---

제1절 연구개요

제2절 전국 교통혼잡지도 구축  
알고리즘 검증 및 고도화

제3절 교통혼잡비용 추정 연구

제4절 전국 교통혼잡지도 시스템 구축

제5절 전국 교통혼잡지표 및  
교통혼잡비용 집계 결과

제6절 결론 및 향후 로드맵





# 제1장 Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구

## 제1절 연구개요

### 1. 연구의 배경 및 목적

- 교통시설의 공급은 한정된 재원과 토지로 인하여 지속적으로 늘어나는 교통수요에 미치지 못하므로 이는 교통 혼잡의 주된 원인이 되며, 교통 혼잡은 국민의 경제활동 및 실생활에 큰 불편을 끼치고 있음
- 전국에서 교통 혼잡으로 발생하는 비용은 약 29조원(2010년 기준)으로 국가예산 291조(2010년 기준)의 10%에 해당되며, 이 중 절반이 승용차 보급의 확대에 인하여 발생하는 혼잡비용임
- 전국 자동차등록대수의 경우도 2011년 기준 1,844만대에서 장래 2015년경에는 2,000만대에 도달할 것으로 예상되어 도로의 교통 혼잡은 갈수록 심해질 것으로 판단됨
- 이러한 교통 혼잡에 대한 객관적이고 정확한 진단을 위하여 과거 「2013년 국가교통수요조사 및 DB구축사업」에서는 2013년 9월 한 달 분량의 내비게이션 자료를 활용하여 전국 교통 혼잡지도를 구축하는 연구가 기 수행되었음
- 본 연구에서는 Car Navigation 자료의 범위를 기존 연구보다 확장(2013년 10월~2014년 1월, 4개월치)하여 보다 더 세밀하고 정확한 전국 교통혼잡지도를 구축하고자 함
- 또한 기 수행되었던 연구를 기반으로 알고리즘 검증 및 고도화, 애플리케이션 고도화, 분석용 Map 구축, WebBook 작성, 교통 혼잡비용 추정, Web service를 위한 GIS 웹 엔진 구축 등을 통하여 전국 교통 혼잡지도의 활용성을 증진하고자 함
- 더불어 도시 전반에 걸친 교통 혼잡상태를 보다 더 효율적으로 나타낼 수 있도록 관련 혼잡지표를 검증 및 보완하며, 이용자 중심의 혼잡지도로서 다양한 분석을 통하여 해당 지역의 교통상태 파악 및 도로망 성능평가 등에 활용될 수 있도록 전국 교통 혼잡지도의 고도화 연구를 수행하고자 함

## 2. 연구의 범위

### 가. 공간적 범위

- 대상범위 : 전국권 대상

### 나. 시간적 범위

- 분석연도 : 과업기간 내 수집된 Car Navigation 자료
  - 데이터 수집기간(4개월) : 2013년 10월 1일 ~ 2014년 1월 31일

### 다. 내용적 범위

- 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화
  - 교통혼잡지표 검증 및 개선
  - 기 개발된 교통혼잡지도의 구축 알고리즘의 정확도 개선
  - 기 개발된 혼잡판정 알고리즘의 도로특성별 검증
  - 혼잡판정 알고리즘 수행속도 개선
  - 혼잡지표의 시/공간적 집계 방법론 고도화
  - 혼잡지표의 이용자 분석 방법론 고도화
- 전국 교통 혼잡지도 혼잡비용 추정 연구
  - 기존 교통혼잡비용 산정방법론에 대한 이론적 진단
  - GPS프로브 기반의 교통량 전수화 방법론 및 알고리즘 개발
  - GPS프로브 기반의 교통량 전수화 알고리즘의 파라미터 최적화
  - 전국 도로구간별 혼잡비용 항목별 산정 방법론
  - 전국 도로구간/도로등급/행정구역별 혼잡비용 집계 방법론 개발

- 전국 교통 혼잡지도 시스템 구축
  - 전국 교통 혼잡지도 애플리케이션 고도화
  - 전국 교통 혼잡지도 분석용 Map 구축
  - 전국 교통 혼잡지도 WebBook 작성
  - 전국 교통 혼잡지도의 Web service를 위한 GIS 웹 엔진 구축
- 전국 교통혼잡지표 및 혼잡비용 집계 산출
  - 혼잡지표의 시/공간적 집계 산출 결과
  - 전국 행정구역별/ 도로별 혼잡비용 추정 결과
  - 기존 교통혼잡비용과의 비교 및 원인 분석
- 향후 로드맵 구축
  - 향후 로드맵 구축

## 제2절 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화

### 1. 개요

- 지난 2013년 한국교통연구원에서 수행한 “2013년 국가교통수요조사 및 DB구축사업 중 Car Navigation 자료를 이용한 교통혼잡지도 연구”의 성과물로 전국 교통혼잡지도 시스템을 개발하였음. 해당 시스템에서 제공하고 있는 교통혼잡지도는 특정한 공간(도로, 지역)이나 특정한 시간대의 혼잡의 정도를 한눈에 파악할 수 있도록 전국 도로의 혼잡상태를 GIS Map상으로 표출한 21세기형 도로망 정보지도로서 활용되고 있음
- 하지만, 지난 과업에서는 3개월이라는 짧은 사업 기간으로 인하여 불가피하게 전국 교통혼잡지도 시스템 개발에만 집중하게 되어 시스템 내 교통혼잡지표를 산출하기 위한 관련 알고리즘의 검증과정이 미미하였음. 또한, 교통혼잡지표를 산출하기 위하여 이용되고 있는 Car Navigation 자료도 2013년 9월 한 달 분량의 데이터만을 이용함에 따라 데이터 수급의 한계도 발생하였음
- 본 연구에서는 먼저 Car Navigation 자료 수급의 범위가 지난 과업보다 확장되어 보다 더 세밀하고 정확한 전국 교통혼잡지도를 구축하기 위한 제반 여건이 갖춰지게 되었음
  - 기존 연구 : 1개월치 자료 수급(2013년 9월 1일 ~ 30일)
  - 본 연구 : 4개월치 자료 수급(2013년 10월 1일 ~ 2014년 1월 31일)
- 확장된 Car Navigation 자료를 이용하여 전국 도로의 혼잡 상태를 보다 더 효율적으로 나타낼 수 있도록 관련 교통혼잡지표 산출 알고리즘을 검증하고, 교통 혼잡에 대한 객관적이고 정확한 진단을 위하여 이용자 중심의 다양한 분석을 통하여 전국 교통 혼잡지도의 고도화 연구를 수행하고자 함
- 본 연구의 전국 교통 혼잡지도 구축 알고리즘 검증 및 고도화에 대한 세부 내용은 다음과 같음
  - 교통혼잡지표 검증 및 개선
  - 기 개발된 교통혼잡지도의 구축 알고리즘의 정확도 개선
  - 기 개발된 혼잡판정 알고리즘의 도로특성별 검증
  - 혼잡판정 알고리즘 수행속도 개선
  - 혼잡지표의 시/공간적 집계 방법론 고도화
  - 혼잡지표의 이용자 분석 방법론 고도화

## 2. 구축 알고리즘의 검증

### 가. 교통혼잡지표 검증 및 개선

#### 1) 기 개발된 교통혼잡지표 검토

- 지난 2013년 한국교통연구원에서 수행한 “2013년 국가교통수요조사 및 DB구축사업 중 Car Navigation 자료를 이용한 교통혼잡지도 연구”의 교통혼잡지표는 혼잡강도( $CI$ ; Congestion Intensity, %),  $CO_2$ 배출량( $COE$ ;  $CO_2$  Emission, g/km/대), 연료소모량( $FC$ ; Fuel Consumption, l/km/대), 지체시간( $CD$ ; Control Delay, 초/대)로 총 4가지의 지표를 기 개발하였음
- 위의 지표들은 일반 승용차를 기준으로 개발하였으며, 각 지표별 세부내용은 다음과 같음

#### ① 혼잡강도( $CI$ ; Congestion Intensity, %)

- 혼잡강도( $CI$ ; Congestion Intensity)는 특정 시간대(시간대, 일, 주중, 주말)에 해당 도로구간을 이용한 총 차량이 경험한 총 통행시간대비 혼잡속도로 주행한 차량의 총 통행시간의 비율(%)로 개발하였음

$$\begin{aligned}
 CI(\%) &= 100 \times \frac{\sum_{i=1}^{v_b} \left[ q_r(i) \cdot \frac{l}{v(i)} \right]}{\sum_{i=1}^{v_b} \left[ q_r(i) \cdot \frac{l}{v(i)} \right] + \sum_{i=v_b}^{v_{\max}} \left[ q_r(i) \cdot \frac{l}{v(i)} \right]} \\
 &= 100 \times \frac{\sum_{i=1}^{v_b} [q_r(i) \cdot tt(i)]}{\sum_{i=1}^{v_b} [q_r(i) \cdot tt(i)] + \sum_{i=v_b}^{v_{\max}} [q_r(i) \cdot tt(i)]} \\
 &\cong 100 \times \frac{\sum_{i=1}^{v_b} [q_s(i) \cdot tt(i)]}{\sum_{i=1}^{v_b} [q_s(i) \cdot tt(i)] + \sum_{i=v_b}^{v_{\max}} [q_s(i) \cdot tt(i)]}
 \end{aligned}$$

여기서,  $v(i)$ :  $i$ 번째 속도(정수),  $1 \leq v(i) \leq v_{\max}$

$v_b$ : 혼잡과 비혼잡의 경계속도(km/시)

$v_{\max}$ : 차량의 최대속도(km/시)

$q_r(i)$ : 속도( $i$ , 정수)로 주행한 차량의 비(0~1.0)

$q_s(i)$ :  $i$ 번째 속도차량의 Probe Sample(건)

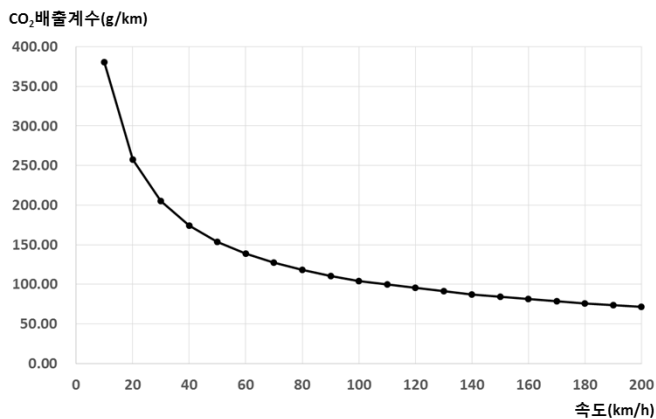
$tt(i)$ : 속도( $i$ )로 주행한 차량의 통행시간(시)

$l$ : 도로구간의 길이(km)

## ② $CO_2$ 배출량( $COE$ ; $CO_2$ Emission, g/km/대)

### a) 속도별 $CO_2$ 배출계수

- 속도별  $CO_2$ 배출계수는 「도로·철도 부문사업의 예비타당성조사 표준지침 수정·보완 연구(제5판)」에서 제시하고 있는 차종별/속도별 오염물질 배출계수 중 승용차 기준의  $CO_2$ 배출계수(g/km)를 이용하여 차량 속도별  $CO_2$ 배출량 모형식을 구축하고 이를 링크별 평균  $CO_2$ 배출량(g/km/대) 산정식에 적용하였음
- 위의 차량 속도별  $CO_2$ 배출량(g/km) 모형식은 아래와 같음



$$FN_C = \frac{1,391.5}{v^{0.56}}$$

여기서,  $FN_C$ : km당  $CO_2$  배출량(g/km)

$v$ : 차량 속도(km/h)

<그림 1-1> 기 개발된 속도별  $CO_2$ 배출계수(승용차 기준)

### b) $CO_2$ 배출량 산정식

- $CO_2$ 배출량( $COE$ ;  $CO_2$  Emission, g/km/대)은 특정 시간대에 해당 도로구간을 이용한 차

량이 배출한 평균  $CO_2$ 배출량으로 개발하였음

$$COE = \sum_i [FN_c(i) \times q_r(i)] \cong \frac{\sum_i [FN_c(i) \times q_s(i)]}{\sum_i q_s(i)}$$

여기서,  $FN_c(i)$ : 속도( $i$ )의 km당  $CO_2$  배출량 함수

$q_r(i)$ : 속도( $i$ )의 차량 비율(0~1.0)

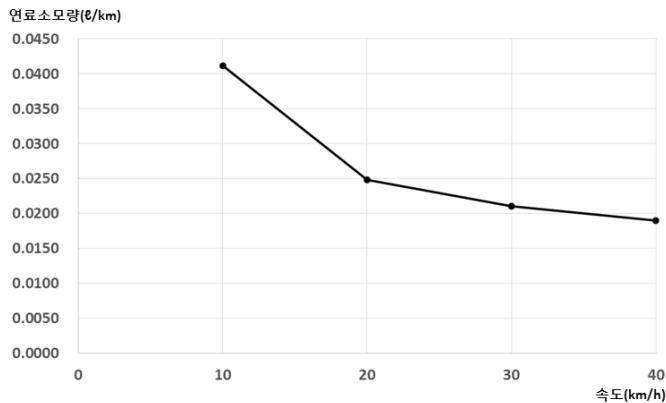
$q_s(i)$ :  $i$ 번째 속도차량의 Probe Sample(건)

$l$ : 도로구간의 길이(km)

### ③ 연료소모량( $FC$ ; Fuel Consumption, l/km/대)

#### a) 속도별 연료소모량 계수

- 속도별 연료소모량 계수는 「도로·철도 부문사업의 예비타당성조사 표준지침 수정·보완 연구 (제5판)」에서 제시하고 있는 차종별/속도별 차량운행비용 중 승용차 기준의 유류비(원/km)를 리터 당 휘발유 단가(1,895.27원)<sup>1)</sup>로 나누어 차량 속도별 연료소모량( $l$ /km) 모형식을 구축하고 이를 링크별 평균 연료소모량( $l$ /km/대) 산정식에 적용하였음
- 위의 차량 속도별 연료소모량( $l$ /km) 모형식은 아래와 같음
  - 차량 속도가 40km 이하인 경우



$$FN_g = 0.0357v^2 - 2.4846v + 61.8401$$

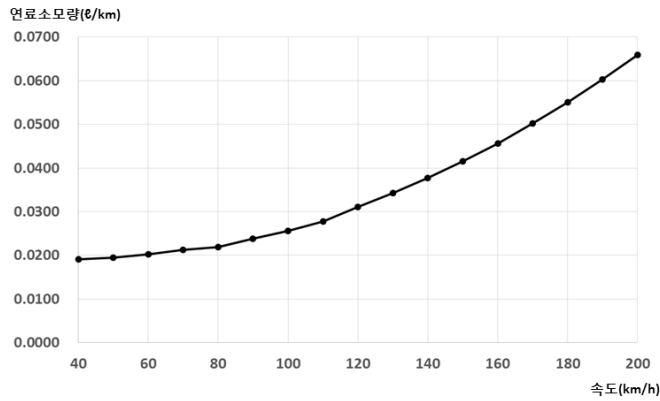
여기서,  $FN_g$ : km당 연료소모량( $l$ /km)

$v$ : 차량 속도(km/h)

<그림 1-2> 기 개발된 속도별 연료소모량 계수(속도 40km 이하, 승용차 기준)

1) 2013. 10. 27 한국석유공사 Opinet 기준

- 차량 속도가 40km 초과인 경우



$$FN_g = 0.0018v^2 - 0.1403v + 22.0199$$

여기서,  $FN_g$ : km당 연료소모량 (l /km)

$v$ : 차량 속도(km/h)

<그림 1-3> 기 개발된 속도별 연료소모량 계수(속도 40km 초과, 승용차 기준)

#### b) 연료소모량 산정식

- 연료소모량( $FC$ ; Fuel Consumption, l /km/대)은 특정 시간대에 해당 도로구간을 이용한 차량(승용차 기준)의 평균 유류 소모량(휘발유 기준)으로 개발하였음

$$FC = \frac{\sum_i [FN_g(i) \times q_s(i)]}{\sum_i q_s(i)}$$

여기서,  $FN_g(i)$ : 속도( $i$ )의 km당 유류소모량 함수

$q_r(i)$ : 속도( $i$ )의 차량 비율(0~1.0)

$q_s(i)$ :  $i$ 번째 속도차량의 Probe Sample(건)

$l$ : 도로구간의 길이(km)

#### ④ 지체시간( $CD$ ; Control Delay, 초/대)

- 링크 지체시간( $CD$ ; Control Delay, 초/대)는 특정 시간대에 해당 도로구간을 이용한 차량의 평균 지체시간으로 개발하였음



$$\begin{aligned}
 CD &= 3,600 \cdot \sum_i \left[ \left( \frac{l}{v(i)} - \frac{l}{v_{85}} \right) \times q_r(i) \right] = 3,600 \cdot \frac{\sum_i [(tt(i) - tt_{85}) \times q_r(i)]}{\sum_i q_r(i)}, \\
 &\cong 3,600 \cdot \frac{\sum_i \left[ \left( \frac{l}{v(i)} - \frac{l}{v_{85}} \right) \times q_s(i) \right]}{\sum_i q_s(i)} = 3,600 \cdot \frac{\sum_i [(tt(i) - tt_{85}) \times q_s(i)]}{\sum_i q_s(i)}
 \end{aligned}$$

여기서,  $0 < v(i) < v_{85}$

$v_{85}$ : 비 혼잡시 누적 85%속도(km/시)

$v(i)$ :  $i$ 번째 속도(km/시)

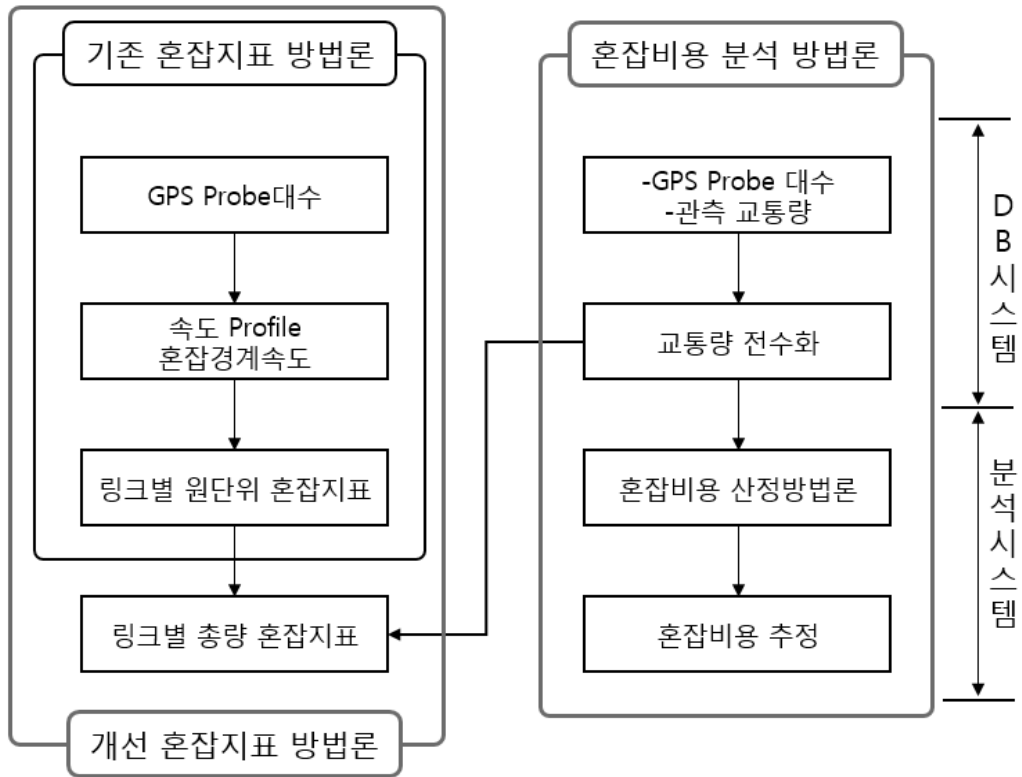
$q_r(i)$ :  $i$ 번째 속도차량의 비율(0~1.0)

$q_s(i)$ :  $i$ 번째 속도차량의 Probe Sample(건)

$l$ : 도로구간의 길이(km)

## 2) 기존 교통혼잡지표의 개선

- 앞서 살펴본 기 개발된 교통혼잡지표는 모두 원단위의 지표로서 Car Navigation 자료의 Probe 1대당 발생하는 교통혼잡지표 만으로 이용되고 있음
  - 혼잡강도(CI, Congestion Intensity, %),
  - CO2배출량(COE, CO2 Emission, g/km/대)
  - 연료소모량(FC, Fuel Consumption, l/km/대)
  - 지체시간(CD, Control Delay, 초/대)
- 위의 원단위 교통혼잡지표 만으로는 전국의 도로구간별 총량 교통혼잡지표를 산출하지 못하는 한계가 발생함. 이는 전국의 모든 도로구간에 대한 구간별 Probe건수가 아닌 실제 차종별 교통량에 대한 정보를 알 수 없기 때문임
- 이러한 한계를 극복하기 위하여 본 연구에서는 해당 도로구간의 Probe건수와 (인접 도로구간의)관측교통량을 이용하여 Probe건수를 차종별 교통량으로 전수화하기 위한 모형을 개발하고, 각 구간별로 추정된 차종별 전수화 교통량을 이용한 본 연구의 총량 교통혼잡지표를 산정하고자 함
- 즉, ①기존의 혼잡지도시스템에서 평가된 속도 Profile과 혼잡 경계속도를 이용하여 링크별 원단위 혼잡지표를 산정하고, ②산정된 원단위 혼잡지표와 혼잡비용시스템에서 추정된 관측 교통량을 이용하여 링크별 총량 혼잡지표를 산정하도록 함



<그림 1-4> 교통혼잡비용 분석시스템과 연동된 교통혼잡지표 산출 흐름

- Car Navigation 자료의 표본과 원단위에 기초하는 기존 교통혼잡지표는 아래의 표와 같이 본 연구의 총량 교통혼잡지표로 개선하며, 각 교통혼잡지표의 개선 내용과 추가지표의 차원을 보여주고 있음

<표 1-1> 교통혼잡지표의 개선 내용

구분		기존	개선
기존 교통혼잡지표	혼잡강도	표본 원단위(%)	총량(%)
	CO2배출량	표본 원단위(g/km/대)	총량(g, kg, ton)
	연료소모량	표본 원단위(l /km/대)	총량(l, kl)
	지체시간	표본 원단위(초/대)	총량(초, 분, 시)
추가 교통혼잡지표	교통혼잡비용	-	총량(억원)

## 3) 교통혼잡지표의 검증

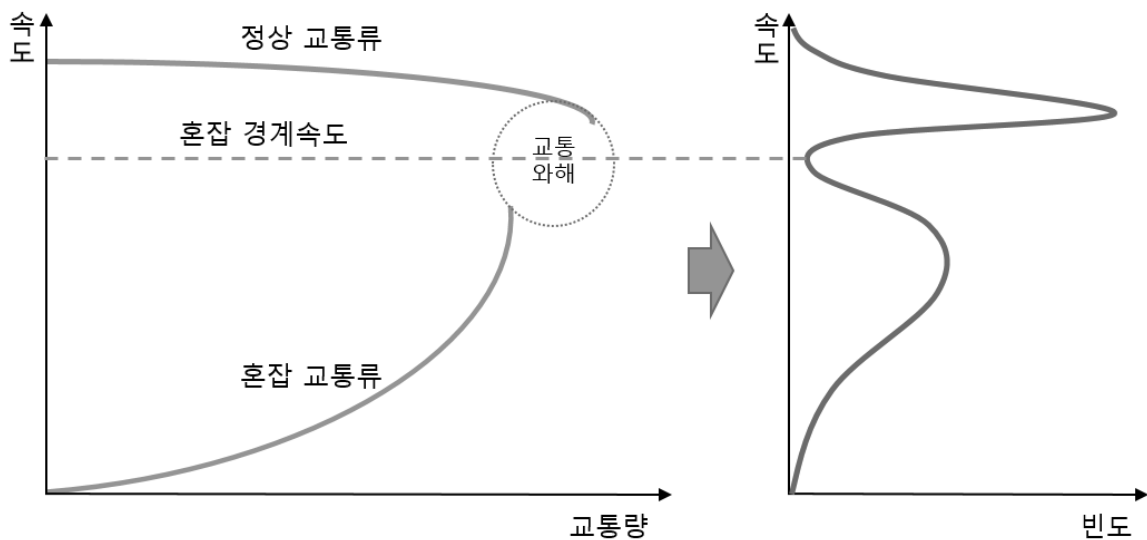
- 본 연구의 혼잡지표의 검증은 Car Navigation 자료의 Probe건수에 따른 각 교통혼잡지표의 정확도를 통하여 수행하였음
- 각 교통혼잡지표의 정확도 검증을 위한 도로의 형태는 ①연속류와 ②단속류의 특성을 가지는 도로구간으로 구분하여 수행하였음
- 이 중 <표 1-2>는 연속류의 특성을 가지는 서울외곽순환도로(송파~서하남 구간)의 Probe건수 표본율(%)에 따른 교통혼잡지표의 정확도를 보여주고 있음
- Probe 표본율이 75~2%(Probe건수 2,544 ~ 81건)까지 감소하였을 경우, 경계속도는 99.3%의 정확도를 보이고 있음
- 또한, 혼잡강도의 경우 평균절대오차(%)는 4.82%로서 95.2%의 정확도를 보이고 있으며, 최소 표본수는 약 80개로 나타나고 있음.

&lt;표 1-2&gt; Probe 대수에 따른 혼잡지표 정확도 분석결과

표본율(%)	100%	75%	50%	25%	10%	5%	3%	2%
유효표본수	3390	2544	1700	840	357	177	116	81
경계속도	60.0	60.0	60.0	61.0	60.0	60.0	60.0	62.0
전체속도평균	80.8	81.1	81.0	80.9	79.8	81.6	79.5	77.8
하위평균속도	44.6	44.8	44.8	45.3	44.9	43.7	44.4	44.0
상위평균속도	89.3	89.3	89.2	89.0	88.5	88.9	88.6	87.6
하위속도비율	0.19	0.18	0.18	0.19	0.20	0.16	0.21	0.22
상위속도비율	0.81	0.82	0.82	0.81	0.80	0.84	0.79	0.78
상위85%속도	105.1	104.8	104.9	104.5	105.1	103.0	104.7	102.9
혼잡강도(%)	32.0	31.0	30.9	31.0	33.2	30.1	34.2	34.4

#### 나. 기 개발된 교통혼잡지도의 구축 알고리즘의 정확도 개선

- 본 연구의 혼잡경계속도는 <그림 1-5>와 같이 교통량과 속도의 관계에서 교통와해(break down)상태에 해당하는 속도로서 혼잡과 정상상태를 구분하기 위한 속도를 의미함
- 또한, 혼잡경계속도는 혼잡강도, CO2배출량, 연료소모량, 지체시간 등과 같은 본 연구의 교통혼잡지표 산출하기 위하여 이용되는 가장 중요한 변수임



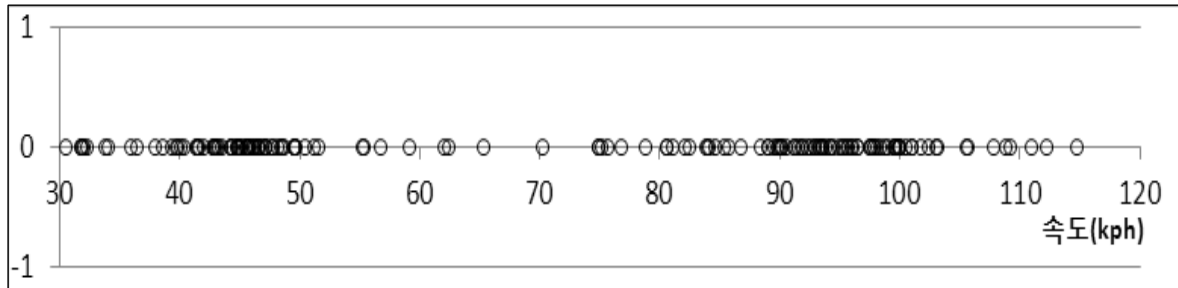
<그림 1-5> 교통량과 속도와의 관계

- 따라서, 기 개발된 구축 알고리즘의 정확도 개선은 무엇보다도 교통혼잡지표 산출을 위한 중요변수인 링크별 혼잡경계속도를 합리적인 값으로 추정해야 함
- 이를 위하여 Car Navigation 속도 자료에 대한 기존의 이상치 제거 방법론을 고도화하여 교통혼잡지도의 구축 알고리즘의 정확도를 개선하고 산출된 혼잡경계속도에 대한 정확도 검증을 수행하고자 함

##### 1) 이상치 제거 알고리즘 개선

- 본 연구의 링크별 혼잡경계속도를 추정하는 알고리즘은 표본수가 높은 경우에는 정확한 경계속도를 산정하지만, 표본수가 낮은 경우에는 정확도가 다소 저하되는 경우가 발생하게 됨
- 낮은 표본 수에서 발생하는 알고리즘 정확도의 Variation 문제는 혼잡경계속도를 산정하기 위하여 이용되는 Car Navigation 속도 자료 중 이상치의 Noise현상으로 판단됨
- 또한, 혼잡과 비 혼잡이 발생한 구간에 대한 수집된 개별 Car Navigation 차량의 속도 분

포는 60kph를 전/후로 혼잡과 비 혼잡 교통류로 극명하게 양분되고 있으며, 전형적인 쌍봉 분포(Two-peak distribution)를 보이고 있음



<그림 1-6> 개별 Car Navigation 차량의 속도 분포

- 이러한 이상치에 대한 영향을 최소화하기 위하여 기존의 정규분포가 아닌 개별 Car Navigation 차량의 속도 분포 즉, 쌍봉 형태의 분포에서 이상치 제거 알고리즘을 개선하였음

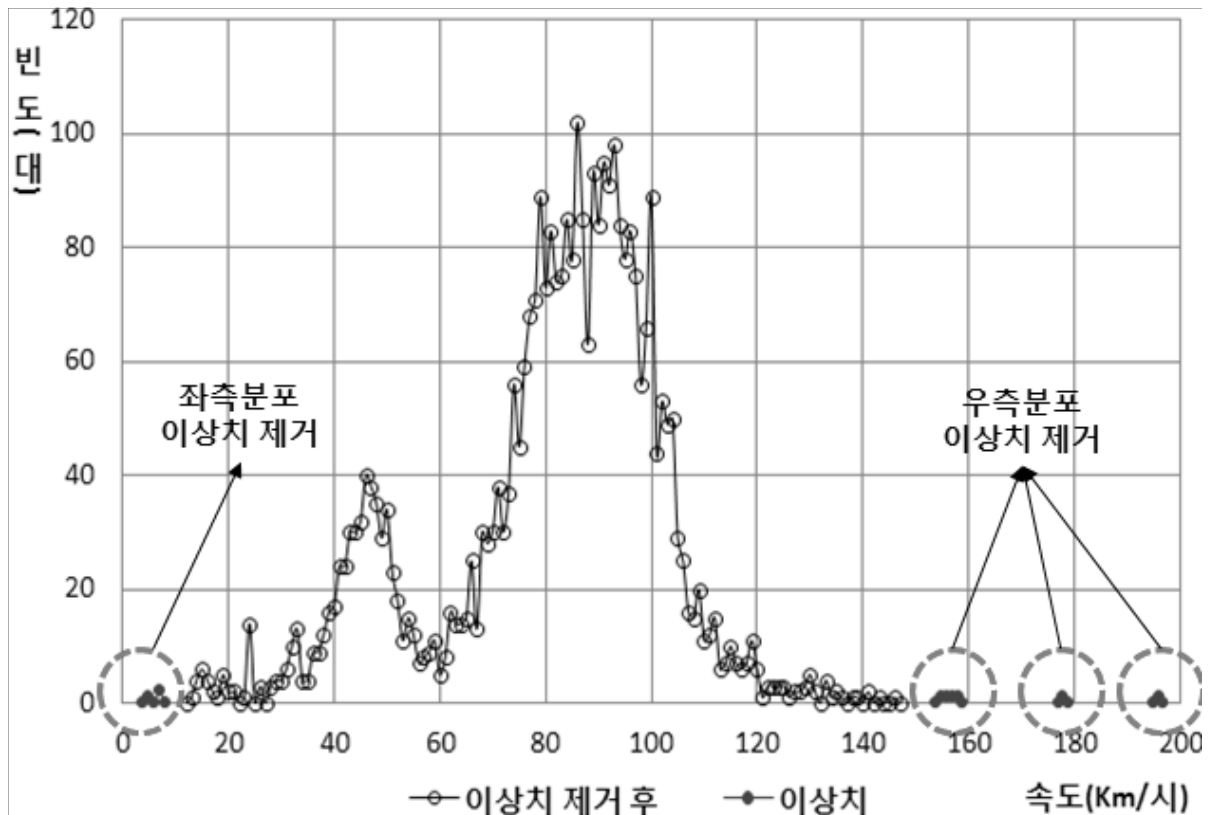
<표 1-3> 이상치 제거 알고리즘 개선

구분	기존 방법(2013년 기 구축)	개선 방법(본 연구)
개선 내용	- 이상치 $\langle \text{평균} - 3\delta \text{ 또는 } \text{평균} + 3\delta \rangle$ 이상치	- 속도 Profile <sup>2)</sup> 의 사건이 발생한 도수 간 최 인접 좌/우측 거리분포의 표준편차 최소화 알고리즘 개발

- 본 연구의 이상치 제거 알고리즘에 대한 단계별 세부 내용은 다음과 같음
  - 주어진 도로구간, 이동류, 시간조건(년, 월, 일, 시간대)에 대하여;
  - Step 1: 초기 밀도분석
    - 1-1. 1km/시의 계급구간으로 구성되는 속도 Histogram 구축
    - 1-2. Paragon Window의 길이( $w_l$ ,  $1 \leq w_l \leq v_{\max}$ ,  $v_{\max}$ =차량의 이상적 속도 최대 값) 최적화를 통한 속도의 밀도 Profile( $SDP_{op}$ )을 구축
    - 1-3. 쌍봉분포 또는 단일분포의 분석 및 첨두에서 속도값 결정
  - Step 2: 저속도 이상치 제거
    - 2-1. 속도 0과  $SDP_{op}$ 의 저속 분포의 첨두시 속도( $v_{lp}$ )\*g구간의 속도 Histogram 구축, 여기서, g=황금비(0.618)
    - 2-2. step 1-2을 수행하여  $SDP_{op}$ 을 재구축

<sup>2)</sup> 1km/시의 계급구간으로 구성되는 속도 Histogram

- 2-3. 쌍봉분포의 분석 및 첨두에서  $tv_{lt}$  값 결정
- Step 3 : 고속도 이상치 제거
  - 3-1. 고속 분포의 첨두시 속도( $v_{up}$ )\*g와  $v_{max}$  구간의 속도 Histogram 구축
  - 3-2. step 1-2을 수행하여  $SDP_{op}$ 을 재구축
  - 3-3. 쌍봉분포의 분석 및 첨두에서  $tv_{rt}$  값 결정
- Step4 : 최종 밀도분석 및 혼잡경계속도 판정
  - 4-1. 1km/시의 계급구간으로 구성되는 속도 Histogram 구축
  - 4-2. 갱신된 좌우측 outlier 제거 비율값( $tv_{lt}$ ,  $tv_{rt}$ )을 이용한 이상치 제거
  - 4-3. Paragon Window의 길이( $w_l$ ,  $1 \leq w_l \leq v_{max}$ ,  $v_{max}$ =차량의 이상적 속도 최대 값) 최적화를 통한 최종 속도의 밀도 Profile( $SDP_{op}$ )을 구축
  - 4-4. 쌍봉분포 또는 단일분포의 분석 및 첨두에서 혼잡/비혼잡 속도값( $v_c$ ,  $c_{nc}$ ), 혼잡과 비혼잡 결정 속도( $v_b$ ) 산정
  - 4-5. 최종  $SDP_{op}$ 와 혼잡/비혼잡 속도값( $v_c$ ,  $c_{nc}$ ), 혼잡과 정상교통류의 경계를 결정하는 속도( $v_b$ ) 산정



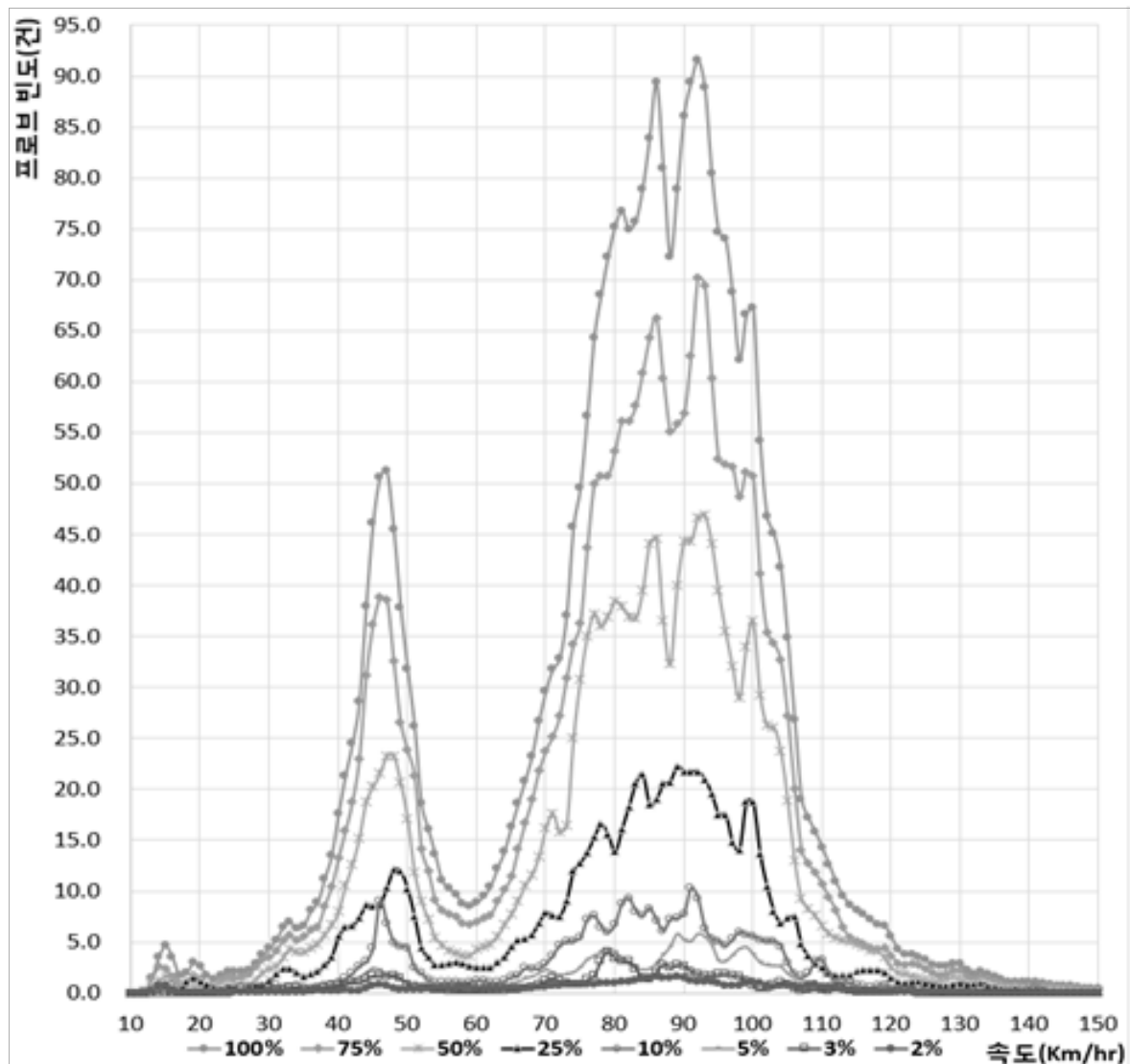
<그림 1-7> 속도 Profile의 이상치 제거(본 연구의 개선 알고리즘)

## 2) 혼잡경계속도의 정확도 검증

- 앞서 설명한 바와 같이 혼잡 경계속도는 GPS 프로수 샘플수가 일정 이상일 경우 합리적이고 동일한 값을 평가함. 그러나 GPS 프로수 샘플수가 매우 적은 경우 추정 경계속도의 Variation은 증가함
- 본 연구에서는 산출된 혼잡 경계속도에 대한 정확도 검증을 수행하기 위하여 해당 구간의 GPS Probe 샘플수에 대한 표본율(%) [100, 75, 50, 25, 10, 5, 3, 2]로 분석 시나리오를 설정하고, 각 분석 시나리오별로 붓스트랩(bootstrap) 방법을 이용하여 표본 값들을 추출하여 정확도를 검증하였음
  - 붓스트랩(bootstrap) 방법은 비모수 통계적 기법으로 실제 조사한 결과를 바탕으로 가상의 샘플링을 수행하고 그 결과를 기반으로 결과의 정확성을 평가하거나 분포를 추정하는 방법임. 즉, 주어진 확률표본  $x_1, x_2, \dots, x_n$ 으로부터 복원 추출방법으로 붓스트랩 표본  $x_1^*, x_2^*, \dots, x_n^*$ 을 구함. 추출가능한 붓스트랩 표본의 총 수는  $n^n$ 개임
- 아래의 표와 그림에서는 서울외곽순환도로(송파~서하남 구간)의 GPS Probe 표본율(%)에 따른 혼잡지표의 정확도를 보여주고 있음
- 해당 지점의 일평균 교통량은 175,000대이며, 샘플의 수집기간은 22일임. 따라서 샘플율  $0.0881 [= 3,390 / (175,000 * 22) * 100]$ 와  $0.00213 [= 81 / (175,000 * 22) * 100]$ 에서 동일한 혼잡경계속도가 추정되고 있음
- 샘플수 3,390개일 경우를 모수로 가정한 경우, 표본율(%) 시나리오 [75, 50, 25, 10, 5, 3, 2]에 대한 표본수 [2,544, 1,700, 840, 357, 177, 116, 81]에 따른 추정 혼잡 경계속도는  $60 \pm 2$ 의 범위에서 안정적으로 추정되고 있으며, 99.3%의 정확도를 보이고 있음. 또한, 혼잡강도의 경우의 평균절대오차(%)는 4.82%로서 95.2%의 정확도를 보이고 있음

&lt;표 1-4&gt; Probe건수에 따른 혼잡지표 정확도 분석결과

구분	표본율								정확도
	100%	75%	50%	25%	10%	5%	3%	2%	
유효표본수	3,390	2,544	1,700	840	357	177	116	81	-
경계속도	60.0	60.0	60.0	61.0	60.0	60.0	60.0	62.0	99.3%
전체속도평균	80.8	81.1	81.0	80.9	79.8	81.6	79.5	77.8	-
하위평균속도	44.6	44.8	44.8	45.3	44.9	43.7	44.4	44.0	
상위평균속도	89.3	89.3	89.2	89.0	88.5	88.9	88.6	87.6	
하위속도비율	0.19	0.18	0.18	0.19	0.20	0.16	0.21	0.22	
상위속도비율	0.81	0.82	0.82	0.81	0.80	0.84	0.79	0.78	
상위85%속도	105.1	104.8	104.9	104.5	105.1	103.0	104.7	102.9	
혼잡강도(%)	32.0	31.0	30.9	31.0	33.2	30.1	34.2	34.4	95.2%



<그림 1-8> Probe 샘플수에 따른 혼잡 경계속도의 안정성

#### 다. 기 개발된 혼잡판정 알고리즘의 도로특성별 검증

- 본 연구의 혼잡경계속도 판정 알고리즘에 대하여 연속류와 단속류와 같이 도로특성별로 수집된 링크별 Probe 표본수와 속도 Profile 자료를 이용하여 본 연구의 혼잡경계속도에 대한 검증을 수행하였음
- 여기서, 기존 혼잡기준속도는 현재 교통혼잡비용 추정 시 이용되고 있는 도로등급별 혼잡기준속도를 적용하였음. 도로 특성에 따른 혼잡경계속도의 검증 및 기존 혼잡기준속도와 비교·분석을 수행한 결과는 다음과 같음

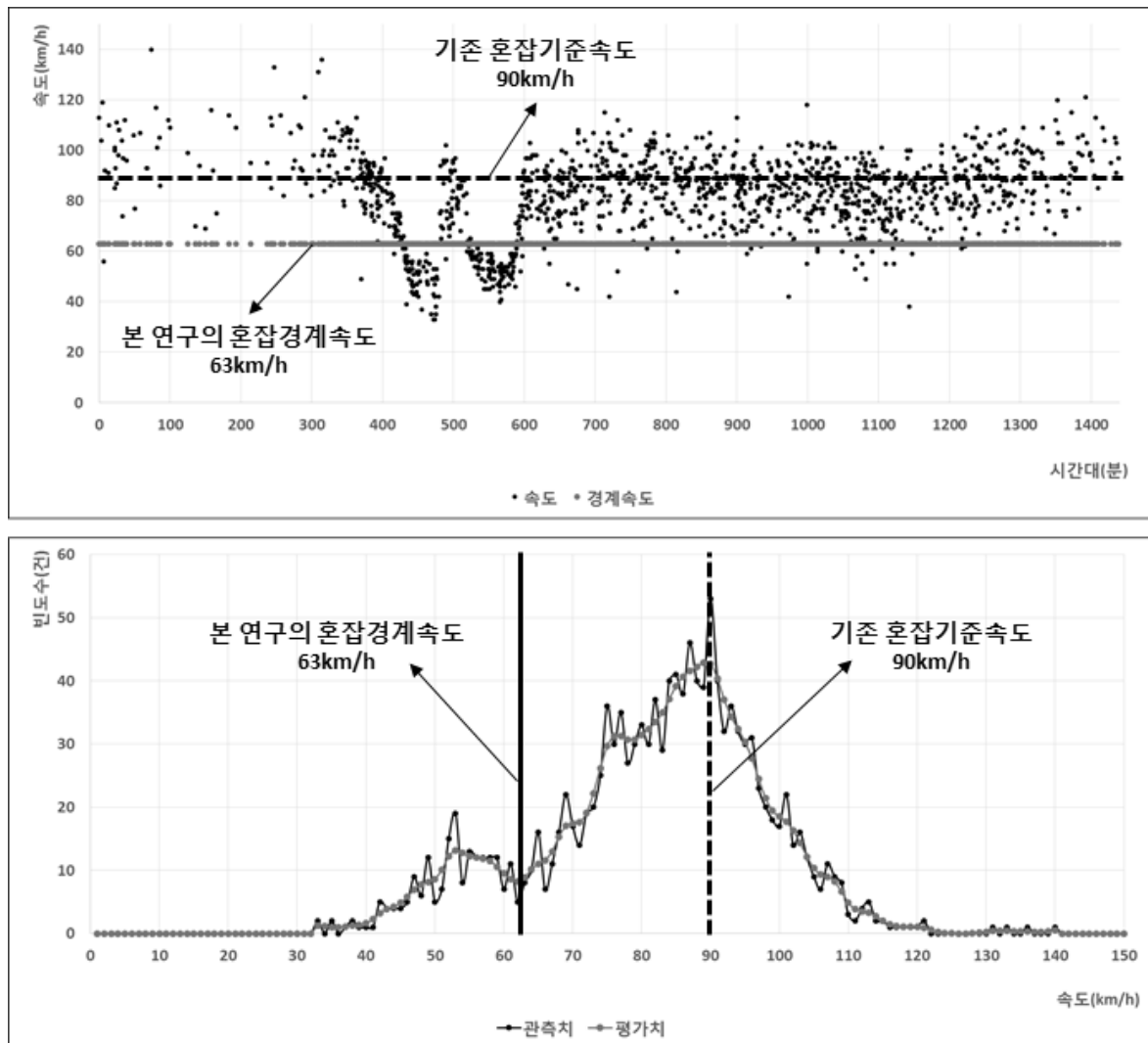
- 데이터 수집일 : 2014년 1월 6일(월), 분 단위 24시간 자료



## 1) 연속류 구간

## ① 서울외곽순환고속도로(서하남IC → 송파IC 구간, 연장 3.90km)

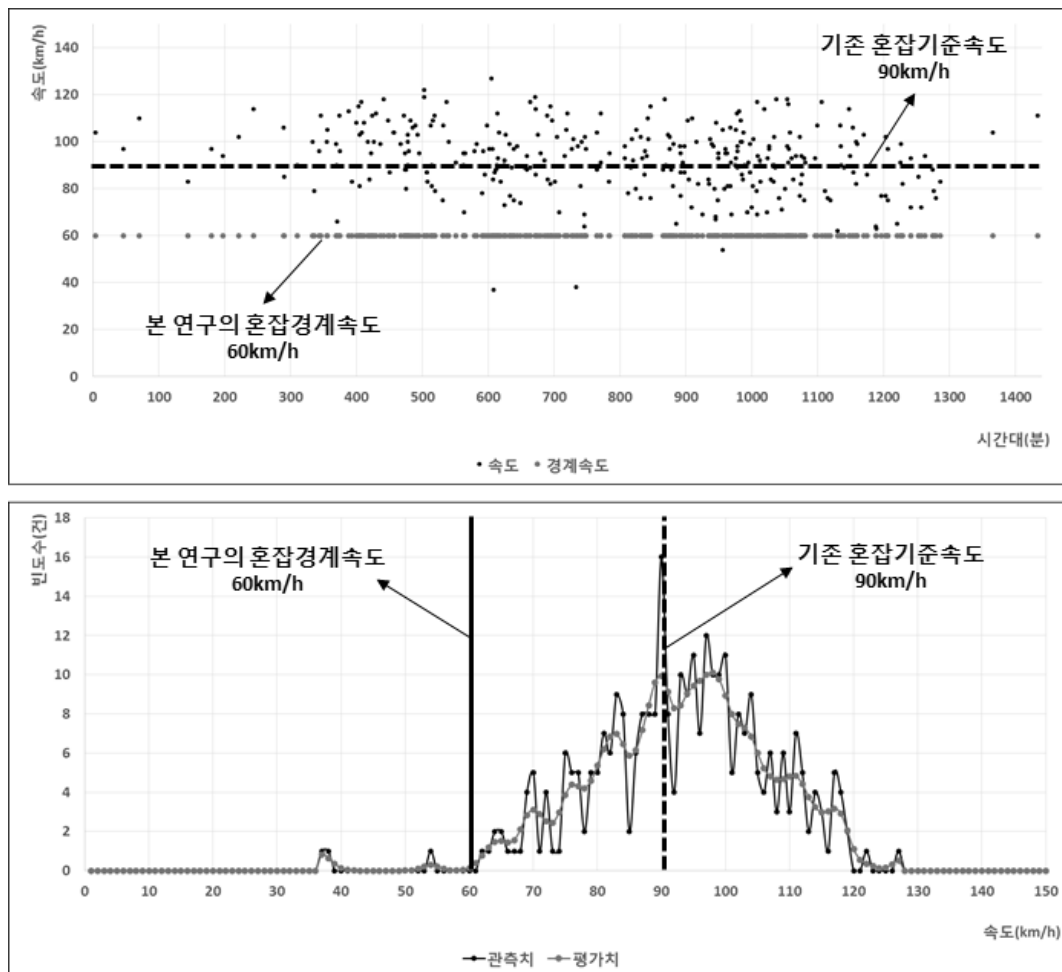
- 본 구간은 혼잡과 비 혼잡 상태가 극명하게 나뉘지는 쌍봉분포의 형태를 띠는 연속류 구간으로 전체 시간대에서 비 혼잡 교통류와 혼잡 교통류를 잘 구분하는 것으로 분석되었음. 즉, 혼잡경계속도가 기존 보다 교통공학적으로 교통량-속도관계의 교통와해(Break down)상태를 잘 설명하고 있음
- 만약 기존 혼잡기준속도를 이용하여 혼잡강도(%)를 산출한다면, 혼잡경계속도를 이용하여 산출한 결과 보다 높게 산출될 것으로 판단되며, 이는 기존의 교통혼잡이 교통와해 상태(서비스 수준 E)이하가 아닌 서비스 수준 C~D이하에서 발생하는 것으로 정의하기 때문임



<그림 1-9> 서울외곽순환고속도로(서하남IC → 송파IC 구간) 혼잡경계속도 검증 결과

② 인천대교 고속도로(송도 → 인천공항 구간, 연장 14.01km)

- 본 구간은 하루 중 교통 혼잡이 거의 발생하지 않는 단봉분포의 형태를 띠는 연속류 구간으로 혼잡과 비 혼잡의 교통류를 구분할 필요는 없으며, 이럴 경우 전체 속도 분포에서 이상치를 제외한 최저 속도를 본 연구의 혼잡경계속도로 지정하였음. 즉, 해당 구간은 교통 수요의 증가로 인한 교통 혼잡이 거의 발생하지 않는 자유류 상태의 구간으로 이러한 속도 분포는 교통 혼잡에 의한 영향이 아닌 저속 또는 고속으로 운전하는 개별 차량의 운전 행태에 의한 변화인 것으로 판단됨
- 만약 기존 혼잡기준속도를 이용하여 혼잡강도(%)를 산출한다면 본 연구의 결과보다 과다 산출될 우려가 있음. 즉, 기존 혼잡기준속도는 도로등급별로 동일한 기준속도를 적용하기 때문에 각 구간별로 다양하게 발생하는 현장의 교통상황을 제대로 반영하지 못하는 것으로 판단됨

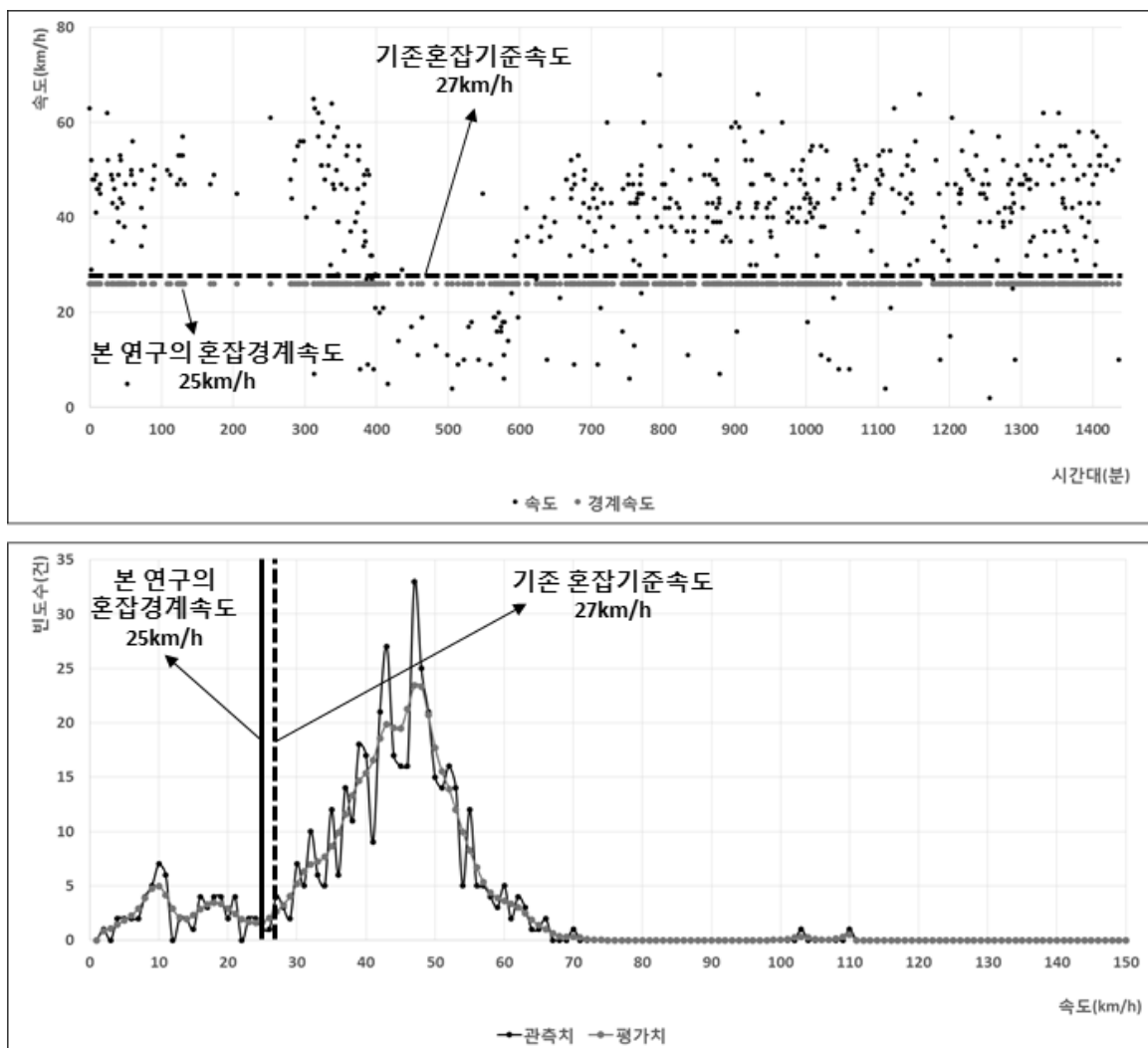


<그림 1-10> 인천대교 고속도로(송도 → 인천공항 구간) 혼잡경계속도 검증 결과

## 2) 단속류 구간

## ① 남부순환도로(낙성대삼거리 → 서울대입구역, 연장 0.19km)

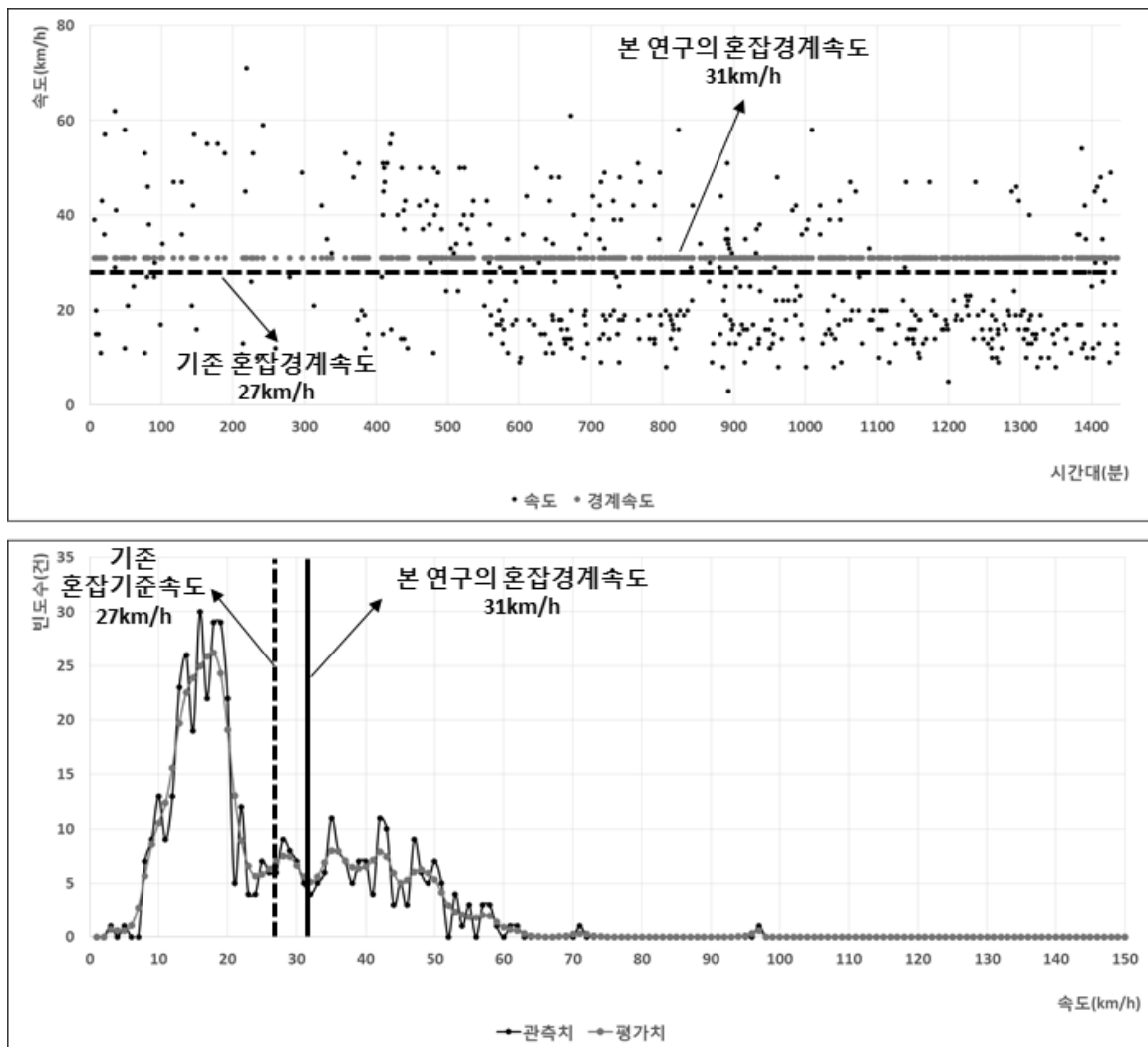
- 본 구간은 혼잡과 비 혼잡 상태가 극명하게 나뉘지는 쌍봉분포의 형태를 띠는 단속류 구간으로 본 연구의 혼잡경계속도와 기존 혼잡기준속도의 차이가 크지 않은 것으로 분석되었음
- 이는 도시부 단속류 구간의 특성에 따라 연속류 구간에 비해 전체 속도의 분포 변화는 심하지 않으며, 제한적인 속도 범위 내에서 본 연구의 혼잡경계속도를 산출하므로 기존 혼잡기준속도와 비슷한 결과가 나오는 것으로 판단됨



<그림 1-11> 남부순환도로(낙성대삼거리 → 서울대입구역) 혼잡경계속도 검증 결과

② 강남대로(교보타워 사거리 → 강남역, 연장 0.56km)

- 본 구간은 혼잡과 비 혼잡 상태가 나눠지는 쌍봉분포의 형태를 띠는 단속류 구간으로 본 연구의 혼잡경계속도와 기존 혼잡기준속도의 차이가 크지 않은 것으로 분석되었음
- 이는 도시부 단속류 구간의 특성에 따라 연속류 구간에 비해 전체 속도의 분포 변화는 심하지 않으며, 제한적인 속도 범위 내에서 본 연구의 혼잡경계속도를 산출하므로 기존 혼잡기준속도와 비슷한 결과가 나오는 것으로 판단됨

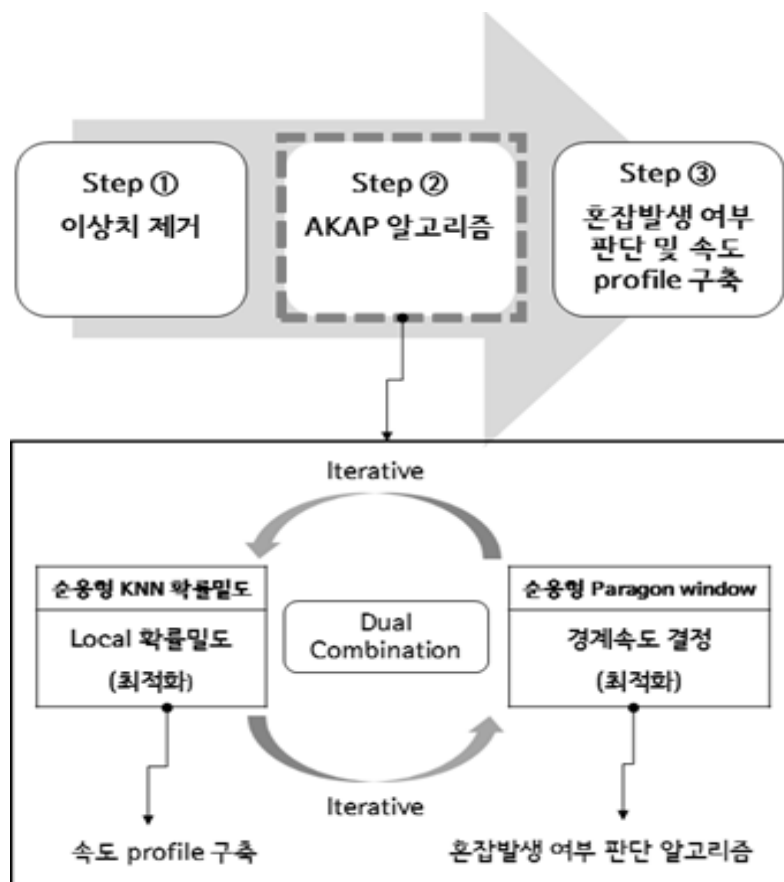


<그림 1-12> 강남대로(교보타워 사거리 → 강남역) 혼잡경계속도 검증 결과

### 3. 구축 알고리즘의 고도화

#### 가. 혼잡판정 알고리즘 수행속도 개선

- 본 연구의 혼잡판정 알고리즘은 AKAP(dual combination of Adaptive K-nearest-neighbor probability density and Adaptive Paragon window)기법을 적용하였으며, 두 알고리즘은 Iterative하게 결합되어 최적의 속도 Profile과 혼잡 경계속도 값을 결정함
- AK(Adaptive K-nearest-neighbor probability density)알고리즘은 KNN(최근린 이웃, K-Nearest Neighbour) 확률밀도를 기반으로 입력 속도 Profile과 오차를 최소화하면서 속도 Profile를 가공함
- AP(Adaptive Paragon window)는 KNN기반 Paragon Window 알고리즘으로서 이전 AK 스텝에서 평가된 속도 Profile를 이용하여 첨두간의 Paragon Window의 사이즈를 극대화하는 지점에서 경계속도를 평가함



<그림 1-13> 본 연구의 혼잡경계속도 판정 알고리즘의 구조

- 여기서, (데이터 입/출력, 속도 Profile 구축 등을 제외한)AKAP 알고리즘의 수행속도는 일

반 PC에서 링크 1개당 0.0013초가 소요되었으며, 250,000개의 링크에 대하여 5.42분이 소요되어 대규모 가로망 적용이 가능하였음

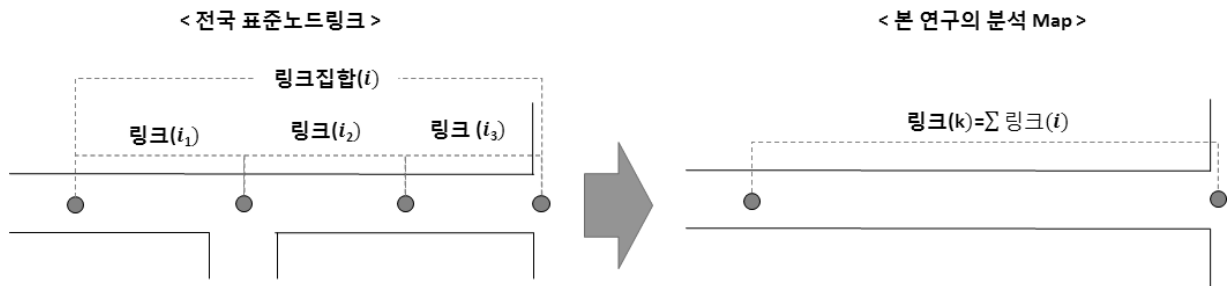
- 하지만, Car Navigation 자료의 확장으로 인하여 전체 250,000개 링크에 대한 4개월의 요일(7일)별 시간대(24시간)별을 수행하는데 약 60.7시간 $[(250,000 \times 4 \times 7 \times 24 \times 0.0013) / 3,600]$ 이 소요되므로 본 연구의 혼잡판정 알고리즘에 대한 수행속도 개선이 필요함
- 따라서, 본 연구의 혼잡판정 알고리즘의 수행속도를 개선하기 위하여 ①Step size 최적화, ②KNN 확률밀도 최대/최소값 최적화, ③Window size 최대/최소값 최적화, ④최대/최소 혼잡경계속도 범위 최적화를 수행하였으며, 각 개선 방안에 대한 연산 회수의 감소효과는 다음과 같음
  - AK 알고리즘의 Step size 최적화: 반복회수 30%감소
  - AK 알고리즘의 KNN 확률밀도 최대/최소값 최적화: 반복회수 39.3%감소
  - AP 알고리즘의 Window size 최대/최소값 최적화: 반복회수 39.3% 감소
  - AP 알고리즘의 최대/최소 혼잡경계속도 범위 최적화: 반복회수 53.3% 감소

#### 나. 혼잡지표의 시/공간적 집계 방법론 고도화

- 기 개발된 혼잡지표는 기본적으로 전국 표준노드링크를 이용하여 개별 링크 단위로 혼잡지표를 산정함. 또한 링크별 특정 시간대(시간대, 일, 주중, 주말)에 따라 해당 지역(광역시/도, 시/군/구)의 공간적 단위로 집계함
- 하지만 본 연구에서는 동일한 도로구간이 여러 개로 분할 되어있는 국가표준노드링크를 보완하고자 자체 분석 Map을 구축하며, 과거 월 단위 내비게이션 자료가 년 단위 내비게이션 자료로 확장됨에 따라 혼잡지표에 대한 새로운 시·공간적 집계방법이 필요함
- 본 연구의 혼잡지표는 아래와 같이 총 4개의 혼잡지표를 사용함
  - 혼잡강도( $CT$ ; Congestion Intensity, %)
  - $CO_2$ 배출량( $COE$ ;  $CO_2$  Emission,  $g/km$  또는  $g$ ,  $kg$ ,  $ton$ )
  - 연료소모량( $FC$ ; Fuel Consumption,  $l/km/대$  또는  $l$ ,  $kl$ )
  - 지체시간( $CD$ ; Control Delay, 초/대 또는 초, 분, 시)

## 1) 분석 Map의 혼잡지표 집계

- 전국 표준노드링크의 개별 링크단위 혼잡지표를 본 연구의 분석 Map의 개별 링크단위 혼잡지표로 집계함
- 해당 자료의 시간적 범위에 따라 단위 시간대별(월/일/시간 단위)로 산출함
- 이는 이용자들의 주어진 사·공간적 조건에 따른 혼잡지표 집계를 위하여 가장 기본이 되는 분석단위임



&lt;그림 1-14&gt; 분석 Map 집계 예시

## ① 혼잡강도(LACI; Link Aggregated Congestion Intensity)

- 혼잡강도의 집계는 분석 Map의 해당 링크( $k$ )에 포함되는 링크집합( $i$ )의 혼잡강도를 각 해당 링크별 표본 수와 링크길이를 이용하여 가중 평균한 값으로 산출함

$$LACI(k) = \frac{\sum_i [CI(i) \times q_n(i) \times l(i)]}{\sum_i [q_n(i) \times l(i)]}$$

여기서,  $LACI(k)$ : 분석 Map의 해당 링크( $k$ )의 혼잡강도(%)

$CI(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별 혼잡강도(%)

$q_n(i) = [q_s(i), \hat{q}(i)]$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크집합( $i$ )를 이용하는 차량의 Probe 차량 수( $q_s(i)$ , 건)와 추정교통량( $\hat{q}(i)$ , 대)

$l(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별 연장(km)

②  $CO_2$ 배출량( $LACO_2I$ ; Link Aggregated  $CO_2$  Emission Intensity)

- $CO_2$ 배출량의 집계는 분석 Map의 해당 링크( $k$ )에 포함되는 링크집합( $i$ )의  $CO_2$ 배출량

(g/km) 합산하여 산출함

$$LACO_2I(k) = \frac{\sum_i [COE(i) \times l(i)]}{\sum_i [l(i)]}$$

여기서,  $LACO_2I(k)$ : 분석 Map의 해당 링크( $k$ )의  $CO_2$ 배출량(g/km)

$COE(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별  $CO_2$ 배출량(g/km)

$l(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별 연장(km)

③ 연료소모량(LAFCI; Link Aggregated Fuel Consumption Intensity)

- 연료소모량의 집계는 분석 Map의 해당 링크( $k$ )에 포함되는 링크집합( $i$ )의 연료소모량( $\ell/km$ ) 합산하여 산출함

$$LAFCI(k) = \frac{\sum_i [FCI(i) \times l(i)]}{\sum_i [l(i)]}$$

여기서,  $LAFCI(k)$ : 분석 Map의 해당 링크( $k$ )의 연료소모량( $\ell/km$ )

$FCI(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별 연료소모량( $\ell/km$ )

$l(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별 연장(km)

④ 지체시간(LACDI; Link Aggregated Control Delay Intensity)

- 지체시간의 집계는 분석 Map의 해당 링크( $k$ )에 포함되는 링크집합( $i$ )의 지체시간을 각 해당 링크별 교통량( $q_n(i)$ )을 이용하여 가중 평균한 값으로 산출함

$$LACDI(k) = \frac{\sum_i [CDI(i) \times q_n(i)]}{\sum_i [q_n(i)]}$$

여기서,  $LACDI(k)$ : 분석 Map의 해당 링크( $k$ )의 지체시간(초/대)

$CDI(i)$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )별 지체시간(초/대)

$q_n(i) = [q_s(i), \hat{q}(i)]$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크집합( $i$ )를 이용하는 차량의 Probe 차량 수( $q_s(i)$ , 건)와 추정교통량( $\hat{q}(i)$ , 대)



## 2) 혼잡지표의 시·공간적 집계

- 혼잡지표의 시/공간적 집계를 위해서는 먼저 이용자가 분석하고자 하는 시/공간적 범위가 설정되어야 하며, 설정된 범위를 기준으로 해당 시/공간적 범위 내 각 혼잡지표를 집계함

### ① 혼잡강도(NCI; Network Congestion Intensity)

- 집계 혼잡강도(NCI; Network Congestion Intensity)는 주어진 시/공간적 범위를 기준으로 해당 범위 내 도로구간을 이용한 총 차량이 경험한 총 통행시간대비 혼잡속도로 주행한 차량의 총 통행시간의 비율(%)로 산정함

$$NCI(k) = \frac{\sum_i [CI(i) \times q_n(i) \times l(i)]}{\sum_i [q_n(i) \times l(i)]}$$

여기서,  $CI(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )별 혼잡강도(%)

$q_n(i) = [q_s(i), \hat{q}(i)]$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )를 이용하는

차량의 Probe 차량 수( $q_s(i)$ , 건)와 추정교통량( $\hat{q}(i)$ ), 대)

$l(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )의 연장(km)

$k$ : 주어진 시/공간적 단위

### ② $CO_2$ 배출량( $NCO_2I$ ; Network $CO_2$ Emission Intensity)

- 집계  $CO_2$ 배출량( $NCO_2I$ ; Network  $CO_2$  Emission Intensity)은 주어진 시/공간적 범위를 기준으로 해당 범위 내 도로구간을 이용한 차량이 배출한 평균  $CO_2$ 배출량(g/km) 또는 총  $CO_2$ 배출량(g)으로 산정함

$$NCO_2I(k) = \frac{\sum_i [COE(i) \times l(i)]}{\sum_i [l(i)]} \quad \text{또는} \quad NCO_2I(k) = \sum_i [COE(i) \times l(i)]$$

여기서,  $COE(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )별  $CO_2$ 배출량(g/km)

$q_n(i) = [q_s(i), \hat{q}(i)]$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )를 이용하는

차량의 Probe 차량 수( $q_s(i)$ , 건)와 추정교통량( $\hat{q}(i)$ ), 대)

$l(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )의 연장(km)

$k$ : 주어진 시/공간적 단위

### ③ 연료소모량(NFCI; Network Fuel Consumption Intensity)

- 집계 연료소모량(NFCI; Network Fuel Consumption Intensity)은 주어진 시/공간적 범위를 기준으로 해당 범위 내 도로구간을 이용한 차량의 평균 연료 소모량(  $l$  /km) 또는 총 연료소모량(  $l$  )으로 산정함

$$NFCI(k) = \frac{\sum_i [FCI(i) \times l(i)]}{\sum_i [l(i)]} \text{ 또는 } NFCI(k) = \sum_i [FC(i) \times l(i)]$$

여기서,  $FC(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )별 연료소모량(  $l$  /km)

$q_n(i) = [q_s(i), \hat{q}(i)]$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )를 이용하는 차량의 Probe 차량 수( $q_s(i)$ , 건)와 추정교통량( $\hat{q}(i)$ , 대)

$l(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )의 길이(km)

$k$ : 주어진 시/공간적 단위

### ④ 지체시간(NCDI; Network Control Delay Intensity)

- 집계 지체시간(NCDI; Network Control Delay Intensity)는 주어진 시/공간적 범위를 기준으로 해당 범위 내 도로구간을 이용한 차량의 평균 지체시간(초/대) 또는 총 지체시간(초)로 산정함

$$NCDI(k) = \frac{\sum_i [CD(i) \times q_n(i)]}{\sum_i q_n(i)} \text{ 또는 } NCDI(k) = \sum_i [CD(i) \times \hat{q}(i)]$$

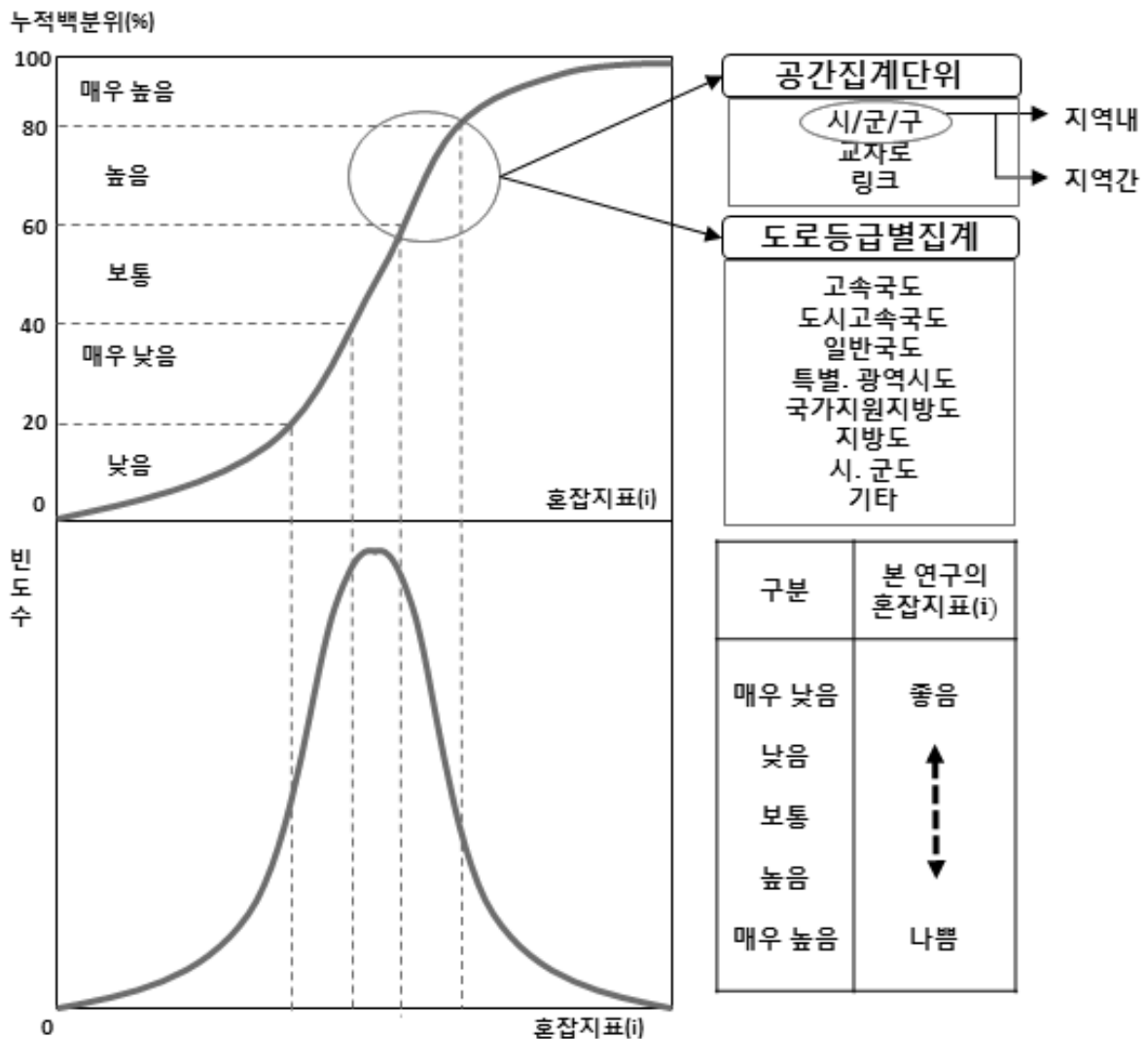
여기서,  $CD(i)$ : 주어진 시·공간적 범위 내 해당 링크 집합( $i$ )별 평균 지체시간(초/대)

$q_n(i) = [q_s(i), \hat{q}(i)]$ : 분석 Map의 해당 링크( $k$ )에 속하는 링크 집합( $i$ )를 이용하는 차량의 Probe 차량 수( $q_s(i)$ , 건)와 추정교통량( $\hat{q}(i)$ , 대)

$k$ : 주어진 시/공간적 단위

#### 다. 혼잡지표의 이용자 분석 방법론 고도화

- 주어진 시/공간적 범위를 기준으로 집계된 각 혼잡지표(i)에 대하여 이용자들이 좀 더 보기 쉽고 효율적으로 분석하기 위한 방법론 구축이 필요함
- 이용자에 의해 설정된 시간적(년/월/일)단위와 공간적 단위(도로구간/도로등급/행정구역)별로 집계된 각 혼잡지표(i)를 0에서부터 100%까지 누적한 백분위 분포를 이용하여 5단계(아주 높음, 높음, 보통, 낮음, 아주 낮음)로 표출하였음
- 이용자가 분석하고자 하는 등급을 선택 할 경우, 선택한 공간적 단위(시/군/구, 교차로, 링크)의 세부 분석결과를 표출하였음



<그림 1-15> 혼잡지표 이용자 분석방법론

### 제3절 교통혼잡비용 추정 연구

#### 1. 개요

- 전국 교통혼잡비용은 2010년 기준 29조원으로 2010년 기준 국가 예산이었던 291조의 9.9%에 해당하며, 승용차 보급이 늘어남에 따라 발생한 비용이 교통혼잡비용의 절반을 차지하고 있음
- 자동차 등록대수는 2004년 1,493만대에서 2014년 1,950만대(1월 기준)로 지속적으로 증가하는 추세를 보이고 있으며, 이로 인해 도로의 교통 혼잡이 갈수록 심해질 것으로 판단됨
- 교통혼잡비용은 교통의 외부불경제 규모를 제시하는 정책지표로, 각종 교통정책(국가기간교통망계획, 도로망계획, 국가철도망계획, 물류기본계획, 대중교통기본계획 등)에서 가장 많이 인용되고 중요한 정책지표로 활용되고 있음
- 따라서 지속적으로 교통혼잡비용에 대한 모니터링/관리/개선이 이루어져야 하나, 교통혼잡과 관련한 기초 데이터가 부족한 이유로 관련 연구 및 대책이 적절히 수립되지 않고 있음
- 과거의 교통혼잡비용 추정 관련 연구는 교통조사의 어려움으로 인해 풍부한 자료를 취득하는 것이 어려워 연구의 범위가 제한적이었음
- 교통혼잡비용을 추정하기 위한 분석자료의 계절적/시간적 특성이 반영되지 못하는 문제가 존재하였으며, 2014년 현재 2012년을 기준으로 한 교통혼잡비용 추정 결과가 제시되고 있어 교통혼잡비용 추정 년도에 대한 시의성 확보가 이루어지지 못함
- 하지만 최근 네비게이션 수요의 증가 및 스마트폰의 보급에 따라 교통분야 빅데이터 이용이 폭발적으로 증가하고 있으며, 차량의 네비게이션에서 제공되는 개별 차량의 이동 궤적 정보를 활용하는 것이 가능해짐
- 즉, 기존 교통데이터의 단편적, 편향적 한계를 극복하여 네트워크 커버리지, 이력/실시간 정보 및 객관성과 일관성을 확보할 수 있게 됨
- 본 연구에서는 네비게이션에서 제공되는 개별 차량 이동궤적 정보를 활용하여 교통혼잡비용을 추정하는 방법론을 개발하고자 하며, 개발된 교통혼잡비용 추정방법론을 기존에 적용해 온 교통혼잡비용 추정방법과 비교/분석 수행하도록 함

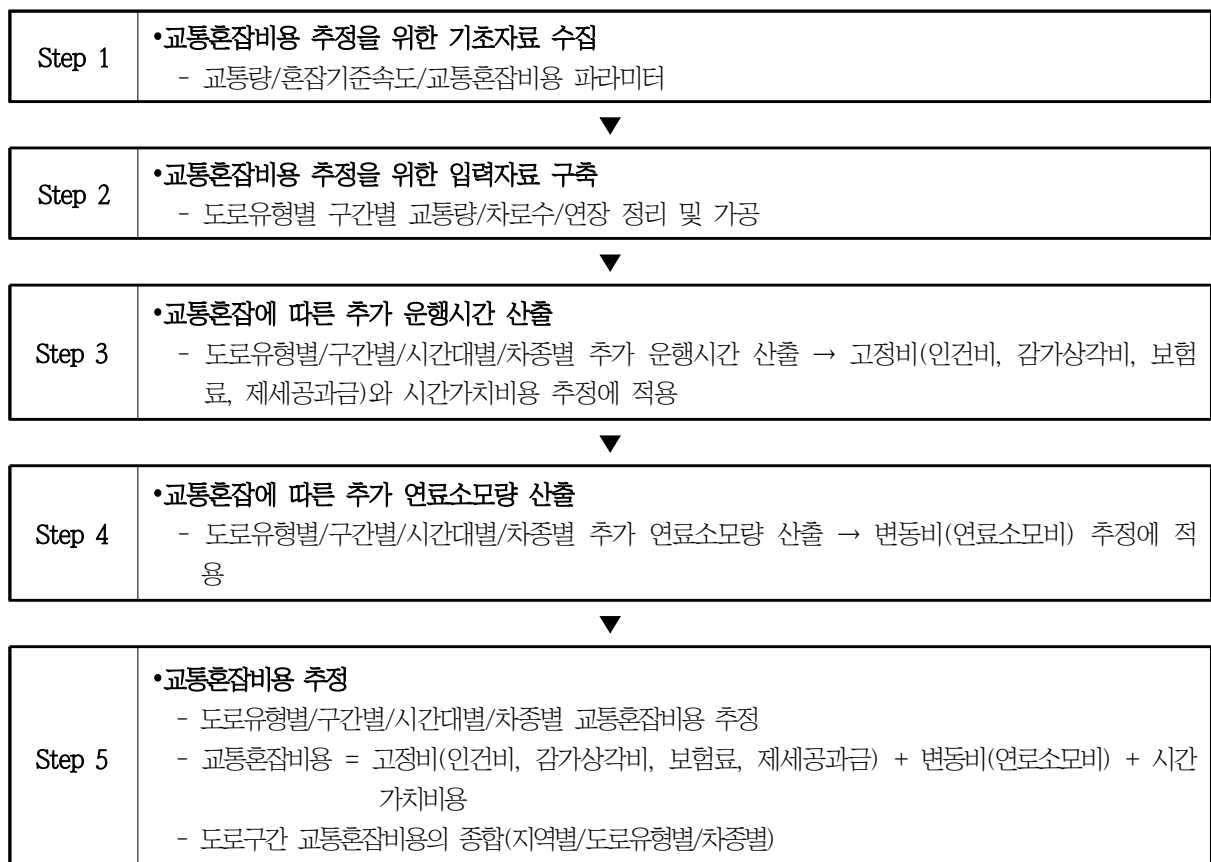
## 2. 기존 교통혼잡비용 추정 방법 및 한계

### 가. 교통혼잡비용의 추정방법

- 교통혼잡비용을 추정하는 방법은 크게 지역간 도로와 도시부 도로에서의 추정방법으로 구분됨
- 모두 혼잡기준속도에 기초하여 추정하는 교통공학적 의미의 교통혼잡비용임

#### 1) 지역간 도로의 교통혼잡비용 추정방법

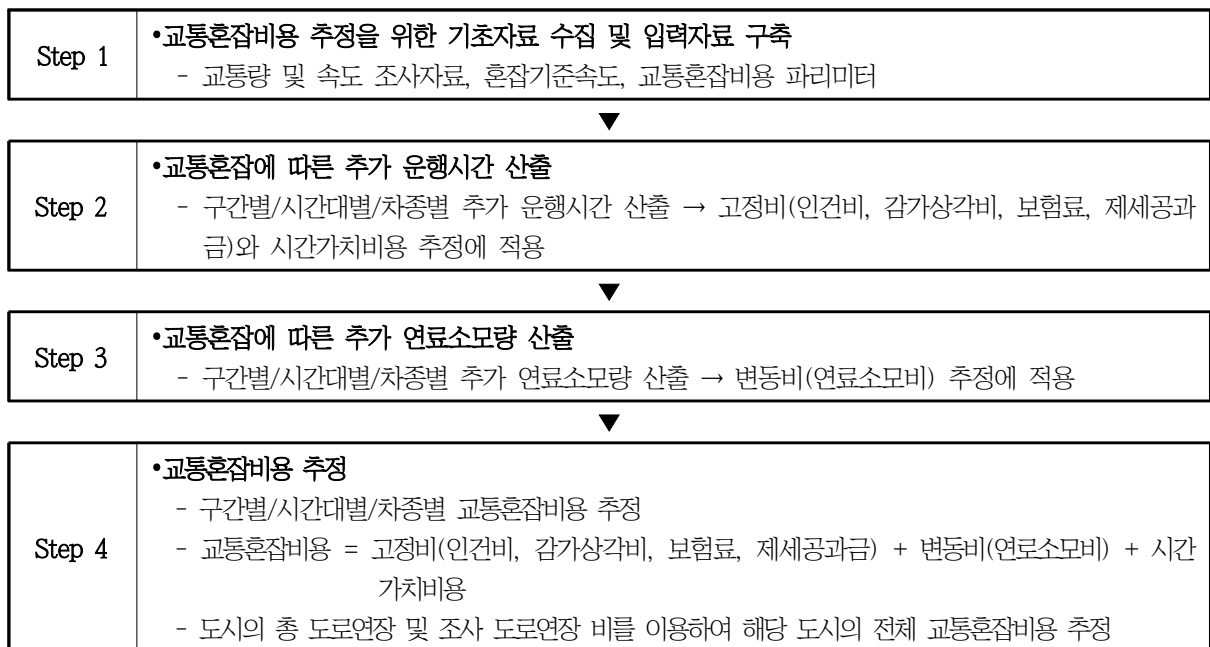
- 지역간 도로의 교통혼잡비용은 도로의 등급별로 혼잡기준속도를 설정한 뒤, 조사지점별 도로용량과 차종별 시간대별 교통량을 이용해 BPR 모형식으로 차량운행속도를 구해 추정하는 방법을 이용함
- 지역간 도로의 교통혼잡비용 프로세스는 <그림 1-16>에 정리하여 제시하였음



<그림 1-16> 지역간 도로의 교통혼잡비용 추정 프로세스

## 2) 도시부 도로의 교통혼잡비용 추정방법

- 도시부 도로의 교통혼잡비용은 지자체별로 교통량과 속도조사가 모두 이루어진 도로에 대하여 교통량, 속도 및 구간 연장 자료를 이용하여 추정하고, 각 지자체별 도시의 총 도로연장과 조사된 구간의 연장 비율을 적용하여 도시의 총 교통혼잡비용을 추정함
- 도시부 도로의 교통혼잡비용 프로세스는 <그림 1-17>에 정리하여 제시하였음



<그림 1-17> 도시부 도로의 교통혼잡비용 추정 프로세스

### 나. 기존 교통혼잡비용 추정방법의 한계 및 시사점

- 기존 교통혼잡비용 추정방법은 분석 데이터 수집의 한계 때문에 여러 가지 제약 조건을 고려하여 혼잡비용을 추정하였으나, 최근 IT기술의 발달로 스마트폰을 비롯하여 차량용 네비게이션을 단말기를 이용하여 개별 차량이동계적 정보를 추출할 수 있게 되어 아래와 같은 과거의 분석 자료(①교통량, ②속도)의 구축에 대한 한계를 극복할 수 있게 되었음
- 교통량 자료의 경우, 기존에는 국토교통부(지역간 도로)와 각 지자체(도시부 도로)의 교통량 조사구간 만을 대상으로 한정적으로 교통혼잡비용을 추정하여 교통량 조사구간 이외의 구간에 대한 교통혼잡비용 추정에 어려움이 존재함. 즉, 해당 기관별 도로교통량 조사지점을 기준으로 교통축 또는 대구간 단위의 대상구간(해당 조사지점의 교통량을 모든 구간에

동일하게 적용함)을 설정하여 각 구간별 교통혼잡비용을 추정하므로 정확도가 떨어지고 전체 전수화된 비용이라고 볼 수 없으며, 또한 관측조사 이외의 구간에 대한 교통량 차이를 제대로 반영하지 못하고 있음

- 속도 자료의 경우, 지역간 도로는 BPR모형식을 이용하여 구간별 통행속도를 추정하기 때문에 실제 현장의 교통상황을 제대로 반영 못하고 있으며, 도시부 도로는 각 지자체에서 상시적으로 조사하고 있는 일부 구간의 평균통행속도 자료만을 이용하여 한정적으로 교통혼잡비용을 추정하고 있어 통행속도 조사구간 이외의 구간에 대한 혼잡비용 추정이 어려움
- 따라서, 현재 개별 차량의 이동궤적 자료를 이용할 수 있게 됨에 따라 교통혼잡비용을 추정하기 위한 분석 데이터의 변화 및 전국 표준노드링크(GIS) 기반의 새로운 교통혼잡비용 추정방법으로의 진화가 필요함

## 1) 분석 데이터 변화

### ① 교통량 및 속도 자료

- 교통혼잡비용의 추정에 있어 기본이 되는 분석 데이터는 교통량과 속도 자료이기 때문에 이 두 자료에 대한 입력 데이터 구축(수집/가공/저장)이 매우 중요하나, 기존 교통혼잡비용 추정 시에는 해당 자료 취득에 한계가 있어 전수화된 교통혼잡비용을 제시하는 데 어려움이 존재하였음
- 하지만, 최근 IT기술의 발달로 인하여 개별 차량이동궤적 자료를 이용할 수 있게 됨에 따라 과거의 분석 자료(교통량, 속도)구축 상의 한계를 극복할 수 있게 되었음
- 교통량 및 속도 자료의 변화에 대한 검토내용은 아래의 <표 1-5>에 제시함

&lt;표 1-5&gt; 교통량 및 속도 자료의 변화에 대한 검토내용

구분	기존 방법	본 연구
교통량	<ul style="list-style-type: none"> <li>- 국토교통부와 각 지자체별 교통량 조사구간을 대상으로만 교통혼잡비용을 추정함</li> <li>· 교통량 조사구간 이외의 구간(연속류의 램프 구간 포함)에 대한 교통혼잡비용 산출이 어려움 → 전수화된 교통혼잡비용의 추정이 어려움</li> </ul>	<ul style="list-style-type: none"> <li>- 기존의 교통량 조사구간 자료 이용</li> <li>- 개별 차량이동궤적 자료 이용 → 구간별 Probe 대수 산출 가능</li> <li>· 본 연구의 교통량 전수화 모형으로 관측 조사구간 이외의 구간에 대한 교통량을 추정함</li> <li>· 교통량 조사구간을 포함하여 그 외의 구간(연속류의 램프 구간 포함)에도 교통혼잡비용 추정이 가능함 → 전수화된 교통혼잡비용의 추정이 가능해짐</li> </ul>
속도	<ul style="list-style-type: none"> <li>- 지역간 도로 : 시간대별 교통량 자료를 이용하여 BPR모형식에 의한 구간별 통행속도를 추정하여 이용함</li> <li>· 분석의 정확도가 떨어짐</li> <li>· 실제 현장의 교통상황을 제대로 반영하지 못함</li> <li>- 도시부 도로 : 각 지자체별 통행속도 조사구간을 대상으로 하는 제한적인 속도자료 이용 → 조사구간 이외의 구간에 대한 자료 취득이 어려움</li> </ul>	<ul style="list-style-type: none"> <li>- 개별 차량이동궤적 자료 이용 → 구간별 통행속도 산출 가능</li> <li>- 모형식이 아닌 실제 통행속도 자료의 이용이 가능</li> <li>· 분석의 정확도 향상이 가능</li> <li>· 실제 현장의 교통상황이 반영된 통행속도 자료 이용이 가능</li> <li>· 조사구간 이외의 구간에 대한 통행속도 자료 취득이 가능</li> </ul>

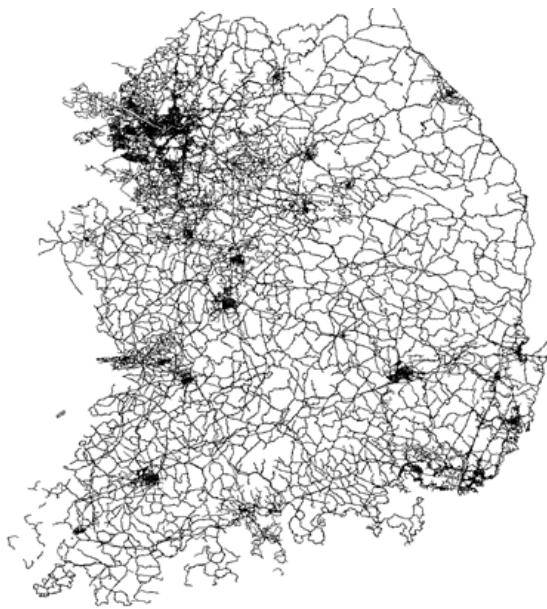
## ② 분석 네트워크 자료

- 기존의 추정 방법을 이용한 교통혼잡비용 추정 시에는 국토교통부(지역간 도로)과 각 지자체(도시부 도로)의 교통량 및 속도 조사구간에 대한 링크 데이터(구간명, 차로수, 연장 등)만을 이용하기 때문에, 별도의 GIS기반 전국 표준노드링크 데이터를 이용할 필요가 없었음
- 하지만, 본 연구에서는 개별 Car Navigation 단말기의 차량이동궤적 정보를 이용하기 위해 전국 표준노드링크 GIS 자료(국토교통부 제공)를 이용해야 함
  - 표준 노드/링크는 교통정보를 수집, 제공하기 위해 교통체계 지능화 사업자가 구축하는 전자전자교통지도로서, 도로관리청이 구축기준에 따라 구축하여 국토해양부장관의 확인을 받은 것을 말함<sup>3)</sup>
  - 노드(Node)는 도로법에 의한 도로로서 교차로, 분기점, 터널입구 등 교통흐름이 변경되는 지점을 전자적으로 형상화한 것을 말함

3) 국토해양부, “지능형교통체계 표준 노드·링크 구축·관리지침”, 2008.



- 링크(Link)는 위의 노드와 노드를 도로선형에 따라 연결하는 도로구간을 전자적으로 형상화한 것을 말함
- 표준 노드/링크체계는 노드정보와 회전정보, 링크정보, 링크부가정보로 구성하며, 각각은 별도의 자료구조를 가지고 있음
- 전국 표준노드링크 자료는 국토교통부에서 2014년 3월에 배포한 자료를 이용하며, 해당 데이터에 관한 설명은 아래의 <그림 1-18>에서 제시함



- 총 링크 255,591개, 총 노드 98,729개로 구성
- 표준 노드/링크 구성체계
  - 노드정보/회전정보/링크정보/링크부가정보
- 8개 Type의 도로등급(ROAD\_RANK)으로 구성
  - 고속국도(101)
  - 도시고속국도(102)
  - 일반국도(103)
  - 특별/광역시도(104)
  - 국가지원지방도(105)
  - 지방도(106)
  - 시/군도(107)
  - 기타(108)

<그림 1-18> 전국 표준노드링크 데이터 설명

## 2) 교통혼잡비용 추정방법의 진화

### ① 개별 차량이동궤적 자료의 이용에 따른 공간적 범위 확대

- 기존에 지역간 도로와 도시부 도로의 교통혼잡비용을 추정하는 방법은 기초자료 취득의 제약으로 인해 두 가지 방법 모두 관측조사 지점 또는 구간에 한해서만 교통혼잡비용을 추정하도록 이루어졌기 때문에, 이는 전수화된 교통혼잡비용이라고 보기 어려움
- 지역간 도로의 경우, 국토교통부의 고속도로/일반국도/지방도 교통량 조사구간에 한하여 교통혼잡비용을 추정함
- 도시부 도로의 경우, 전국 7개의 특별·광역시를 대상으로 각 지자체별로 교통량 및 속도조사가 모두 이루어지고 있는 도로구간 또는 교통축에 한하여 교통혼잡비용을 추정함
- 즉, 교통량 및 속도 조사구간 이외 구간들에 대한 교통혼잡비용의 추정이 어려우며, 특히

(도시)고속도로 등과 같은 연속류 램프 구간에 대한 교통혼잡비용의 추정 또한 어려움

- 그러나, 개별 차량이동계적 자료와 더불어 관측교통량 추가 조사자료를 이용함에 따라 기존의 추정 방법보다 교통혼잡비용을 추정할 수 있는 공간적 범위를 확대하여 전수화에 가까운 교통혼잡비용의 추정이 가능해질 것으로 판단됨
- 교통혼잡비용의 추정을 위한 기초 조사자료와 이에 따른 공간적 범위에 대해 기존 방법과 비교한 내용은 아래의 <표 1-6>에서 제시함

<표 1-6> 기초 조사자료에 따른 공간적 범위 비교

구분	기존 방법	본 연구
지역간 도로	<ul style="list-style-type: none"> <li>○ 교통혼잡비용 추정을 위한 분석자료               <ul style="list-style-type: none"> <li>- 교통량 자료 : 전체 3,390개 관측조사 자료 &lt;국토교통부&gt;                   <ul style="list-style-type: none"> <li>· 고속도로 : 451개</li> <li>· 일반국도 : 1,469개(상시조사 298개, 수시조사 1,171개)</li> <li>· 지방도(국지도 포함) : 1,470개</li> </ul> </li> <li>- 속도 자료 : BPR모형식으로 산출</li> </ul> </li> <li>↓</li> <li>○ 공간적 범위               <ul style="list-style-type: none"> <li>- 대상지역 : 전국(관측조사 구간)</li> <li>- 대상구간 : 고속도로, 일반국도, 지방도</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>○ 교통혼잡비용 추정을 위한 분석자료               <ul style="list-style-type: none"> <li>- 교통량 자료 : Car Navigation의 링크별 Probe 대수, 전체 4,085개 관측조사 자료 &lt;국토교통부&gt;                   <ul style="list-style-type: none"> <li>· 고속도로 : 451개</li> <li>· 일반국도 : 1,469개(상시조사 298개, 수시조사 1,171개)</li> <li>· 지방도(국지도 포함) : 1,470개</li> </ul> </li> <li>- 속도 자료 : Car Navigation의 링크별 통행속도 자료</li> </ul> </li> <li>↓</li> <li>○ 공간적 범위               <ul style="list-style-type: none"> <li>- 대상지역 : 전국</li> <li>- 대상구간 : 고속도로, 일반국도, 지방도</li> </ul> </li> </ul>
도시부 도로	<ul style="list-style-type: none"> <li>○ 교통혼잡비용 추정을 위한 분석자료               <ul style="list-style-type: none"> <li>- 교통량 자료(총 248개), 속도 자료(총 263개)                   <ul style="list-style-type: none"> <li>· 서울 : 교통량(94개구간), 속도(94개구간)</li> <li>· 부산 : 교통량(21개측), 속도(36개측)</li> <li>· 대구 : 교통량(24개측), 속도(24개측)</li> <li>· 인천 : 교통량(40개측), 속도(42개측)</li> <li>· 광주 : 교통량(20개측), 속도(19개측)</li> <li>· 대전 : 교통량(18개측), 속도(18개측)</li> <li>· 울산 : 교통량(31개측), 속도(30개측)</li> </ul> </li> </ul> </li> <li>↓</li> <li>○ 공간적 범위               <ul style="list-style-type: none"> <li>- 대상지역 : 전국 7개의 특별·광역시</li> <li>· 서울/부산/대구/대전/광주/인천/울산</li> <li>- 대상구간 : 각 지자체별 주요 도로구간 또는 교통축</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>○ 교통혼잡비용 추정을 위한 분석자료               <ul style="list-style-type: none"> <li>- 교통량 자료 : Car Navigation의 링크별 Probe 대수, 전체 4,085개 관측조사 자료 &lt;한국교통연구원&gt;                   <ul style="list-style-type: none"> <li>· 주요 도시별 코트라인 조사 : 1,362개 자료 &lt;서울특별시&gt;</li> <li>· 상시 교통량 조사 : 192개 자료 &lt;전국 광역시(서울특별시 제외)&gt;</li> <li>· 광역시별 조사자료 : 총 4,149개</li> </ul> </li> <li>- 속도 자료 : Car Navigation의 링크별 통행속도 자료</li> </ul> </li> <li>↓</li> <li>○ 공간적 범위               <ul style="list-style-type: none"> <li>- 대상지역 : 전국 7개의 특별·광역시, 중·소도시</li> <li>· 서울/부산/대구/대전/광주/인천/울산/중·소도시</li> <li>- 대상구간 : 지자체별 전체 도시고속도로 및 간선도로</li> </ul> </li> </ul>

## ② 상대적인 개념의 혼잡경계속도 적용

- 기존 교통혼잡비용의 추정방법에서 교통혼잡은 교통공학적 의미에서 도로의 서비스수준(지역 간 서비스수준 C, 도시부 서비스수준 D)기준에 따른 지역별 도로등급별로 일률적인 혼잡기준속도를 설정하고, 해당 혼잡기준속도 이하로 운행하는 차량들에 의해서 추가적으로 발생하는 사회적 비용을 교통혼잡비용으로 추정하고 있음
- 기존의 교통혼잡비용 추정 방법에서 적용되고 있는 혼잡기준속도는 아래와 같음
  - 지역간 도로
    - 고속국도 : 4차로 이상 90km/h, 2차로 이상 75km/h
    - 일반국도 : 4차로 이상 80km/h, 2차로 이상 70km/h
    - 지방도 : 60km/h
  - 도시부 도로
    - 자동차전용도로 : 75km/h, 일반도로 : 27km/h
- 기존 혼잡기준속도는 전국의 모든 도로에 대하여 도로등급별로 일률적인 값이 적용되기 때문에 도로구간별 교통 및 운영조건 등에 의해 상대적으로 발생하는 교통혼잡의 상황을 제대로 반영하지 못함
  - 예1: 수도권 도시부(서울시 강남대로)와 지방 도시부(창원시 창원대로)를 대상으로 동일한 도로등급과 차로수 기준의 각 도로구간을 통과하는 이용차량들이 느끼는 혼잡의 정도는 서로 다를 것이기 때문에, 기존 도시부 일반도로의 일률적인 혼잡기준속도 27km/h를 적용하는 것은 무리가 있음
    - 즉, 서울시 도시부 간선도로의 평균통행속도를 감안할 때 매일 강남대로를 통과하는 차량 이용자들은 오히려 혼잡기준속도 27km/h를 혼잡하지 않다고 느낄 수 있으나, 그와는 반대로 창원시 창원대로의 경우에는 혼잡기준속도 27km/h를 많이 혼잡하다고 느낄 수 있기 때문임
  - 예2: 도로의 기하구조 상 물리적으로 속도를 높일 수 없는 구간(대관령, 미시령 등 오르막 경사 구간)을 이용하는 차량들의 통행속도는 다른 평지 도로구간에 비해 전체적으로 현저히 낮은 상태이기 때문에, 기존의 도로등급별 혼잡기준속도를 적용할 경우 교통혼잡비용이 과다 추정될 우려가 있음
- 따라서, 본 연구에서는 Car Navigation의 보급을 통한 개별 차량의 이동궤적 자료를 이용할 수 있게 됨에 따라, 기존의 도로등급별 혼잡기준속도가 아닌 도로구간별 기하구조 및 교통운영조건 등을 반영할 수 있는 상대적인 개념의 혼잡경계속도 도입이 필요함

### ③ 교통혼잡비용 추정방법에 대한 진단 및 개선방향

#### a) 지역간 도로

- 지역간 도로의 교통혼잡비용은 도로유형에 따라 구간별/시간대별/차종별로 주행하는 차량들 중 혼잡기준속도 미만으로 주행한 차량들에 한하여 추가적으로 발생하는 사회적 비용(고정비+변동비+시간가치비용)에 해당하는 교통혼잡에 따른 추가 운행시간과 추가 연료소모량 값을 산출한 후, 각 혼잡비용 항목별 원단위를 곱하여 추정함
- 추가 운행시간은 도로유형에 따른 구간별/시간대별/차종별로 혼잡기준속도 미만으로 운행하는 조사교통량의 운행시간에서 혼잡기준속도로 운행하는 차량의 운행시간을 뺀 산출값을 의미함
- 추가 연료소모량은 도로유형에 따른 구간별/시간대별/차종별로 혼잡기준속도 미만으로 운행하는 조사교통량의 연료소모량에서 혼잡기준속도로 운행하는 차량의 연료소모량을 뺀 산출값을 의미함
- 추가 운행시간과 연료소모량은 지역간 교통혼잡비용을 추정할 때 가장 기본이 되는 입력변수로서 보다 더 정확한 값으로 산출할 필요가 있으며, 이에 기존 방법에 대해 이론적으로 진단하고 개선방향을 제시하고자 함

&lt;표 1-7&gt; 기존 교통혼잡비용 추정방법의 이론적 진단 및 개선방향(지역간 도로)

구분		기존 방법	본 연구
교통 혼잡 비용 추정	추가 운행시간 산출 방법	<ul style="list-style-type: none"> <li>구간별 조사 운행시간은 BPR모형식으로 추정함</li> <li>- 실제 현장의 자료가 아닌 모형식으로 추정하므로 정확도가 떨어짐</li> <li>- 구간별로 1시간 단위의 교통량과 용량으로 통행속도를 추정하므로 실제 현장의 개별 차량들의 운행행태를 제대로 반영하기 어려움</li> </ul>	<ul style="list-style-type: none"> <li>구간별 조사 운행시간은 기존의 BPR모형식이 아닌 실제 현장의 개별 Car Navigation 속도 자료를 이용함</li> <li>- 실제 현장의 자료를 이용하므로 기존 방법보다 정확도가 높아짐</li> <li>- 개별 Car Navigation 속도 자료를 이용하여 수집 건별로 모두 운행시간을 산출하여 적용하므로 개별 차량의 운행행태가 반영되어 있음</li> </ul>
		<ul style="list-style-type: none"> <li>혼잡기준속도에 따른 운행시간은 도로 등급별 일률적인 혼잡기준속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 제대로 반영하지 못함</li> </ul>	<ul style="list-style-type: none"> <li>혼잡기준속도에 따른 운행시간은 본 연구에서 개발한 구간별 상대적인 개념의 혼잡경계속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 반영할 수 있음</li> </ul>

&lt;표계속&gt;

구분		기존 방법	본 연구
교통 혼잡 비용 추정	추가 연료소모량 산출 방법	<ul style="list-style-type: none"> <li>구간별 조사 연료소모량은 해당 구간연장에서 BPR식으로 추정한 운행시간을 나눈 속도 값을 차종별 속도-연료소비모형식에 적용하여 산출함</li> <li>- 실제 현장의 자료가 아닌 모형식으로 추정하므로 정확도가 떨어짐</li> </ul>	<ul style="list-style-type: none"> <li>구간별 조사 연료소모량은 개별 Car Navigation 속도 자료를 차종별 속도-연료소비모형식에 적용하여 산출함</li> <li>- 실제 현장의 자료를 이용하므로 기존 방법보다 정확도가 높아짐</li> </ul>
		<ul style="list-style-type: none"> <li>혼잡기준속도에 따른 연료소모량은 도로 등급별 일률적인 혼잡기준속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 제대로 반영하지 못함</li> </ul>	<ul style="list-style-type: none"> <li>혼잡기준속도에 따른 연료소모량은 본 연구에서 개발한 구간별 상대적인 개념의 혼잡경계속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 반영할 수 있음</li> </ul>

b) 도시부 도로

- 도시부 도로의 교통혼잡비용은 해당 지역에 따라 구간별/시간대별/차종별로 주행하는 차량들 중 혼잡기준속도 미만으로 주행한 차량들에 한하여 추가적으로 발생하는 사회적 비용(고정비+변동비+시간가치비용)에 해당하는 교통혼잡에 따른 추가 운행시간과 연료소모량 값을 산출한 후, 각 혼잡비용 항목별 원단위를 곱하여 추정함
  - 추가 운행시간은 해당 지역에 따라 구간별/시간대별/차종별로 혼잡기준속도 미만으로 운행하는 조사교통량의 운행시간에서 혼잡기준속도로 운행하는 차량의 운행시간을 뺀 산출값을 의미함
  - 추가 연료소모량은 해당 지역에 따라 구간별/시간대별/차종별로 혼잡기준속도 미만으로 운행하는 조사교통량의 연료소모량에서 혼잡기준속도로 운행하는 차량의 연료소모량을 뺀 산출값을 의미함
- 위에서 제시한 추가 운행시간과 연료소모량은 도시부 교통혼잡비용을 추정할 때 가장 기본이 되는 입력변수로서 보다 더 정확한 값으로 산출할 필요가 있기 때문에, 이에 기존 방법에 대해 이론적으로 진단하고 개선방향을 제시하고자 함

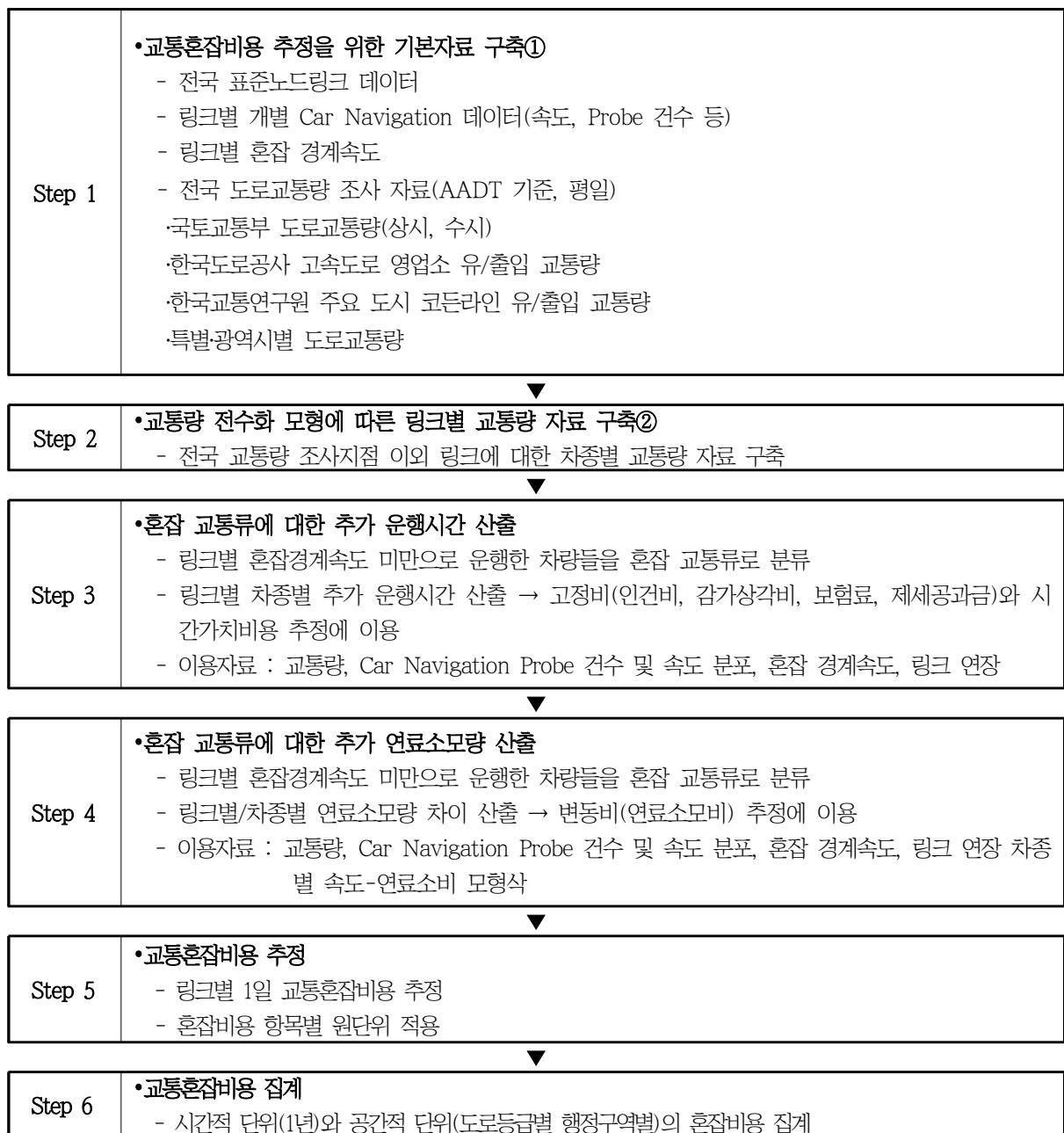
&lt;표 1-8&gt; 기존 교통혼잡비용 추정방법의 이론적 진단 및 개선방향(도시부 도로)

구분		기존 방법	본 연구
교통 혼잡 비용 추정	추가 운행시간 산출 방법	<ul style="list-style-type: none"> <li>•구간별 조사 운행시간은 해당 지역의 구간별 1시간 평균 조사속도를 이용하여 운행시간을 산출함</li> <li>- 1시간 단위의 집계 운행시간이므로 개별 차량들의 운행속도 행태를 반영하기 어려움</li> <li>- 도로 및 교통축 단위로 조사된 일률적인 평균속도를 이용하므로 내부 구간들에 대한 운행 행태를 제대로 반영하지 못함</li> </ul>	<ul style="list-style-type: none"> <li>•구간별 조사 운행시간은 기존의 BPR모형식이 아닌 실제 현장의 개별 Car Navigation 속도 자료를 이용함</li> <li>- 실제 현장의 자료를 이용하므로 기존 방법보다 정확도가 높아짐</li> <li>- 개별 Car Navigation 속도 자료를 이용하여 수집 건별로 모두 운행시간을 산출하여 적용하므로 개별 차량의 운행행태가 반영되어 있음</li> </ul>
		<ul style="list-style-type: none"> <li>•혼잡기준속도에 따른 운행시간은 도로등급별 일률적인 혼잡기준속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 반영하지 못함</li> </ul>	<ul style="list-style-type: none"> <li>•혼잡기준속도에 따른 운행시간은 본 연구에서 개발한 구간별 상대적인 개념의 혼잡경계속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 반영할 수 있음</li> </ul>
	추가 연료소모량 산출 방법	<ul style="list-style-type: none"> <li>•구간별 조사 연료소모량은 해당 지역 조사된 구간별 평균속도 값을 차종별 속도-연료소비모형식에 적용하여 산출함</li> <li>- 도로 및 교통축 단위로 대상구간을 설정하고 있어 내부 구간들에 대한 개별 연료소모량 배출 행태를 반영하지 못함</li> </ul>	<ul style="list-style-type: none"> <li>•구간별 조사 연료소모량은 개별 Car Navigation 속도 자료를 차종별 속도-연료소비모형식에 적용하여 산출함</li> <li>- 실제 현장의 자료를 이용하므로 기존 방법보다 정확도가 높아지며, 개별 차량의 연료소모량 배출 행태도 반영할 수 있음</li> </ul>
		<ul style="list-style-type: none"> <li>•혼잡기준속도에 따른 연료소모량은 도로등급별 일률적인 혼잡기준속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 반영하지 못함</li> </ul>	<ul style="list-style-type: none"> <li>•혼잡기준속도에 따른 연료소모량은 본 연구에서 개발한 구간별 상대적인 개념의 혼잡경계속도를 적용하여 산출함</li> <li>- 구간별 도로/교통/운영조건 등에 의해 발생하는 상대적인 교통혼잡의 차이를 반영할 수 있음</li> </ul>

### 3. Car Navigation 자료 기반의 교통혼잡비용 추정 방법론 개발

#### 가. 교통혼잡비용 추정 프로세스 정립

- 본 연구에서의 교통혼잡비용 추정은 지역간 도로와 도시부 도로의 구분 없이 하나의 교통혼잡비용 추정 방법론을 이용하며, 전체적인 프로세스는 아래의 <그림 1-19>와 같이 정리함



<그림 1-19> 본 연구의 교통혼잡비용 추정 프로세스



## 나. 단계별 교통혼잡비용 추정 세부내용

### 1) 교통혼잡비용 추정을 위한 기본자료 구축

- 본 연구에서 교통혼잡비용 추정을 위해 필요한 기본자료는 ①전국 표준노드링크 데이터, ② 링크별 개별 Car Navigation 데이터(속도, Probe 건수 등), ③링크별 혼잡 경계속도, ④전국 도로교통량 조사자료(AADT 기준, 평일)를 구축해야 함
- 각 항목별 특성 및 형태는 아래와 같음

#### ① 전국 표준노드링크 데이터

- 전국 표준노드링크 데이터는 2014년 3월에 국토교통부에서 배포한 자료를 이용함
  - 표준 노드/링크 구성
    - 총 링크개수 (255,591개), 총 노드개수 (98,729개)
    - 전체 노드링크 정보 : 노드정보/회전정보/링크정보/링크부가정보
  - 도로등급(ROAD\_RANK) : 8개 Type으로 구성
    - 고속국도 (101), 도시고속국도 (102), 일반국도 (103), 특별/광역시도 (104), 국가지원지방도 (105), 지방도 (106), 시/군도 (107), 기타(108)

#### ② 링크별 개별 Car Navigation 데이터

- 본 연구의 혼잡경계속도 판정을 위해서는 개별 Car Navigation 차량 속도 데이터를 이용하며, 자료 형태는 아래와 같음
  - 시간적 범위 : 2013년 10월 1일 ~ 2014년 1월 31일, 4개월치 이력자료
  - 공간적 범위 : 전국
  - 수집 Column(총 5개) : ①auth\_key/②link\_time/③from\_link/④to\_link/⑤speed

Column	설명	Data Type
auth_key	단말기 ID	Datetime
link_time	수집 시간	Integer
from_link	진입 링크	Integer
to_link	진출 링크	Integer
speed	속도	Integer

### ③ 링크별 혼잡경계속도

- 본 연구에서 적용하는 혼잡경계속도는 개별 Car Navigation 데이터를 이용하여 현재 고도와 중인 전국 교통혼잡지도 시스템 내 링크별 혼잡경계속도 판정 알고리즘을 수행하여 구축함
- 평일 기준 링크별 혼잡경계속도 산출을 통해 혼잡비용 추정을 위한 입력자료로 이용함

### ④ 전국 도로교통량 조사 자료

- 본 연구에서의 교통혼잡비용을 추정하는 데 있어 이용하는 전국 도로교통량 조사자료는 아래 기관에서 제공하며, 링크별 차종별 평일 기준 AADT(연평균 일교통량) 값을 적용함
  - 국토교통부 도로교통량 조사
    - 고속도로 : 451개 자료
    - 일반국도 : 1,469개 자료(상시조사 298개, 수시조사 1,171개)
    - 지방도(국지도 포함) : 1,470개 자료
  - 한국도로공사 고속도로 영업소 유/출입 조사
    - 총 조사자료 : 695개 자료(민자 영업소 40개 포함)
  - 한국교통연구원 코든라인 조사
    - 총 조사자료 : 1,362개 자료
  - 특별·광역시별 도로교통량 조사
    - 서울시 : 총 192개 자료
    - 그 외 기타 광역시 : 총 4,149개 자료

## 2) 교통량 전수화 모형에 따른 링크별 교통량 자료 구축

### ① 교통량 전수화 개요

- 본 연구에서는 해당 링크의 Car Navigation Probe 대수와 (인접 도로구간의) 관측교통량 자료를 이용하여, 교통량이 관측되는 링크 이외의 미관측 링크 구간에 대한 차종별 교통량을 추정하는 교통량 전수화 모형을 개발하고자 함
- 즉, Car Navigation Probe가 수집되는 전국 표준노드링크의 전체 링크 중 1일 수집되는 표본 수가 25개 미만인 링크를 제외하고 나머지 링크들에 대하여 교통량 전수화(평일기준 AADT)를 수행하고자 함

- 본 연구의 교통량 전수화는 지역간 고속도로와 그 이외의 도로로 구분하여 수행하며, 이를 통해 지역간 고속도로와 기타 도로의 네비게이션 이용률 차이를 고려할 수 있음
- 교통량과 Car Navigation Probe의 행태는 혼재된 상태(Mixed state)일 가능성이 매우 높기 때문에, 교통량 전수화 모형은 전국 도로의 전역해가 아닌 도로구간별 해를 탐색함

## ② 교통량 전수화 모형 개발

- 본 연구의 교통혼잡비용을 추정하기 위해, 개별 Car Navigation 속도 데이터가 수집되는 전국 표준노드링크 구간에서 관측교통량 링크구간을 제외한 링크구간에 대하여 인접 링크의 관측교통량과 해당 링크의 Car Navigation Probe 대수를 이용하여 교통량 전수화를 수행함

## ③ 교통혼잡모형 추정을 위한 입력자료 구축

- 교통혼잡모형 추정을 위한 입력자료는 총 6가지로 분류됨
  - 링크 데이터(5개 Column) : 진입 링크 ID, 차로수, 도로등급, 노선번호, 연장(km)
  - 링크별 혼잡경계속도(1개 Column) : 혼잡 경계속도(km/h)
  - 링크별 교통량(4개 Column) : 승용차, 버스, 트럭, 전체 교통량(대/일)
  - 링크별 Probe 데이터(151개 Column) : 프로브 대수(건/일), 속도 분포건수(건/일)
  - 링크별 행정구역 구분(3개 Column) : 광역시/도, 시/군/구, 읍/면/동 단위 ID
- 링크별 교통량 전수화 자료는 본 연구에서 적용하는 교통량 전수화 알고리즘을 이용하여 관측조사지점 구간을 제외한 링크구간(JC와 IC 램프구간 포함)에 대해 차종별 전수화 교통량 자료를 구축하도록 함

&lt;표 1-9&gt; 교통혼잡비용 추정을 위한 링크별 입력데이터 테이블 정의서

Column		설명	Data Type
링크데이터	LINK_ID	진입 링크 ID	Varchar
	LANES	차로수	Integer
	ROAD_RANK	도로등급	Integer
	ROAD_NO	노선번호	Integer
	LENGTH	연장(km)	Double
혼잡경계속도	Vb_speed	혼잡 경계속도(km/h)	Integer
교통량 (평일기준, AADT)	Auto_vol	승용차 교통량(대/일)	Double
	Bus_vol	버스 교통량(대/일)	Double
	Truck_vol	트럭 교통량(대/일)	Double
	Total_vol	전체 교통량(대/일)	Double
Probe (평일기준,AADT)	Porbe_cnt	프로브 대수(건/일)	Double
	1~150km/h	속도 분포건수(건/일)	Integer
행정구역 구분	Si_do	광역시/도 단위 ID	Varchar
	Si_gun_gu	시/군/구 단위 ID	Varchar
	Eup_myeon_dong	읍/면/동 단위 ID	Varchar

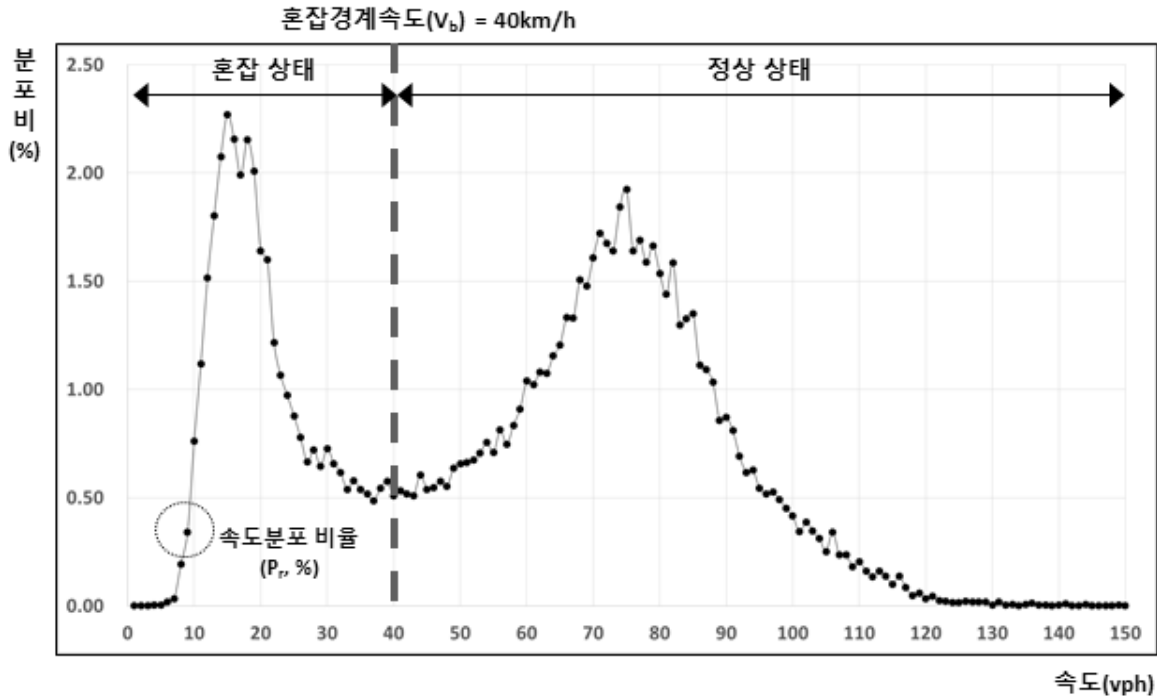
### 3) 링크별 교통혼잡비용 추정 방법론 개발

#### ① 개발 개요

- 기존의 연구방법은 관측지점을 제외한 지점의 교통량 및 속도에 관한 정보를 취득할 수 없었기 때문에, 관측지점에 있는 교통량 및 속도 정보를 동일구간 전체로 확장하여 교통혼잡비용을 추정하는 방법을 이용하였음
- 고속도로의 경우 구간단위가 IC 혹은 JCT로 구분되기 때문에 교통정보 확장에 대한 편차가 적을 수 있으나, 일반국도 이하 등급이 낮은 도로는 교차로 및 신호체계 등이 큰 영향을 미치기 때문에 대구간 단위의 교통혼잡비용 추정 시 각 세부구간별로 편차가 존재할 가능성이 높음
- 본 연구방법에서는 세부구간별 교통량 및 속도자료를 취득할 수 있으므로, 기존에 비해 교통상황을 좀더 현실적으로 반영할 수 있는 세부구간 교통혼잡비용 추정방법을 제시함

## ② 혼잡교통류의 추가 운행시간 및 연료소모량 산출

- 본 연구의 교통혼잡비용 추정을 위해 우선적으로 링크별 Car Navigation Probe 차량의 통행속도 프로파일 분포도가 필요함



<그림 1-20> 해당 링크의 통행속도 분포도

- 교통혼잡으로 인해 추가적으로 발생하는 운행시간은, 해당 링크에서 혼잡경계속도 미만으로 운행하는 차량의 속도분포를 이용해 산출한 링크 운행시간에서 혼잡경계속도의 운행시간을 뺀 값으로 산출함
- 해당 링크의 추가 운행시간 = SUM[차량 속도분포( $j$ )의 통행시간 - 해당 링크 혼잡 경계속도( $V_b$ )의 통행시간]
  - 차량 속도분포( $j$ ) < 혼잡 경계속도( $V_b$ )
  - 혼잡교통류 차량 속도분포( $j$ )의 통행시간( $TT_j$ ) = 해당 링크길이(L)/차량 속도분포( $j$ )
  - 혼잡 경계속도의 통행시간( $tt_{V_b}$ ) = 해당 링크길이(L)/해당 링크의 혼잡 경계속도( $V_b$ )
  - 추가 운행시간 산출

$$TT_{addi} = \sum_{j=1}^{v_b-1} [(TT_i^j - tt_{v_b^i}) \times P_{ri}^j]$$

$$TT_i^j = L_i / j_i, \quad tt_{v_b^i} = L_i / V_{bi}$$

여기서,  $TT_{add}$ : 교통혼잡에 따른 추가 운행시간(시)

$TT$ : 속도 분포( $j$ )에 따른 운행시간(시)

$tt_{v_b}$ : 혼잡경계속도( $v_b$ )의 운행시간(시)

$P_r$ : 속도 분포에 따른 프로브차량 비율(0~100%)

$L$ : 링크연장(km),  $i$ : 링크( $i$ ),  $j$ : 평일기준 1일 속도분포(1~150km/h)

- 교통혼잡으로 인해 추가적으로 발생하는 연료소모량은, 해당 링크에서 혼잡경계속도 미만으로 운행하는 차량의 속도분포를 이용해 산출한 링크별 연료소모량에서 혼잡경계속도의 링크별 연료소모량을 뺀 값으로 산출함
- 해당 링크의 추가 연료소모량 = SUM[차량 속도분포( $i$ )의 연료소모량 - 해당 링크 혼잡 경계속도( $V_b$ )의 연료소모량]
- 차종별 속도분포( $i$ ) < 혼잡경계속도
- 차종별 속도분포( $i$ )의 연료소모량 =  $f_{\text{연료-소비모형식}}(\text{차종별 속도분포}(i))$
- 혼잡 경계속도( $V_b$ )의 연료소모량 =  $f_{\text{연료-소비모형식}}(\text{혼잡 경계속도}(V_b))$

$$FC_{addi,k} = \sum_{j=1}^{v_b-1} [(FC_{i,k}^j - fc_{v_b^i,k}) \times P_{ri,k}^j]$$

$$FC_{i,k}^j = \beta_{1,k} / j_{i,k} + \beta_{2,k} \times (j_{i,k})^2 + c_k,$$

$$fc_{v_b^i,k} = \beta_{1,k} / v_{bi,k} + \beta_{2,k} \times (v_{bi,k})^2 + c_k$$

여기서,  $FC_{add}$ : 혼잡에 따른 추가 연료소모량( $l$ )

$FC$ : 차량 속도분포에 따른 연료소모량( $l$ )

$fc_{v_b}$ : 혼잡경계속도( $v_b$ )에 따른 연료소모량( $l$ )

$k$ : 차종(승용차, 버스, 트럭)

$\beta_1, \beta_2, c$ : “차종별 차량운행비의 산정”(한국교통연구원, 1998)산출계수값 적용

### ③ 차량운행비용(고정비+변동비) 추정

## a) 고정비(시간당운행비)

- 차량운행비용은 고정비 및 변동비로 구분되며, 비용의 각 항목들은 분석 기준년도에 대한 각 항목의 시간당 금액인 원단위 값으로 환산하여 산출함. 차량운행비용 중 고정비는 도로 구간 상에서 교통 혼잡 여부와는 관계없이 항상 일정하게 지출되는 비용을 의미함
- 버스와 화물차의 경우, 도로 구간에 교통혼잡이 발생함에 따라 차량 운행속도가 감소할 때에는 그에 따라 더 많은 차량을 투입해야 할 필요가 있기 때문에 추가적인 고정비(감가상각비, 보험료, 제세공과금, 운전사 인건비)가 발생함. 그러나, 승용차의 경우는 교통혼잡이 발생함에 따라 이루어지는 대응이 버스 및 화물차와 다르기 때문에 승용차의 고정비는 제외함
- 결과적으로 고정비는 기존 교통혼잡비용 추정 방법에서 적용하는 원단위를 동일하게 적용하되, 소비자 물가지수를 고려하여 2013년 기준 고정비 항목별 원단위 값으로 환산해서 적용함
  - 고정비(시간당운행비) = 인건비 + 감가상각비 + 보험료 + 제세공과금
  - 차종별 고정비 = 차종별 추가 운행시간 × 차종별 고정비(인건비, 감가상각비, 보험료, 제세공과금)의 원단위 × 차종별 교통량
  - 단, 고정비는 버스와 트럭의 비용으로 산정하도록 하며, 승용차의 고정비는 제외함

## b) 변동비(연료소모비)

- 차량운행비용의 대부분은 변동비로, 연료비, 유지정비비, 엔진오일비, 타이어 마모비 등이 포함됨. 그러나 연료비를 제외한 항목은 교통혼잡 발생 시 비용의 변화폭이 미미하기 때문에, 본 연구의 교통혼잡비용 추정에서는 이를 제외하고 연료비만을 분석대상으로 함
- 교통혼잡의 발생에 따라 차량의 통행속도가 감소하게 되면 차량 연료소비는 증가하게 되므로, 변동비는 차량 운행속도와 유류 가격의 영향을 받음
- 본 연구의 연료소모비는 기존 교통혼잡비용 산정방법에서 적용하는 「차량운행비 추정에 관한 연구」의 속도-연료소비모형(한국교통연구원, 1998)을 동일하게 적용하여 산출함
  - 차종별 변동비(연료소모비) = 차종별 추가 연료소모량 × 차종별 연료소모비 단가 원단위 × 차종별 교통량 × 링크길이

## ④ 시간가치비용 추정

- 시간가치비용은 교통 혼잡이 발생함에 따라 수행하지 못하게 되는 경제활동의 손실분을 의미하며, 대상 분석구간별로 차량이 교통 혼잡을 경험함에 따라 추가적으로 발생하는 통행시간을 시간가치로 환산하여 산정함
- 차종별 이용자(승용차, 버스)의 시간가치비용은 한국개발연구원의 「도로·철도 부문사업의 예비타당성조사 표준지침 수정·보완 연구(제5판)」에서 제시한 차종별 차량 1대당 평균 통행시간가치 원단위를 적용하되, 소비자 물가지수를 고려하여 2013년 기준의 시간가치 원단위 값으로 환산하여 적용함
- 혼잡 통행시간은 해당 링크에서 시간대별 차종별로 교통 혼잡을 경험한 차량들의 통행시간을 의미하며, 기준 통행시간은 해당 링크의 혼잡경계속도를 기준으로 하여 산출한 통행시간을 의미함
- 차종별 시간가치비용 = 차종별 운행시간 차이 × 차종별 통행시간가치 × 차종별 교통량

#### ⑤ 링크단위 교통혼잡비용 추정

- 링크단위의 1일 교통혼잡비용 추정은 차량운행비용의 고정비(시간당운행비)와 교통 혼잡이 발생함에 따라 발생하는 추가적 통행시간의 시간비용 및 연료소모비용의 합을 통해 이루어짐

$$TCC = \sum_{i=1}^n \sum_{k=1}^3 V_{i,k} \times [(FCC_{i,k} \times L_i \times \max(FC_{addi,k}, 0)) + ((OC_{i,k} + ATV_{i,k}) \times TT_{addi,k})]$$

여기서,  $TCC$ : 일단위 교통혼잡비용,  $ATCC$ : 년단위 교통혼잡비용

$V$ : 링크별 (전수화)교통량(대),  $FCC$ : 차종별유류비

$L$ : 링크연장(km),  $OC$ : 시간당 운행비

$ATV$ : 차량당 평균시간가치

- 1년 단위의 링크단위 교통혼잡비용을 추정하는 경우, 구간별 1일 교통혼잡비용 추정치에 주말 및 공휴일을 제외한 평일 일수(270일)를 곱하여 산출함

$$ATCC = TCC \times 270\text{일}$$

#### 라. 교통혼잡비용 집계방법론 개발

- 위에서 추정된 링크단위 1일 교통혼잡비용을 주어진 공간적 범위(도로구간/도로등급/행정구



역별)에 따라 집계하는 방법론을 개발하고자 함

- 따라서, 본 연구의 링크별 교통혼잡비용의 공간적 집계 방법(NTCC, Network Traffic Congestion Cost)은 아래와 같이 두 가지 형태로 개발함
  - 해당 공간적 범위 내 전체 1일 혼잡비용을 산정할 경우

$$NTCC(k) = \sum_{i=1}^n [TCC(i)]$$

여기서,  $NTCC(k)$ : 주어진  $k$  단위의 공간적 범위에서 발생하는 총 교통혼잡비용(원/일)

$TCC(i)$ : 해당 범위 내 링크  $i$ 에서 발생하는 교통혼잡비용(원/일)

- 해당 공간적 범위 내 차량 1대당 1일 혼잡비용을 산정할 경우

$$NTCC(k) = \frac{\sum_{i=1}^n [(TCC(i) \times q(i))]}{\sum_{i=1}^n [q(i)]}$$

여기서,  $NTCC(k)$ : 주어진  $k$  단위의 공간적 범위에서 발생하는 차량당 평균 교통혼잡비용(원/일/대)

$TCC(i)$ : 해당 범위 내 링크  $i$ 에서 발생하는 차량당 평균 교통혼잡비용(원/일/대)

$q(i)$ : 해당 범위 내 링크  $i$ 의 교통량(대/일)

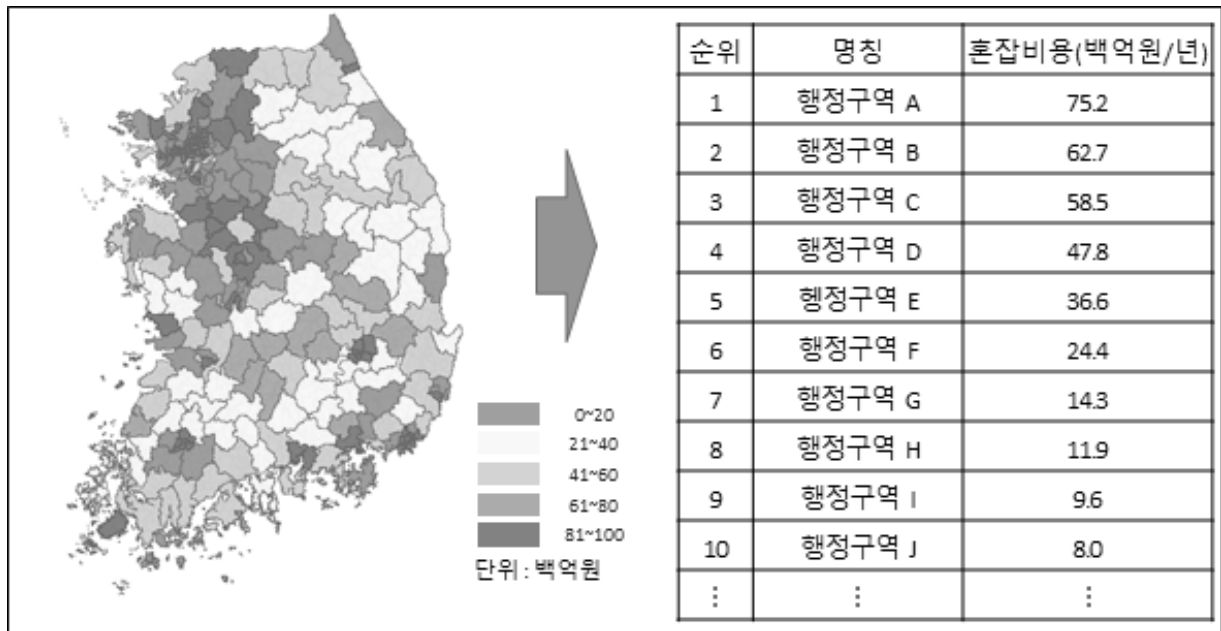
- 기존 교통혼잡비용의 경우, 1년 단위(270일, 평일 기준)로 행정구역(광역시/도, 시/군/구)별 도로등급별 차종별 혼잡비용 산출결과를 정기적으로 제시하고 있음
- 본 연구 역시 기존의 경우와 같이 산출한 링크단위 1일 교통혼잡비용을 시/공간적 범위 내에서 1년 단위(270일, 주말 및 공휴일 제외 평일기준)로 집계하여 행정구역별/도로별 혼잡비용을 산출함

$$NTCC(k) = \sum_{i=1}^{270} \sum_{j=1}^n (TCC_{ij})$$

여기서,  $NTCC(k)$ : 주어진  $k$  단위의 공간적 범위(행정구역별/도로별)에서 발생하는 총 교통혼잡비용(원/년)

$TCC_{ij}$ : 해당 범위 내  $i$ 일에 링크  $j$ 에서 발생하는 1일 교통혼잡비용(원/일)

$i$  : 일 단위( $i=1, \dots, 270$ )

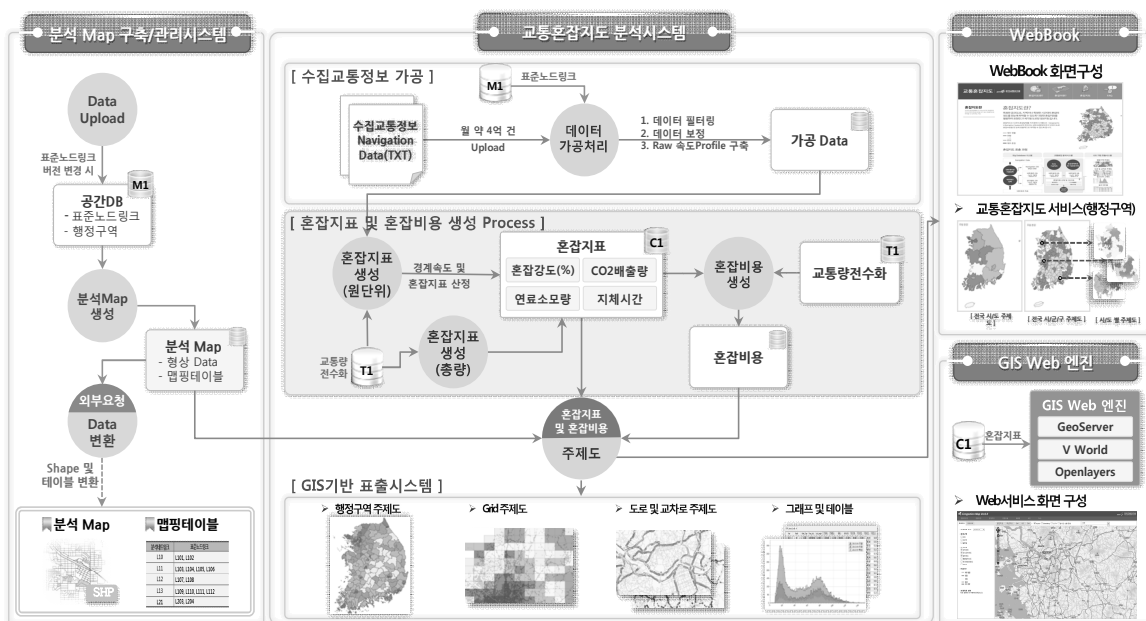


<그림 1-21> 전국 행정구역별/도로별 혼잡비용 집계 결과(예시)

## 제4절 전국 교통혼잡지도 시스템 구축

### 1. 전체시스템 구성

- 전국교통혼잡지도 시스템은 분석맵 관리시스템, 교통혼잡지도 분석시스템, Web service부분으로 구성
- 분석맵 관리시스템은 분석맵을 구축하는 편집기능, 검증기능, 관리기능으로 구성
- 교통혼잡지도 분석시스템은 데이터 가공 프로세스, 혼잡지표 및 혼잡비용 생성 프로세스, GIS기반 주제도 표출시스템으로 구성
  - 데이터 가공 프로세스는 수집교통정보의 단말기 별 1차 필터링, 단말기 별 데이터 보정, Raw 속도 Profile 생성 단계이며, 표준노드링크 단위로 Output이 생성되도록 시스템 개발
  - 혼잡지표 및 혼잡비용 생성 프로세스는 수집 데이터의 Probe 및 관측교통량을 기반으로 교통량 전수화 후 혼잡지표(총량) 및 혼잡비용을 산정하며, 분석맵 단위로 Output 생성
  - 주제도 표출시스템은 표준노드링크, 분석맵, 행정구역, Grid 단위로 표출하도록 개발
- 전국교통혼잡지도 분석시스템에 의하여 생성된 다양한 콘텐츠는 GIS Web 엔진을 기반으로 Web 서비스함



<그림 1-22> 전체 시스템 구성

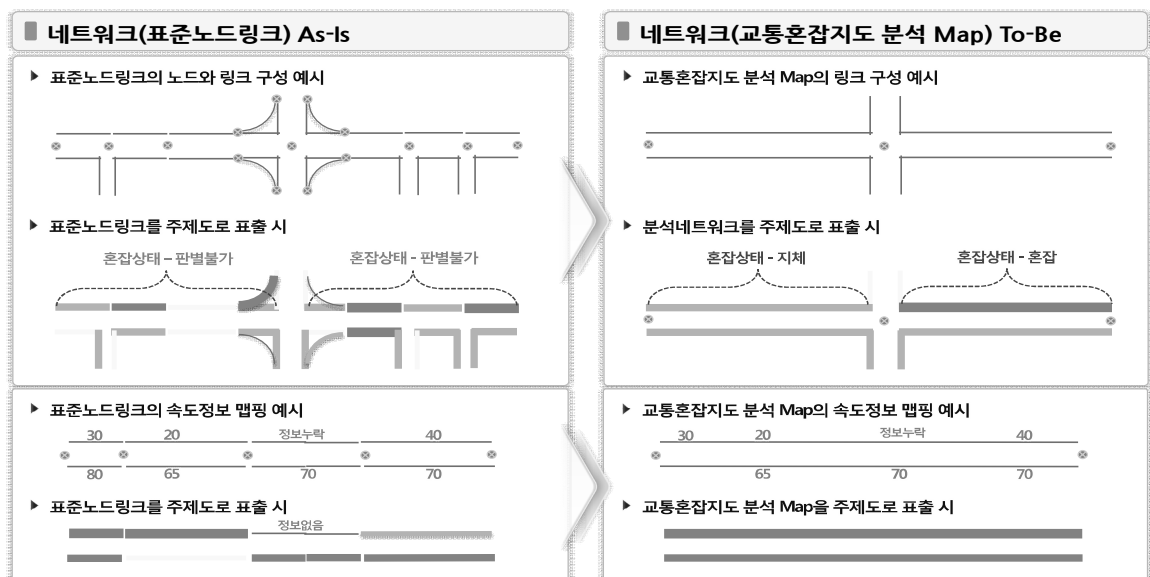
## 2. 전국교통혼잡지도 분석 Map 구축

### 가. 교통혼잡지도 분석 Map 개요

- 전국교통혼잡지도 분석 Map이란? 표준노드링크가 가지고 있는 분석용 네트워크로서의 문제점 보완과 교통혼잡지도의 분석 및 활용성의 극대화를 위하여 구축한 Map을 전국 교통혼잡지도 분석 Map이라 함

#### 1) 분석 Map의 필요성

- 표준노드링크는 동일도로구간이 교차로, 속성변환지점(속도, 차선수 등), 행정경계등으로 인하여 링크가 상세하게 분할되어 있음
- 동일도로구간에 링크가 여러 개로 분할되어 있고, 해당 링크의 혼잡지표의 값이 각기 차이가 발생하기 때문에 동일도로구간의 혼잡여부를 판별하기 힘들
- 주요도로의 등급이 별도로 구분되지 않아 주요구간 위주의 분석이 힘들
- 동일도로구간에 대한 혼잡여부를 인지하기 쉽도록, 동일도로구간에서 여러 개로 분할되어 있는 링크를 병합하여 주요 도로구간의 혼잡여부를 한눈에 파악 할 수 있도록 구축



<그림 1-23> 표준노드링크 및 분석 Map 비교

## 2) 구축 범위

## ○ 공간적 범위

- 대상범위 : 전국권
- 좌표 계 : TM 중부원점 (타원체 GRS80)



&lt;그림 1-24&gt; 구축 범위

## ○ 내용적 범위

- 교통혼잡지도 분석 Map 편집 및 관리 시스템 개발
- 주요도로 이상의 표준노드링크의 링크ID와 교통혼잡지도 분석 Map 링크 ID의 맵핑테이블 구축

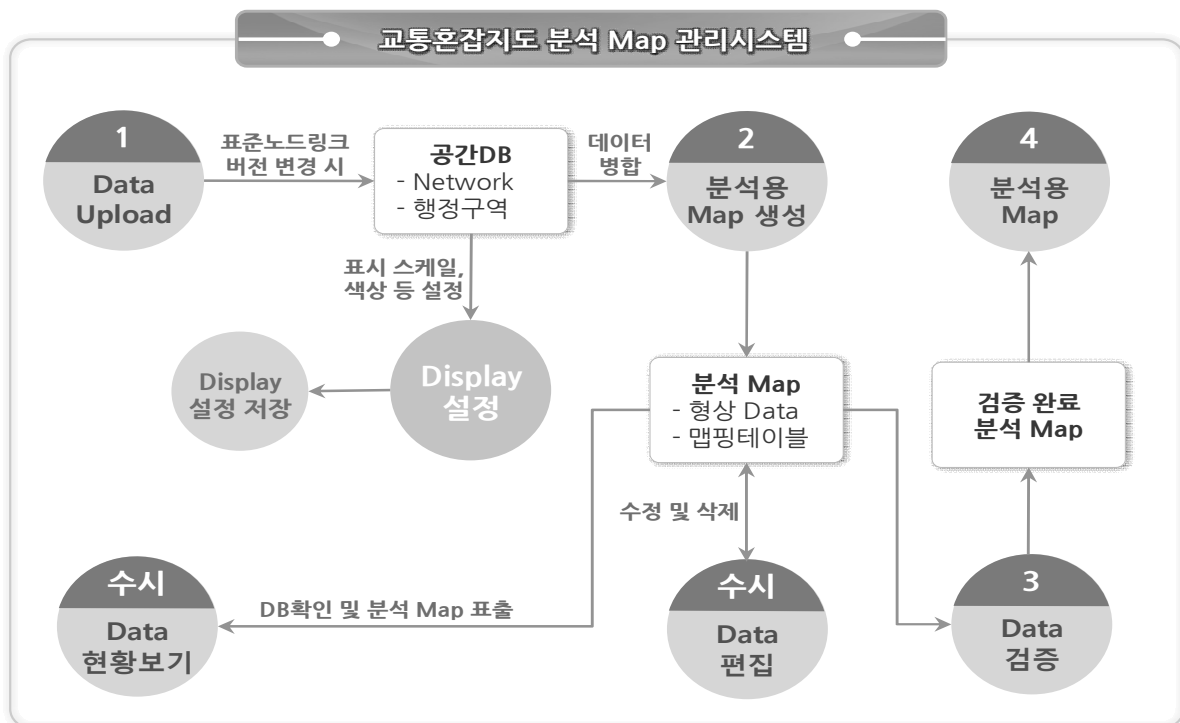
&lt;표 1-10&gt; 도로등급 별 작업 내역

도로등급 구분	작업내역
고속국도	<ul style="list-style-type: none"> <li>- 고속도로의 연결로(IC,JC)와 연결로 사이 링크 병합</li> <li>- 고속도로 본선과 연결로 분리</li> </ul>
도시고속국도	<ul style="list-style-type: none"> <li>- 도시고속국도의 연결로(램프)와 교차지점 사이, 또는 연결로와 연결로 사이 링크 병합</li> <li>- 도시고속국도 본선과 연결로 분리</li> </ul>
일반국도	<ul style="list-style-type: none"> <li>- 일반국도 이하 도로는 주요교차로와 주요교차로 사이 링크 병합</li> <li>- 본선과 연결로 분리</li> <li>- 특별·광역시도 및 시·군도는 주요도로(새주소 기준의 대로와 로 이상의 도로 구축)와 그 외 도로 구분</li> </ul>
국가지원지방도	
특별·광역시도	
시·군도	
도로시설물	<ul style="list-style-type: none"> <li>- 도로시설물 별도</li> </ul>

### 3) 분석 Map 관리시스템 개발

#### ① 시스템 구성

- 교통혼잡지도 분석Map 관리시스템은 데이터 구축 시 링크병합 및 맵핑테이블 자동생성이 가능한 편집기능, 작업자가 효율적으로 작업할 수 있도록 환경 셋팅 및 작업지역 표시 등이 지원되는 데이터 관리 기능, 데이터의 정확도를 위한 검증 기능으로 구성



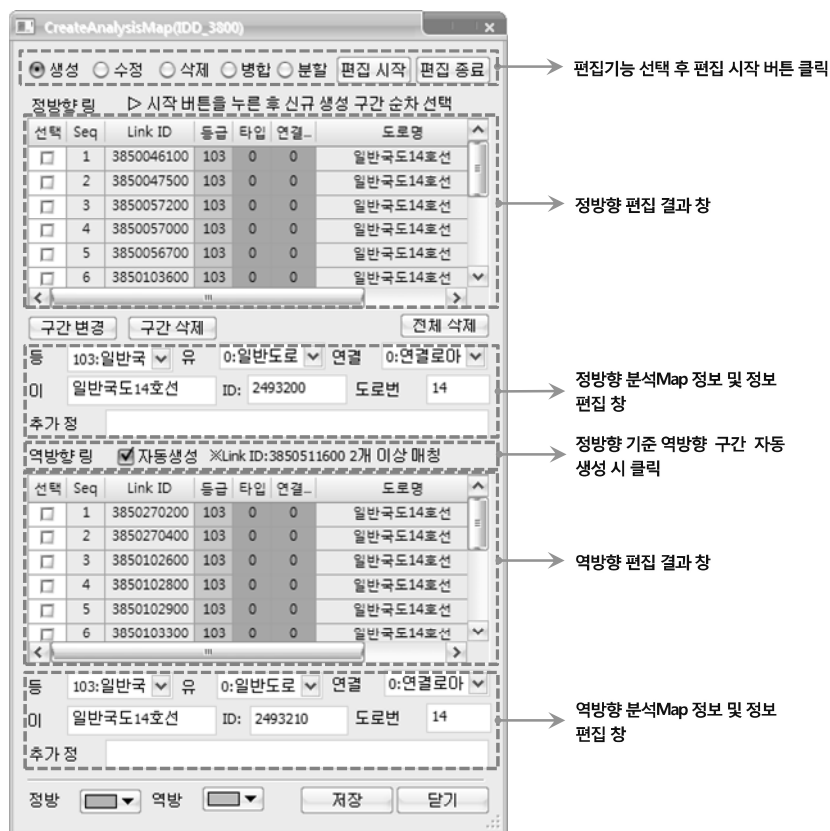
<그림 1-25> 분석 Map 관리시스템 구성

#### ② 분석Map 관리 기능

- 작업자가 효율적으로 작업할 수 있도록 표준노드링크와 분석Map을 각각 레이어 제어할 수 있도록 기능 개발
  - 링크는 도로등급 별 색상 및 두께 정의가 가능하도록 기능 개발
  - 노드는 노드유형 별 색상, 크기, 심볼 정의가 가능하도록 기능 개발
  - 화면상의 표시레벨(표출할 화면 스케일 정의)설정을 사용자가 정의할 수 있는 기능과 , 레이어 디스플레이 On/Off 기능을 개발
- 분석용 네트워크 현황을 확인과 DB 조회 및 표출 할 수 있도록 기능 개발

### ③ 분석Map 편집 기능

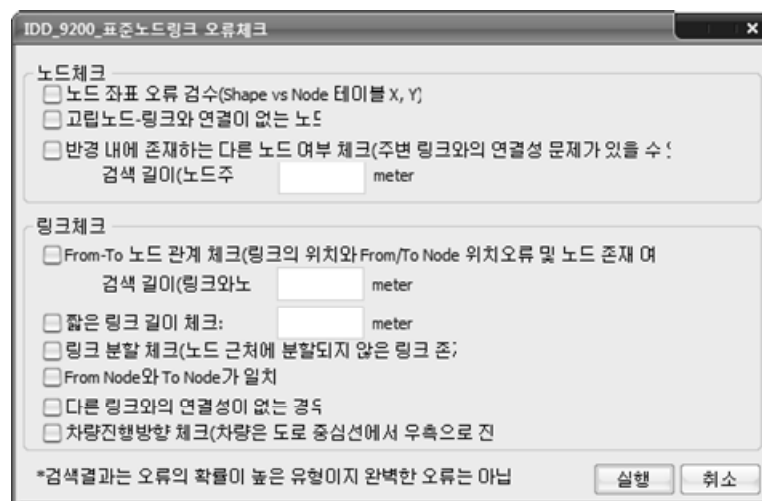
- 분석Map 편집기능은 도로구간 생성, 수정, 삭제가 가능하도록 기능 개발
  - 도로구간 생성은 해당구간의 링크를 모두 선택하지 않아도 시작링크 선택 후 종료 링크에 마우스 포인트를 가져가면 경로탐색 알고리즘에 의해서 시작링크와 종료링크 사이의 링크들이 선택되며 선택된 데이터가 맞으면 완료 버튼 클릭하여 도로구간이 자동생성 되도록 기능 개발
  - 정방향 도로구간 생성 후 이를 역방향 도로구간도 자동생성 되도록 기능 개발
  - 도로구간을 생성하면 표준노드링크와 분석 네트워크 간의 Mapping Table이 자동으로 생성되며, 이를 관리 및 수정이 가능하도록 기능 개발
  - 데이터 편집 시 선택한 링크의 정보보기 및 도로구간 속성편집이 용이하도록 화면UI 설계 및 개발
- 분석Map은 표준노드링크를 주요도로 구간별로 연결한 데이터로 분석Map 기본정보 테이블 및 분석Map과 표준노드링크의 상세 맵핑 테이블로 구성함



<그림 1-26> 분석Map 편집 UI

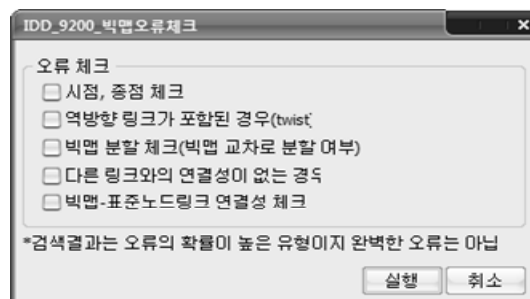
#### ④ 분석Map 검증 기능

- 분석Map을 구성하는 하위 레벨 데이터인 표준노드링크의 노드/링크의 오류 체크
  - 링크와 연결이 없는 고립노드 체크
  - 노드의 중복 및 사용자가 정의한 반경 내의 노드 체크
  - 링크의 시작노드/종료노드 정보와 실제 노드와의 관계 체크
  - 다른 링크와의 연결성이 없는 고립 링크 체크
  - 링크의 방향성 체크



<그림 1-27> 표준노드링크 오류체크 UI

- 분석Map의 노드/링크의 오류 체크
  - 분석맵 분할 여부 체크 (분석맵의 교차로 분할 여부 체크)
  - 표준노드링크와의 연결성 체크
  - 시점, 종점 체크 및 다른 링크와의 연결성 여부 체크



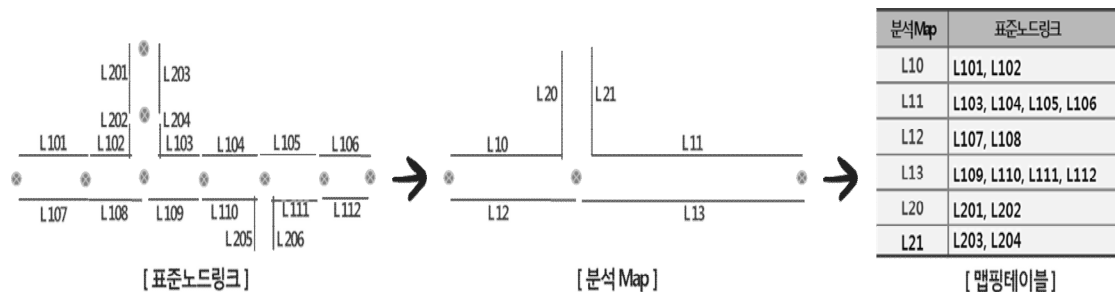
<그림 1-28> 분석맵 오류체크 UI



## 4) 분석 Map 구축

## ① 분석Map 구축 기준

- 분석 Map은 표준노드링크와 동일한 도로등급으로 구분하며, 도로유형과 연결로 코드도 표준 노드링크와 동일하게 적용하여 구축
- 고속국도에서 ~ 지방도로는 연결성이 없는 도로를 제외하고는 모두 분석맵을 생성하였으며, 특별 광역시도와 시군도로는 인터넷 지도의 주요도로 및 새주소의 ‘대로’, ‘로’ 이상의 주소정보를 참조하여 분석 Map을 생성
- 분석 Map은 표준노드링크와 분석 네트워크 간의 Mapping Table을 분석 Map이라 하며, Mapping Table의 정보를 이용하여 화면상에 디스플레이 함



&lt;그림 1-29&gt; 맵핑테이블 생성 예시

## ② 분석Map 구축을 위한 표준노드링크 오류 수정

- 분석 Map 관리시스템에서 개발한 표준노드링크 오류 검증 기능을 이용하여 다음과 같이 표준노드링크 오류 수정

&lt;표 1-11&gt; 표준노드링크 오류 수정 내역

오류 항목	오류 건수	처리 건수	기타 오류 체크 현황
노드 위치 중복 검수 (설정: 반경 2m이내)	34	34	<ul style="list-style-type: none"> <li>▪ 링크와 연결이 없는 노드 체크 → 283건 발생 (삭제하지 않음)</li> <li>▪ 링크 분할 여부 체크 → 1504건 중 49건 수정 → IC 및 고가도로/지하차도가 대부분으로 분석맵 생성도로만 검수 및 수정 함</li> </ul>
링크속성의 From/To 노드 관계 체크 (위치오류, 노드유무)	1270	1270	
링크속성의 From/To 노드 일치	13	13	
링크의 방향성 체크	469	469	

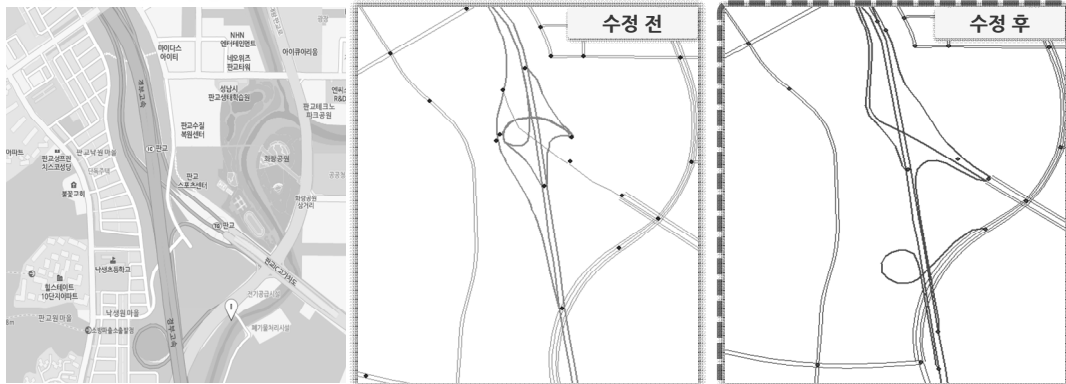
- 고속도로 IC/JC/TG 의 형상 및 연결성 체크 및 수정

<표 1-12> 고속도로 오류 수정 내역

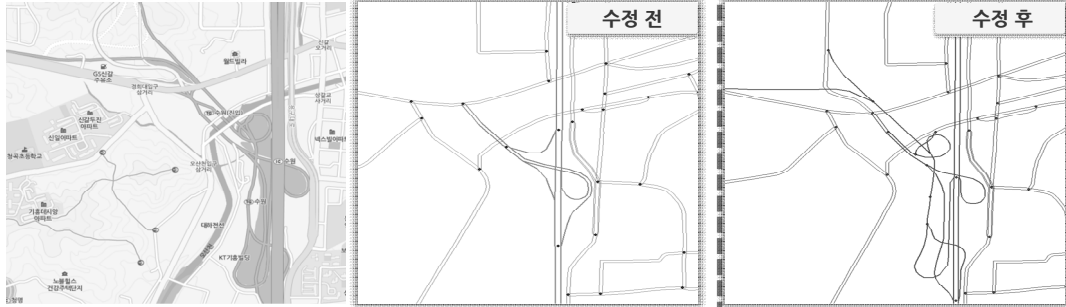
오류 항목	전체 개수	생성 건수	수정 건수	비고
IC	378	19	70	■ 고속도로와 연결성을 위하여 12지역 주변도로 생성
JC	73	1	8	
톨게이트	332	27	48	

- 고속도로의 오류유형은 IC의 램프 누락, 톨게이트 형상 누락이 가장 많았으며, 속성정보 오류, IC부근의 주요도로 누락도 다수 발생하여 모두 처리 완료

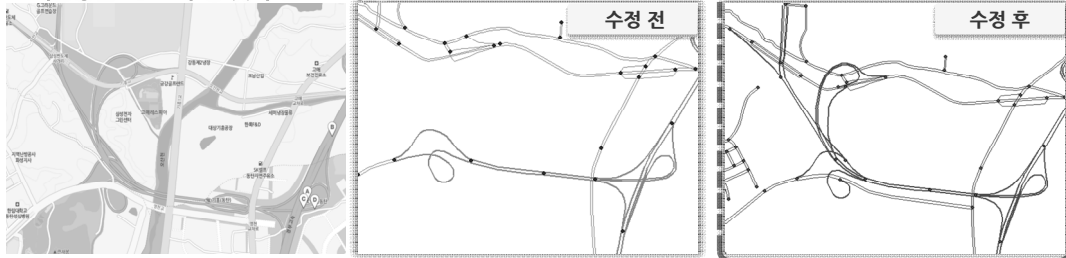
▶ 대왕판교 IC/TG 수정



▶ 수원 IC/TG 수정 및 주변도로 생성



▶ 기흥동탄 IC/TG 수정 및 주변도로 생성



<그림 1-30> 고속도로 형상 수정 결과 화면

## ③ 분석 Map 구축결과

- 전국 분석 Map 구축 및 검수 완료

&lt;표 1-13&gt; 분석 Map 도로등급 별 구축 현황

도로 등급	표준노드링크 개수 (전체 : 254,896개)	분석 Map 개수 (전체 : 53,541개)	비고
고속국도	6,296개	5,564개	■ 표준노드링크의 기타도로 (108)은 단지내 도로로 생성하지 않음
도시 고속국도	2,203개	1,375개	
일반국도	44,665개	12,563개	
국가지원지방도	9,963개	2,762개	
지방도	31,231개	6,315개	
특별·광역시도	35,548개	6,961개	
시·군도	122,422개	1,8001개	

## - 분석 Map 구축 결과 화면

▶ 서울시 분석 Map 구축 현황



▶ 서울시 표준노드링크



&lt;그림 1-31&gt; 서울시의 표준노드링크 및 분석 Map 비교 화면

▶ 부산시 분석 Map 구축 현황



▶ 부산시 표준노드링크



&lt;그림 1-32&gt; 부산시의 표준노드링크 및 분석 Map 비교 화면

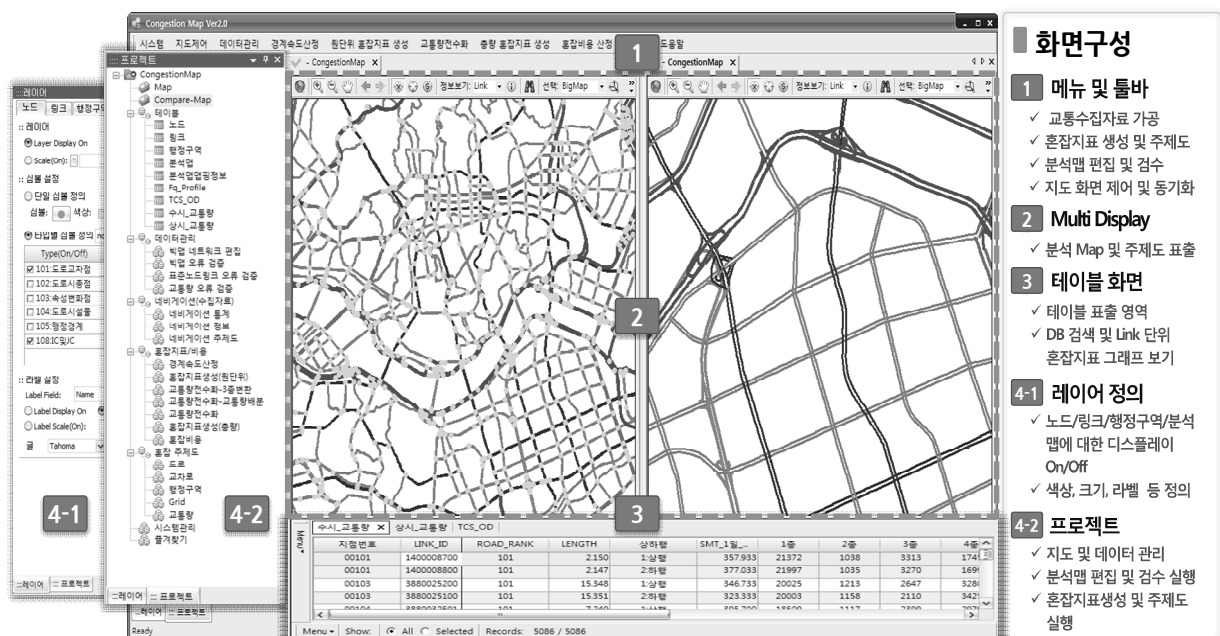
## ④ 분석 Map 구축의 한계점

- 표준노드링크를 기반으로 생성되는 분석 Map은 표준노드링크의 형상오류, 주요도로 누락 발생 구간은 표준노드링크와 동일하게 분석 Map에도 적용되어 향후 개선이 필요함
- 현재 고속도로는 표준노드링크를 수정하여 분석Map에 적용. 점차적으로 일반국도, 지방도, 주요도로 순으로 향후 개선이 필요함
- 표준노드링크는 지속적으로 유지보수가 진행됨으로 수정한 현재 표준노드링크와의 호환성 주의 필요

### 3. 전국교통혼잡지도 분석시스템 개발

#### 가. 화면구성

- 교통혼잡지도 분석시스템 화면구성은 메뉴 및 툴바, 지도화면, 테이블 화면, 프로젝트 관리, 레이어 정의 영역으로 구성
  - 메뉴는 지도제어, 네트워크 편집 및 오류검증, 혼잡지표 생성, 혼잡지표 주제도, Tools, 주제도, 도움말로 구성됨
  - 툴바는 지도 확대/축소/이동과 지도화면 동기화, 정보보기, 찾기, 영역선택으로 기본 구성
  - 지도 화면은 주제도 및 데이터 비교를 할 수 있도록 동적 분할 윈도우 형태로 개발하여 편리한 비교분석이 가능하도록 개발
  - 테이블 화면은 선택한 데이터의 내역과 기본 정보를 제공하며, 데이터 검색 및 지도 상에 위치 표출 및 이동 등의 기능으로 구성
  - 프로젝트 관리는 지도 및 데이터 관리 기능 및 데이터 가공부터 분석까지의 기능을 실행 순서에 맞추어 배치하여 실행할 수 있도록 구성
  - 레이어 정의는 디스플레이 On/Off, 색상/크기/두께 설정, 라벨 정의 등으로 구성



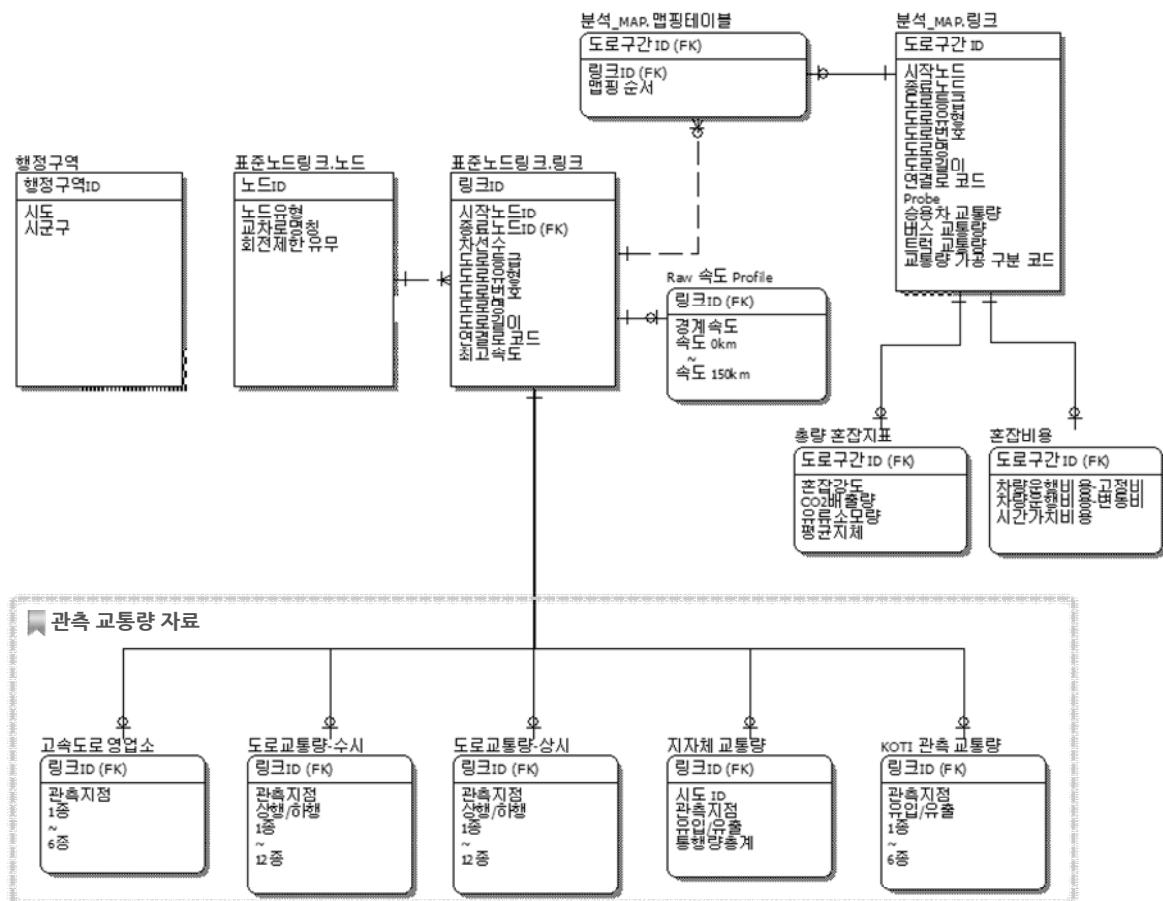
<그림 1-33> 시스템 화면 구성

## 나. 데이터베이스 설계

- 전국교통혼잡지도 데이터베이스는 점차적으로 증가하는 대용량 데이터를 고려하여 안정적이고 효율적으로 관리할 수 있도록 데이터베이스 설계를 개선함
- 변화되는 요구사항에 부합되도록 확장 가능한 설계를 하여 다양한 자료를 효과적으로 구축할 수 있는 구조로 설계
- 데이터베이스 설계 방법론에 입각한 현행 시스템 및 데이터를 분석하여 표준화 방안 등을 마련하여 시스템에서 요구하는 데이터 분석이 가능한 구조로 설계

### 1) Entity Relation Diagram

- 전국교통혼잡지도 분석 시스템의 주요 DB구성은 네트워크(표준노드링크, 분석맵), 관측교통량, 혼잡지표, 혼잡비용, 시스템 관리 항목 등으로 구성됨



<그림 1-34> 전국교통혼잡지도 시스템 ERD

## 2) 테이블 정의

### ① 표준노드링크 및 분석 맵

- 본 연구의 분석 맵은 기존의 전국표준노드링크가 가지고 있는 분석용 네트워크로서의 문제점 보완과 교통혼잡지도의 분석 및 활용성 극대화를 위하여 구축한 Map을 전국교통혼잡지도 분석 Map이라 함. 또한 도로속성으로 인하여 상세하게 분할되어 있는 링크를, 동일구간을 기준으로 링크를 병합 및 관리할 수 있도록 구축하여 도로구간의 혼잡여부를 쉽게 판단할 수 있도록 함
- 분석 Map 구성은 표준노드링크의 링크와 분석 Map의 링크간의 맵핑 테이블과 이를 관리하는 관리테이블로 구성
- 표준노드링크 테이블 정의서는 다음과 같음

&lt;표 1-14&gt; 표준노드링크의 링크 테이블

No	Column	Type	설명	코드	코드정보
1	LINK_ID	Integer	분석 Map 링크 ID	-	-
2	F_NODE	Integer	시작노드	-	-
3	T_NODE	Integer	종료노드	-	-
4	ROAD_RANK	Char	도로등급	101 102 103 104 105 106 107	고속국도 도시고속국도 일반국도 특별·광역시도 국가지원지방도 지방도 시·군도
5	ROAD_TYPE	Char	도로유형	000 001 002 003 004	일반도로 고가차도 지하차도 교량 터널
6	ROAD_NO	Char	도로번호	-	-
7	NAME	Char	도로명	-	-
8	LENGTH	Integer	도로길이	-	-
9	CONNECT	Char	연결로 코드	000 101 102 103 104 105 106 107	연결로 아님 고속국도 연결로 도시고속국도 연결로 일반국도 연결로 특별·광역시도 연결로 국가지원지방도 연결로 지방도 연결로 시·군도 연결로
10	LANES	Integer	차선수	-	-
11	MAX_SPD	Integer	최고속도	-	-
12	LENGTH	Double	링크길이(km)	-	-
13	EMD_ID	Integer	읍면동 ID	-	-



&lt;표 1-15&gt; 표준노드링크의 노드 테이블

No	Column	Type	설명	코드	코드정보
1	NODE_ID	Integer	표준노드링크의 노드 ID	-	-
2	NODE_TYPE	Char	노드유형	101 102 103 104 108	도로교차점 도로시·종점 속성변화점 도로시설물 IC및JC
3	NODE_NAME	Char	교차로 명칭	-	-
4	X	Double	X 좌표	-	
5	Y	Double	Y 좌표	-	
6	SIDO_ID	Integer	시도 ID		
7	SIGUN_ID	Integer	시군 ID		
8	SIGUNGU_ID	Integer	시군구 ID		
9	EMD_ID	Integer	읍면동 ID		

&lt;표 1-16&gt; 분석맵의 링크 테이블

No	Column	Type	설명	코드	코드정보
1	K_LINK_ID	Integer	분석 Map 링크 ID	-	-
2	F_NODE	Integer	시작노드	-	-
3	T_NODE	Integer	종료노드	-	-
4	ROAD_RANK	Char	도로등급	101 102 103 104 105 106 107	고속국도 도시고속국도 일반국도 특별·광역시도 국가지원지방도 지방도 시·군도
5	ROAD_TYPE	Char	도로유형	000 001 002 003 004	일반도로 고가차도 지하차도 교량 터널
6	ROAD_NO	Char	도로번호	-	-
7	ROAD_NAME	Char	도로명	-	-
8	LENGTH	Integer	도로길이	-	-
9	CONNECT	Char	연결로 코드	000 101 102 103 104 105 106 107	연결로 아님 고속국도 연결로 도시고속국도 연결로 일반국도 연결로 특별·광역시도 연결로 국가지원지방도 연결로 지방도 연결로 시·군도 연결로
10	auto_wd	Double	승용차 평일 교통량		
11	auto_h	Double	승용차 휴일 교통량		
12	auto_tot	Double	승용차 전체 교통량		
13	bus_wd	Double	버스 평일 교통량		
14	bus_h	Double	버스 휴일 교통량		
15	bus_tot	Double	버스 전체 교통량		
16	truck_wd	Double	트럭 평일 교통량		
17	truck_h	Double	트럭 휴일 교통량		
18	truck_tot	Double	트럭 전체 교통량		
19	index	Double	교통량 추정 유무	1 2	정상 추정
20	traffic_type	Integer	관측교통량 구분코드	10 20 30 31 40 50	건기원 상시 건기원 수시 한국도로공사 톨게이트 한국도로공사 본선 톨게이트 KOTI 코든 서울시
21	traffic_check	Integer	교통량 데이터 이상여부 체크 코드	1 2 3	정상 검토대상 제외대상

- 분석맵의 테이블 정의서는 다음과 같음

<표 1-17> 분석맵의 링크 테이블

No	Column	Type	설명	코드	코드정보
1	K_LINK_ID	Integer	분석 Map 링크 ID	-	-
2	F_NODE	Integer	시작노드	-	-
3	T_NODE	Integer	종료노드	-	-
4	ROAD_RANK	Char	도로등급	101 102 103 104 105 106 107	고속국도 도시고속국도 일반국도 특별·광역시도 국가지원지방도 지방도 시·군도
5	ROAD_TYPE	Char	도로유형	000 001 002 003 004	일반도로, 고가차도 지하차도 교량 터널
6	ROAD_NO	Char	도로번호	-	-
7	ROAD_NAME	Char	도로명	-	-
8	LENGTH	Integer	도로길이	-	-
9	CONNECT	Char	연결로 코드	000 101 102 103 104 105 106 107	연결로 아님 고속국도 연결로 도시고속국도 연결로 일반국도 연결로 특별·광역시도 연결로 국가지원지방도 연결로 지방도 연결로 시·군도 연결로
10	auto_wd	Double	승용차 평일 교통량		
11	auto_h	Double	승용차 휴일 교통량		
12	auto_tot	Double	승용차 전체 교통량		
13	bus_wd	Double	버스 평일 교통량		
14	bus_h	Double	버스 휴일 교통량		
15	bus_tot	Double	버스 전체 교통량		
16	truck_wd	Double	트럭 평일 교통량		
17	truck_h	Double	트럭 휴일 교통량		
18	truck_tot	Double	트럭 전체 교통량		
19	index	Double	교통량 추정 유무	1 2	정상 추정
20	traffic_type	Integer	관측교통량 구분코드	10 20 30 31 40 50	건기원 상시 건기원_수시 한국도로공사_톨게이트 한국도로공사_본선_톨게이트 KOTI 코든 서울시
21	traffic_check	Integer	교통량 데이터 이상여부 체크 코드	1 2 3	정상 검토대상 제외대상

&lt;표 1-18&gt; 분석맵과 표준노드링크의 맵핑테이블

No	Column	Type	설명	코드	코드정보
1	k_link_id	Integer	분석 Map 링크 ID	-	-
2	seq	Integer	링크 순서	-	-
3	link_id	Integer	표준노드링크 링크 ID	-	-

## 2) 전국 교통량 조사자료

- 전국 교통량 조사 자료는 교통량 전수화 및 혼잡지표 및 혼잡비용 산정을 위하여 사용되며 본 과업에서 사용되는 관측 교통량은 다음과 같음

&lt;표 1-19&gt; 도로교통량 상시 테이블

No	Column	Type	설명	코드	코드정보
1	SPOT_ID	Varchar	관측지점 명		
2	LINK_ID	Integer	링크 ID		
3	UP_DOWN	Integer	상행/하행		
4	traffic_type	Integer	관측 교통량 코드	10	건기원 상시
5	V_TOT_wd	Double	통행량 총계 (평일)		
6	V_1_wd	Double	차종_1종 (평일)		
7	V_2_wd	Double	차종_2종 (평일)		
8	~	Double	차종_3종~10종 (평일)		
9	V_11_wd	Double	차종_11종 (평일)		
10	V_12_wd	Double	차종_12종 (평일)		
11	V_TOT_h	Double	통행량 총계 (휴일)		
12	V_1_h	Double	차종_1종 (휴일)		
13	V_2_h	Double	차종_2종 (휴일)		
14	~	Double	차종_3종~10종 (휴일)		
15	V_11_h	Double	차종_11종 (휴일)		
16	V_12_h	Double	차종_12종 (휴일)		

&lt;표 1-20&gt; 도로교통량 수시 테이블

No	Column	Type	설명	코드	코드정보
1	SPOT_ID	Varchar	관측지점 명		
2	LINK_ID	Integer	링크 ID		
3	UP_DOWN	Integer	상행/하행		
4	traffic_type	Integer	관측 교통량 코드	20	건기원 수시
5	V_TOT_wd	Double	통행량 총계 (평일)		
6	V_1_wd	Double	차종_1종 (평일)		
7	V_2_wd	Double	차종_2종 (평일)		
8	~	Double	차종_3종~10종 (평일)		
9	V_11_wd	Double	차종_11종 (평일)		
10	V_12_wd	Double	차종_12종 (평일)		

&lt;표 1-21&gt; 고속도로 영업소 테이블

No	Column	Type	설명	코드	코드정보
1	SPOT_ID	Varchar	관측지점 명		
2	LINK_ID	Integer	링크 ID		
3	IN_OUT	Integer	진입/진출		
4	traffic_type	Integer	관측 교통량 코드	30 31	한국도로공사-톨게이트 한국도로공사-본선-톨게이트
5	V_TOT_wd	Double	통행량 총계 (평일)		
6	V_1_wd	Double	차종_1종 (평일)		
7	V_2_wd	Double	차종_2종 (평일)		
8	~	Double	차종_3종~5종 (평일)		
10	V_6_wd	Double	차종_6종 (평일)		
11	V_TOT_h	Double	통행량 총계 (휴일)		
12	V_1_h	Double	차종_1종 (휴일)		
13	V_2_h	Double	차종_2종 (휴일)		
14	~	Double	차종_3종~5종 (휴일)		
15	V_6_h	Double	차종_6종 (휴일)		

&lt;표 1-22&gt; KOTI-코든라인 테이블

No	Column	Type	설명	코드	코드정보
1	SPOT_ID	Varchar	관측지점 명		
2	LINK_ID	Integer	링크 ID		
3	UP_DOWN	Integer	상행/하행		
4	traffic_type	Integer	관측 교통량 코드	40	KOTI - 코든라인
5	V_TOT_wd	Double	통행량 총계 (평일)		
6	V_1_wd	Double	차종_1종 (평일)		
7	V_2_wd	Double	차종_2종 (평일)		
8	~	Double	차종_3종~5종 (평일)		
9	V_6_wd	Double	차종_6종 (평일)		

## 3) Car Navigation 및 혼잡지표/혼잡비용 자료

- 본 연구의 링크별 Car Navigation 및 혼잡 경계속도 자료는 기 구축된 전국 교통혼잡지도 분석 시스템 내 1차 가공한 Raw Probe DB를 이용함. 즉, Car Navigation 수집 정보를 월/일/시간 단위의 속도 프로파일로 생성하기 위하여 1차 가공한 데이터로 이를 속도별로 재분류한 자료임

&lt;표 1-23&gt; Car Navigation 차량 수집 데이터 테이블

No	Column	Type	설명	코드	코드정보
1	AUTH_KEY	Integer	단말기 ID		
2	FROM_LINK	Integer	진입 링크		
3	TO_LINK	Integer	진출 링크		
4	LINK_TIME	DateTime	수집 시간		
5	SPEED	Double	속도		

&lt;표 1-24&gt; Raw Probe DB 테이블(Car Navigation 데이터 + 혼잡 경계속도)

No	Column	Type	설명	코드	코드정보
1	YEAR	Varchar	수집 년		
2	MONTH	Varchar	수집 월		
3	FROM_LINK_ID	Integer	진입 링크 ID		
4	WEEK_TYPE	Integer	월/화/수/목/금/토/일		
5	TIME	Integer	시간		
6	VEHICLE_COUNT	Double	probe 대수		
7	BOUNDARY_SPEED	Double	경계속도		
8	SPD_0km_h	Double	속도 0km 빈도수		
9	~	Double	~ (1단위로 0부터 150km까지 컬럼 생성)		
10	SPD_150km_h	Double	속도 150km 이상 빈도수		

- 혼잡지표는 원단위 혼잡지표와 총량 혼잡지표로 구분하며, 원단위 혼잡지표는 표준노드링크 기반, 총량 혼잡지표와 혼잡비용은 분석맵 기반으로 테이블이 다음과 같이 구성됨

&lt;표 1-25&gt; 원단위 혼잡지표 테이블

No	Column	Type	설명	코드	코드정보
1	Link_id	Integer	표준노드링크 링크 ID		
2	net_con_intensity	Double	혼잡강도		
3	net_co2_intensity	Double	CO2 배출량		
4	net_fuel_intensity	Double	유류소모량		
5	net_delay_intensity	Double	평균지체		
6	old_net_con_intensity	Double	기존 혼잡강도		
7	net_counting_sample	Double	표본수		

&lt;표 1-26&gt; 총량 혼잡지표 테이블

No	Column	Type	설명	코드	코드정보
1	K_Link_id	Integer	분석맵 링크 ID		
2	net_con_intensity	Double	혼잡강도		
3	net_co2_intensity	Double	CO2 배출량		
4	net_fuel_intensity	Double	유류소모량		
5	net_delay_intensity	Double	평균지체		
6	old_net_con_intensity	Double	기존 혼잡강도		
7	net_counting_sample	Double	표본수		

&lt;표 1-27&gt; 혼잡비용 테이블

No	Column	Type	설명	코드	코드정보
1	K_Link_id	Integer	분석맵 링크 ID		
2	week_type	Integer	구분(주중/주말)	1 2	주중 주말
3	TCC	Double	혼잡비용		



## 다. 전국교통혼잡지도 분석시스템 주요기능 개발

### 1) 수집 교통정보 가공

#### ① Car Navigation 데이터 분석

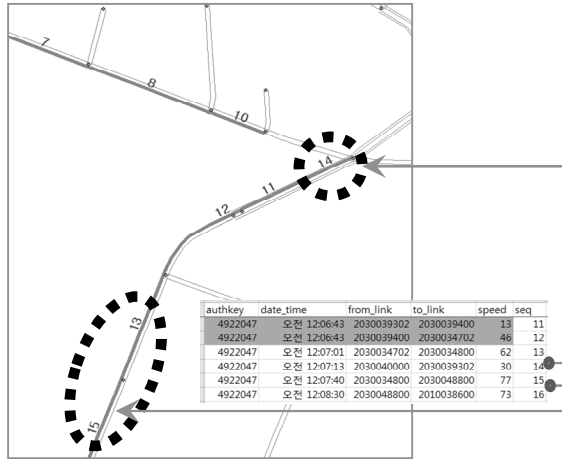
- 현대 엠앤소프트사에서 제공한 Car Navigation 데이터 현황은 다음과 같음
  - 진입링크, 진출링크, 수집시간, 속도로 구성
  - 수집한 데이터는 1일단위로 취합하여 Text 파일로 제공

<표 1-28> Car Navigation 수집 현황

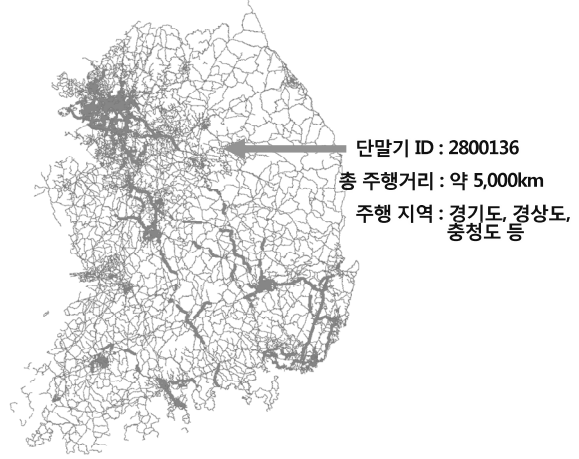
구분 \ 기간	2013년			2014년
	10월	11월	12월	1월
1일 평균 운행 대수	184,709	180,452	162,959	143,397
이벤트 건수	440,395,473	445,596,480	401,343,793	310,960,262
Link 매칭정보 (전체 256,160개)	155,830	186,144	215,122	216,236
Coverage	60.83%	72.09%	83.98%	84.41%

- 수집 대상(지역, 도로등급)을 점차적으로 확대하여 2014년 1월 기준으로 전국 84.41% 영역의 데이터가 수집됨
- 데이터 분석 결과 특이사항
  - 2014년 3월 기준의 표준노드링크와 매칭한 결과 수집되지 않은 도로구간 발생
  - 교차로 구간에서 차량의 우회전 또는 좌회전 시 진행 방향의 링크가 수집되지 않고 직진방향의 링크가 수집되는 경우가 발생
  - 수집시간정보로 차량의 주행 경로를 표출 시 차량의 주행 경로가 앞뒤가 엇갈리게 표출됨. 수집시간으로 데이터를 정렬하지 않고 데이터의 입력 순서로 정렬 시 차량의 주행경로가 정상적으로 표출
  - 단말기 ID 중복 발생 (동일 단말기에서 여러 대의 차량이 주행한 데이터가 발생. 예측근거는 1일1대의 차량의 주행거리가 5,000Km를 초과)

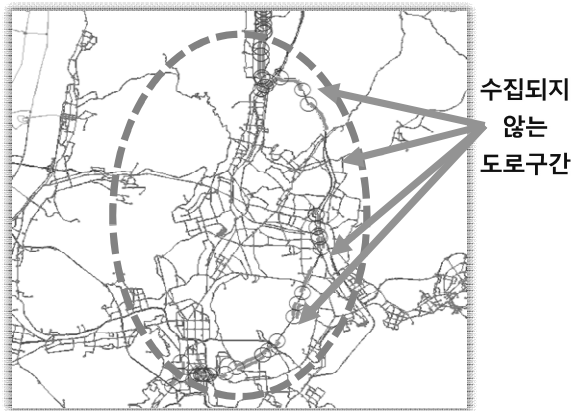
▶ 수집시간정보가 차량의 이동경로 순서와 일치하지 않는 경우



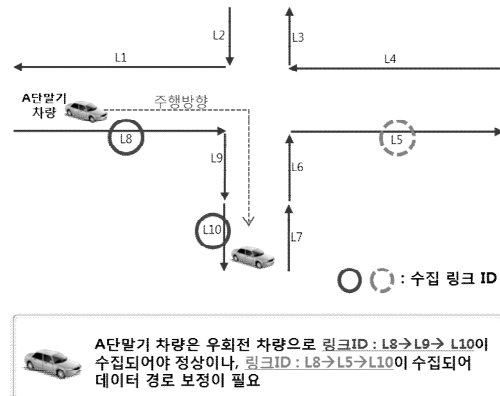
▶ 단말기 ID 중복 발생 (구형 단말기 일부)



▶ 수집되지 않는 도로구간 발생



▶ IC 및 교차로 구간에서의 수집 오류 유형



<그림 1-35> Car Navigation 수집 유형(특이사항)

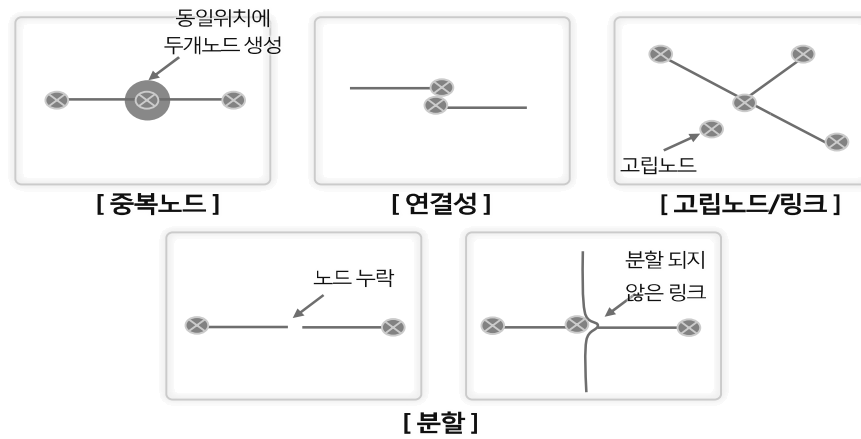
② 수집교통정보 가공 프로세스

- 수집 데이터 누락 구간 및 데이터 연결성 등을 보완하기 위하여 경로 보정 프로그램 개발

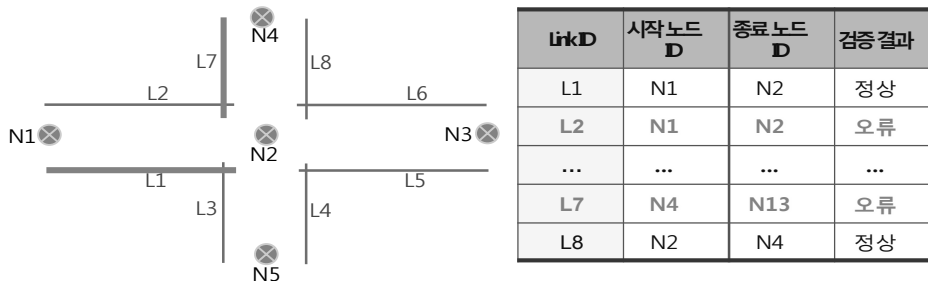
a) 경로보정을 위한 표준노드링크 오류 검증 및 수정

- 링크와 연결성이 없는 고립 노드, 노드 중복, 링크의 속성정보와 실제 노드와의 연결성 체크, 다른 링크와의 연결성이 없는 고립 링크 체크, 링크의 방향성 등에 대하여 표준노드링크 수정
- 표준노드링크 오류를 수정하기 위하여 다음과 같이 검증기능 개발

○ 표준노드링크의 오류 유형에 대한 예시 화면



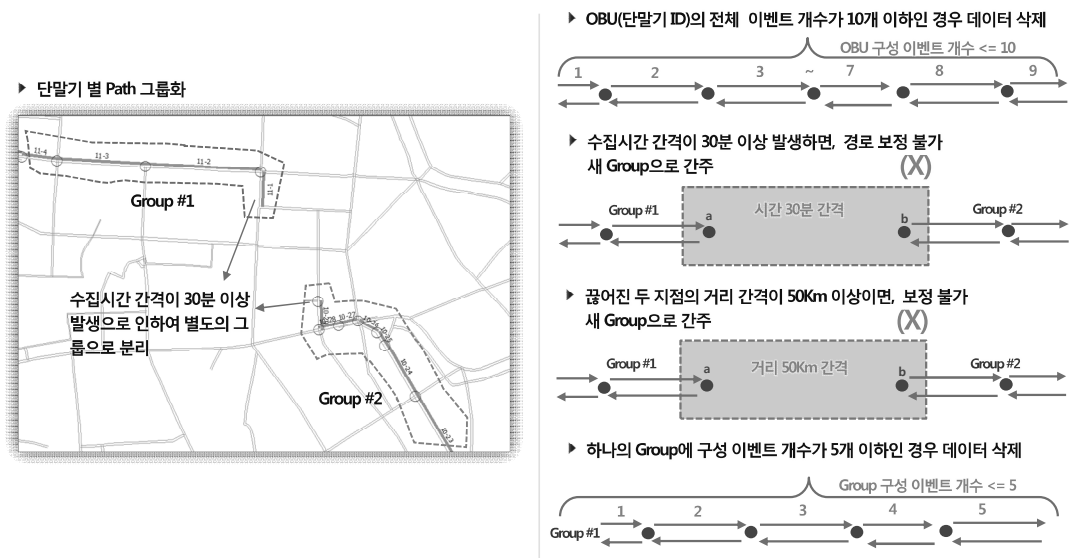
▪ 링크의 시작노드/종료노드의 속성정보 오류 체크



<그림 1-36> 표준노드링크 오류 예시

b) 경로탐색을 이용한 경로 보정

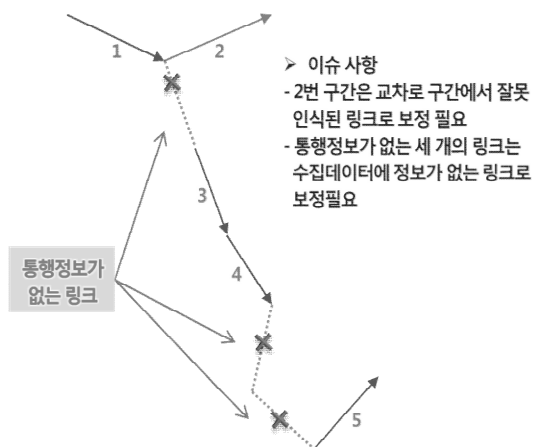
- 수집교통정보의 네트워크와 본 과제에서 사용하는 네트워크(표준노드링크 2014년 4월 배포 데이터)와의 차이로 인하여 발생하는 수집데이터 누락 구간 및 데이터 연결성 문제 등을 보완하기 위하여 프로그램 개발
  - 수집데이터는 단말기 별 데이터를 확인하여 이벤트 개수가 특정 개수 이하이면, 데이터를 필터링하여 불안정한 데이터를 제거함
  - 동일 단말기 내에서 이벤트 수집시간 간격이 30분 이상 발생 시에는 별도의 그룹으로 데이터를 분류함
  - 동일 단말기 내에서 이벤트의 누락 구간의 거리의 차가 최대 50Km이상 발생 시에는 별도의 그룹으로 데이터를 분류함



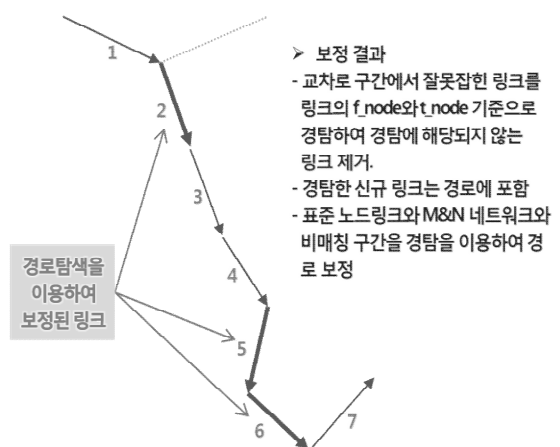
&lt;그림 1-37&gt; 통행 그룹 생성

- 별도의 그룹으로 분류된 데이터는 데이터 로딩 순서대로 경로탐색을 반복 수행하여 데이터의 연결성 검수 및 데이터를 보정함
- 경로탐색 방법은 링크의 from\_node\_id(시작노드)와 다음 진행 링크의 to\_node\_id(종료노드)를 기준으로 경로탐색하여 해당구간의 링크를 검색함
- 검색한 링크 리스트 중에서 새로운 링크를 검색 시에는 이를 그룹 내의 링크로 추가하여 누락구간 부분을 보정함
- 검색한 링크 리스트 중에서 그룹 내의 링크가 해당하지 않으면 데이터를 제거함

## ▶ 경로 보정 전



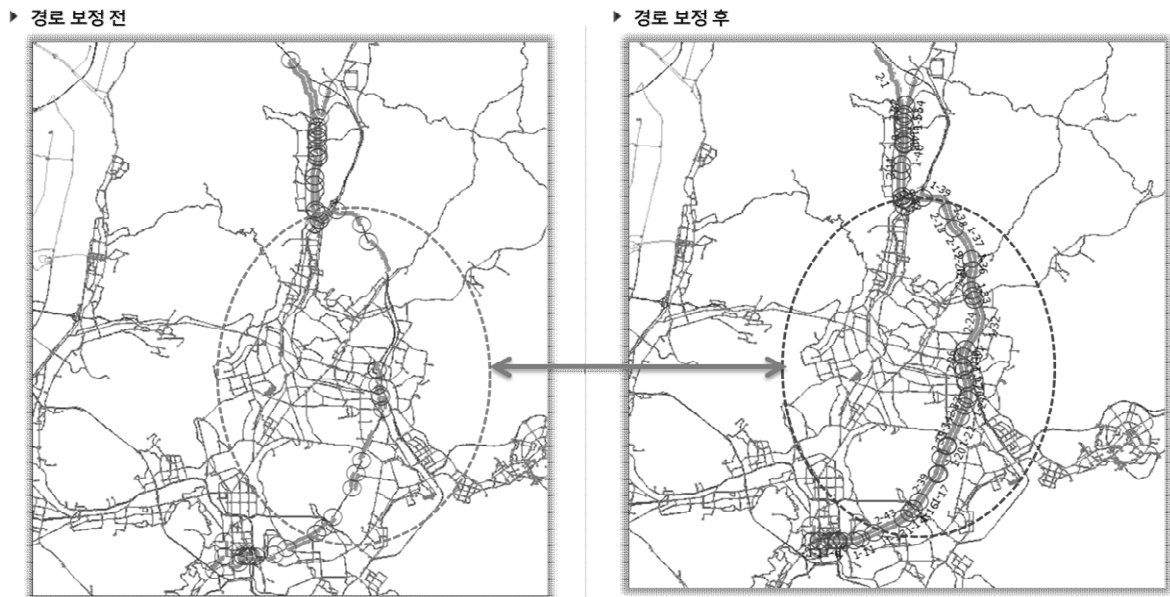
## ▶ 경로 보정 후



&lt;그림 1-38&gt; 통행 그룹 경로 보정

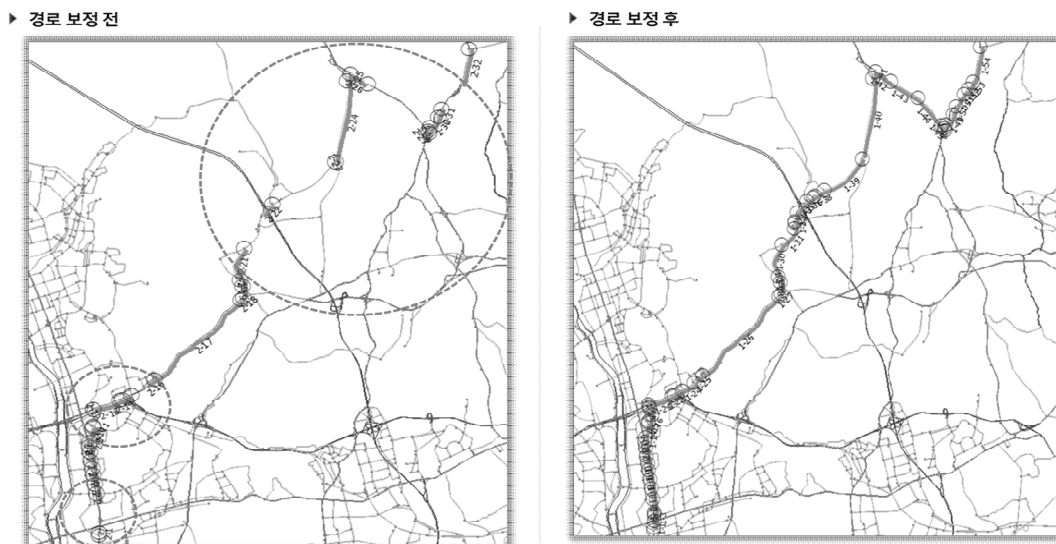
○ 수집교통정보 가공을 통한 경로 보정 결과 비교

- 2013년10월2일 주행한 'A' 단말기 차량의 경로 보정 전과 보정 후의 결과 비교



<그림 1-39> 'A' 단말기의 경로 보정 결과 화면

- 2013년 10월 2일 주행한 'B' 단말기 차량의 경로 보정 전과 보정 후의 결과 비교



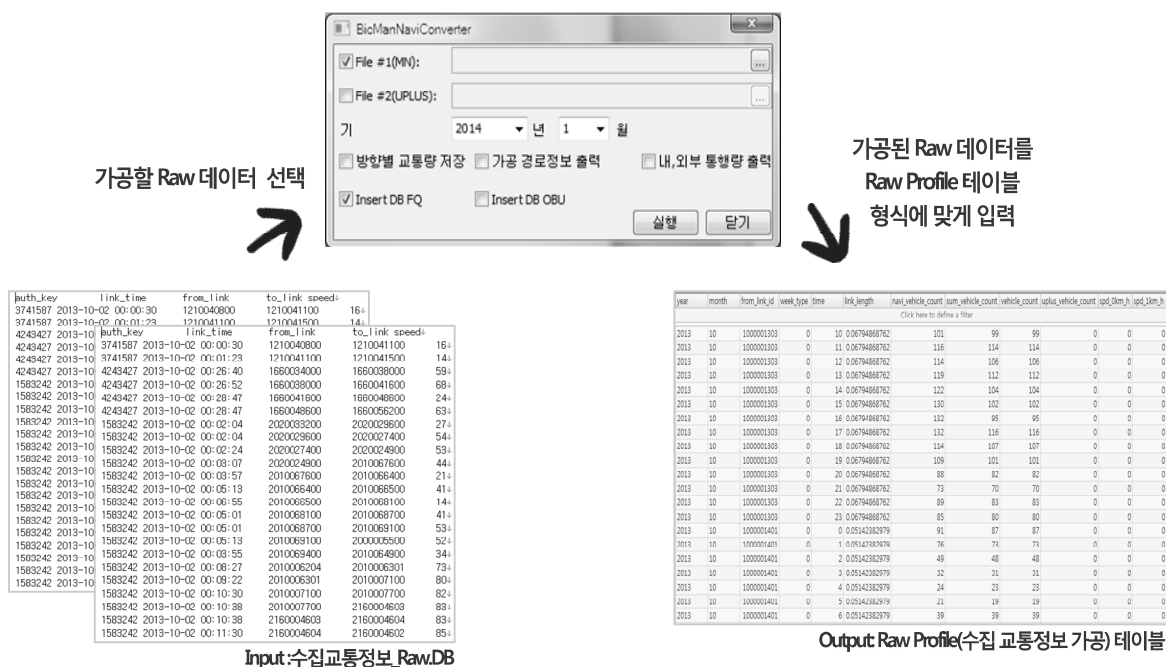
<그림 1-40> 'B' 단말기의 경로 보정 결과 화면

c) 경로 보정 데이터의 Raw 속도 프로파일 생성

○ 수집교통정보 가공은 원시 데이터 로딩→경로탐색을 이용한 경로 보정→Raw 속도 프로파일

생성 순으로 자동처리 하도록 개발함

- Raw 속도 프로파일 Output 포맷은 링크ID, 수집 년, 수집 월, 수집 요일, 수집 시간(3시간 단위), 0~150km 별 통행량 정보로 구성되어 있음



<그림 1-41> 수집교통정보 가공 및 Raw Profile 생성

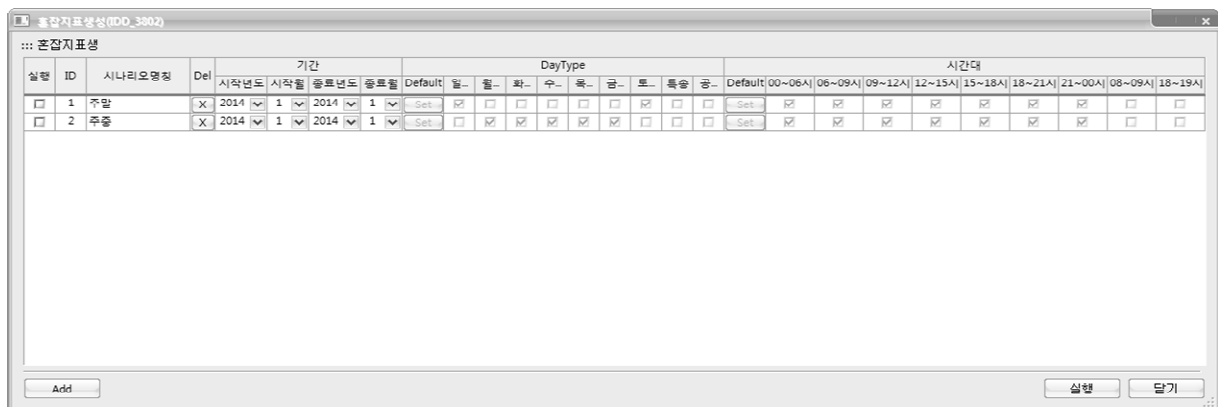
- 수집 교통정보 가공 후 가공전과 가공후의 데이터 결과는 다음과 같음

<표 1-29> 수집 교통정보 가공 후 데이터 결과 비교

구분 \ 기간		2013년			2014년
		10월	11월	12월	1월
1일 평균 운행 대수		184,709	180,452	162,959	143,397
가공 전	이벤트 건수	440,395,473	445,596,480	401,343,793	310,960,262
	Link 매칭정보 (전체 256,160개)	155,830	186,144	215,122	216,236
	Coverage	60.83%	72.09%	83.98%	84.41%
가공 후	이벤트 건수	609,807,948	618,102,595	550,804,284	422,274,073
	Link 매칭정보 (전체 256,160개)	188,689	207,951	231,514	232,428
	Coverage	73.66%	81.18%	90.38%	90.74%

## 2) 경계속도 및 혼잡지표 생성

- 경계속도 산정은 Raw Speed Profile를 기준으로 사용자가 기간을 선택 후 실행하면 Output 테이블에 표준노드링크의 링크 단위로 경계속도가 입력되도록 개발
- 원단위 혼잡지표생성은 사용자가 기간, 요일, 시간을 선택할 수 있도록 UI설계. 원단위 혼잡지표는 표준노드링크의 링크 단위로 경계속도가 입력. 총량 혼잡지표는 분석맵의 링크 단위로 입력되도록 개발
- 생성된 혼잡지표는 분석맵, 행정구역, Grid단위의 주제도 및 테이블로 결과 확인이 가능하도록 기능구현



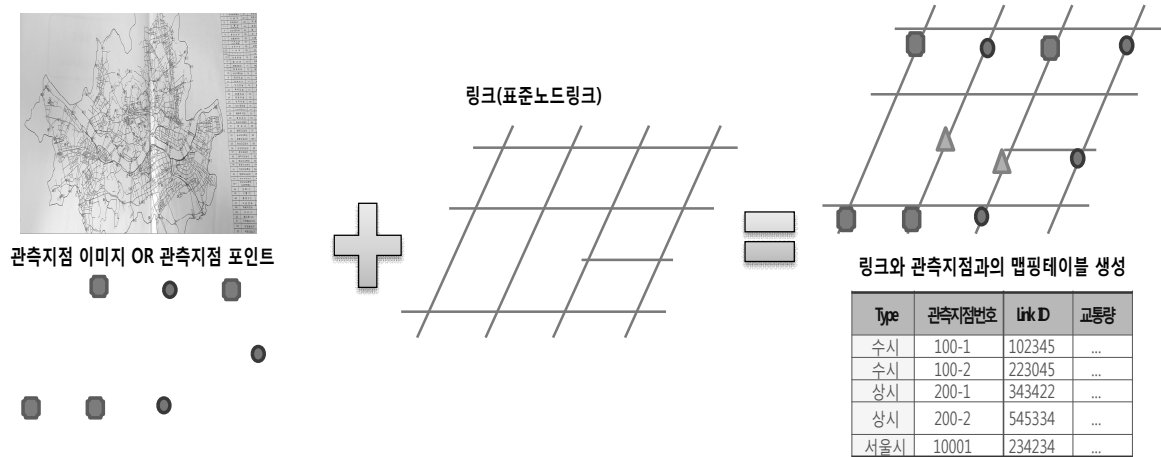
<그림 1-42> 원단위 혼잡지표 생성 UI

## 3) 교통량 전수화

### ① 교통량 전수화를 위한 관측 교통량 가공

- 전국 교통량 조사 자료는 교통량 전수화 및 혼잡지표 및 혼잡비용 산정을 위하여 본 시스템에서 사용하고 있는 표준노드링크 ID와 맵핑
  - 관측지점은 위치(x,y)정보 또는 이미지로 제공되어 링크(표준노드링크)와 지점별 맵핑하여 테이블 생성
  - 맵핑테이블 구축대상은 표준노드링크가 기 구축된 도로에 한해서만 맵핑 완료
  - 수시, 상시, KOTI-코든라인은 제공된 위치 포인트를 기준으로 맵핑테이블 구축
  - 한국도로공사는 톨게이트명 및 진출입 정보와 표준노드링크의 도로명 또는 지도사이트를 참조하여 맵핑테이블 구축

- 서울시는 교통량 조사자료 책자의 이미지를 참조하여 맵핑테이블 구축

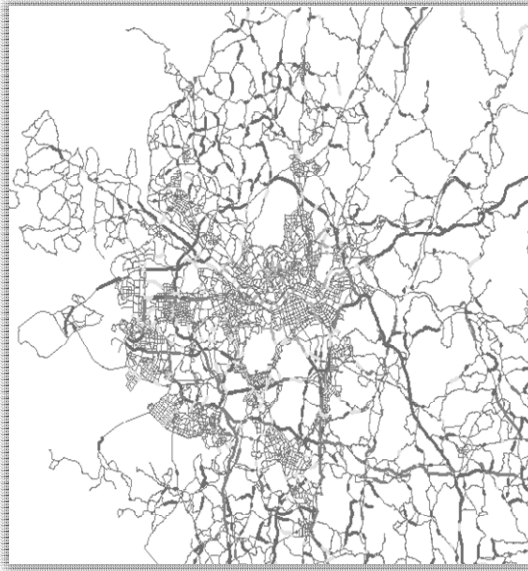


<그림 1-43> 관측지점 및 표준노드링크 맵핑테이블 생성

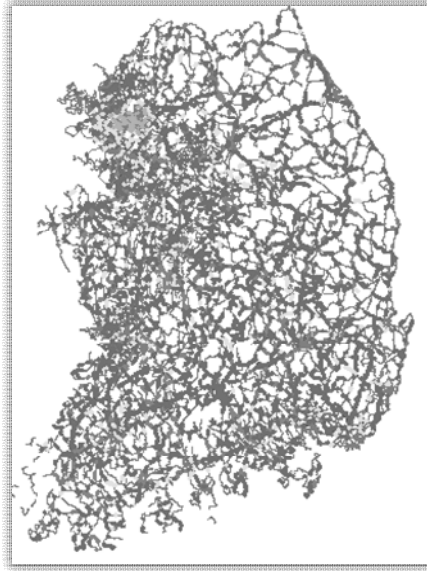
- 표준노드링크와의 맵핑결과는 다음과 같음
  - 도로교통량 상시조사지점 : 769개소, 12개 차종별 교통량, 평일/휴일로 구분
  - 도로교통량 수시조사지점 : 2,543개소, 12개 차종별 교통량
  - 고속도로 영업소 : 349개소, 6개 차종별 교통량, 평일/휴일로 구분
  - 서울시 교통량 : 96개소, 전체 교통량(차종구분 없음), 평일/휴일로 구분
  - 전국 코든라인(한국교통연구원) : 681개소, 6개 차종별 교통량



## ▶ 수도권 관측지점 분포



## ▶ 전국 관측지점 분포



- : 한국도로공사
- : 상시 교통량
- : 수시 교통량
- : KOTI-코든라인
- : 서울시 교통량

&lt;그림 1-44&gt; 수도권 및 전국 관측지점 분포도

## ② 교통량 전수화 프로세스

- 교통량 전수화 프로세스는 분석 맵 링크를 기준으로 관측교통량 별 입력→대표 Probe 입력→3종 교통량 변환→교통량 추정 단계로 구현되도록 시스템 개발
- 교통량 전수화는 전국 교통량 조사지점 이외 링크에 대한 교통량을 교통량 전수화 모형을 이용하여 산정하도록 개발. 대상은 전국 기준으로 고속도로, 자동차전용도로, 일반국도, 국가지원지방도, 지방도까지 생성되며, 특별광역시도로는 서울특별시만 생성



&lt;그림 1-45&gt; 교통량 전수화 UI

## 4) 혼잡비용

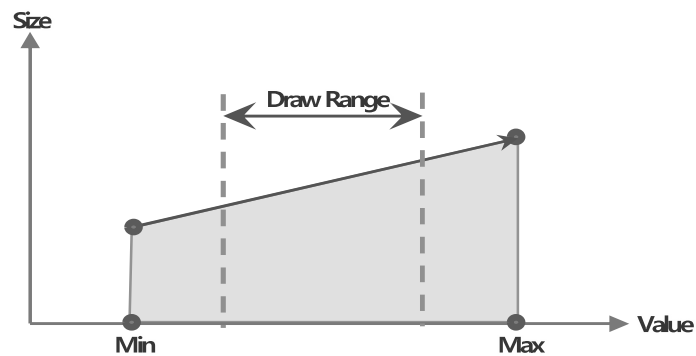
- 교통혼잡비용은 링크, 경계속도, 교통량, Probe 데이터, 행정구역 구분에 대한 데이터를 입력받아 차량운행비용(고정비+변동비), 시간가치비용에 대한 혼잡비용 모형을 이용하여 추정

할 수 있도록 개발

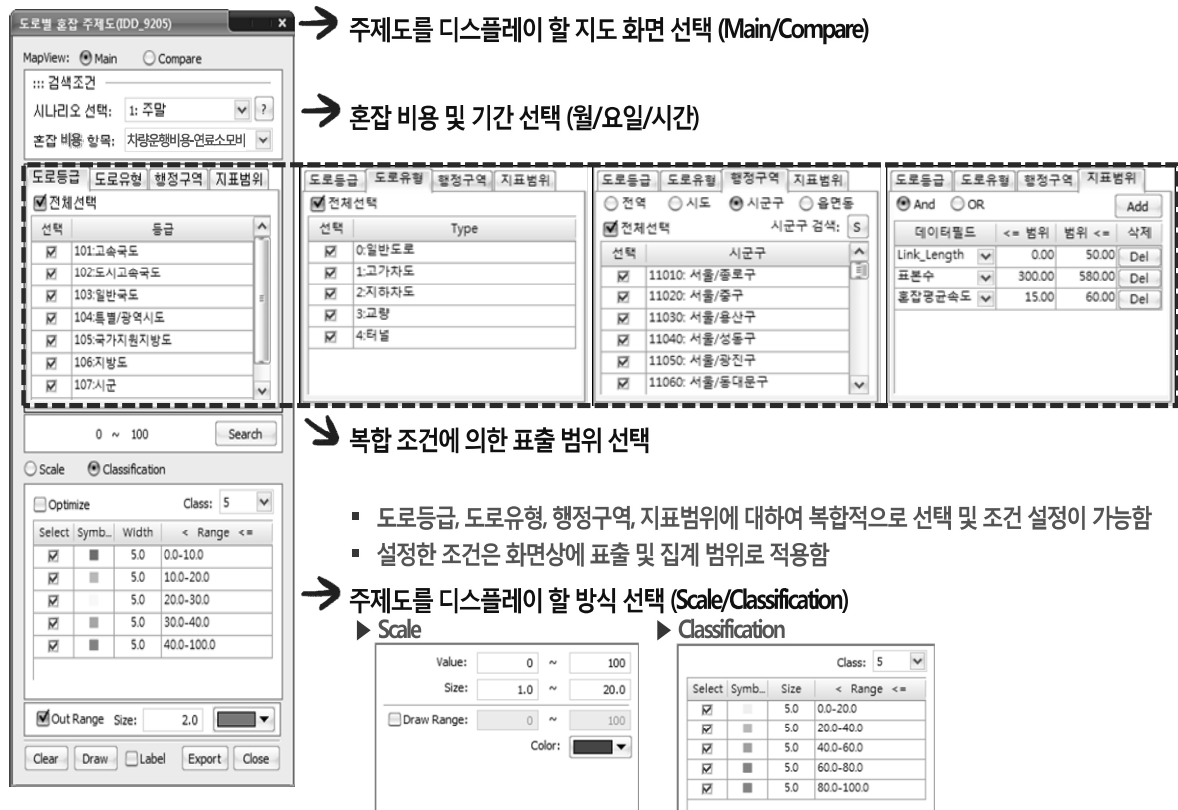
- 생성된 혼잡비용은 분석맵, 행정구역, Grid단위의 주제도 및 테이블로 결과 확인이 가능하도록 기능구현

##### 5) 주제도 기능 강화

- 주제도는 표준노드링크, 분석맵, 행정구역, Grid 단위로 다양한 주제도 표출이 가능하며, 각각의 주제도는 사용자가 복합적인 조건을 설정할 수 있도록 하여 분석하고자 하는 용도에 맞게 텍스트 파일 및 화면상에 디스플레이 할 수 있도록 구현
  - 주제도는 디스플레이할 지도 화면(Main/Compare)선택이 가능하며, 각 화면 별 주제도 설정을 사용자가 달리하여 비교분석 할 수 있도록 기능 구현
  - 복합검색 조건은 도로등급, 도로유형, 행정구역, 지표범위에 대하여 복합적으로 선택 및 조건 설정이 가능
  - 주제도는 Scale 또는 Classification방식으로 디스플레이 가능
  - Out Range는 결과 값이 없는 링크의 경우도 Out Range에서 정의한 칼라와 두께로 화면상에 디스플레이 됨
  - Scale로 주제도 표출 시 사용자가 지정하는 Value 값에 따라 화면상에 표출되는 주제도 스케일 크기를 조정할 수 있도록 구현



<그림 1-46> 스케일 설정에 따른 크기 변화



&lt;그림 1-47&gt; 주제도 UI

### ○ 도로 주제도

- 도로 주제도는 표준노드링크 또는 분석맵 네트워크로 주제도 표출이 가능하도록 개발
- 표준노드링크와 분석맵은 도로등급, 도로유형, 도로정보가 동일한 포맷으로 구성되어 있음
- 도로 주제도를 설정하는 순서는 1. 시나리오 선택(사용자가 기간, 시간 등을 설정 후 혼잡 비용을 생성한 데이터 명칭), 2. 혼잡비용 항목 선택, 3. Search 버튼 클릭, 4. 도로등급, 도로유형, 도로 축별, 행정구역, 지표범위 설정, 5. Scale 또는 Classification 선택, 6. 표출할 색상 및 두께 설정 후 Draw 버튼 클릭
- 도로등급은 고속국도, 도시고속국도, 일반국도, 특별/광역시도, 국가지원지방도, 지방도, 시군구로 구성되어 있으며, 선택된 항목에 대해서만 표출
- 도로유형은 일반도로, 고가차도, 지하차도, 교량, 터널로 구성되어 있으며, 선택된 항목에 대해서만 표출
- 행정구역은 시도/시군구/읍면동으로 구분되며 선택한 행정구역에 대해서만 도로 네트워크가 표출됨. 전역은 전국 도로 네트워크가 표출되도록 구성
- 도로축은 고속도로의 본선을 기준으로 도로축 데이터를 구축한 데이터로 축을 선택하면 선

택한 도로만 화면상에 표출

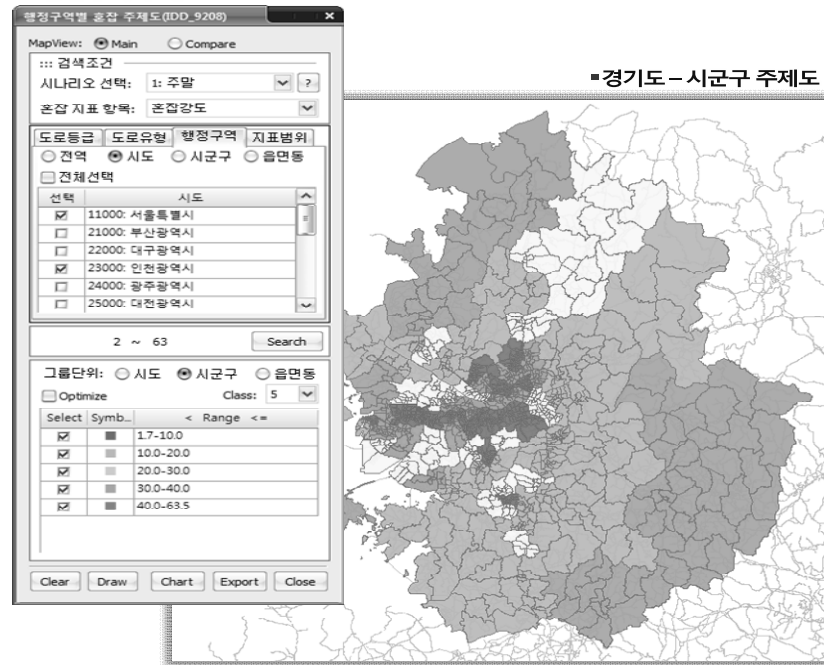


<그림 1-48> 도로 및 교차로 UI

#### ○ 행정구역 주제도

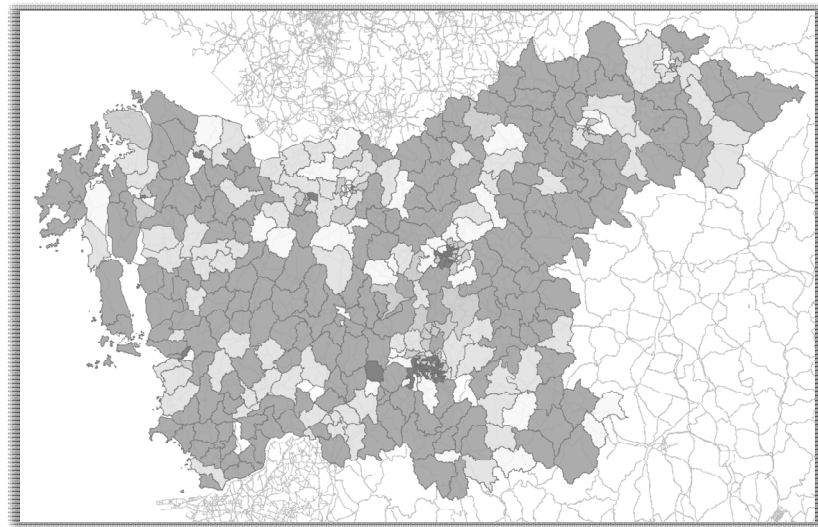
- 행정구역 주제도는 도로주제도와 동일하게 도로등급, 도로유형, 행정구역, 지표범위를 사용자가 복합 조건을 설정할 수 있도록 기능구현
- 행정구역 표출단위는 사·도, 시·군·구, 읍·면·동 단위의 주제도 디스플레이가 가능하며, 사용자의 선택에 따라 전국 및 사용자가 선택한 지역만 디스플레이 할 수 있도록 기능구현

## ▶ 행정구역 주제도 UI



&lt;그림 1-49&gt; 행정구역 주제도 UI 및 시군구 주제도 예시 화면

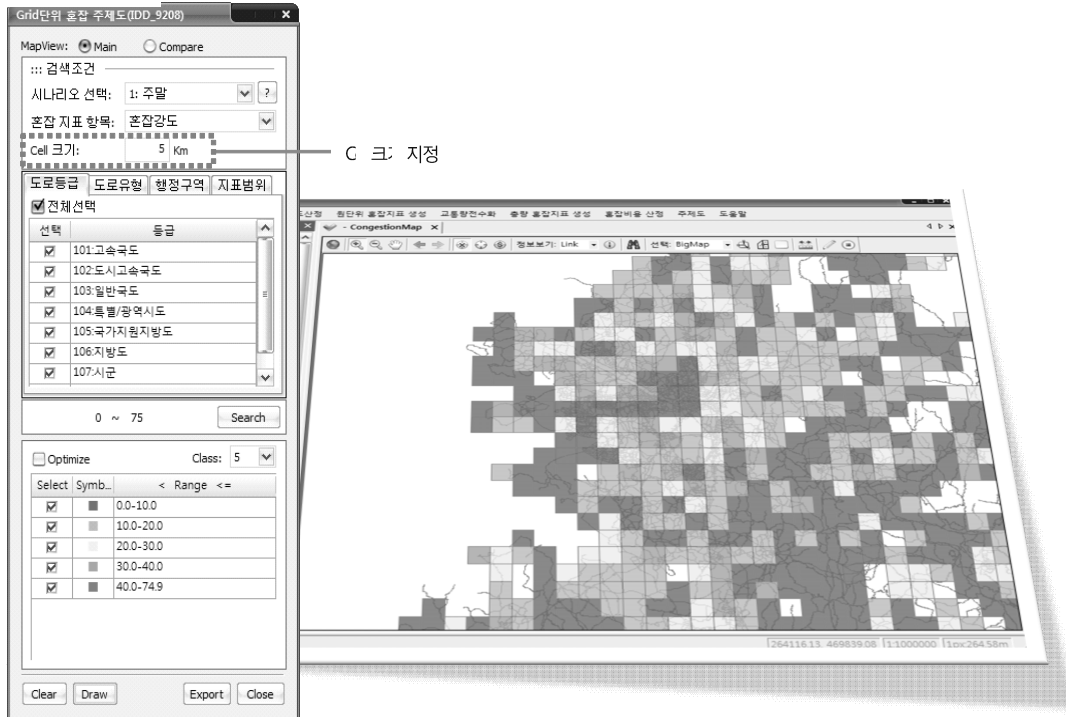
## ■ 충청도-읍면동 주제도



&lt;그림 1-50&gt; 읍면동 - 충청도 지역만 선택 시 디스플레이 결과

- 행정구역으로 공간 분석 시 지역별 영역의 범위와 도로 밀집도에 따라 정확한 주제도 분석이 힘들. 분석하기에 최적화된 최소단위를 사용자가 지정하여 그리드 단위의 분석이 가능하도록 기능구현

### ▶ Gri 주제! UI



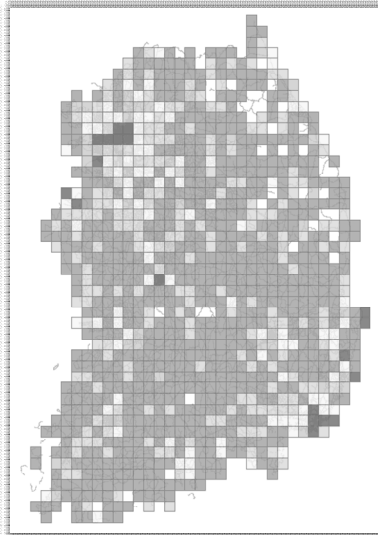
<그림 1-51> Grid 주제도 UI

○ Cell 사이즈를 5km/10km/20km 설정시의 주제도 예시화면은 다음과 같음

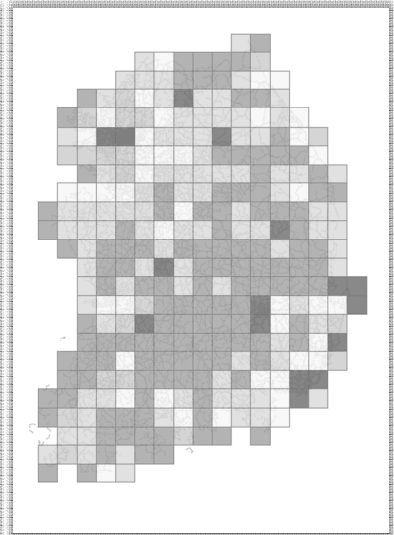
▪ Cell Size : 5km



▪ Cell Size : 10km



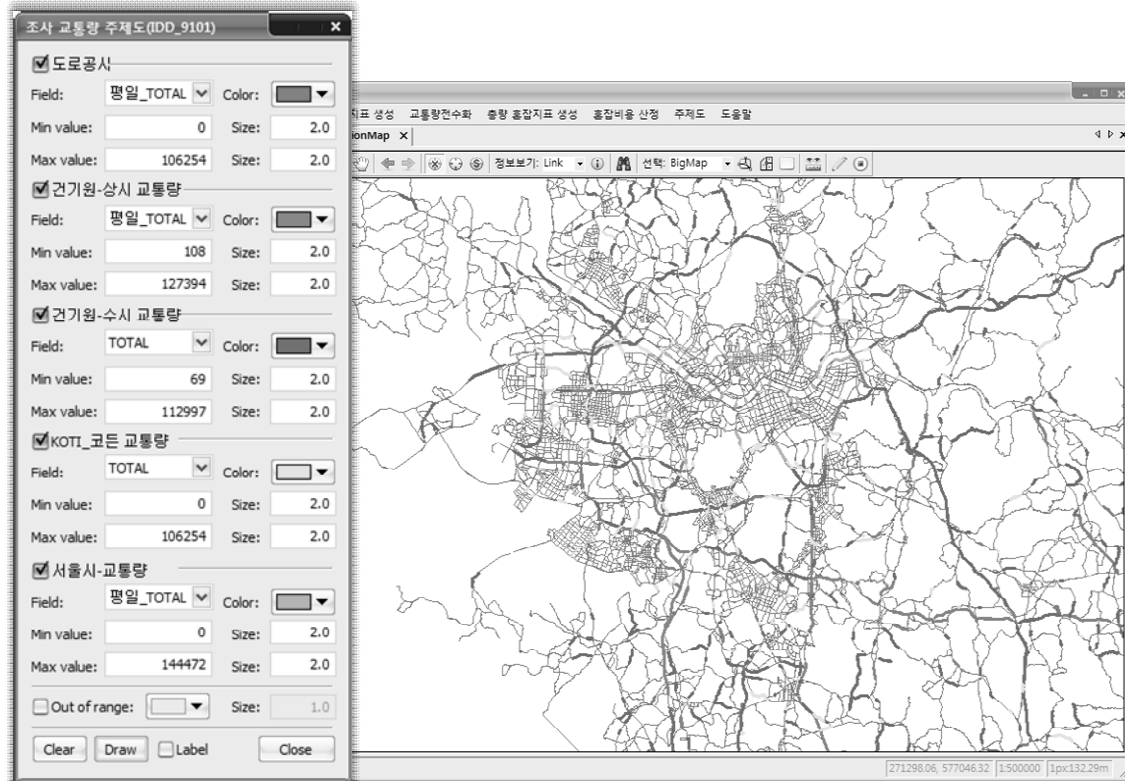
▪ Cell Size : 20km



<그림 1-52> 그리드 주제도 예시 화면

○ 표준노드링크와 맵핑한 관측 교통량 별 주제도 UI

## ▶ 관측 교통량 주제도 UI



&lt;그림 1-53&gt; 관측교통량 UI 및 예시 화면

## 6) 사용자 관리 기능 강화

- 사용자 검색 조건에 따라 다양하게 표출되는 주제도를 비교 분석 할 수 있도록 Multi Map 지원 기능 개발
- 지도의 스케일과 위치 정보를 상호 동기화 할 수 있도록 개발하여 서로 다른 주제도의 분석 결과를 쉽게 비교 분석할 수 있도록 함
  - 지도 스케일 및 위치 동기화는 한쪽의 지도를 이동했을 때 다른 편의 지도도 함께 동일 스케일 및 위치로 이동하는 기능
- 레이어 정의는 노드/링크, 행정경계, 분석맵을 사용자가 레이어를 셋팅 할 수 있도록 기능 개발



<그림 1-54> 레이어 정의 UI

- 수집한 네비게이션에 대한 이벤트 건수, Converge 등에 대한 기본 정보 제공

네비게이션 통계 정보(DD\_3840)

Navi Type: MN 기간: 2013 년 12 월

월단위 합계

OBU개수	이벤트건수	Link매칭정보	Coverage	보정후 이벤트건수	보정후 Link매칭정보	보정후 Coverage
4,888,797	401,343,793	215,122 / 216,496 / 256,160	83.98%	550,804,284	231,514 / 256,160	90.38%

지역별 이벤트건수(시도)시간대별 이벤트건수

지역	이벤트건수	비율	시간대	이벤트건수	비율
서울특별시	86,087,879	21.45%	00	10022836	02.50%
부산광역시	26,422,238	06.58%	01	7830490	01.95%
대구광역시	16,144,907	04.02%	02	5820215	01.45%
인천광역시	20,879,538	05.20%	03	4382874	01.09%
광주광역시	13,368,090	03.33%	04	3993154	00.99%

일단위 현황

일자	OBU개수	이벤트건수	Link매칭정보	Coverage	보정후 이벤트	보정후 Link매	보정후 Covera
1	156,988	12,547,496	165,143 / 166,4...	64.47%	17,228,858	183,763 / 256,1...	71.74%
2	174,868	13,816,571	168,671 / 169,9...	65.85%	19,058,133	186,291 / 256,1...	72.72%
3	170,954	14,023,060	170,202 / 171,4...	66.44%	19,311,210	188,003 / 256,1...	73.39%
4	170,508	13,950,151	172,494 / 173,7...	67.34%	19,229,694	189,634 / 256,1...	74.03%
5	170,395	14,118,458	173,347 / 174,5...	67.67%	19,422,851	190,102 / 256,1...	74.21%
6	177,052	14,793,016	176,611 / 177,8...	68.95%	20,461,801	194,136 / 256,1...	75.79%
7	174,724	14,806,726	177,414 / 178,6...	69.76%	20,506,095	195,092 / 256,1...	76.16%

Export닫기

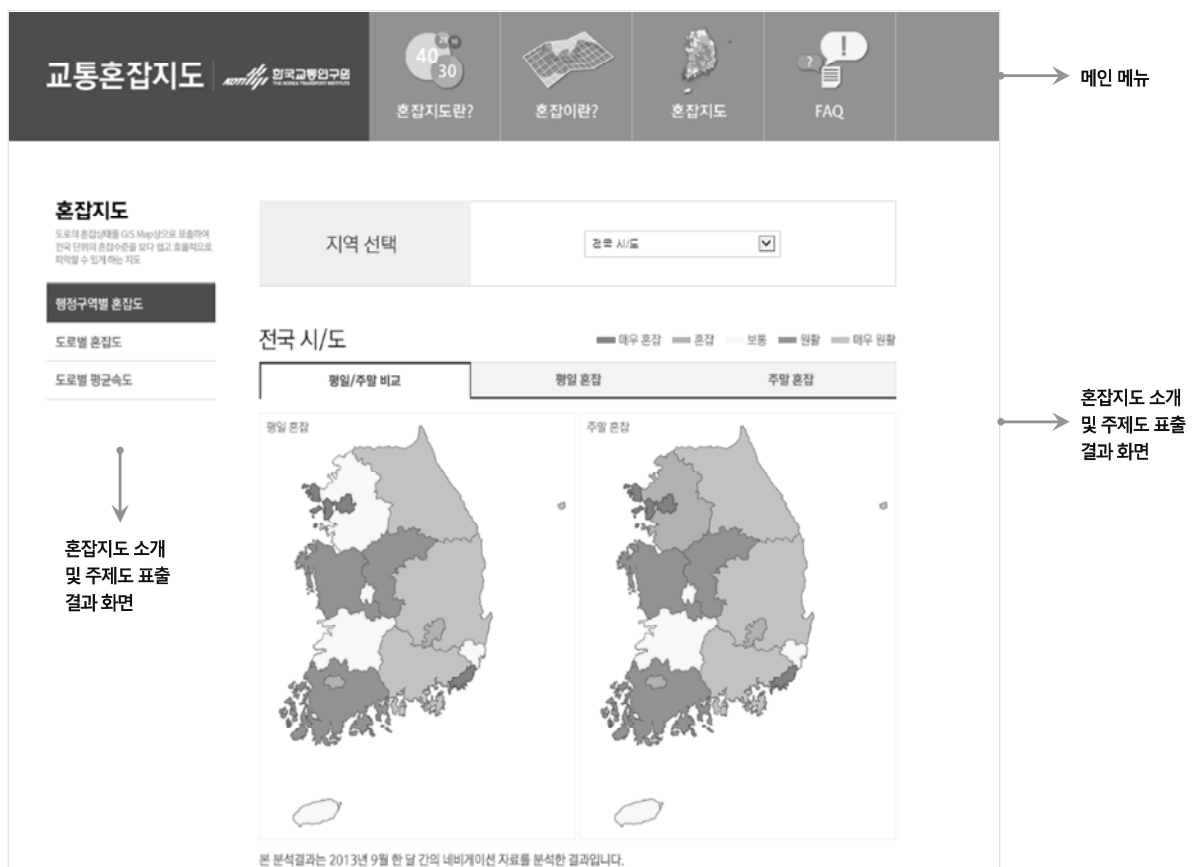
<그림 1-55> 2013년 12월 네비게이션 통계 결과



## 4. 전국교통혼잡지도 WebBook 작성

### 가. WebBook 구성

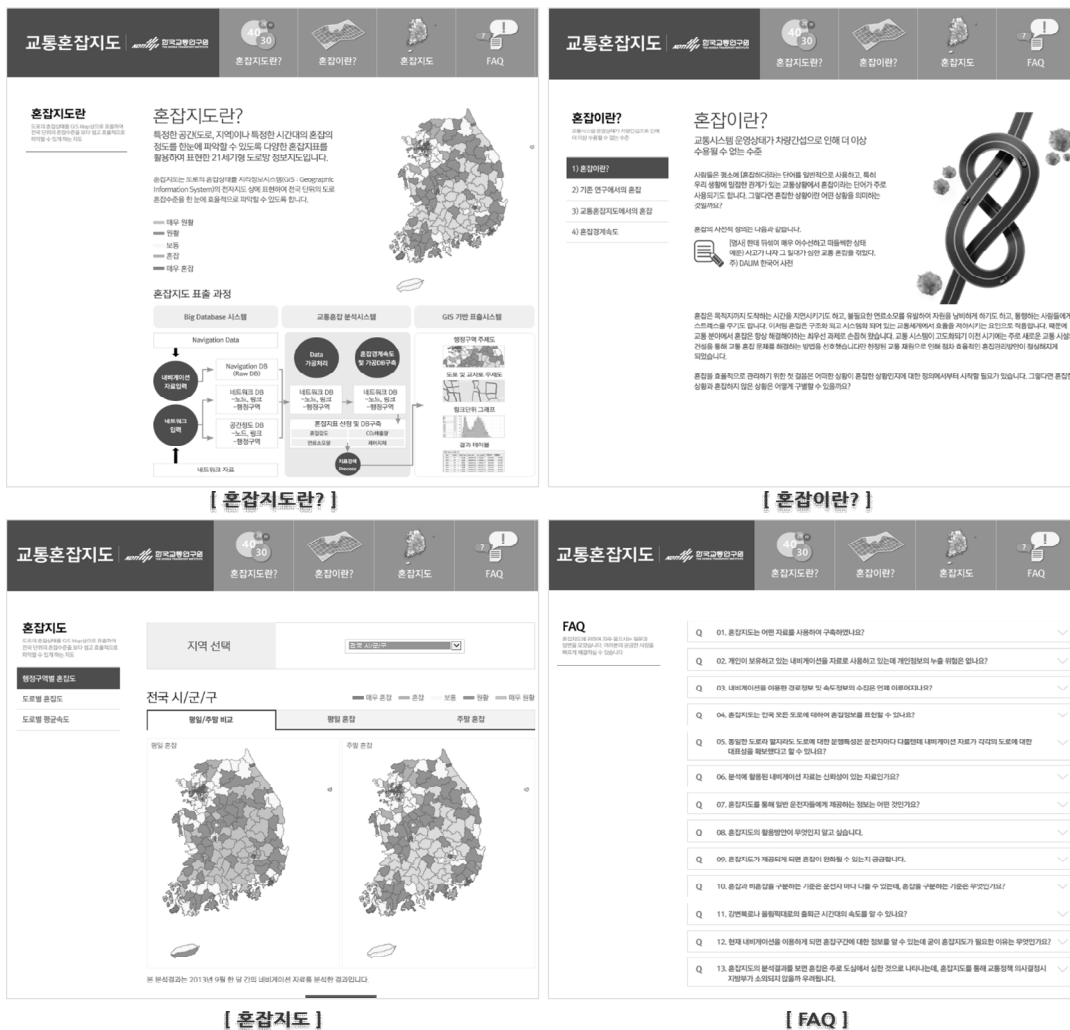
- WebBook을 통한 전국교통혼잡지도 서비스 내용은 크게 교통혼잡지도 소개 및 혼잡에 대한 결과를 서비스함
  - 분석 결과는 2013년 9월 한 달 간의 네비게이션 자료를 분석한 결과임
- 전국 교통혼잡지도 및 혼잡 소개 시 일반사용자도 쉽게 이해할 수 있도록 설명
- 혼잡도에 대한 서비스는 행정구역별, 도로별로 나뉘며, 광역시·도/시·군·구 별에 대한 평일/주말로 구분하여 혼잡도를 서비스함
- 화면구성은 메인 메뉴, 서브 메뉴, 주제도 결과 및 혼잡 소개 화면으로 구성
  - 메인 메뉴는 혼잡지도란, 혼잡이란, 혼잡지도, FAQ 로 구성



<그림 1-56> WebBook 화면구성

○ 각 메뉴별 화면구성은 다음과 같음

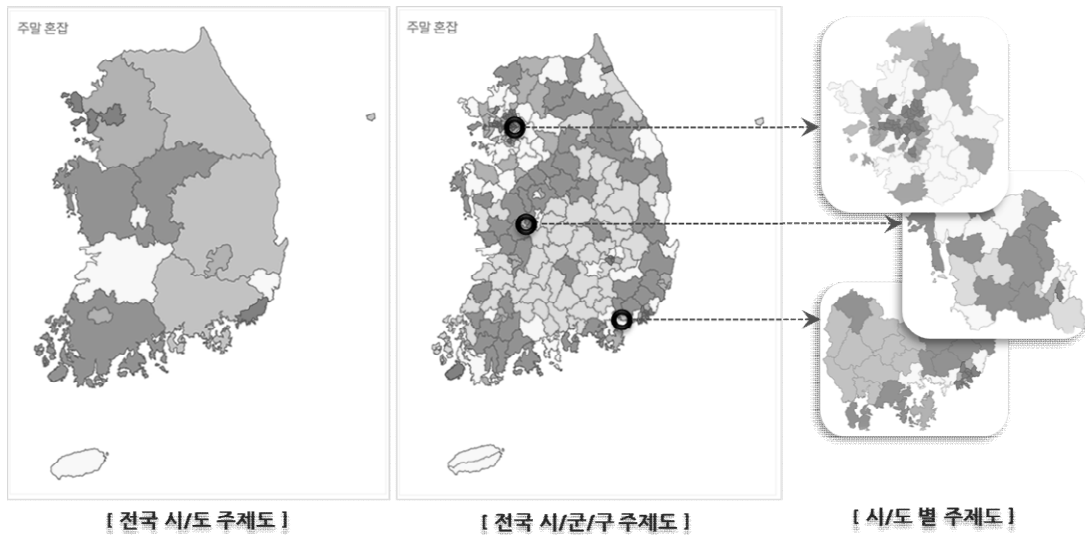
- 혼잡지도란 - 혼잡지도 소개 및 혼잡지도 표출 과정, 혼잡지표의 종류 기대효과 및 활용방안을 간략하게 설명
- 혼잡이란 - 전국교통혼잡지도 시스템에서 정의한 혼잡의 기준을 간략하게 정의
- 혼잡지도 - 행정구역/도로 별 혼잡도를 주제도로 표출
- 혼잡지표는 행정구역 별, 도로별 혼잡도와 평균속도를 주제도로 표출
- 주제도 표출 지역은 전국 단위, 시/도 단위, 시/군/구 단위, 읍/면/동 단위로 지역을 구분하여 평일/주말 주제도로 결과를 표출
- FAQ - 혼잡지도에 관하여 자주 묻는 질문과 답변을 게시함



<그림 1-57> 메뉴 별 화면 구성

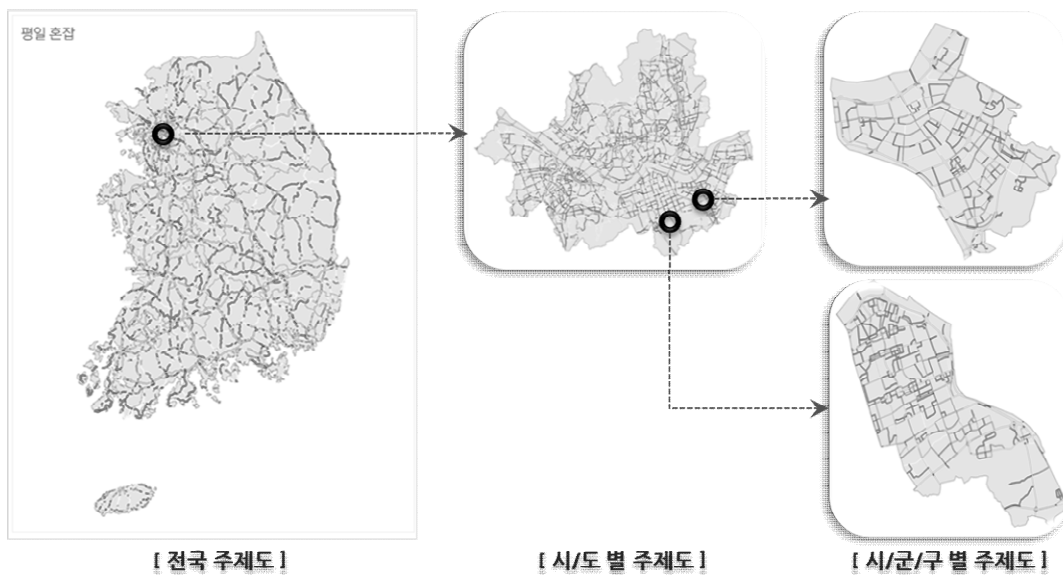
## 나. 혼잡지도 주제도 서비스

- 주제도는 크게 행정구역별 혼잡도, 도로별 혼잡도, 도로별 평균속도로 나뉘며, 혼잡의 표출 단계는 매우혼잡, 혼잡, 보통, 원활, 매우 원활 5단계로 구분함
- 행정구역별 혼잡도는 전국 시/도, 전국 시/군/구, 시/도 별 단위로 서비스함



<그림 1-58> 행정구역별 혼잡도

- 도로별 혼잡도/평균속도는 전국, 시/도 별, 시/군/구 별 단위로 서비스함

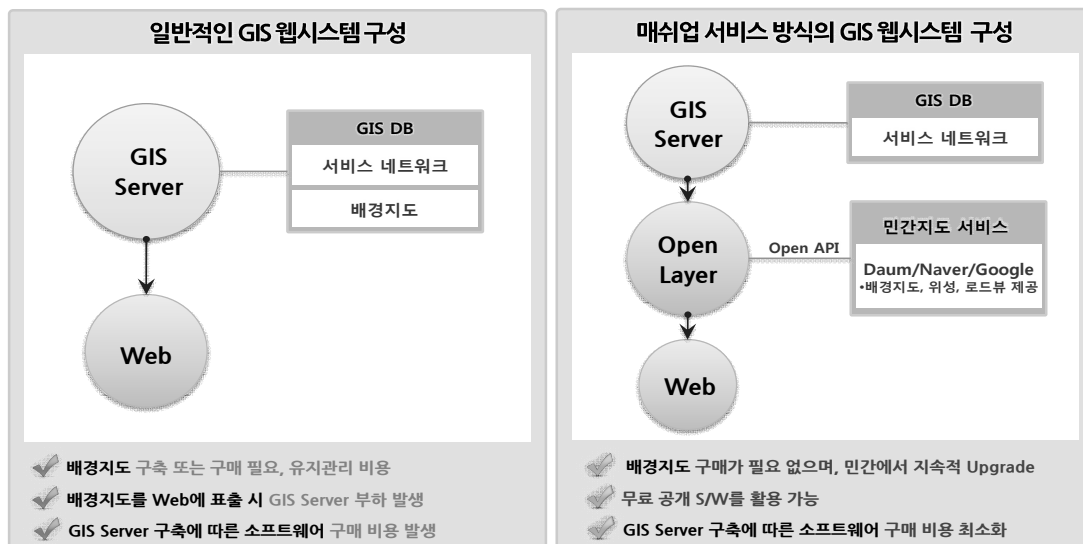


<그림 1-59> 도로별 혼잡도/평균속도

## 5. 교통혼잡지도 Web Service를 위한 GIS 웹 엔진 개발

### 가. GIS 웹 엔진 개발 방안

- 상용 S/W를 활용한 GIS 웹 엔진 구성은 고가의 소프트웨어 구매 비용 발생 및 배경지도 관리 비용이 발생하는 문제점이 있음. 대형 포털의 Open API 서비스를 활용할 경우 주제도 서비스 속도 저하 문제가 있음
- 대형 포털의 Open API 서비스를 통한 배경지도 서비스와 Open Source를 활용한 GIS 서버 구성을 통한 매쉬업 서비스 방식을 통해 구축 비용 절감 및 빠른 주제도 서비스가 가능할 수 있도록 구성함
  - 대형 포털은 배경지도 서비스 외에 부가적인 서비스를 지속적으로 업그레이드 하기 때문에 포털의 신뢰도 높은 서비스를 제공하고 있음. 대형 포털의 배경지도를 활용하기 때문에 배경 지도 서비스에 대한 구축 및 유지보수 비용이 발생하지 않음
  - 안정적인 주제도 서비스를 위해서는 GIS 서버구성이 필요하며, GIS 서버는 Open Source를 활용하여 추가적인 비용이 발생하지 않도록 함



<그림 1-60> GIS 시스템 구성 비교

- 국토교통부, 안전행정부, 지자체 Web GIS 활용사례를 살펴 보면 배경지도를 소유하고 있는 단체의 경우 자체적으로 서비스하는 방식으로 서비스 하며, 또한 Open API를 통해 자체의 지도를 서비스할 수 있도록 Open 되어있음. GIS Server는 일반적으로 GeoServer를 Web

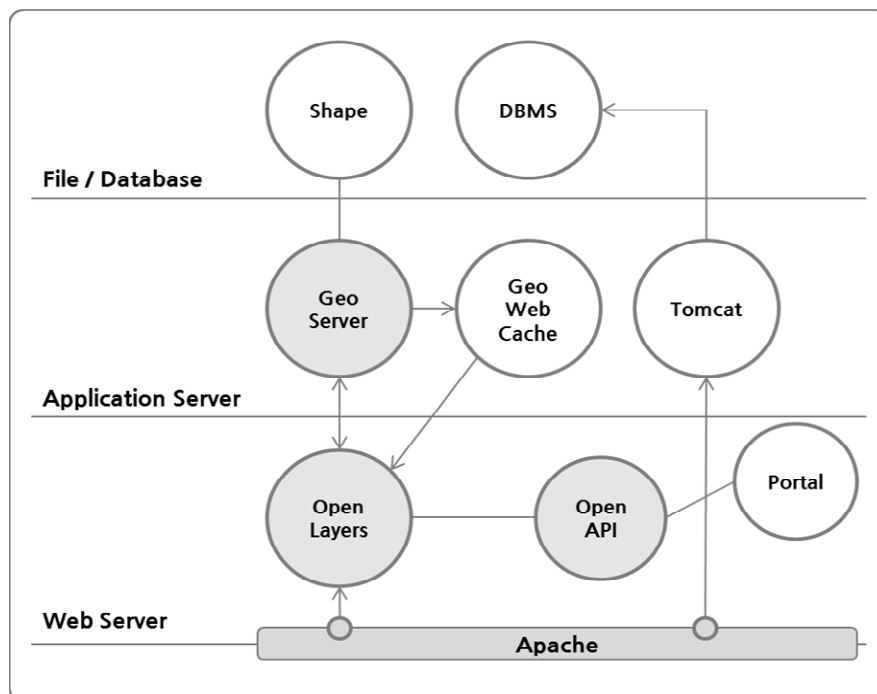
Client는 OpenLayers를 활용하여 구축

- 안전행정부의 생활공감지도의 경우 별도의 GIS Server를 구성하지 않고 Web Client를 통해 직접 주제도를 표현하기 때문에 속도 문제가 발생하여 서비스의 한계가 있음

<표 1-30> 포털지도서비스/Open Source 활용 사례

사례	배경지도	GIS Server	Web Client
서울교통정보센터 (서울시)	Daum	GeoServer	OpenLayers
국가공간정보통합체계 (국토교통부)	자체	-	OpenLayers
생활공감지도 (안전행정부)	자체	-	OpenLayers
국토지리정보원	자체	GeoServer	OpenLayers
한국항공우주연구원	자체	GeoServer	OpenLayers
모바일 예보관 기상분석 (기상청)	-	GeoServer	OpenLayers
국가교통센터(모바일교통정보)	Google	GeoServer	OpenLayers
청주시 버스정보시스템	Daum		Deaum
부산시 버스정보시스템	Daum/Naver		Daum, Naver

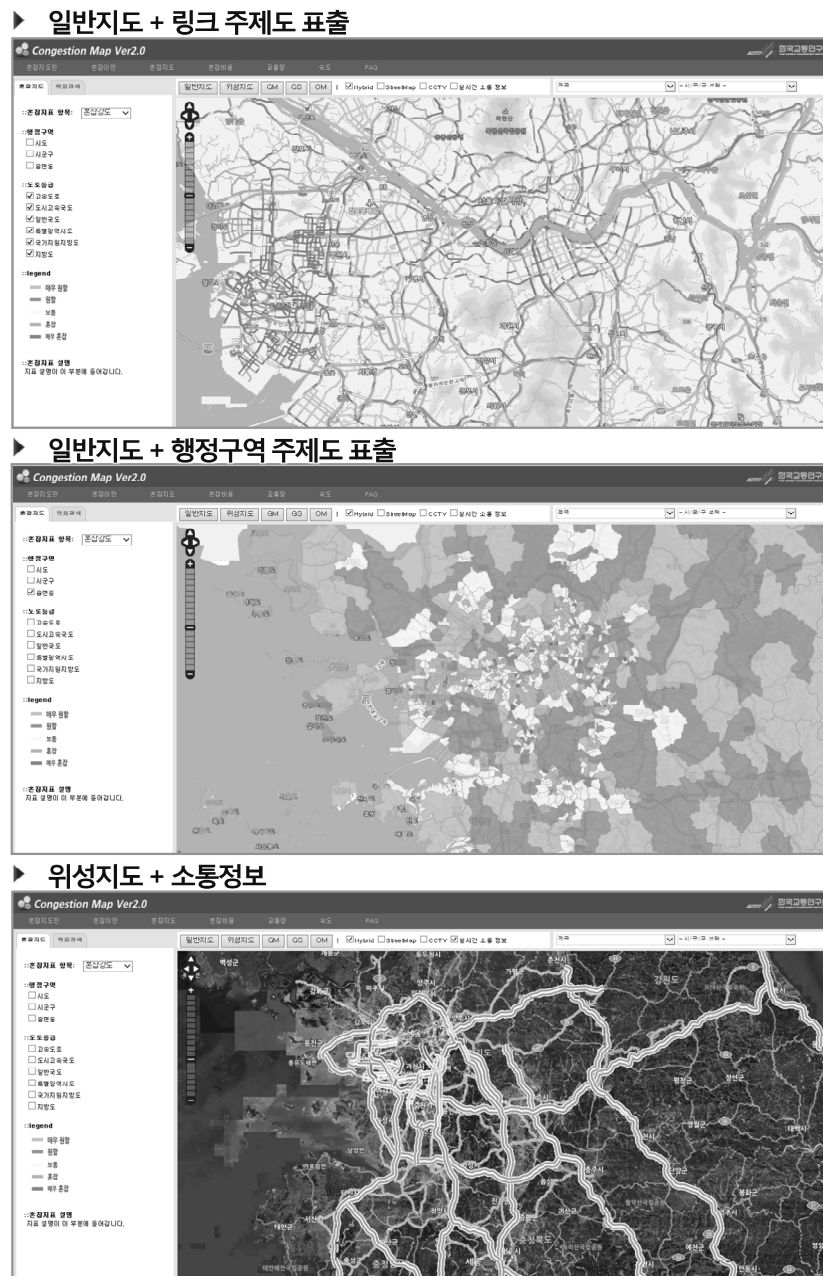
- 혼잡지도 GIS 시스템은 Open Layers를 통한 포털 배경지도 서비스, 주제도서버인 GeoServer를 연동하여 Web GIS Service를 구성



<그림 1-61> 혼잡지도 Web 시스템 구성도

## 나. 혼잡지도 Web 시스템 구성 적용

- 혼잡지도 Web 시스템은 Open Layers 자바스크립트 웹 클라이언트를 통한 VWorld/Google/ OpenStreet Map 배경지도 서비스, Geo Server를 통한 혼잡지도 주제도 서비스, Open Data API를 통한 데이터 서비스로 구성
- 웹서비스 화면은 다음과 같음



<그림 1-62> 웹 서비스 예시 화면

## 제5절 전국 교통혼잡지표 및 혼잡비용 추정결과 집계

### 1. 개요

- 본 연구에서는 확장된 Car Navigation 자료를 이용하여 관련 교통혼잡지표 산출 알고리즘을 검증하였으며, 교통 혼잡에 대한 객관적이고 정확한 진단을 위하여 전국 교통 혼잡지도의 고도화 연구를 수행하였음
- 이러한 개발모형은 ①학문적 합리성, ②(미시적 수준에서) 다양한 혼잡관련 지표의 산정, ③시스템 적용성, ④운영자 편리성, ⑤개발 모형의 확장가능성과 같이 다섯 가지의 요구조건을 고려하여 개발모형의 성능을 고도화하였음
  - ①학문적 합리성 : 개발모형에 의하여 추정된 혼잡경계속도는 교통공학적인 교통류 이론을 기반으로 합리적인 경계속도를 도출해야 하며, 대상 교통류는 고속도로 교통류(연속류, Uninterrupted traffic flow), 신호교차로 교통류(단속류, Interrupted traffic flow), 기타 복합적 요인에 의한 교통류에 대하여 거시적으로 합리적인 결과를 도출해야 함
  - ②다양한 혼잡관련 지표의 산정 : 개발모형으로 산정된 경계속도와 추정된 속도의 분포는 미시적 수준에서 다양한 혼잡관련 지표를 정밀하게 산정해야 하며, 이를 위해서 기존의 집계방식이 아닌 비집계 방법론을 기반으로 도로구간별 속도특성을 파악할 수 있도록 해야 함
  - ③시스템 적용성: 표준 노드-링크체계의 링크 개수는 약 25만개이며, 링크별 도로의 위계 및 교통류의 형태는 명확하지 않음. 따라서 도로의 위계 및 교통류의 행태가 구분되지 않아도 적용이 가능해야 함. 표준 노드-링크체계의 링크 개수는 약 25만개이며, 본 연구에서 이용된 9월 한 달 동안의 개별 GPS 링크 이동류별 속도자료는 약 5억 건에 달함. 따라서 자료관리시스템(Data management system)을 기반으로 빠른 연산 수행속도를 가지도록 모형의 알고리즘을 설계해야함. 자료관리 시스템과 연동되어 개발 시스템에 탑재가 용이해야 함. 향후 지속적인 시스템 업데이트를 고려하여 모형의 구조적인 알고리즘 변경과 개선이 용이해야 함
  - ④운영자 편리성 : 시스템에 탑재된 모형이 비록 매우 정교한 모형일지라도 지속적인 파라미터 갱신이 필요하다면 모형의 효율성은 시스템 구축 후 현저히 떨어지게 되며, 파라미터의 Update에 많은 시간이 소요됨. 따라서 개발모형의 입력자료인 해당 링크별 개별 GPS 속도자료 이외에 모형의 파라미터를 최소화하여 운영시 효율성을 극대화해야 함. 도로별 특

성과 교통류별 특성이 변화해도 시스템에 탑재된 모형의 알고리즘 구조변경과 파라미터의 재 정산 과정없이 적용이 가능해야 함. 개발모형의 수행결과는 명료하고 직관적인 분석결과를 제공함으로써 교통전문가 뿐만 아니라 타 분야 전문가 및 일반인도 쉽게 이해할 수 있어야 함

- ⑤확장 가능성 : 국내의 경우 GPS속도자료는 다양한 형태로 존재함. 따라서 일부 GPS속도자료에 의존하지 않고 범용적으로 적용이 가능한 프레임워크로 모형을 개발해야 함. 개발 모형은 혼잡 경계속도 뿐만 아니라 향후에 적용되는 혼잡예보, 여러 시간대에 걸친 교통소통상태 예측 등에 필요한 입력자료 및 예측 지원기능을 수행할 수 있도록 확장 가능성을 고려해야 함
- 특히 혼잡경계속도 판정 알고리즘의 경우, 국/내외적으로 혼잡기준속도에 대한 정의가 매우 다양하므로 교통 혼잡에 대한 개념을 정립하고 이를 판정하기 위한 범용적인 모형으로 개발하였음
- 위의 과정을 통하여 고도화한 개발모형을 이용하여 다양한 시간적 범위(주중, 주말, 시간)에 따라 아래와 같이 총 다섯 가지의 교통혼잡지표(원단위 또는 총량)를 산출하였음
  - 혼잡강도( $CI$ ; Congestion Intensity, %)
  - $CO_2$ 배출량( $COE$ ;  $CO_2$  Emission,  $g/km/대$  또는  $g, kg, ton$ )
  - 연료소모량( $FC$ ; Fuel Consumption,  $l/km/대$  또는  $l, kl$ )
  - 지체시간( $CD$ ; Control Delay, 초/대 또는 초, 분, 시)
  - 교통혼잡비용( $TCC$ ; Traffic Congestion Cost, 억원/년)
- 따라서, 본 장에서는 먼저 원단위 지표가 아닌 총량의 지표를 산출하기 위하여 개발된 본 연구의 교통량 전수화 모형에 대한 평가를 수행하고, 이어 전국 링크 단위로 산출된 교통혼잡 지표(혼잡강도, 교통혼잡비용)를 도로등급과 지역 단위로 집계하여 그 결과를 제시하였음
- 여기서, 교통혼잡비용의 경우는 지역간 도로와 도시부 도로로 구분하여 교통혼잡비용의 집계 결과를 제시하였으며, 도시부 도로는 도시 내 내부도로의 관측교통량 조사가 상시적으로 수행되고 있는 서울특별시를 대상으로 교통혼잡비용 집계 결과를 제시하였음



## 2. 교통량 전수화 평가 결과

- 개별 Car Navigation 속도 데이터가 수집되는 ITS 표준노드링크 구간에서 대상 링크와 인접한 링크의 관측교통량과 대상 링크의 Car Navigation Probe 대수를 이용하여 교통량 전수화 모형을 개발하여 전국 관측교통량 링크구간을 제외한 모든 링크 구간의 차종별 교통량을 추정하였음
- 본 연구의 전수화 교통량에 대한 평가는 2013년 기준전국 관측교통량 조사지점 매칭 링크를 대상으로 실제 관측교통량과 추정된 전수화 교통량에 대한 오차율(%)로 평가를 수행하였으며, 세부내용은 아래와 같음
  - 전국 표준링크와 관측교통량 조사지점 매칭 : 총 3,408개소(6,816개 링크)
  - 국토교통부 상시조사지점 : 전국 총 769개소(1,538개 링크)
  - 국토교통부 수시조사지점 : 전국 총 2,543개소(5,086개 링크)
  - 서울시 교통량 : 총 96개소(192개 링크)
  - 평균절대비율오차(MAPE: Mean Absolute Percentage Error)

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

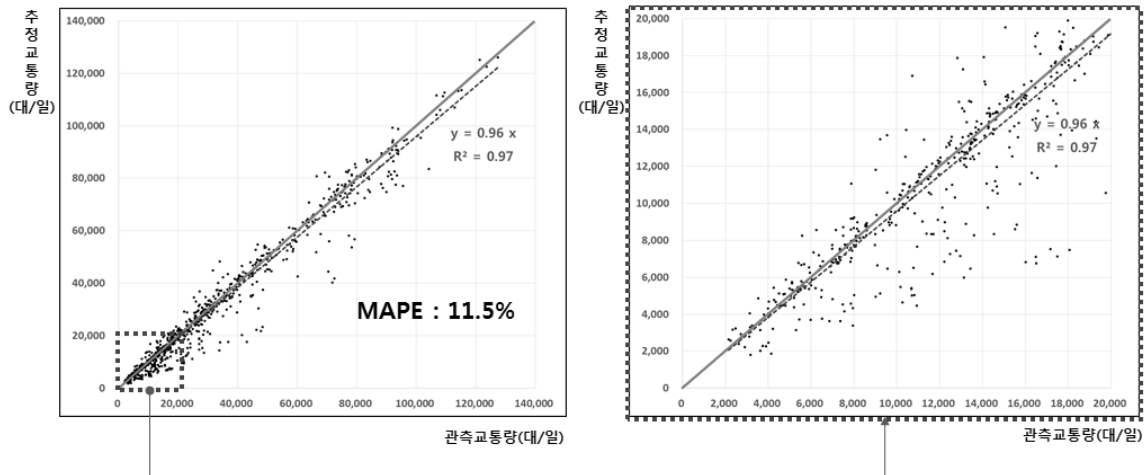
여기서,  $V_i$  = 링크  $i$ 의 관측교통량(대/일)

$\hat{V}_i$  = 링크  $i$ 의 전수화 추정교통량(대/일)

### 가. 전국 지역간 도로

#### 1) 고속도로

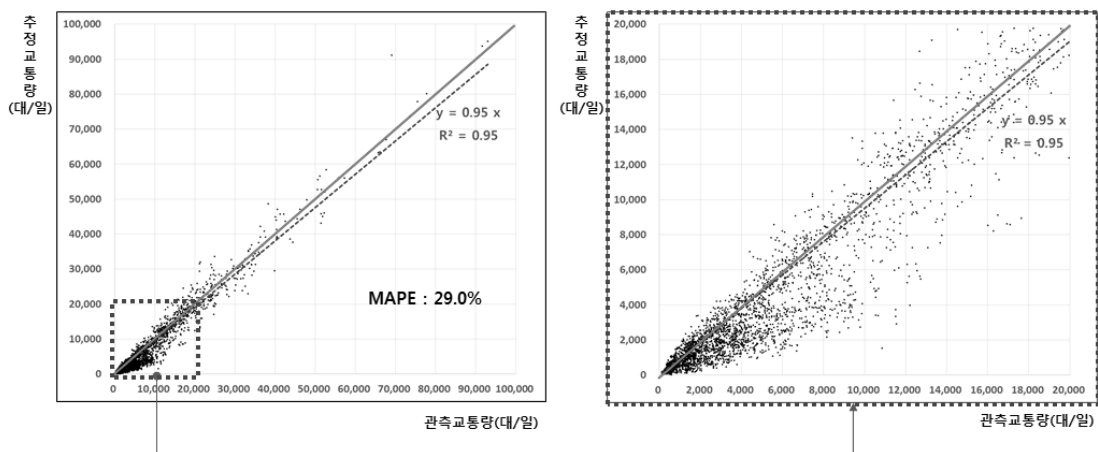
- 전국 고속도로의 관측교통량 조사지점 총 509개소(1,018개 링크 매칭)에 대하여 실제 관측교통량과 전수화 교통량에 대한 평가결과는 <그림 1-63>과 같으며, 전체 관측교통량 조사지점을 기준으로 MAPE 11.5%,  $R^2=0.97$ 로 전체적으로 모형의 평가가 우수한 것으로 분석되었음
- 또한, 고속도로의 경우는 일반국도와 지방도에 비하여 전수화 교통량의 평가결과가 우수하게 나타났으며, 이는 고속도로의 링크별 관측교통량과 Probe건수의 상관관계가 좀 더 선형에 가까운 것으로 판단됨. 즉, 본 연구의 링크별 고속도로 전수화 모형을 이용하여 관측교통량 조사지점 이외의 링크구간에 대한 교통량 전수화가 가능할 것으로 판단됨



<그림 1-63> 전국 고속도로 전수화 교통량 평가결과(전체, 평일기준)

## 2) 일반국도

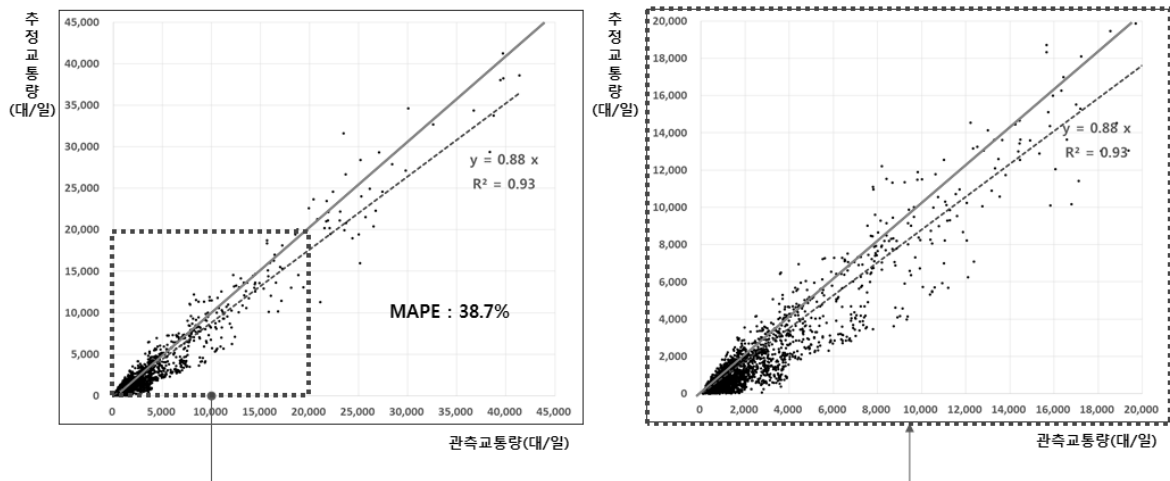
- 전국 고속도로의 관측교통량 조사지점 총 1,198개소(2,396개 링크 매칭)에 대하여 실제 관측 교통량과 전수화 교통량에 대한 평가결과는 <그림 1-64>와 같으며, 전체 관측교통량 조사지 점을 기준으로 MAPE 29.0%,  $R^2=0.95$ 로 앞서 고속도로의 경우 보다 추정교통량의 다소 오 차가 높은 것으로 분석되었음
- 이는 링크별 Probe건수(건/일)의 경우, Car Navigation 자료의 수집기간인 2013년 3개월 (2013년 10월 ~ 2013년 12월)동안의 링크별 평균 Probe건수이므로 2013년 기준 전체 1년 단위의 평균 Probe건수가 아님. 즉, 이러한 오차는 관측교통량과 Probe건수의 수집시점에 대한 시간적인 gap으로 인하여 발생하는 variation의 영향이 크므로 1년 이상의 Car Navigation 자료에 대한 확장이 필요할 것으로 판단됨



<그림 1-64> 전국 일반국도 전수화 교통량 평가결과(전체, 평일기준)

## 3) 지방도

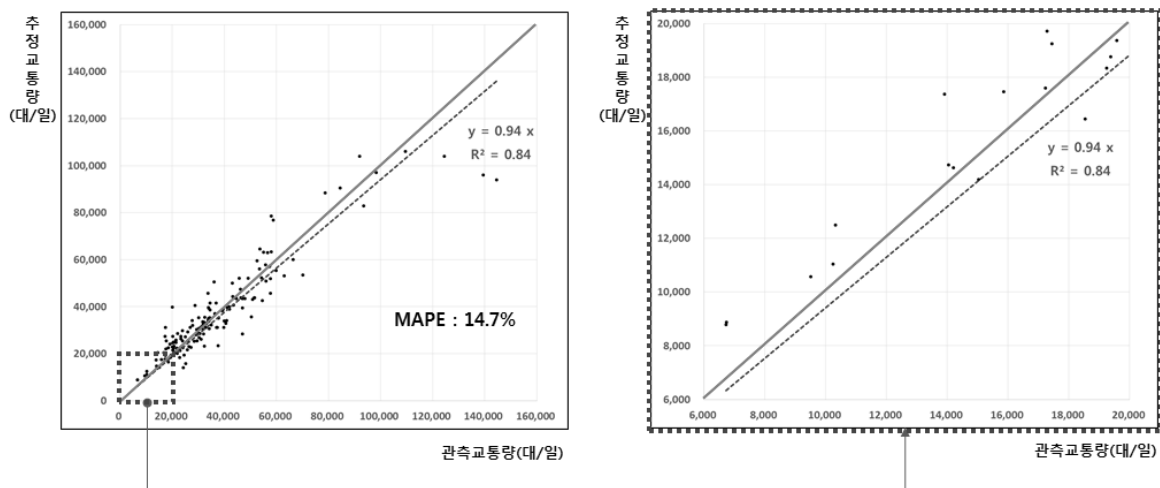
- 전국 고속도로의 관측교통량 조사지점 총 1,494개소(2,988개 링크 매칭)에 대한 평가결과는 <그림 1-65>와 같으며, 전체 관측교통량 조사지점을 기준으로 MAPE 38.7%,  $R^2=0.93$ 로 앞서 고속도로와 일반국도의 경우보다 추정교통량의 오차가 다소 높은 것으로 분석되었음



<그림 1-65> 전국 지방도 전수화 교통량 평가결과(전체, 평일기준)

## 나. 도시부 도로 : 서울특별시

- 서울특별시의 관측교통량 조사지점 총 96개소(192개 링크 매칭)에 대한 평가결과는 <그림 1-66>과 같으며, 전체 관측교통량 조사지점을 기준으로 MAPE 14.7%,  $R^2=0.84$ 로 전체적으로 모형의 평가가 우수한 것으로 분석되었음



<그림 1-66> 서울특별시 전수화 교통량 평가결과(전체, 평일기준)

## 2. 교통혼잡지표 분석 결과

- 전국 링크별로 산출된 교통혼잡지표 중 혼잡강도(%)에 대하여 광역시/도, 시/군/구 단위의 지역별로 집계하여 아래와 같이 각 CASE별로 분석 결과를 제시하였음
- CASE1(고속도로 대상), CASE2(고속도로를 제외한 나머지 도로 대상), CASE3(전체 도로 대상)

### 가. 전국 광역시/도 단위 분석결과

#### 1) 고속도로의 경우(CASE1)

- 전국 고속도로를 대상으로 광역시/도 단위의 집계를 수행한 결과, 주중 기준 전국 평균 11.7% 보다 혼잡한 광역시/도는 총 6개 지역으로 서울특별시, 인천광역시, 부산광역시 등의 순으로 분석되었음
- 주말의 혼잡강도(%)는 고속도로의 특성 상 도시지역 보다는 지방의 관광지역을 많이 포함하는 강원도, 경상북도, 충청북도를 제외하고 지역별로 주중 보다 낮은 것으로 분석되었음

<표 1-31> 전국 광역시/도 단위 혼잡강도(%) 분석 결과(CASE1)

순위	지역명	혼잡강도(%)		
		주중①	주말②	증감(②-①)
1	서울특별시	36.3	33.1	-3.3
2	인천광역시	24.1	14.4	-9.8
3	부산광역시	23.3	20.6	-2.7
4	경기도	15.0	12.5	-2.5
5	광주광역시	14.7	10.8	-3.9
6	대구광역시	10.2	9.7	-0.5
7	경상남도	9.9	8.8	-1.0
8	대전광역시	8.2	6.7	-1.5
9	울산광역시	8.2	7.0	-1.2
10	충청남도	7.7	6.6	-1.1
11	강원도	6.8	12.7	5.8
12	전라남도	5.9	4.5	-1.4
13	경상북도	4.8	5.0	0.3
14	전라북도	4.3	3.6	-0.7
15	충청북도	4.0	4.4	0.4
16	세종특별자치시	3.5	2.5	-1.0
평균		11.7	10.2	-1.5

주 : 지역별 순위는 주중 혼잡강도 기준 내림차순임.

## 2) 기타 도로의 경우(CASE2, 고속도로 제외)

- 전국 고속도로를 제외한 나머지 도로를 대상으로 광역시/도 단위의 집계를 수행한 결과, 앞서 주중 기준 고속도로의 경우보다 전체적으로 혼잡강도가 높게 분석되었음. 이는 지역간 고속도로를 제외한 지역별 도시 내부도로들에 대한 교통혼잡이 포함되어 있기 때문임
- 또한, 지역별로 혼잡강도(%)를 살펴보면 주중 기준 전국 평균 29.0% 보다 혼잡한 광역시/도는 총 7개 지역으로 인천광역시, 대구광역시, 부산광역시 등의 순으로 분석되었음. 이는 지역간 고속도로를 제외한 나머지 도로의 혼잡강도(%)를 집계하므로 지역 간 통행 보다는 해당 지역의 내부통행량이 높은 지역일수록 다른 지역보다 상대적으로 혼잡강도가 높은 것으로 판단됨
- 주중 대비 주말의 혼잡강도(%)는 지역별 모두 낮게 분석되었으며, 이는 지역별로 주중보다 주말 도시 내부도로의 전체 통행량이 감소하기 때문인 것으로 판단됨

&lt;표 1-32&gt; 전국 광역시/도 단위 혼잡강도(%) 분석 결과(CASE2)

순위	지역명	혼잡강도(%)		
		주중①	주말②	증감(②-①)
1	인천광역시	45.8	41.8	-3.9
2	대구광역시	43.6	42.3	-1.4
3	부산광역시	40.4	38.9	-1.5
4	서울특별시	40.1	36.8	-3.3
5	광주광역시	36.1	33.7	-2.4
6	대전광역시	35.0	33.5	-1.5
7	경기도	29.0	25.4	-3.6
8	울산광역시	27.7	25.0	-2.7
9	전라북도	27.1	24.7	-2.4
10	경상남도	23.6	21.1	-2.5
11	제주특별자치도	22.7	20.9	-1.8
12	충청북도	22.3	18.5	-3.8
13	경상북도	22.1	19.9	-2.2
14	강원도	21.2	17.8	-3.4
15	충청남도	21.0	18.4	-2.6
16	전라남도	20.2	17.6	-2.6
17	세종특별자치시	15.9	15.1	-0.8
평균		29.0	26.5	-2.5

주 : 지역별 순위는 주중 혼잡강도 기준 내림차순임.

## 3) 전체 도로의 경우(CASE3)

- 전국 전체 도로를 대상으로 광역시/도 단위의 집계를 수행한 결과, 전체적으로 CASE1과 CASE2의 중간 정도의 혼잡강도로 분석되었음. 이는 지역간 고속도로와 지역별 내부도로의 교통혼잡이 모두 포함되어 있기 때문임
- 또한, 지역별로 혼잡강도(%)를 살펴보면 주중 기준 전국 평균 25.7% 보다 혼잡한 광역시/도는 총 7개 지역으로 서울특별시, 부산광역시, 대구광역시 등 인구밀도와 자동차 등록대수가 높은 대도시 순으로 분석되었음
- 주중 대비 주말의 혼잡강도(%)는 지역별 모두 낮게 분석되었으며, 이는 지역별로 주중보다 주말 도시 내부도로의 전체 통행량이 감소하기 때문인 것으로 판단됨

&lt;표 1-33&gt; 전국 광역시/도 단위 혼잡강도(%) 분석 결과(CASE3)

순위	지역명	혼잡강도(%)		
		주중①	주말②	증감(②-①)
1	서울특별시	40.0	36.7	-3.3
2	부산광역시	38.3	36.5	-1.8
3	대구광역시	35.0	33.6	-1.5
4	인천광역시	34.7	27.9	-6.8
5	광주광역시	32.9	30.1	-2.8
6	대전광역시	30.4	28.6	-1.8
7	경기도	26.3	22.9	-3.4
8	울산광역시	25.4	22.7	-2.8
9	전라북도	23.1	20.5	-2.6
10	제주특별자치도	22.7	20.9	-1.8
11	경상남도	20.9	18.3	-2.6
12	경상북도	18.9	17.3	-1.6
13	충청남도	18.5	15.9	-2.5
14	강원도	18.4	16.6	-1.8
15	충청북도	18.4	15.6	-2.8
16	전라남도	16.9	14.5	-2.5
17	세종특별자치시	15.4	14.4	-0.9
평균		25.7	23.1	-2.5

주 : 지역별 순위는 주중 혼잡강도 기준 내림차순임.

## 나. 전국 시/군/구 단위 분석결과

### 1) 고속도로의 경우(CASE1)

- 전국 고속도로를 대상으로 시/군/구 단위의 집계를 수행한 결과, 주중 기준 전국 평균 10.6% 보다 혼잡한 시/군/구는 전체 166개 지역 중 64개 지역(전체 지역의 39.8%)으로 경기 부천시 원미구, 서울 서초구, 부산 북구 등의 순으로 분석되었음. 이는 전국 고속도로를 대상으로 혼잡강도(%)를 집계하므로 해당 지역의 고속도로 비중이 높고 고속도로의 통행량이 높은 지역일수록 다른 지역보다 상대적으로 혼잡강도가 높은 것으로 판단됨
- 주말의 혼잡강도(%)는 고속도로의 특성 상 도시지역 보다는 일부 지방의 관광지역(경북 문경시, 경북 경주시 등)에서 주중 보다 높은 것으로 분석되었음

<표 1-34> 전국 시/군/구 단위 혼잡강도(%) 분석 결과(CASE1)

순위	지역명	혼잡강도(%)		
		주중①	주말②	증감(②-①)
1	경기 부천시 원미구	47.6	34.3	-13.3
2	서울 서초구	47.2	45.0	-2.2
3	부산 북구	37.6	28.2	-9.5
4	서울 금천구	36.2	35.6	-0.7
5	서울 양천구	32.3	27.6	-4.7
6	인천 남동구	28.2	11.4	-16.8
7	인천 서구	27.8	10.6	-17.3
8	인천 부평구	27.1	22.1	-5.0
9	대구 서구	27.0	26.7	-0.4
10	전남 고흥군	26.9	20.1	-6.7
⋮	⋮	⋮	⋮	⋮
157	경북 경산시	2.5	1.0	-1.5
158	경남 창원시 의창구	2.5	2.7	0.2
159	경기 고양시 일산동구	2.5	3.5	1.1
160	경북 문경시	2.4	2.7	0.3
161	경남 사천시	2.3	2.3	0.0
162	충북 음성군	2.3	2.0	-0.3
163	경북 경주시	2.2	4.0	1.8
164	전북 정읍시	2.2	1.4	-0.8
165	경북 구미시	1.9	2.2	0.3
166	전북 익산시	1.3	1.1	-0.3
평균		10.6	9.4	-1.2

주 : 지역별 순위는 주중 혼잡강도(%) 기준 내림차순임.

## 2) 기타 도로의 경우(CASE2, 고속도로 제외)

- 전국 고속도로를 제외한 나머지 도로를 대상으로 시/군/구 단위의 집계를 수행한 결과, 앞서 주중 기준 고속도로의 경우보다 전체적으로 혼잡강도가 높게 분석되었음. 이는 지역간 고속도로를 제외한 지역별 도시 내부도로들에 대한 교통혼잡이 포함되어 있기 때문임
- 또한, 지역별로 혼잡강도(%)를 살펴보면 주중 기준 전국 평균 27.0% 보다 혼잡한 시/군/구는 전체 245개 지역 중 111개 지역(전체 지역의 45.7%)으로 부산 서구, 인천 동구, 부산 부산진구 등의 순으로 분석되었음. 이는 지역간 고속도로를 제외한 나머지 도로의 혼잡강도(%)를 집계하므로 지역 간 통행량 보다는 해당 지역의 내부통행량이 높은 지역일수록 다른 지역보다 상대적으로 혼잡강도가 높은 것으로 판단됨
- 주말의 혼잡강도(%)는 지역별 도시지역 보다는 일부 지방의 관광지역(강원 화천군, 강원 인제군, 전남 신안군 등)에서 주중 보다 높은 것으로 분석되었음

&lt;표 1-35&gt; 전국 시/군/구 단위 혼잡강도(%) 분석 결과(CASE2)

순위	지역명	혼잡강도(%)		
		주중①	주말②	증감(②-①)
1	부산 서구	64.6	64.5	0.0
2	인천 동구	60.9	57.2	-3.7
3	부산 부산진구	60.5	56.7	-3.8
4	대구 중구	58.3	57.6	-0.7
5	인천 중구	57.5	53.4	-4.1
6	경기 부천시 오정구	56.3	49.3	-7.0
7	부산 중구	55.3	59.8	4.6
8	인천 남동구	53.5	53.0	-0.6
9	인천 서구	51.5	41.0	-10.4
10	경기 광명시	49.1	45.1	-3.9
⋮	⋮	⋮	⋮	⋮
236	전남 강진군	9.2	7.1	-2.1
237	강원 화천군	8.7	11.1	2.4
238	전남 영광군	8.6	8.0	-0.5
239	경남 산청군	7.9	7.2	-0.7
240	경북 울진군	7.4	8.0	0.6
241	강원 영월군	6.9	3.7	-3.2
242	경북 예천군	6.6	4.7	-1.9
243	강원 인제군	6.5	10.4	3.9
244	전남 고흥군	5.5	5.0	-0.5
245	전남 신안군	2.4	11.0	8.6
평균		27.0	24.8	1.0

주 : 지역별 순위는 주중 혼잡강도(%) 기준 내림차순임.



## 3) 전체 도로의 경우(CASE3)

- 전국 전체 도로를 대상으로 시/군/구 단위의 집계를 수행한 결과, 전체적으로 CASE1과 CASE2의 중간 정도의 혼잡강도로 분석되었음. 이는 지역간 고속도로와 지역별 내부도로의 교통혼잡이 모두 포함되어 있기 때문임
- 또한, 지역별로 혼잡강도(%)를 살펴보면 주중 기준 전국 평균 23.9% 보다 혼잡한 시/군/구는 전체 245개 지역 중 114개 지역(전체 지역의 46.5%)으로 부산 서구, 인천 동구, 부산 부산진구 등의 순으로 분석되었음. 이는 전국 전체 도로의 혼잡강도(%)를 집계하므로 지역 간 통행량과 해당 지역의 내부통행량이 모두 높은 지역일수록 다른 지역보다 상대적으로 혼잡강도가 높은 것으로 판단됨
- 주말의 혼잡강도(%)는 지역별 도시지역 보다는 일부 지방의 관광지역(강원 평창군, 충북 괴산군, 전남 신안군 등)에서 주중 보다 높은 것으로 분석되었음

&lt;표 1-36&gt; 전국 시/군/구 단위 혼잡강도(%) 분석 결과(CASE3)

순위	지역명	혼잡강도(%)		
		주중①	주말②	증감(②-①)
1	부산 서구	64.6	64.5	0.0
2	인천 동구	60.9	57.2	-3.7
3	부산 부산진구	60.5	56.7	-3.8
4	대구 중구	58.3	57.6	-0.7
5	부산 중구	55.3	59.8	4.6
6	서울 은평구	47.8	47.5	-0.3
7	경기 부천시원미구	46.9	38.1	-8.8
8	부산 북구	46.0	42.7	-3.3
9	서울 서대문구	45.9	43.7	-2.2
10	부산 동구	45.6	40.5	-5.1
⋮	⋮	⋮	⋮	⋮
236	강원 평창군	6.1	21.6	15.5
237	전남 고흥군	5.9	5.3	-0.6
238	충남 공주시	5.6	4.6	-1.0
239	전남 곡성군	5.6	4.7	-0.9
240	경남 함양군	5.4	3.8	-1.6
241	전북 임실군	5.1	6.2	1.1
242	충북 괴산군	4.6	5.3	0.6
243	전남 구례군	4.3	3.8	-0.5
244	충북 보은군	4.0	3.8	-0.1
245	전남 신안군	2.4	11.0	8.6
평균		23.9	21.8	-2.1

주 : 지역별 순위는 주중 혼잡강도(%) 기준 내림차순임.

### 3. 교통혼잡비용 추정 결과

- 교통혼잡비용의 추정은 지역간 도로와 도시부 도로로 구분하여 추정하였으며, 지역간 도로의 경우는 도로등급별(고속도로, 일반국도, 국지도, 지방도)로 교통혼잡비용을 추정하고 도시부 도로의 경우는 서울특별시 내부도로를 대상으로 교통혼잡비용을 추정하였음
- 교통혼잡비용 추정 시, 기존 선행연구에서 적용되는 도로등급별 혼잡기준속도와 본 연구의 혼잡경계속도를 각각 별도로 적용하여 혼잡비용을 추정하였으며, 그 결과는 고정비를 포함한 혼잡비용과 포함하지 않는 혼잡비용으로 구분하여 제시하였음
- 지난 2014년 2월 한국교통연구원에서 수행한 “2011, 2012년 전국 교통혼잡비용 추정과 추이 분석”이라는 선행 연구에서 제시한 평일 기준 교통혼잡비용의 추정 결과와 비교·분석을 수행하였음
- 추가적으로 본 연구에서는 선행연구에서 추정하지 못한 주말에 대한 교통혼잡비용의 추정결과도 제시하였음

#### 가. 전국 교통혼잡비용 추정 결과

##### 1) 지역간 도로

- 도로등급별 지역간 도로의 추정 결과는 지역별 도심부를 많이 통과하는 일반국도의 경우가 다른 도로들에 비해 교통혼잡비용이 높게 추정되었음
- 또한, 모든 도로에서 본 연구의 혼잡경계속도를 적용한 경우 보다 기존의 도로등급별 혼잡기준속도를 적용한 경우가 교통혼잡비용이 높게 추정되었음
- 이는 기존 혼잡기준속도의 교통혼잡은 서비스 수준 C~D이하이며, 본 연구의 혼잡경계속도는 서비스 수준 E이하로 교통혼잡의 기준이 서로 상이하기 때문임
- 선행 연구에서는 분석자료 구축의 한계로 인하여 평일 기준으로만 교통혼잡비용을 추정하고 있어 주말에 발생하는 교통혼잡에 대한 혼잡비용을 추정할 수가 없었음. 하지만 본 연구에서는 전국 표준노드링크 기반의 개별 Car Navigation 자료를 이용함에 따라 평일과 주말에 대한 각각의 교통혼잡비용을 추정하였음
- 따라서, 전국 도로등급별 전체 교통혼잡비용은 기존의 평일 기준 혼잡비용이 아닌 평일과 주말의 혼잡비용을 합한 비용으로 추정하였음

&lt;표 1-37&gt; 전국 교통혼잡비용 추정 결과(지역간 도로)

(단위: 억원)

구분		평일			주말			전체		
		기존	본연구	증감	기존	본연구	증감	기존	본연구	증감
고정비 포함	고속도로	32,035	11,020	-21,015	10,357	3,961	-6,396	42,392	14,981	-27,411
	일반국도	98,083	41,196	-56,887	37,543	15,018	-22,525	135,626	56,214	-79,412
	국지도	13,662	5,517	-8,145	4,742	1,733	-3,009	18,404	7,250	-11,154
	지방도	23,180	8,452	-14,728	8,734	2,818	-5,916	31,913	11,270	-20,643
	합계	166,960	66,185	-100,775	61,376	23,530	-37,846	228,335	89,715	-138,620
고정비 제외	고속도로	24,431	8,590	-15,841	8,932	3,448	-5,484	33,363	12,038	-21,325
	일반국도	77,496	33,059	-44,437	30,836	12,481	-18,355	108,332	45,540	-62,792
	국지도	10,687	4,421	-6,266	3,884	1,435	-2,449	14,571	5,856	-8,715
	지방도	17,691	6,606	-11,085	7,124	2,322	-4,802	24,815	8,928	-15,887
	합계	130,305	52,676	-77,629	50,776	19,686	-31,090	181,081	72,362	-108,719

주 1: 기존은 선행연구의 혼잡기준속도를 적용한 경우이며, 본 연구는 본 연구의 혼잡경계속도를 적용한 경우임.

주 2: 본 연구는 4개월 분량의 내비게이션 자료를 사용하여 추정한 값으로 참값이 아니며, 본 연구의 방법론을 적용한 추정값임. 향후 내비게이션 자료의 확대 구축 및 방법론의 고도화를 통해 보다 정밀한 교통혼잡비용의 추정이 필요함.

## 2) 도시부 도로(서울특별시)

- 도시부 도로의 경우는 서울특별시의 내부도로에 한하여 교통혼잡비용을 추정하였으며, 혼잡비용 추정과정은 지역간 도로와 동일함
- 또한, 전체 교통혼잡비용은 본 연구의 혼잡경계속도를 적용한 경우 보다 기존의 도로등급별 혼잡기준속도를 적용한 경우가 더 높게 추정되었음

&lt;표 1-38&gt; 전국 교통혼잡비용 추정 결과(도시부 도로, 서울특별시)

(단위: 억원)

구분	평일			주말			전체		
	기존	본연구	증감	기존	본연구	증감	기존	본연구	증감
고정비 포함	32,616	27,253	-5,363	11,304	9,381	-1,923	43,920	36,634	-7,286
고정비 제외	27,684	23,179	-4,505	9,405	7,825	-1,580	37,089	31,004	-6,085

주 1: 기존은 선행연구의 혼잡기준속도를 적용한 경우이며, 본 연구는 본 연구의 혼잡경계속도를 적용한 경우임.

주 2: 본 연구는 4개월 분량의 내비게이션 자료를 사용하여 추정한 값으로 참값이 아니며, 본 연구의 방법론을 적용한 추정값임. 향후 내비게이션 자료의 확대 구축 및 방법론의 고도화를 통해 보다 정밀한 교통혼잡비용의 추정이 필요함.

#### 나. 기존 교통혼잡비용과의 비교 분석

- 앞서 추정한 평일 기준 본 연구의 교통혼잡비용과 선행 연구의 교통혼잡비용 추정결과와의 비교는 <표 1-39>와 같음
- 먼저 지역간 도로의 경우, 선행 연구결과 대비 기존 혼잡기준속도를 적용한 경우는 교통혼잡비용이 높게 추정된 반면 본 연구의 혼잡경계속도를 적용한 경우에는 교통혼잡비용이 낮게 추정되었음. 이러한 추정 결과의 차이에 대한 원인은 아래와 같음
  - 기존 연구에서 통행속도 산출 시 이용되고 있는 BPR모형의 한계
    - 통행속도 추정 결과는 교통량 변화에 민감함
    - 파라미터( $\alpha$ ,  $\beta$ ) 최적화 문제
    - 도로의 교통운영 상황(신호교차로 등)을 제대로 반영하지 못함
  - 대상구간 설정 방법에 따른 교통량과 속도 입력변수 적용의 차이
    - 선행 연구 : 대구간 단위의 대상구간 설정 → 해당 구간 내 교통량과 속도를 모든 구간에 동일하게 적용
    - 본 연구 : 전국 표준노드링크 기반의 세부 링크별로 교통량과 속도 데이터 적용
  - 교통혼잡비용 추정 시, 혼잡기준속도의 적용의 차이(본 연구의 혼잡경계속도 < 기존 혼잡기준속도)
    - 선행 연구 : LOS C ~ D 수준(도로등급별 일률적인 교통혼잡 기준 적용)
    - 본 연구 : LOS E 수준(시스템 운영 기반 상대적 개념의 교통혼잡 기준 적용)
- 도시부 도로(서울특별시)는 기존 혼잡기준속도와 본 연구의 혼잡경계속도를 적용한 경우 모두 선행 연구결과에 비하여 교통혼잡비용이 낮게 추정되었음. 이러한 추정 결과의 차이에

대한 원인은 아래와 같음

- 교통혼잡비용 추정 시 이용되는 속도 자료의 차이
  - 선행 연구 : 대상 도로의 전체 평균 속도자료 이용 → 링크별 상대적인 교통상황을 반영하지 못함
  - 본 연구 : 전국 표준노드링크 기반의 개별 Car Navigation 속도자료 이용 → 링크별 교통상황이 개별 Car Navigation 속도 데이터에 반영이 되어 있음
- 대상 공간적 범위 내 교통혼잡비용 전수화 방법의 차이
  - 선행 연구 : 관측 구간에 대한 교통혼잡비용 추정 → 해당 구간의 연장에 대한 전체 도로의 연장 비율을 적용하여 전수화 수행
    - ※ 해당 관측 구간의 교통혼잡비용이 높게 추정될 경우, 전체 교통혼잡비용은 더욱 과다 추정될 우려가 존재함
  - 본 연구 : 전국 표준노드링크 기반의 링크 단위 교통혼잡비용 추정 → 대상 공간적 범위 내 링크별 교통혼잡비용 집계
- 교통혼잡비용 추정 시, 혼잡기준속도의 적용의 차이(본 연구의 혼잡경계속도 < 기존 혼잡기준속도)
  - 선행 연구 : LOS C ~ D 수준(도로등급별 일률적인 교통혼잡 기준 적용)
  - 본 연구 : LOS E 수준(시스템 운영 기반 상대적 개념의 교통혼잡 기준 적용)

<표 1-39> 선행 연구의 교통혼잡비용 추정결과와의 비교(평일 기준)

(단위: 억원)

구분		선행 연구 (2012년)①	본 연구 방법		증감	
			기존 혼잡기준속도②	본 연구의 혼잡경계속도③	②-①	③-①
지역간 도로 (전국)	고정비 포함	111,296	166,960	66,185	55,664	-45,111
	고정비 제외	86,339	130,305	52,676	43,966	-33,663
도시부 도로 (서울특별시)	고정비 포함	84,144	32,616	27,253	-51,528	-56,891
	고정비 제외	72,402	27,684	23,179	-44,718	-49,223

주1: 선행 연구의 교통혼잡비용 추정결과는 “2011, 2012년 전국 교통혼잡비용 추정과 추이분석, 한국교통연구원, 2014. 2.”의 연구결과임.

주2: 본 연구는 4개월 분량의 내비게이션 자료를 사용하여 추정한 값으로 참값이 아니며, 본 연구의 방법론을 적용한 추정값임. 향후 내비게이션 자료의 확대 구축 및 방법론의 고도화를 통해 보다 정밀한 교통혼잡비용의 추정이 필요함.

## 제6절 결론 및 향후 로드맵

### 1. 결론

- IT기술의 발달과 더불어 전 세계적으로 대용량 자료(Big data)에 대한 많은 연구가 수행되고 있으며, 교통분야의 ITS분야 또한 기존의 루프 검지기와 같은 고정식 지검검지체계에서 GPS 및 DSRC(단거리 이동통신)을 기반으로 하는 이동식 개별검지로 교통정보의 수집영역이 확대됨으로 인하여 방대한 양의 자료를 수집하고 있음
- 이러한 ITS분야의 정형 빅 데이터를 활용한 사례로 지난 2013년 한국교통연구원에서 수행한 “2013년 국가교통수요조사 및 DB구축사업 중 Car Navigation 자료를 이용한 교통혼잡지도 연구”에서는 방대한 양의 차량이동궤적 정보를 활용한 전국 교통혼잡지도 시스템을 개발하였음
- 하지만, 지난 과업에서는 짧은 사업 기간으로 인하여 불가피하게 시스템 개발에만 집중하게 되어 교통혼잡지표를 산출하기 위한 관련 알고리즘에 대한 검증과정이 미미하였음. 또한, 교통혼잡지표 산출을 위한 Car Navigation 자료도 2013년 9월 한 달 분량의 데이터만을 이용함에 따라 데이터 수급의 한계도 발생하였음
- 본 연구에서는 Car Navigation 자료를 기존 연구보다 확장(2013년 10월 ~ 2014년 1월, 4개월치)하고 이를 이용하여 전국 교통혼잡지도 구축을 위한 교통혼잡지표 산출 알고리즘의 검증 및 고도화 연구를 수행하였음. 또한, Car Navigation 자료의 활용으로 기존 교통혼잡비용 추정 방법의 한계를 극복할 수 있는 교통량 전수화 모형, 링크별 교통혼잡비용 추정방법, 지역 단위 교통혼잡비용 집계방법 등의 방법론을 제시할 수 있었음
- 향후 개발모형의 정확도를 더욱 높이기 위해서는 4개월이 아닌 1년 이상의 축적된 Car Navigation 자료가 필요할 것으로 판단되며, 이에 따른 본 연구의 활용 방안은 아래와 같음
  - 도로의 운영현황을 효율적으로 파악 과학적인 도로 운영체계의 수립 지원
  - 다양한 도로망 성능평가 : 지역, 도시, 축, 도로별 위계 등
  - 비 혼잡 교통류 상태에서 속도분포 최고/최저 제한속도 등 운영규제
  - 교통수요분석 도구의 통행비용함수에 활용 가능
  - 신호교차로 운영상태(제어지체, 연동효과, 이동류별 구성비)를 지속적으로 모니터링 가능
  - 속도의 밀도 Profile 이력 DB 혼잡예보 및 다중시간대 동적 통행시간 예측에 활용 가능

## 2. 향후 로드맵 구축

<표 1-40> 단계별 향후 로드맵

구분	단기(2014년)		중기(2015~2016년)	장기(2017~2018년)
	거시적		----->	미시적
혼잡지도	• 교통혼잡지도시스템 구축 (이력)		• 교통혼잡지도시스템 구축 및 운영(이력, 실시간)	• 시스템 운영 및 유지 보수
	• 교통혼잡지도시스템 DB 제공 서비스(이력)		• 교통혼잡지도시스템 DB 제공 서비스(이력, 실시간) • 인터넷 포털 사이트 콘텐츠 제공 • 비즈니스 인텔리전스(BI) 애플리케이션 활용	
	• 교통혼잡지도 제공 서비스(이력) - 혼잡 및 추가 지표 - 평일/주말 일단위 - 전국 광역시/도, 시/군/구, 도로등급별, 링크구간 단위		• 교통혼잡지도 제공 서비스(이력, 실시간) - 혼잡 및 추가 지표 - 년/분기/월/일/시 - 전국 광역시/도, 시/군/구, 읍/면/동, 도로(등급)별, 교차로, 링크구간 단위	
SOC 투자 사업 종합 진단 시스템	진단 시스템	-	• SOC투자사업 종합 진단을 위한 지표 개발 • SOC투자사업 종합 진단 분석방법론 개발 • SOC투자사업 종합 진단시스템 구축	• SOC투자사업 종합 진단시스템 운영
	진단 구성 요소 기술 개발	-	• 공간 클러스터링 방법론 개발 • 사회·경제지표 분석 클러스터링 • 도로이용 노출도 지표 개발 • 통행기반 수형망 지표 분석 연계방법론 개발 • 시각화 분석 지원 도구 개발	-
	도로 안전 진단	-	• 도로안전진단 방법론 개발 • 도로안전진단을 위한 기초자료 제공	• 도로안전진단 시스템 구축 - 기존도로 교통사고 안전진단 모니터링 • 도로안전진단을 위한 DB 제공





## 제2장 빅데이터 기반 교통예보를 위한 핵심 요소기술 개발

---

제1절 연구개요

제2절 교통예보 관련 현황 분석

제3절 교통예보를 위한 기초자료 구축

제4절 교통예보를 위한 소통상태 예측  
핵심 방법론 개발

제5절 교통예보를 위한 시스템 구축 방안

제6절 사례분석 및 평가

제7절 결론 및 향후 연구



## 제2장 빅데이터 기반 교통예보를 위한 핵심 요소기술 개발

### 제1절 연구개요

#### 1. 연구의 배경 및 목적

- 현재 우리나라는 지속적인 경제성장으로 인하여 삶의 수준이 향상되면서 도시의 출·퇴근 통행의 증가와 더불어 여가통행의 비중도 꾸준히 증가하고 있음
- 특히 여가를 목적으로 하는 통행의 경우는 그만큼 운전자가 이동하는 날짜와 시간대에 이용하고자 하는 도로들에 대한 교통상황 정보(원활, 서행, 정체)에 민감할 수밖에 없으며, 대부분의 운전자들은 출발 전에는 스마트폰과 각종 인터넷 포털사이트 등에서 제공하는 교통상황 정보에 의존하고, 출발 후에는 차량 내비게이션에 의해 제공되는 이동경로 및 교통상황 정보에 의존하게 됨
- 따라서, 운전자들에게 제공되고 있는 교통상황 정보는 21세기 정보화 시대에 중요한 정보임과 동시에 운전자들이 신뢰할 수 있도록 정확한 교통상황 정보를 제공하는 것이 무엇보다도 중요하다고 볼 수 있음
- 이러한 측면에서 20세기 말 국내에 정보수집 → 정보가공 → 정보제공이라는 일련의 지능형 교통체계(Intelligent Transport System, 이하 ITS)가 도입된 후, 지속적인 기술 발전으로 최근에는 첨단 IT기술을 기반으로 하는 정보수집 매체를 통하여 다양한 정보제공 서비스를 실현하고 있음
- 하지만, 현재 ITS는 대부분 실시간 정보를 이용한 교통상황 정보만을 제공하고 있으며, 장래 교통상태 예측정보(교통예보)의 경우는 일부의 민간 교통정보 제공서비스 업체와 한국도로공사를 제외하면 미미한 수준임. 즉, 21세기 IT기술의 급속한 발달과 더불어 시스템 분야의 하드웨어와 다양한 정보제공 콘텐츠 개발에 따른 기술적인 발전에도 불구하고 정보제공 서비스의 핵심요소인 장래 교통상태 예측분야는 이러한 기술적 발전과 병행하지 못하고 있는 실정임
- 또한, 일부 ITS의 장래 교통상태 예측방법은 대부분 현재시각 기준 5~15분 이후의 장래 교통상태를 예측하는 단기예측(Short-term forecasting)기법에 의존하고 있으며, 단기예측 모형의 입력 값으로 이용되는 자료의 시간적인 차원도 현행(Current)의 실시간 수집정보로 한정

시켜왔음. 즉, 기존의 단기에측 방법은 방대한 양의 이력자료 관리 및 활용이라는 측면을 고려할 필요가 없었음

- 이러한 단기에측 기법은 익일 이상의 중장기 장래 교통소통 정보를 예측하고 평가를 수행할 수 없는 모형의 구조적 단점을 가지고 있으며, 기존 모형으로 개발할 경우 시스템의 지원능력도 방대한 이력자료의 활용이라는 요구를 충족하지 못하였음
- 이러한 현실 속에서 ITS에서의 장래 교통정보 예측분야가 한걸음 더 도약하기 위해서는 현재의 단기에측 기술이 넘지 못하고 있는 예측의 시/공간적 범위(Forecasting horizon)를 극복해야 함
- 그 동안 선행 연구들에서 수많은 단기에측 모형이 개발이 되었으나, van Hinsbergen et al.(2007)이 지적한 바와 같이 타 모형에 비하여 매우 우수한 성능을 가진 모형은 없다고 지적하였음. Chang et al.(2012)가 지적한 바와 같이 결국 21세기의 ITS 예측분야의 개척 영역은 수용 가능한 수준 이내에서 예측영역의 시/공간적 확장이라 판단됨
- 또한 최근 내비게이션과 스마트폰의 보급 확대 등과 같이 Big Data로 표현할 수 있는 21세기 첨단 IT기술과 정보화는 기존 ITS 예측기법의 변화를 요구하고 있음. 즉, 스마트폰을 선두로 하는 첨단 개인단말은 다양한 교통정보 Needs의 증대와 동시에 실시간 교통정보를 포함하여 과거의 방대한 이력자료를 활용한 익일 이상의 중장기 장래 교통상태 예측정보를 요구하고 있음
- 더불어 최근의 발전된 Database관리 능력과 검색 엔진도 교통 Big Data를 활용한 장래 교통예보정보 생성에 있어 새로운 가능성을 제시해주고 있음
- 따라서 본 연구의 목적은 개별 차량의 내비게이션 단말에서 수집되는 교통 빅데이터(Probe건수, 속도)를 이용하여 아래와 같이 익일의 장래 통행속도 예측이 가능한 빅데이터 기반의 통행속도 예측 방법론을 개발하는데 있음
  - 빅데이터 기반으로 하는 장래 통행속도 예측을 위한 핵심 요소기술 개발
  - 개발된 요소기술의 검증을 통한 실용화 가능성 검토
  - 개발 기술의 개선방향 및 실용화를 위한 시스템 요구사항 제시

## 2. 연구의 범위 및 방법

### 가. 시간적 범위

- 분석연도 : 2014년(→ 과업기간 내 수집된 Car Navigation 자료 이용)
- 데이터 수집기간(4개월) : 2013년 10월 1일 ~ 2014년 1월 31일

### 나. 공간적 범위

- 대상범위 : 서울특별시 강남 일대(→ 전국 모든 도로에 적용이 가능한 방법론 개발)

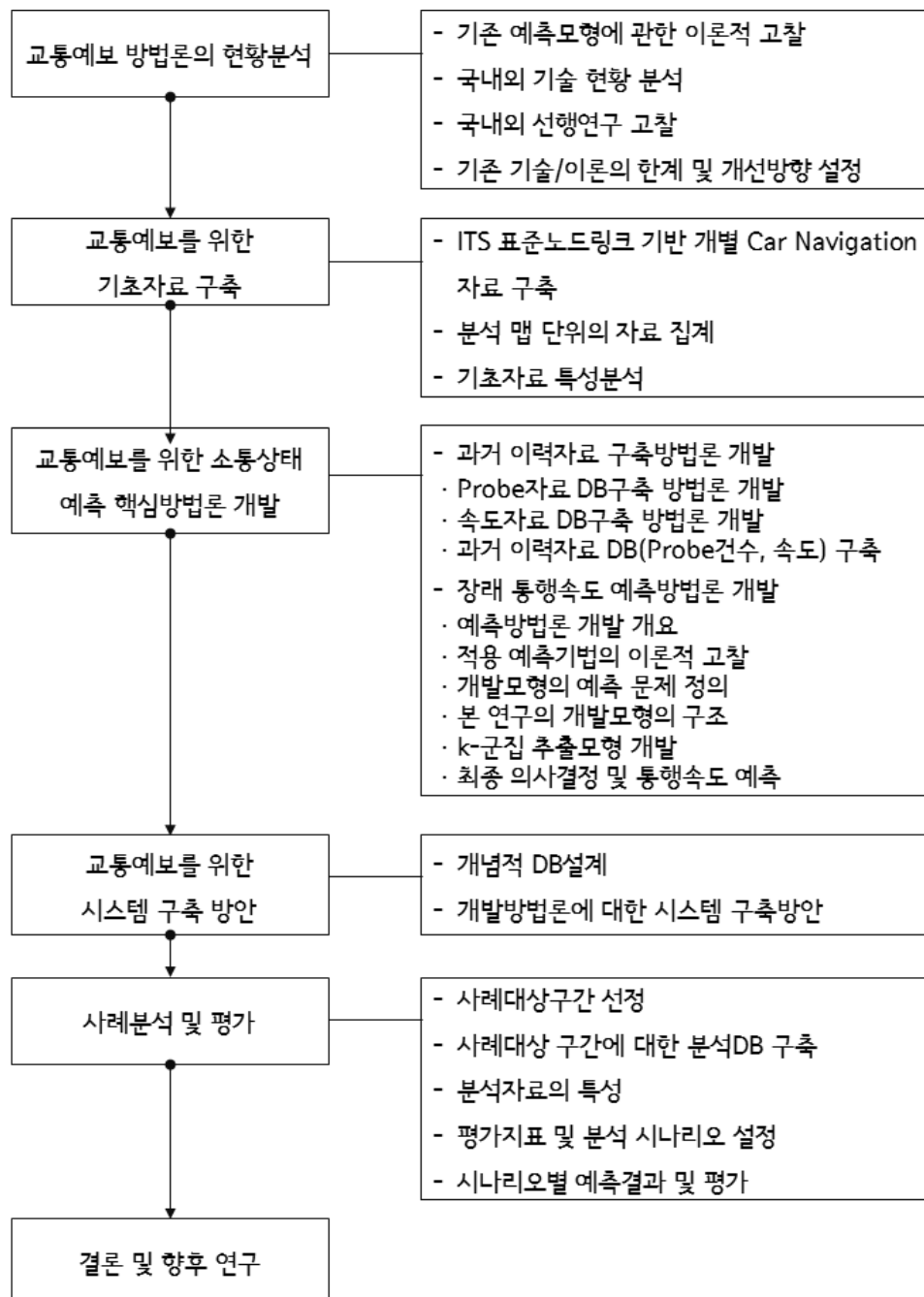
### 다. 연구의 방법

- 본 연구의 목표는 향후 교통예보 시스템에 탑재되기 위한 빅데이터 기반의 장래 통행속도를 예측하는 방법론을 개발하는데 있음. 즉, ITS 표준노드링크 기반의 Car Navigation 자료 (Probe건수, 속도)를 이용하여 일일의 링크별/시간대별 통행속도를 예측하기 위한 핵심방법론을 개발하는 것으로 전체 연구의 방법은 아래와 같이 크게 4단계로 구성됨
- Step 1 : ITS 표준노드링크 기반의 개별 Car Navigation 기초자료 구축
- Step 2 : 과거 이력자료 구축 방법론 개발
  - Probe건수 DB구축 방법론 개발
  - 속도자료 DB구축 방법론 개발(속도 유전자 지도, 장래 구간 통행속도 산출)
  - 과거 이력자료 DB(probe건수, 속도) 구축
- Step 3 : 구축된 이력자료 기반의 장래 통행속도 예측 방법론 개발
  - KNN<sup>1)</sup> 단계 : k-군집 추출모형 개발
  - 예측 단계 : 최종 의사결정 및 통행속도 예측모형 개발
- Step 4 : 방법론 적용 및 평가
  - 평가지표 및 분석 시나리오 설정
  - 시나리오별 예측결과 및 평가

<sup>1)</sup> k개의 최근린 이웃(k-nearest neighbor, 이하 KNN)

### 3. 연구의 수행 체계

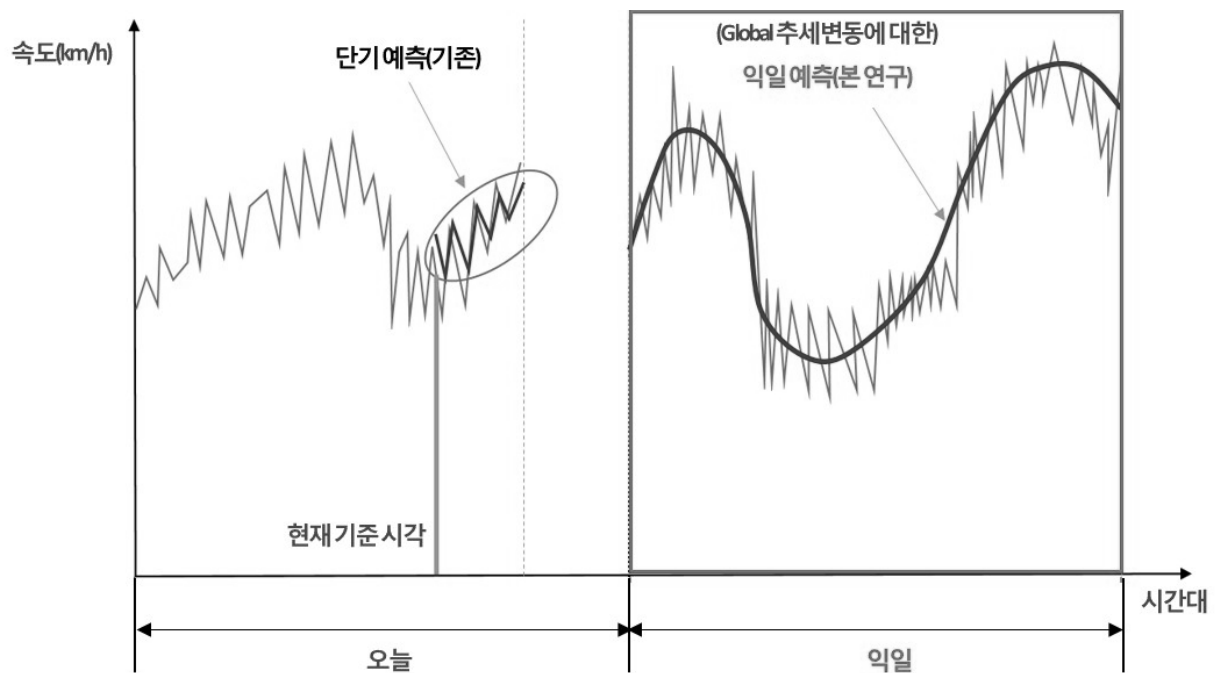
- 본 연구는 교통예보 관련 현황분석, 교통예보를 위한 기초자료 구축, 교통예보를 위한 소통상태 예측 핵심방법론 개발, 교통예보를 위한 시스템 구축 방안, 사례분석 및 평가, 결론 및 향후 연구로 수행되며, 연구의 수행체계는 아래의 그림과 같음



<그림 2-1> 연구의 수행체계

#### 4. 본 연구의 교통예보 개념

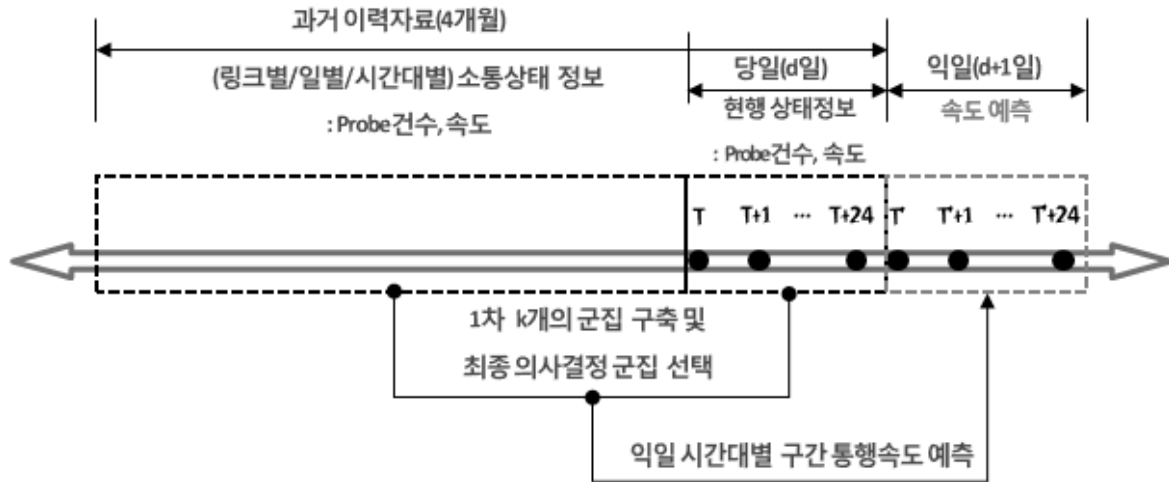
- 본 연구의 교통예보는 기존의 ITS의 예측분야 관점에서 단기예측(Short-term forecasting)이 아닌 중/장기 예측에 해당하는 익일 이후의 장래 소통상태(통행속도)를 예측하는 것임
  - 단기 예측 : 해당 시간대 간격(1~15분)으로 과거 5~15분 이전의 실시간 소통상태 정보를 이용하여 현재시각 기준 5~15분 이후의 장래 소통 상태를 예측
  - 중/장기 예측(본 연구) : 해당 시간대 간격(15분~1시간)으로 당일의 실시간 소통상태 정보를 이용하여 익일 이후(+n일)의 장래 소통 상태를 예측함 → 즉, 본 연구에서는 익일의 소통 상태를 예측



<그림 2-2> 본 연구의 교통예보 개념(1)

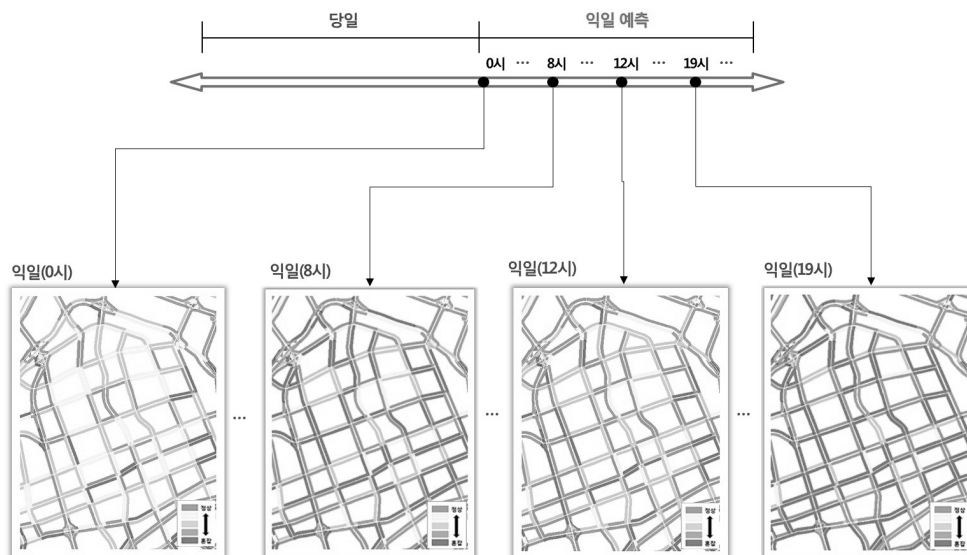
- 본 연구에서는 장래 교통예보를 익일의 링크별/시간대별 통행속도 예측으로 정의하며, 이를 위한 분석 자료는 본 연구의 시/공간적 범위에서 수집된 ITS 표준노드링크 기반의 개별 Car Navigation Probe 데이터를 시간 단위로 집계하여 구축한 링크별/일별/시간대별 Probe 건수와 속도 자료를 이용함
- 익일의 링크별/시간대별 통행속도를 예측하기 위해서는 현재시각 기준으로 당일의 링크별/시간대별 실시간 소통상태 정보(Probe건수, 속도)가 필요하나, 현재는 교통예보시스템을 구축하기 이전이므로 실시간 소통상태 정보를 과거 어느 시점에서의 이력자료로 대체함. 즉, 과거의

이력자료 중 임의로 선택된 일자(d일)를 당일로 지정하고, 해당 일자의 링크별/시간대별 소통 상태 정보를 이용하여 익일(d+1일)의 링크별/시간대별 통행속도를 예측하는 것임



<그림 2-3> 본 연구의 교통예보 개념(2)

- 향후 장래 교통예보 시스템이 구축될 경우, 익일의 교통예보는 링크별/시간대별 통행속도 예측결과를 교통 혼잡 상태(정상 ↔ 혼잡)로 구분하여 아래의 그림과 같이 표출하고자 함



<그림 2-4> 장래 교통예보시스템 정보표출 예시



## 제2절 교통예보 관련 현황 분석

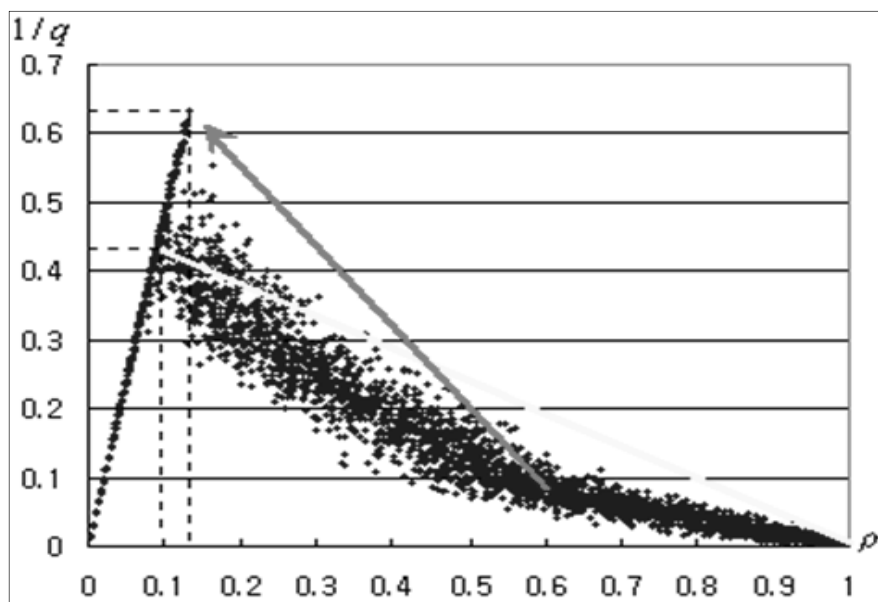
### 1. 분석 개요

- 본 연구의 교통예보를 위한 핵심 요소기술을 개발하기에 앞서 교통예보와 관련된 현황 분석을 수행하여 관련 분야의 현재 이론 및 기술적인 수준과 한계점을 파악하고 이를 토대로 장래 예측모형을 개발하기 위한 시사점과 개선방향을 도출하고자 함
- 교통예보 관련 현황 분석의 구성은 크게 ①기존 예측모형에 관한 이론적 고찰, ②국내외 관련 기술현황 분석, ③국내외 관련 선행연구 고찰, ④기존 기술/이론의 한계 및 개선방향 설정으로 구성됨
- 먼저, 기존 예측모형에 관한 이론적 고찰은 교통량, 속도, 밀도로 구성되는 거시적 교통류 변수들의 특성을 파악하고, 관련 분야의 선행연구들을 통하여 기존 예측 모형들의 시·공간적 정의, 입력벡터 차원과 예측문제의 정의, 예측모형의 분류 및 특성에 관한 내용을 다루고자 함. 특히 기존의 예측모형 중 본 연구의 장래 예측모형에 적용되는 k-최근린 이웃(K-Nearest Neighbor)기법과 군집(Clustering) 기법에 대하여 상세하게 다루고자 함
  - 기존 예측 모형
    - 정적(Naive)기법 : 동시(Instantaneous)기법, 과거평균(Historical average)기법, 군집(Clustering)모형
    - 모수(Parametric)기법 : 교통류 모의실험 모형, 시계열(Time Series)모형
    - 비모수(Nonparametric)기법 : 퍼지 이론(Fuzzy Logic), 신경망(Neural Network), k-최근린 이웃(K-Nearest Neighbor)
  - 본 연구의 적용 모형
    - k-최근린 이웃(K-Nearest Neighbor)기법
- 국내외 관련 기술현황 분석은 교통예보 관련하여 기술적인 측면에서 국내외의 공공기관과 민간업체들에 대한 현재 기술 수준을 파악하며, 국내외 관련 선행연구 고찰은 관련 분야의 이론적인 측면에서 국내외 선행연구들을 통하여 현재 이론적인 수준을 파악하고자 함
- 마지막으로 기술/이론의 한계 및 개선방향 설정은 관련 분야의 현재 기술 및 이론적인 측면에서의 한계점을 도출하고 이를 극복하기 위한 본 연구의 개발모형에 대한 시사점 및 개선방향을 설정하고자 함

## 2. 기존 예측모형에 관한 이론적 고찰

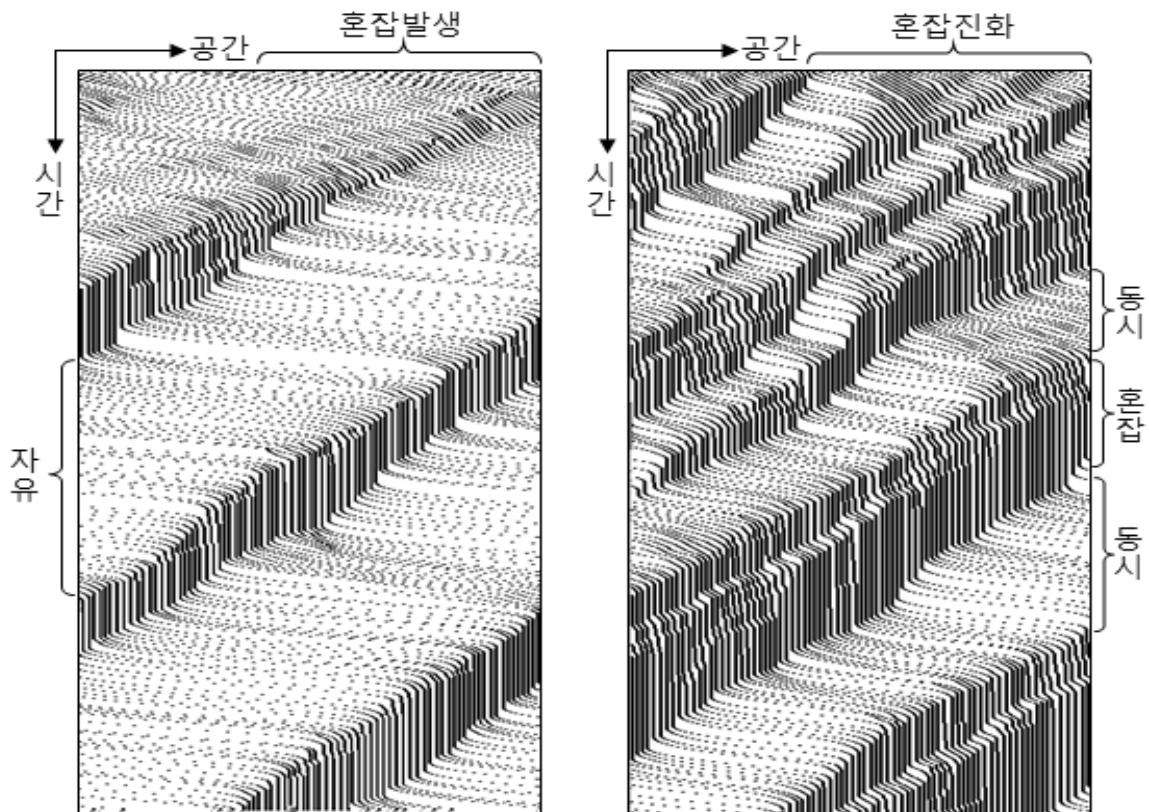
### 가. 거시적 교통류 변수의 특성

- 일반적으로 예측은 장래의 불확실성(Uncertainty)을 극복하기 위한 일련의 노력으로 정의할 수 있으며, 특정 시스템의 상태를 예측하기 위해서는 상태의 시/공간적 진화(Evolution)에 대한 깊은 이해가 필요함
- 따라서, 기존 예측모형의 이론과 선행연구의 고찰에 앞서 거시적 교통류 시스템의 상태에 대하여 알아보고 이러한 거시적 교통류의 상태의 특성에 초점을 둔 이유는 일반적으로 교통상태 예측은 거시적 변수를 대상으로 하기 때문임. 먼저 거시적 교통류 변수는 교통량( $q$ , 대/단위시간), 속도( $v$ , 단위거리/단위시간), 점유율( $k$ , 대/단위거리)로 구성되며, 속도와 관련하여 통행시간( $t$ , 단위시간)도 포함됨
- $q-k$ 의 관계는 <그림 2-5>과 같이 역  $\lambda$ -형태의 분포를 보여주고 있으며, 자유교통류 상태에서  $q-k$ 의 관계는 매우 균질적인(Homogeneous)한 상태이지만, 혼잡상태의 교통류는 임계밀도와 정체 교통류 간에 비 균질적(Non-homogeneous) 관계를 보이며, 정체 교통류는 다시 균질적 상태를 보임(장현호 등., 2004; Chang et al., 2007; Shah, et al., 2008)



<그림 2-5> 거시적 교통량과 밀도의 관계(장현호 등., 2004)

- 연속류의 경우 2개의 불안정 상태(Meta-stable State)가 존재하며, 3개 현시(Phase)의 자유 교통류, 동시교통류 그리고 혼잡교통류로 구분될 수 있음
- 이러한 교통류의 시·공간적 상태 진화(Evolution)는 Kerner(1996)의 동적인 현시 교통류로 설명할 수 있으며, 이러한 현시 교통류에서의 교통류는 자유교통류(F), 동시교통류(S), 혼잡 교통류(J)로 구분됨. 즉, 교통류의 진화 형태 3개 현시의 조합 {F, S, J}의 관계로 나타남
- 예를 들어 현시의 전환은  $F \rightarrow J \rightarrow F$ ,  $F \rightarrow S \rightarrow J \rightarrow F$ ,  $F \rightarrow S \rightarrow J \rightarrow S \rightarrow J \rightarrow S \rightarrow F$ 으로 발달하면서 긴 이동 혼잡(Wide-moving jam)으로 발전하며, <그림 2-6>와 같이 혼잡의 발생과 긴 이동 혼잡으로의 발달과정을 보여주고 있음



<그림 2-6> 교통류 상태의 시/공간적 진화

- 앞서 살펴본 교통류 상태의 진화 행태는 교통류 상태 예측에 매우 주요한 이론적 근거를 제시하며, 이는 대다수의 교통류 예측모형이 교통류 상태의 시·공간적 진화행태를 근거로 개발되었기 때문임
- Disbro and Frame(1989)는 교통류의 특성을 무질서 상태(Chaotic state)라고 정의하였으며, Smith et al.(2002)는 교통류의 패턴은 확률적(Stochastic) 특성이 강하다고 지적하였음. 이는 확

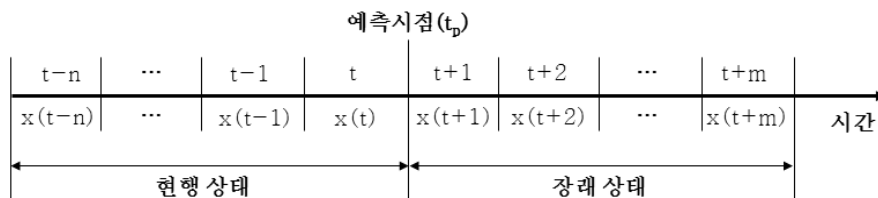
률과정을 기반으로 하는 계절 ARIMA(Seasonal Autoregressive Intergrated Moving Average, 이하 SARIMA)과 같은 기존의 통계기반 시계열 모형의 우수성이 높다는 것을 의미함

- 반면, Vlahogianni et al.(2006)은 교통류 상태 진화의 특성을 비고정성(Non-stationarity) 상태, 비선형성(Non-linearity) 상태, 초기 결정성(Initial deterministic) 상태, 혼돈(Chaos) 상태, 현시전환이동(Transitional movements) 상태로 설명하였음
- 또한, 교통류 상태의 시계열적 진화 상태는 비선형성, 초기 결정성과 혼돈 상태가 강한 혼합 상태(Mixed state) 시스템이라고 지적하였으며, 인공신경망(Artificial Neural Network)과 같은 비선형 모형이 장래 상태의 예측에 합리적이라고 언급하였음
- 따라서 기존의 확률과정을 기반으로 하는 통계적 모형의 장래 상태 예측력은 낮다고 언급하였으며, 비선형 모형의 예측력이 우수하다고 지적하였음
- Chang et al. (2012a, 2012b)는 최근린 이웃(k-nearest neighbor, 이하 KNN)을 기반으로 하는 비모수 회귀(Non-parametric regression, 이하 NPR) 모형의 장래 및 과거 교통상태 설명력이 SARIMA보다 우수함을 보였으며, 이는 교통상태의 시계열적 진화 상태는 다양한 상태가 혼재된 혼합 상태(Mixed state)라고 언급하였음
- 또한, van Hinsbergen et al.(2007)이 지적한 바와 같이 특정 모형이 장래 교통 상태의 설명력이 뛰어나게 우수하지 않다고 지적하였음
- 이상에서 고찰한 교통상태의 사·공간적 진화 행태는 아직까지 지속적인 연구가 수행되고 있는 영역이며, 공통된 사실은 교통류의 진화 행태는 시간과 공간적으로 매우 다양하게 변화하는 동적 혼합 상태라는 것을 알 수 있음

## 나. 선행연구에 의한 기존 예측모형 고찰

### 1) 예측의 시·공간 정의

- 일반적으로 시스템의 상태는 연속적(Continuous)이나, 이를 시계열적으로 분석하기 위해서는 단위 시간의 길이(Length of time interval)로 구분된 이산적 상태(Discrete state)로 전환해야 함
- 이러한 이산적 상태의 시계열 자료(Time-series data)는 예측의 영역에 있어 시간 영역을 정의하는 상태벡터(State vector)로 이용되며, <그림 2-7>은 현재 시각( $t$ )에서 예측의 시간적 정의를 보여주고 있음
- 여기서, 상태는 예측 시점(Prediction point,  $t_p$ )을 기준으로 하는 시간의 전, 후로 단위 시간의 길이에 따라 구분하며, 이러한 단위 시간의 길이는 보통 30초~5분 정도로 설정하게 되며, 기존 통행시간 예측의 경우 주로 5분을 이용함
- 예측 시점을 기준으로 구분되는 예측 영역은 각각 시간대(Time interval)으로 정의하며, 이러한 시간대는  $t_p$ 보다 과거 시간대의 경우  $(t-n), \dots, (t-1), t$ 로 정의하며 현행 영역으로 정의함. 그리고  $t_p$ 보다 미래의 시간대는  $(t+1), (t+2), \dots, (t+m)$ 으로 장래 영역으로 정의함
- 현행 영역의 시계열적 자료는 주로 개별 현행 시간대의 상태 평균값을 이용하여 상태벡터( $X = [x(t), x(t-1), \dots, x(t-n)]$ )로 정의됨
- 따라서, 장래 예측을 위하여 현행 상태벡터를 설정해야 하므로 현행 상태의 개수 즉, lag의 개수는 최적화를 수행하여 이에 대한 파라미터 값을 정산해야 함. 그리고 장래 상태의 경우,  $m=1$ 인 경우 단일 시간대 예측,  $m>1$ 인 경우 다중 시간대 예측으로 정의함
- 또한 단일 시간대와 다중 시간대 예측을 포함하여 이를 구분하지 않은 경우는 단기예측(Short-term prediction)으로 함



<그림 2-7> 교통류 상태의 시간적 진화

## 2) 입력벡터의 차원과 예측문제의 정의

- 현행 상태벡터(또는 입력 상태벡터)를 구성하는 시계열 자료의 개수인 lag의 개수 (Embedding size,  $l$ )는 최적화를 수행하여 파라미터 정산을 해야 하지만 주로 Takens(1981)의 정의에 의한 lag의 개수에 의존하는 경우가 많으며, 주어진 독립변수의 개수  $L$ 차원에서 다음의 식과 같음

$$l \geq 2L+1$$

- 예측 문제(Forecasting problem)는 주어진 예측 시점( $t$ )에서 현행 상태벡터( $X_c$ )를 이용하여 장래 시간대  $t+m$ 의 상태인  $\hat{x}(t+m)$ ,  $m \geq 1$ 을 예측하는 문제로 정의될 수 있으며, 다음의 식과 같이 표현함

Given  $X_c(t)$  with  $l$

Predict  $\hat{x}(t+m)$

- 일반적으로 독립변수인 상태벡터가 한 개인 경우 단일 변수(Uni-variate), 두 개 이상인 경우 다중 변수(Multi-variate)로 구성됨
- 예측 문제에서 Input-Output의 관계는 [단일 변수 - 단일 변수], [다중 변수 - 단일 변수], [다중 변수 - 다중 변수]이며, [단일 변수 - 단일 변수], [다중 변수 - 단일 변수]의 관계가 주를 이룸
- Takens(1981)의 정의에서 단일 독립변수( $L=1$ )인 경우 lag의 개수( $l$ )은 3개 이상에서 최적의 상태벡터를 구성하며, 2개의 독립변수( $L=2$ )인 경우 lag의 개수( $l$ )은 5개 이상에서 최적의 상태벡터를 구성할 수 있음
- 반면, Packard et al.(1980)은 최적의 상태벡터의 정의를 아래의 식으로 제시하였으나, Chang et al.(2012)는 최적의 lag의 개수를 정의하는 일반적인 정의는 존재하지 않는다고 지적하였음

$$l \leq 2L+1, l > 0$$

### 3) 예측모형의 분류 및 특성

- 장래 소통상태의 예측 문제를 해결하기 위하여 다양한 예측모형들이 개발되었으며, <표 2-1>은 장래 예측을 위하여 주로 이용되는 예측모형들의 분류를 보여주고 있음. 예측 모형의 분류는 정적기법(Naive)과 동적 모형인 모수(Parametric) 모형과 비모수(Non-parametric) 모형으로 구분됨

<표 2-1> 예측모형의 분류

구분	예측모형
정적(Naive)기법	동시(Instantaneous)기법
	과거평균(Historical average)기법
	군집(Clustering)모형
모수(Parametric)기법	교통류 모의실험 모형
	시계열(Time Series)모형
비모수(Nonparametric)기법	퍼지 이론(Fuzzy Logic)
	신경망(Neural Network)
	k-최근린 이웃(K-Nearest Neighbor)

#### ① 정적기법

- 정적기법은 현장에서 널리 적용되는 방법이지만, 모형에 대한 가정 없이 현상 상태의 변동을 고려하지 않음. 주로 이용되고 있는 정적기법으로는 동시(Instantaneous) 기법, 과거평균(Historical average) 기법, 군집(Clustering) 기법 등이 있음

#### ◦ 동시기법

- 동시기법은 주로 동시 통행시간(Instantaneous Travel Time, 이하 ITT)을 이용한 예측기(Predictor)로 적용되며, 장래에도 현행 상태가 지속된다고 가정함
- 따라서 모형의 연산이 매우 빠르며, 현장 적용성은 뛰어난 반면 예측 성능은 현저히 낮음 (Lint et al., 2002; Park and Rilett, 1998; Park et al., 1999; Rilett and Park, 2001)

#### ◦ 과거평균기법

- 과거평균기법은 구축된 과거 이력자료의 평균속도 자료를 이용하며, 최근 고도화된 예측모형에 비하여 과거평균기법의 예측력은 낮은 편임. 하지만 다중시간대 예측의 경우, 반복적 통행패턴이 나타나는 도로구간의 예측력은 타 모형에 비하여 우수한 성능을 보일 수 있음 (Kamarianakis and Prastacos, 2003; Park and Rilett, 1998; Park et al., 1999; Rilett and Park, 2001; Wu et al., 2004)

- 앞서 설명한 동시기법과 과거평균기법을 조합한다면, 보다 긴 장래 예측 시간대를 예측할 수 있으나, 그다지 높은 예측력을 보이지 않음(Park and Rilett, 1999). 그리고 반복적 특성이 강한 도시부의 연속류 도로구간에 적합하며, 비반복적 특성이 강한 지역 간 연속류 구간의 경우는 예측력이 떨어짐

#### ○ 군집기법

- 군집기법은 과거 이력자료 중 유사한 교통 패턴을 기반으로 해당 요일, 시간대에 대한 특정 그룹의 평균값을 적용하는 방법으로 예측모형의 입력 자료를 전처리하기 하기 위하여 이용됨(Chrobok et al, 2004)
- 군집기법의 경우, 앞서 동시기법, 과거 평균기법이나 선형 회귀모형보다 우수한 성능을 보이는 경우가 많음

#### ② 모수기법

- 모수기법은 교통류 모의실험 모형과 통계를 기반으로 하는 시계열(Time series)기법으로 구분할 수 있음

#### ○ 교통류 모의실험 모형

- 교통류 모의실험 모형은 교통류 행태에 대한 지식을 모형의 구조에 포함시켜 장래 상태를 추정하며, 교통사고와 같은 비 반복적 행태를 평가할 수 있는 장점이 있어 동적 교통관리의 의사결정을 위한 수단으로 이용됨
- 일반적으로 교통류 모의실험 모형은 “워드롭의 제1원칙(Wardrop's First Principle)”에 의한 네트워크 균형의 개념을 이용한 기/종점(Origin-Destination, 이하 OD) 통행량을 배정한 교통 수요를 이용하므로 교통류 모의실험 모형의 예측문제는 동적 기/종점 통행량 예측으로 귀결됨
- 이러한 교통류 모의실험 모형은 거시적(Macroscopic), 중시적(Mesosopic), 미시적(Mircroscopic) 교통류 모형으로 구분할 수 있음
- 거시적 모형은 거시적 교통류 변수인 밀도, 속도, 그리고 교통량이 이용됨(Cremer, 1995). 교통수요를 예측하기 위한 통행배정기법, 즉, 정적 통행배정과 동적 통행배정이 이용됨
  - 정적 통행배정은 전체 O/D 통행량의 균형상태만 고려되므로 출발시간을 고려하지 않으며, 앞막힘(Spillback) 현상 등으로 인한 경로우회 등을 설명할 수 없음. 이러한 단점에도 불구하고, 모형이 심플하고 연산속도가 타 모의실험 모형에 비하여 빠르기 때문에 현장에서 많이 적용되고 있음



- 동적 통행배정은 시간의 흐름에 따른 교통수요의 변화를 고려하며, 유체역학을 기초로 한 거시적 교통류 해석 모형을 이용하여 시간의 흐름에 따른 교통량과 밀도의 전과과정을 해석한 통행배정 모형임(Hounsell and Ishtiaq, 1997; Lin et al. 2004).
- 미시적 모형은 도로망에서 개별 차량 간의 상호작용을 고려하면서 모든 개별 차량에 대하여 모의실험을 수행한 후, 각 교통류 상태를 해석하는 모형임
  - Nagel et al. (1992)에 의해 개발된 Cellular Automata(CA)기반의 차량모형을 기초로 하는 다양한 차량모형이 각종 모의실험에 적용되었으며, 이러한 CA기반의 차량모형은 타 모형에 비해 연산 수행속도가 빠름. 차량이 가다서다하는 현상과 동시 교통류를 잘 설명할 수 있는 장점이 있음
  - Dailey et al. (2002)은 조사된 O/D통행량을 이용하여 장래 교통상태를 예측하였으며, Chrobok et al. (2004)는 링크 회전비를 이용하여 장래 교통상태를 예측하였음
  - 장현호(2002)는 CA기반의 차량모형을 이용하여 단속류에 대한 장래 교통상태를 예측하였으며, 이러한 CA기반의 차량모형을 이용한 장래 교통상태 예측은 교통사고와 같은 비반복적 상태의 장래 교통상태 예측에 이용되고 있음(Chang et al., 2007; Kim et al., 2008; Shah et al., 2008)
- 중시적 모형은 거시적 모형과 미시적 모형을 결합하여 적용한 모형으로서 먼저 거시적 모형을 이용하여 통행배정을 수행한 후, 개별 차량은 산정된 거시적 교통류 변수에 기초하여 네트워크에 부하됨. 중시적 모형의 장점은 차량 대기행렬을 구현할 수 있다는 것이나, 미시적 모형 수준에서 각 개별 차량의 통행행태를 설명할 수는 없음. 대표적 중시적 모형으로는 DynaMIT이 있으며, 이력 자료와 실시간 정보를 이용하여 O/D통행량을 추정하고 중시적 차량모형을 이용하여 해당 네트워크에 차량을 부하함(Ben-Akiva et al. 2001)
- 시계열 모형
  - 시계열 분석(Time series analysis)을 위한 모형은 과거 이력자료의 관측값과 오차항(Error term)의 함수를 이용하여 장래 예측 변수를 모형화하는 과정으로서 이는 정적 상태를 가정함. 관련 모형으로는 선형 회귀(Linear regression), 국부 가중회귀(Locally weighted regression), ARIMA(Autoregressive integrated moving average), 칼만필터링(Kalman filtering), 가우시안 최우도 추정(Gaussian maximum likelihood) 등 다양한 분석기법이 존재함
  - 선형 회귀모형은 변수의 공변량(Covariate)의 선형 조합을 가정한 예측함수(Forecasting function)로서 모형이 간단하기 때문에 연산 수행속도가 빠르며, 어느 정도 만족할 만한 정확도를 보이는 것으로 판단됨(Chrobok et al., 2004; Kwon et al., 2000; Lan and

- Miaou, 1999; Lingras et al., 2002; Rice and Zwet, 2004; Sun et al., 2003; Zhang and Rice, 2003; Zhong et al., 2005)
- 국부 가중회귀모형은 국부적인 회귀(Local regression)를 이용하며, 각 예측시점의 예측 오차는 현행 관측 값에 가까울수록 가중됨. 선형 회귀모형에 비하여 우수한 결과를 보였으며, 연산 수행속도도 빠른 장점이 있음(Sun et al., 2003; Zhong et al., 2005)
  - ARIMA모형은 가장 일반적인 통계적 기법으로 Box-Jenkins(1976) 모형을 기반으로 함. ARIMA모형의 예측력을 향상시키기 위하여 그 동안 다양한 형태의 ARIMA모형들이 개발되었음. 아래의 ARIMA모형들은 서로 간에 연구결과가 달라 어떤 계열의 모형이 정확도가 우수한 지에 대한 논란이 있으나, 대부분 Seasonal ARIMA가 타 ARIMA모형들에 비하여 우수한 것으로 분석됨
    - (1) seasonal ARIMA(Guo, 2005; Smith et al., 2003; Williams et al., 1998)
    - (2) subset ARIMA(Lee and Fambro, 1999)
    - (3) Kohonen ARIMA(Voort et al., 1996)
    - (4) ARIMAX(Williams, 1999)
    - (5) VARMA, STARMA(Kamarianakis and Prastacos, 2003)
    - (6) 지수평활화(Crobok et al., 2004; Part et al., 1998)
  - 칼만필터링 모형은 Kalman(1960)에 의해 개발되었으며, 현행 상태벡터와 바로 직전 상태, 그리고 백색소음과 오차의 잔차를 이용하여 장래 교통상태를 예측함. 연산 수행속도가 매우 빠르고 시계열의 변동을 효과적으로 설명하기 때문에 많은 연구와 실제 시스템에 탑재되고 있음(Ben-Akiva et al., 1992; Chen and Grant-Muller, 2001; Chien and Kuchipudi, 2003; Kuchipudi and Chien, 2003; Okutani and Stephanedes, 1984; Whittaker et al., 1997;)
  - 가우시안 최우도 추정모형은 과거 평균에서 가능한 적게 예측값은 벗어나며, 예측되는 증가는 과거의 증가에서 가능한 적게 벗어난다는 2가지 가정을 기반으로 개발되었음. 예측력은 타 시계열 분석 모형보다 우수한 것으로 보고됨(Nicholson and Swann, 1974)

### ③ 비모수 모형

- 비모수(Non-parametric) 모형은 변수(Parameter)가 없음 의미하지 않으며, 변수의 개수와 내재적 특성이 가변적이고 고정되지 않음을 의미함. 비모수 모형의 구조 및 변수는 이용되는 데이터 즉, 예측대상 시스템의 상태 변화에 따라 결정되며, 통계학적 모형이 아닌 알고리즘 기반의 문제해결 접근법임
- 따라서, 시계열 분석과 같은 모수적 방법에 비하여 많은 데이터가 요구되지만, 시스템 상태의 시간적 변화에 대한 깊은 이해와 지식을 필요로 하지 않음. 즉, 알려지지 않은 변수

(Unknown parameter)와 특성으로 설명되는 내재된 특성을 동적으로 설명할 수 있다는 장점이 있음

- 이러한 비모수 모형의 종류로는 퍼지 이론(Fuzzy theory), 신경망(Neural Network) 그리고 k-최근린 이웃(k-nearest neighbors)으로 구분할 수 있음

#### ○ 퍼지 이론

- 퍼지 모형은 IF-THEN rule의 집합으로 구성되는 의사결정 논리라고 할 수 있으며, 현행 상태는 하나 또는 여러 개의 If에 의해 구성되고 장래 상태는 THEN에 의하여 예측됨
- 퍼지 이론을 기반으로 하는 예측모형은 만족할 수준의 결과를 도출하였지만(Coufal and Turumen, 2004; Li et al., 2006), 신경망과 결합한 모형은 더 우수한 예측력을 보이는 것으로 분석됨(Huisken, 2003)
- 즉, 예측 분야에서 퍼지 이론은 대부분 신경망과 같은 모형과 결합되어 이용되고 있으며, 퍼지 이론 자체만으로 예측 문제를 해결하는 데에는 한계가 존재함

#### ○ 신경망

- 신경망 모형은 비선형적인 교통상태의 변화과정(Non-linear process)을 모델링할 수 있기 때문에 ARIMA, 칼만 필터링 모형과 더불어 교통상태 예측문제에 가장 많이 이용되고 있는 모형임
- 즉, 신경망의 학습과정(Training process)을 현행 상태를 학습시킨 후, 학습된 지식을 통하여 장래 교통상태를 추정하게 됨. 이를 Black box 알고리즘으로 불리기도 함
- 신경망의 구조, 효과적인 훈련과정을 통한 학습, 그리고 신경망 알고리즘의 훈련과정(Training process)에서 요구되는 연산시간을 단축하여 예측 정확도를 향상시키기 위한 일련의 노력으로 다양한 형태의 신경망 모형들이 개발되었음
- 신경망 기반의 예측모형은 전통적으로 학습과정, 신경망의 내부 구조, 입력 자료의 전처리, 시공간(Temporal and spatial) 패턴으로 구분할 수 있음

#### ○ k-최근린 이웃

- k-최근린 이웃 기법은 시계열 분석 기법과 신경망 모형과 달리 현행 상태와 유사한 과거 상태를 탐색하여 현재 상태와 유사한 과거 상태인 k개의 최근린 이웃을 구축한 후, 직접 평균 또는 상태 간 거리의 역수로 가중한 평균 상태를 이용하여 장래를 예측함(Chang et al., 2012b)

- k-최근린 이웃 관련 연구들은 앞서 설명한 정적예측 기법보다는 우수한 성능을 보였음 (Smith and Demetsky, 1996; Smith et al., 2003; Rice and Zwet, 2004).
- k-최근린 이웃 기법은 교통류 상태의 진화에 대한 이해를 방대한 양의 과거 상태 즉, 과거 자료의 다양성과 질에 전적으로 의존함(Chang et al., 2012a, 2012b)
- 기존 몇몇의 연구에서는 k-최근린 이웃 기법의 데이터 요구수준을 충족하지 않은 상태에서 해당 모형을 Benchmark 모형으로 적용하여 결과를 도출한 경우가 많다는 견해가 있음 (Yoon and Chang, 2014)

#### ④ 기존 예측모형의 이론적 고찰의 소결

- 앞서 기존 예측모형을 정적기법(Naive), 모수(Parametric) 모형, 비모수(Non-parametric) 모형으로 구분하여 선행 연구들을 통하여 이론적 고찰을 수행하였음

##### ○ 정적기법

- 정적기법(Naive)은 모수 및 비모수 모형에 비하여 낮은 예측력을 보이지만, 실제 적용에 있어 현재로는 가장 현실적이며 보편적인 방안 중 하나임
- 현재 중/장거리 소통상태의 정보제공에 있어 유력한 대안이라 할 수 있으나, 중/장거리 소통 상태 정보 예측에 적용할 수 있을 정도로 단기 예측모형의 예측의 시간적 범위가 수용할 만한 수준에서 확장되지 않았기 때문에 중/장거리 예측에 적용하기에는 무리가 있음

##### ○ 모수(Parametric) 모형

- 먼저 모의실험 모형들은 교통사고 등의 비 반복적 상태와 다양한 교통제어 전략을 포함하는 동적 교통류 관리에 적합하나, 교통류 및 차량 모형의 파라미터 정산이 쉽지 않고 장래 예측이 아닌 동적 O/D통행량 추정에 더 적합함. 현재까지 동적 O/D통행량 추정은 지속적인 연구 분야이며, 모형의 연산 수행 속도가 타 차량모형에 비하여 상대적으로 느리다는 단점이 있음
- 시계열 분석 기법들은 확률적 상태이론을 기반으로 시계열적으로 정적(Stationary) 상태를 가정하지만 실제 시스템은 동적 시스템임. 일반적으로 시계열 모형은 계절성(Seasonality)이 필요하며, 교통류 이론보다는 통계적 함수에 의존함. 특히 ARIMA모형은 평균값에 의존하는 모형의 본질적 특성으로 인하여 상태변화가 안정정일 때 예측력은 우수한 반면, 교통상태가 크게 변화하는 전환점에서 다음 상태를 예측하지 못 하는 한계가 있음(Vlahogianni et al., 2005). 특히 교통상태가 급격하게 변하는 단속류의 행태를 정확히 설명하지는 못함(Chang et

al., 2012b; Yoon and Chang, 2014)

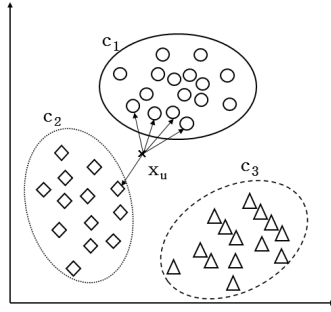
○ 비모수(Non-parametric) 모형

- 신경망 모형은 그 동안 많은 연구에도 불구하고, 한편의 논문(Park and Rilett, 1998)을 제외하면 네트워크 차원의 적용이 이루어지지 않았음. 이는 대다수의 연구가 단일 지점 또는 단일 구간의 자료를 이용하여 수행되었기 때문이며, ITS의 시스템에서 볼 때 신경망 모형의 연산 수행속도는 동적 교통류 관리 및 정보 제공에 불리하기 때문임
- 현재 ITS의 자료수집 기술이 발달하고, 대용량 자료를 지원하게 되면서 비모수 기반의  $k$ -최근린 이웃 기법에 대한 연구가 증가하고 있음(Chang et al., 2012a). Chang et al.(2010)과 Kim et al.(2011)은 해당 모형이 타 모형에 비해 통행시간의 예측에 우수한 예측 정확도를 보였으며, Chang et al.(2012a, 2012b)는 교통량의 예측과 추정에 있어 다중 시간대 예측에도 불구하고 단일 시간대 예측에 주로 이용되는 Seasonal ARIMA, 칼만필터링 모형의 정확도 보다 우수한 것으로 분석됨
- 앞서 살펴본 단기 예측의 모형들은 수용할 만한 예측오차의 수준에서 예측영역을 확장하는데에 실패한 반면, 비모수 기반의  $k$ -최근린 이웃 기법은 예측영역의 확장에 성공한 사례를 보고하고 있음(Smith and Demetsky, 1996; Chang et al., 2010, 2012b). 또한 과거 시계열 자료의 변동이 극심한 단속류의 상태 예측에 있어서도 Seasonal ARIMA와 칼만필터링 모형 보다 훨씬 우수한 성능을 보이는 것으로 분석됨(Yoon and Chang, 2014)

#### 다. 본 연구의 적용기법에 대한 이론적 고찰

- $k$ -최근린 이웃(K-Nearest Neighbor)기법<sup>2)</sup>은 앞서 설명한 비모수(Non-parametric) 모형 중에서 가장 많이 이용되고 있는 방법이며,  $k$ 개의 최근접하는 이웃을 이용한다는 의미에서 붙은 명칭으로 학습 데이터 집합에 있는 표본 간의 유사도에 따라 라벨이 붙지 않은 표본들을 분류하는 매우 직관적인 방법이라고 할 수 있음

<sup>2)</sup> 한빛아카데미(2014), “패턴인식 개론”, p.233



<그림 2-8> k-최근린 이웃의 전략

- 즉, 라벨이 없는 표본  $x_u \in R^D$ 이 주어질 경우, 학습 데이터 집합에서 가장 가까운  $k$ 개의 라벨 표본을 찾고,  $k$ 개의 부분집합 내에 가장 빈도가 많이 나타나는 클래스에  $x_u$ 를 할당하는 방법으로 단지 상수  $k$ , 라벨이 있는 학습데이터 집합의 표본과 거리 척도만 필요함
- 여기서, k-최근린 이웃기법의 거리 척도는 주로 유클리드 거리(Euclidean distance)를 이용하며, 유클리드 거리를 구하는 식은 아래와 같음
  - 데이터 간의 유클리드 거리( $ED$ ) 측정

$$ED(x, y) = \left( \sum_{k=1}^d |x_k - y_k|^m \right)^{1/m} \quad m=2, \quad k \leq d$$

- <그림 2-8>과 같이 3개의 클래스( $c_1, c_2, c_3$ )가 있을 경우, 미지의 표본  $x_u$ 에 대한 라벨을 결정하여 이 표본이 속하는 클래스를 찾고자 함. 앞서 설명한 유클리드 거리를 척도로 이용할 경우  $k$ 를 5로 하고 인접한 표본에 대한 거리를 각각 계산하면, 이들 중 4개가 클래스  $c_1$ 에 속하고 나머지 1개가 클래스  $c_2$ 에 속하므로  $x_u$ 는 클래스  $c_1$ 으로 할당됨

### 3. 국내외 기술 현황 분석

#### 가. 국내 기술현황

- 국내 교통예보 관련 기술현황은 한국도로공사, 서울시 토피스(TOPIS) 등 주요 공공기관에서 수집된 과거 교통 이력자료를 바탕으로 제한적인 교통예측 정보를 제공 중에 있음
- 또한, SK플래닛, KT 등 이동통신사를 중심으로 한 일부 민간기업에서 휴대폰 어플리케이션을 이용하여 실시간 빠른 길 찾기, 예측 통행시간 정보 등을 제공하고 있음

<표 2-2> 국내 교통예보 관련 기술현황

기관	대상도로	예측수준	비고
한국도로공사	고속도로	주간예보, 월간예보	과거 2~4주의 평균값을 이용 ※ 정확도 낮은 실정이며, 타 도로에 적용이 불가능함
서울시 교통정보센터 (TOPIS)	서울시 주요도로, 도시고속도로	-	실시간 교통소통정보 위주 (예측시스템은 현재 개발 중)
엔나비/ SK T Map	전국도로	일중	일중 예측정보는 제공하고 있으나, 익일 이상의 교통예보는 제공하지 않음

#### 1) 한국도로공사 교통정보(ROAD PLUS)

- 통행시간 예측 시, 단순히 과거 2~4주의 이력자료에 대한 평균값을 이용하여 주요 도시 간 고속도로 예상 소요시간 정보를 제공하고 있으며, 익일 이상의 중장기 교통예보 측면보다는 현 상태의 상황 또는 일중의 예측 정보제공의 측면이 강함
  - 익일 이상의 중장기 교통예보의 수준 아님
  - 방대한 양의 고속도로 과거 자료는 이용하지 않음
  - 정확도 낮은 실정이며, 타 도로에 적용이 불가능함
  - 주요도시 간 예상 소요시간을 현재의 교통상황에 맞춰 제공
  - 주간예보 및 월간예보 제공

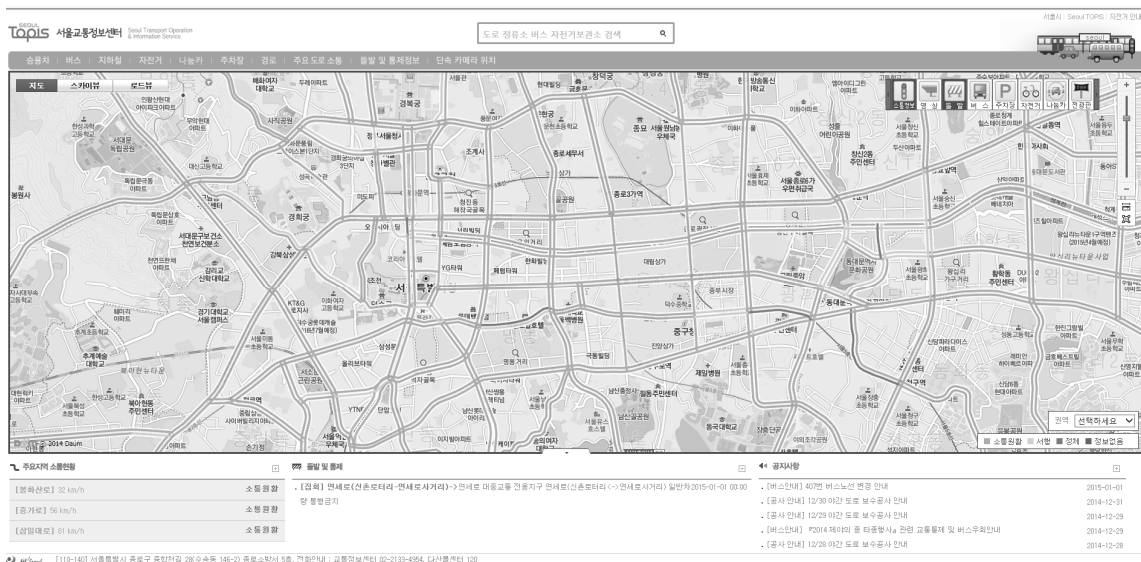


<그림 2-9> 한국도로공사 교통정보(ROAD PLUS)

자료 : <http://www.roadplus.co.kr/main.do?op=main>

## 2) 서울시 교통정보센터(TOPIS)

- 서울시 교통정보센터에서 주로 제공하는 정보는 실시간 교통소통정보(소통원활, 서행, 정체) 위주이며, 단기 또는 중장기 교통소통상태 예측정보는 제공하고 있지 않음
  - 현재 장래 교통소통상태 예측 시스템은 개발 중에 있음
- 그 외 대중교통(버스, 지하철), 자전거, 나눔카, 주차장, 경로, 돌발 및 통제, 단속 카메라 위치 등을 제공 중에 있음



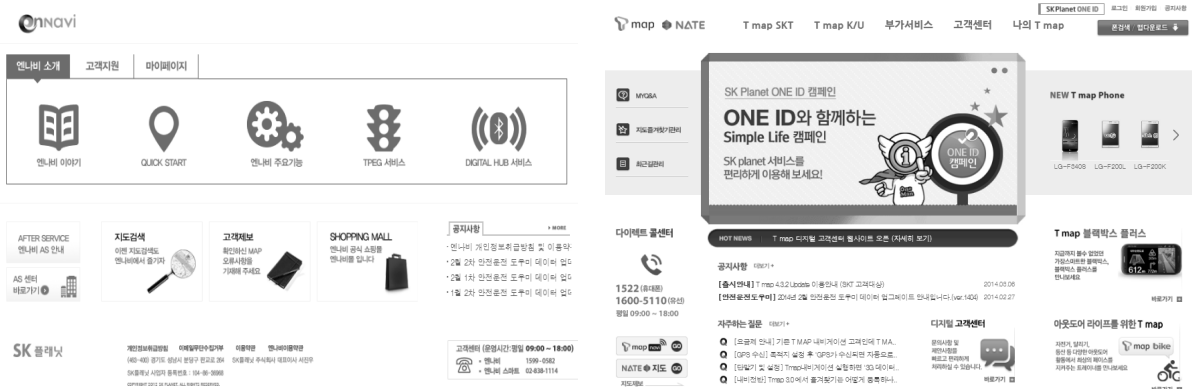
<그림 2-10> 서울시 교통정보센터(TOPIS)

자료 : <http://topis.seoul.go.kr/index.jsp>



## 3) 엔나비/SK T-Map

- SK T-map의 경우, ‘T-map 4.0’ 서비스에서 미래 시점의 출발 및 도착 예측이 가능한 ‘타임머신’ 기능이 추가됨
  - 타임머신 : 이용자가 원하는 요일 및 공휴일별 출발시각과 도착시각 지정 후 출발지로부터 도착지까지의 통행시간 정보를 제공받을 수 있음 → 요일별 패턴 DB 및 실시간 교통정보를 융합하여 교통정보를 제공함
- ‘T-map 4.4’에서는 기존 타임머신 기능이 강화된 ‘언제갈까?’ 서비스가 추가됨
  - 출발지/목적지 설정 시 실시간/패턴 교통정보가 반영된 소요시간 예측
  - 지금 출발, 한 시간 후 출발, 두 시간 후 출발 등 출발시간 별 소요시간을 예측 및 비교 가능
- 해당 업체에서는 해당 당일의 소요시간 예측정보만 제공하고 있으며, 익일 이상의 중·장기의 예측정보는 제공하고 있지 않음



&lt;그림 2-11&gt; 엔나비/SK T-Map 교통정보

자료 : <http://www.ennavi.co.kr/index.jsp>

<http://www.tmap.co.kr/tmap2>

## 나. 국외 기술현황

- 미국의 Beat the Traffic, IBM, INRIX, 영국의 Highway Agency는 대표적인 해외 교통정보 서비스 제공기관임
- 패턴화된 교통정보를 가공하여 주요 간선도로 및 고속도로를 대상으로 교통정보를 제공함

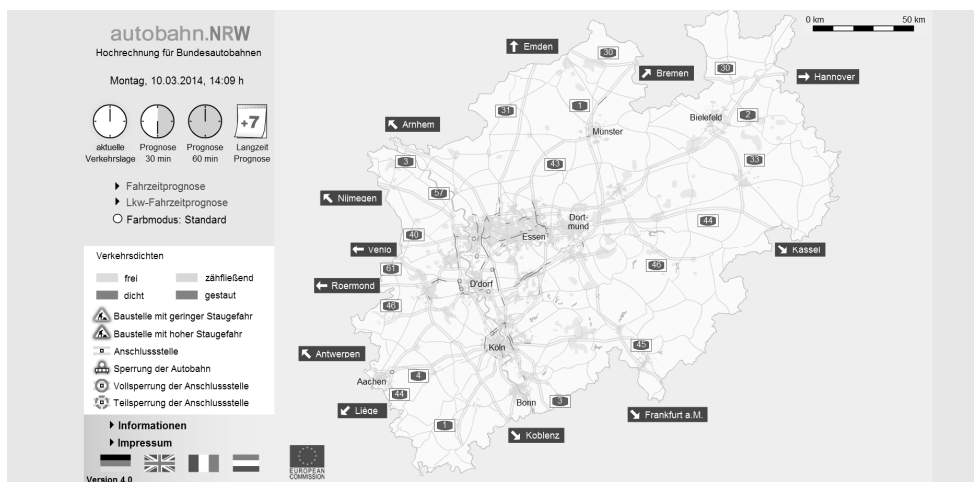
<표 2-3> 국외 교통예보 관련 기술현황

기관	대상도로	예측수준	비고
OLSIM(독일)	고속도로	30분~1시간	Autobahn 고속도로 대상
Help2Travel(영국)	Midland 지역	1주후, 2주후 주기	행사/여행정보 제공 ※ 도로별 교통 예측정보 제공하지 않음
Traffic England(영국)	주요 도시별 고속도로 및 국도	10시간	현재 개발 중
INRIX Traffic(미국)	고속도로, 간선도로, 보조간선도로	12시간	-
Beat the Traffic(미국)	고속도로, 간선도로, 보조간선도로	2시간	-

### 1) OLSIM(독일)

- 독일의 Autobahn 고속도로를 대상으로 예측 교통정보(30분과 1시간 후 고속도로 소통정보)를 제공함
  - 높은 인구밀도에 따라 교통량이 증가하여 발생하는 Autobahn 고속도로의 교통체증 해결을 위하여 개발함
  - 디지털 도로지도 상에 색상차이를 이용해 자유교통류 및 혼잡교통류 등의 교통상황을 표현함
  - OLSIM 시뮬레이션 모델은 기존의 CA기법을 확장한 형태, 기존 CA에 비해 Cell의 크기를 작게 구현하여 모든 경험적 교통상황 표현 가능하도록 구축함
  - 차량을 승용차 및 트럭으로 구분, 차량추종 및 차로변경 방법을 상이하게 구축함
- 고속도로를 대상으로 실시간 교통상황을 고려한 해당 일의 장래 소통 상태를 예측함
  - 단기 소통상황 예측이며, 중·장기의 혼잡 예보는 아님

- 도시고속도로 및 일반도로 제외
- 입력정보 : 4,000개의 고속도로 루프검지기 ⇨ 고속도로 한정
- 예측방법 : On-line 시뮬레이션
- OLSIM 시뮬레이션 방법은 아래와 같음
  - 직선구간, 램프, 감/가속구간 등의 도로 Topology를 디지털 도로지도 상에 상세히 표현함
  - 고속도로 검지기를 통해 매분마다 업데이트 되는 데이터를 이용함
  - 교통상황은 실시간 교통정보가 추가된 고속도로 교통류 시뮬레이션을 통해 계산되어 표현됨
  - 차량의 움직임(감/가속, 좌/우회전, 추월 등)이 시뮬레이션 모델을 통하여 재현됨
- OLSIM은 미국 나사에서 1990년대 초에 개발된 TRANSIMS를 기반으로 개발됨
- 시뮬레이션 시 기상상황, 도로공사, 사고 등의 장애를 고려하는 것이 어려움
- 그 외 도로공사 및 통제 정보 등을 제공 중에 있음
  - 공사정보는 공사기간에 따라 일 단위 및 2주 단위로 수집하여 제공함
  - 공사정보의 경우 교통상황에 미치는 영향에 따라 Low jam risk, High jam risk로 구분하여 표출함

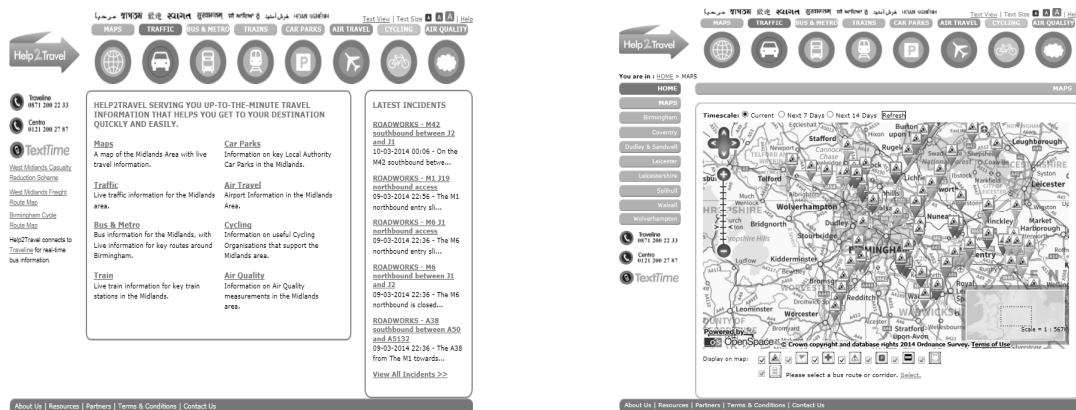


<그림 2-12> 독일 OLSIM 교통정보

자료 : <http://www.autobahn.nrw.de/>

## 2) Help2Travel(영국)

- MATTISSE 시스템 기반 통합 여행정보를 제공하고 있으며, 영국 Midland 지역의 행사/여행 정보를 제공함(1주후, 2주후 주기)
  - 도로별 교통 예측정보 제공하지 않음
  - 도로공사 등으로 인한 제약 상황 관련 정보 제공
  - 웹페이지를 통해 VMS 메시지의 확인이 가능하며, 해당 VMS 메시지의 표출 위치 및 내용 정보 또한 제공
- 그 외 돌발 상황에 관한 정보, 대중교통 및 주차 공항정보, 자전거에 대한 정보를 제공하며 대기오염에 관한 정보도 함께 제공 중에 있음
  - 대중교통(버스, 지하철, 전철)의 경우 실시간 정보, 최적 노선 및 정거장 등의 정보를 제공하고 있음
  - 주차장 위치 및 총 주차면수 정보 제공

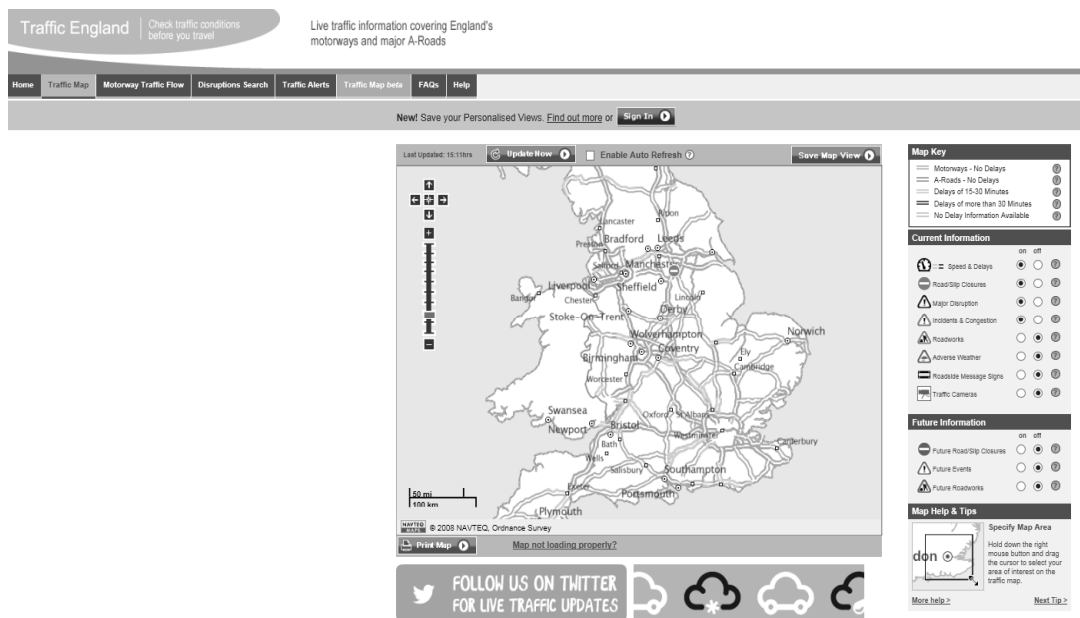


<그림 2-13> 영국 Help2Travel 교통정보

자료 : <http://www.help2travel.co.uk/mattisse/index.htm>

## 3) Traffic England(영국)

- 영국의 도로관리청(Highway Agency)의 교통정보 웹사이트로 이용되고 있음
  - National Traffic Control Center에서 운영 중
  - 월 기준 350,000명 이용
- 주요 도시별 고속도로 및 국도에 대한 실시간 교통정보를 제공
  - 교통량, 평균속도, 지체, 구간 속도, 이벤트 정보 등 제공
  - 도로 폐쇄, 사고 및 혼잡구간, 공사, 날씨, VMS 및 CCTV 정보 등 제공
- 미래 도로 폐쇄 및 이벤트, 공사정보 등 제공
- 과거 이력 데이터를 이용한 예측결과와 현재 속도 비교 기능 제공
- 교통정보 수집은 루프검지기를 이용함
- 교통예측 정보 메뉴인 Traffic Map Beta 개발 중임
  - 현재부터 10시간 후까지 예측결과 표출 기대
- 그 외 돌발 상황에 관한 정보도 제공 중에 있음



&lt;그림 2-14&gt; 영국 Traffic England 교통정보

자료 : <http://www.trafficengland.com/>

#### 4) INRIX Traffic(미국)

- 미국 민간업체 INRIX사의 실시간 교통소통정보 App임
  - 고객 차량의 GPS 데이터, Historical 교통데이터, Fleet 데이터 등의 도로 및 날씨정보 등을 수집
  - 자체 솔루션을 이용해 가공하여 이용자들에게 교통정보 제공
- 실시간 교통흐름정보 및 유고정보 등을 텔레매틱스 단말을 통해 5분마다 제공하는 서비스 실시 중에 있음
- Historical 교통 데이터 기반으로 매일 12시간까지 교통 예측정보 제공
  - Traffic Coverage : 고속도로, 간선도로, 보조간선도로
  - GPS 위성정보를 이용하는 프로브 차량 및 디바이스를 통해 차량의 위치 및 속도정보를 수집 및 가공하여 속도자료 구축
- 이용자 기종점 입력 시 최적 출발시간 예측 후 정보 제공 및 CCTV 정보 제공(유료서비스)

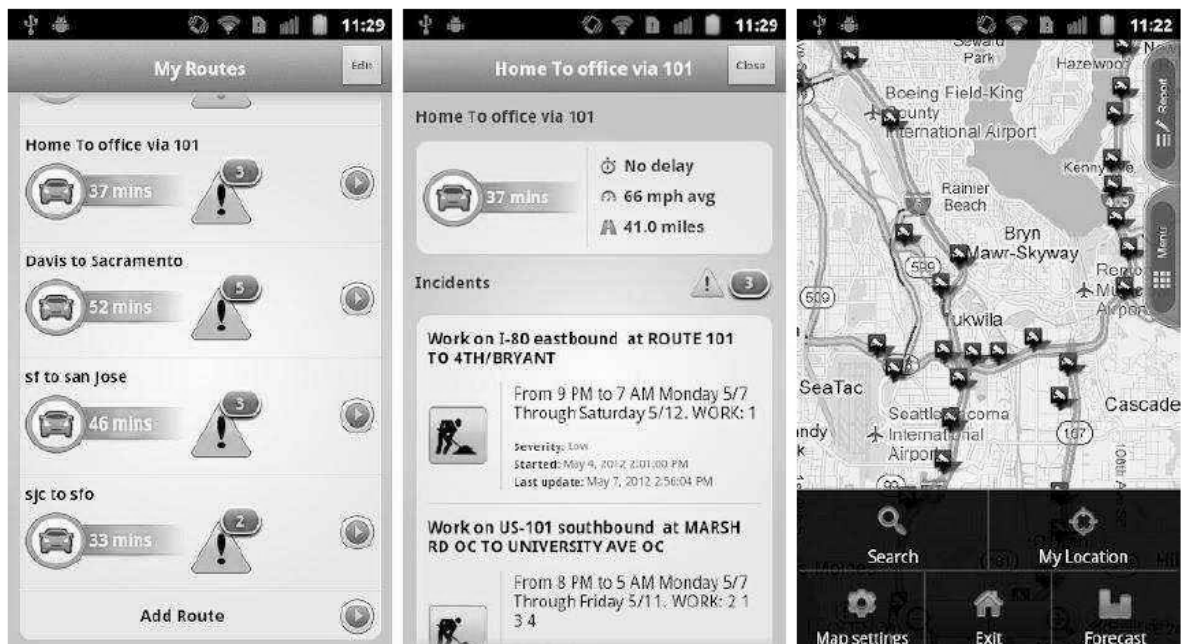


<그림 2-15> 미국 INRIX Traffic 교통정보

자료 : <http://www.inrixtraffic.com>

## 5) Beat the Traffic(미국)

- 미국의 Triangle Software사에서 개발한 App임
  - 미국 및 캐나다 지역의 실시간 소통 정보를 구글맵 기반으로 제공
  - 실시간 교통카메라 이미지 및 최대 2시간 후까지의 교통예측 정보 제공
- 과거 데이터를 기반으로 2시간 후까지의 예측 교통정보 제공
  - Traffic Coverage : 고속도로, 간선도로, 보조간선도로
- 그 외에 이용자가 지정한 기종점 간 노선의 통행시간 및 사고정보 등 표출 및 기상상황 정보 제공



&lt;그림 2-16&gt; 미국 Beat the Traffic

자료: <http://www.inrixtraffic.com/>

#### 4. 국내외 선행연구 고찰

##### 가. 국내 선행연구

- 장래 교통소통상태(통행시간 등) 예측과 관련한 국내 선행연구들은 현재시각 기준 1~15분 후의 단기 예측의 수준에 머물러 있으며, 본 연구와 같이 익일의 시간대별 구간 통행속도를 예측하는 중/장기 예측에 관한 연구는 미미한 실정임

<표 2-4> 국내 선행연구 고찰 결과

구분	대상구간	교통변수	예측주기	데이터 수집	적용 방법론
강정규외 1명(2002년)	고속도로	통행시간	10분/15분	TCS <sup>1)</sup>	신경망모형
이의은외 1명(2002년)	고속도로	통행시간	5분	TCS	신경망모형
오세창외 2명(2003년)	고속도로	통행시간	1분	루프검지기 영상검지기 자석식검지기	칼만필터링
이승재외 1명(2003년)	네트워크	교통량, 통행시간	5분	영상검지기	CA모형 칼만필터링
남궁성(2005년)	고속도로	경로통행시간	Estimation	TCS	전진반복 전후방탐색법
이영인외 2명(2005년)	일반국도	통행시간	5분	영상검지기 AVI <sup>2)</sup>	칼만필터링
김재진외 2명(2006년)	고속도로	통행시간	5분	TCS	베이지안 추론
심상우외 1명(2006년)	단속류	통행시간	-	GPS	대기행렬 이론
여태동외 2명(2009년)	단속류	통행시간	-	VDS <sup>3)</sup> , AVI	히스토리컬 프로파일 모형
심상우외 3명(2013년)	고속도로	통행시간	5분	하이패스 DSRC	PBM/LBM

주1 : 고속도로 통행료수납시스템(TCS, Toll Collection System)

주2 : 주행차량자동인식(AVI, Automatic Vehicle Identification)

주3 : 교통정보수집장치(VDS, Vehicle Detection System)



## 나. 국외 선행연구

- 장래 교통소통상태(통행시간 등) 예측과 관련한 국외 선행연구도 국내 선행연구와 유사하게 시간대의 길이는 1~15분이며, 장래 예측 시간대의 개수는 1~4개로 분석되어 단기 예측의 수준에 머물러 있음을 알 수 있음
- 이는 시간대의 길이를 증가하거나 장래 예측 시간대의 길이를 연장할 경우, 장래에 대한 불확실성이 증가하기 때문에 예측의 정확도는 낮아지기 때문임

<표 2-5> 국외 선행연구 고찰 결과

연번	저자	도로 유형	시간대 길이(분)	예측 시간대개수	자료수집 방법	모형
1	Innamaa (2009)	M	5	4	D	신경망
2	Jintanakul et al. (2009)	M	5	1	S	베이지안
3	van Hinsbergen et al. (2009)	M	5	3	D	신경망
4	Zou et al. (2009)	M	5	1	D	신경망
5	Innamaa (2009)	M	1	1	D	신경망
6	Yang et al. (2010)	M	15	1	A	통계기반
7	Abu-Lebdeh and Singh (2011)	A	5	1	S	Hybrid
8	Bustillos and Chiu (2011)	M	15	1	S	통계기반
9	Fei et al. (2011)	M	1	1	D	베이지안
10	Khosravi et al. (2011)	M	5	1	D	Hybrid
11	Li and Rose (2011)	M	10	6	A	신경망
12	Myung et al. (2011)	M	5	1	D/A	통계기반
13	Oh and Park (2011)	M	1	1	A	신경망
14	Simroth and Zehle (2011)	M	1	1	GPS	통계기반
15	Soriguera and Robusté (2011)	M	5	1	D/A	통계기반
16	Xia et al. (2011)	M	15	1	D	통계기반
17	Wang and Shi (2012)	M	1	1	D	베이지안
18	Mu et al. (2012)	M	1	1	S	통계기반
19	Du et al. (2012)	M/A	2	1	S	통계기반
20	Haworth and Cheng (2012)	M/A	5	1	A	통계기반
21	Khan (2012)	M	5	1	G	베이지안
22	Lu (2012)	M	5	1	D	베이지안
23	Ma et al. (2012)	M	1	1	S	통계기반
24	Qiao et al. (2012)	M	5	1	B	통계기반
25	Zheng and Van Zuylen (2012)	A	1	1	G	신경망

주1 : M(고속도로), A(간선도로)

주2 : D(검지기), A(AVL), G(GPS), S(모의실험)

## 5. 기존 기술/이론의 한계 및 개선방향 설정

### 가. 국내외 기술의 한계 분석

- 국내의 대표적인 3대 교통정보 제공기관인 한국도로공사, 서울시, 그리고 SK T-Map의 교통 정보 예보는 다음의 한계를 가짐
  - 현실적으로 익일 이상의 교통예보를 수행하기 위한 기술은 개발되고 있지 않음
  - 가까운 과거의 평균 또는 단순 과거 패턴 이용
  - 링크기반의 예측모형 ⇨ Network의 교통소통을 고려하지 않음 ⇨ 낮은 예측력
  - 교통수요의 변동을 고려하지 않음
  - 특정 도로 시스템 또는 행정단위에 편중되어 전국 도로에 적용하는 데에는 한계가 있음

<표 2-6> 국내 기술의 한계 분석

기관	대상 도로	예보수준	한계 및 문제점
한국도로공사	고속도로	주간예보, 월간예보	-과거 2~4주의 평균값을 이용 ⇨ 장래 상태에 대한 변화의 고려 불가 -정확도 낮은 실정 -고속도로 본선 대상 ⇨ IC/JC 연결로 적용 불가능 ⇨ 타 도로에 적용이 불가능
엔나비/ SK T Map	전국도로	일중	-명절연휴 외 주간/일중 예보는 제공하지 않음 -대국민 서비스가 아닌 고객 대상
서울시 토피스 (TOPIS)	서울시 주요도로, 도시고속도로	-	〈예측시스템 개발 중〉 -주요 도로에 한정됨 -주요 변수는 속도 ⇨ 교통수요의 변화를 고려하지 않음 -방법론이 검증되지 않은 실정 -본선 위주의 정보 제공 ⇨ IC/JC의 연결로 정보생성 안됨 ⇨ 교차로의 이동류별 정보생성 안됨 -서울시 도로에 한정

- 국외 선진 3국(독일, 영국, 미국)의 교통정보 예보는 다음의 한계를 가짐
  - 독일 OLSIM의 경우, 30분~60분, 1~7일 이후의 시간대별 고속도로 소통상태 예보를 제공함. 그러나 국도 및 시가지 도로의 정보를 제공할 수 없음
  - 국내의 경우, 위의 OLSIM과 같은 시뮬레이션 기법을 도입/적용하기 위해서는 장기적인 연구와 도로망 모니터링 시스템의 확충이 요구됨
  - OLSIM외 타 시스템은 과거 패턴자료와 현행자료를 융합한 모형을 이용하여 2~10시간까지 통행시간 또는 속도 정보를 제공하고 있음. 그러나 장기 교통예보를 위해서는 최소 24시간(1일) 이후의 통행시간 또는 속도를 예측해야 함
  - 사실상, 국내의 경우 ①지자체 ITS 시스템 정보수집, ② 차량 네비게이션 등에서 수집되는 Probe 자료(Probe 건수, 속도), ③한국도로공사의 DSRC, ④도로교통공단의 UTIS 등 교통 정보 수집관련 시스템은 그 어떤 국가에 비하여 높은 수준에 도달해 있음

&lt;표 2-7&gt; 국외 기술의 한계 분석

기관	예보 지역	예보수준	한계 및 문제점
OLSIM (독일)	고속도로	30분~1시간 1~7일	-패턴 O/D와 검지기 정보를 이용 ⇒ 고속도로에 한정 -장기간에 걸친 시뮬레이션 방법론 개발 ⇒ 국내의 경우, 매우 낙후된 기술임 ⇒ 즉각적인 실행의 어려움
Help2Travel (영국)	Midland 지역	1주후, 2주후 주기	행사/여행정보 제공 ※ 도로별 교통 예측정보 제공하지 않음
Traffic England (영국)	주요 도시별 고속도로 및 국도	10시간	-패턴기반 일중 예측 ⇒ 일기반의 예보에 적합하지 않음
INRIX Traffic (미국)	고속도로, 간선도로, 보조간선도로	12시간	-패턴기반 일중 예측 ⇒ 일기반의 예보에 적합하지 않음
Beat the Traffic (미국)	고속도로, 간선도로, 보조간선도로	2시간	-패턴기반 일중 예측 ⇒ 일기반의 예보에 적합하지 않음

#### 나. 국내외 이론의 한계 분석

- 기존 ITS에서 장래 교통소통상태 예측에 관한 연구는 단기에측(Short-term prediction)의 패러다임으로 발전되어 왔음
- 교통류 시스템이 ①Stochastic 시스템, ②Chaos 시스템(Initial deterministic), ③ ①과②가 혼재된 Mixed 시스템 중 어떠한 시스템에 해당되는 지에 대하여 많은 논의가 진행되어 왔으며, 최근 들어 Chaos 시스템 또는 Mixed 시스템에 가깝다는 연구결과가 나오고 있음 (Vlahogianni et al., 2005; Chang et al., 2012a,b, 2014).
- 교통류 시스템의 상태변화는 ①일반적으로 비선형(Nonlinearity)과 비고정성(non-stationarity)을 가지며, ②자유교통류 상태와 혼잡상태에서 비선형과 비고정성은 2~3시간 정도이며, ③통계적으로 시간종속성을 가짐(Vlahogianni et al., 2006).
- 따라서 기존의 확률적인(Stochastic) 상태를 가정하는 대부분의 단기 예측을 위한 통계기반 모형(이동평균, Smoothing, 선형회귀, ARIMA, Local Linear Regression, Kalman Filtering, State Space, Wavelet, Support Vector 등)의 예측력은 비선형 모형(NN, Neural Network 등), 비모수회귀모형(NPR, Non-Parametric Regression), NN 또는 NPR 기반의 Hybrid 모형 등에 비하여 낮음. 그 이유는 기존의 통계기반 모형이 상태의 진화방향이 변화하는 시점에서 상태의 진행방향과 정도(Directionality and Variation)를 고려할 수 없기 때문임
- 즉, 현재의 단기 예측에 대한 시간적 예측정도는 해당 시간대 단위(1~60분)로 과거 1~6개 이전 시간대를 이용한 예측의 시간적 범위는 현재시각 기준 3~60분 후 정도로 보고되고 있음 (Chang et al., 2012)
- **[한계]** 대부분의 선행 연구가 현행자료(Current data)를 이용한 링크 또는 지점의 교통류 상태변수를 예측하기 때문에 (ITS 예측 분야의 관점에서)긴 장래의 상태를 설명할 수 없음
- 또한, 단기 예측에서 일변량(univariate) 기반의 짧은 현행 상태의 변화는 중·장기의 장래 상태를 고려할 수 없으며, 예측의 시간적 범위를 확장하기 위하여 시간대의 길이를 확장한다고 하더라도 장래 상태의 불확실성은 감소하지 않음(Chang et al., 2010)
- 따라서 대부분 선행 연구의 예측시간 범위는 본 연구의 익일 교통예보에 적합하지 않음
- **[시사점]** 본 연구는 기존의 연구와는 달리 장래 교통상태 예측을 위하여 ①Network기반의 상태벡터 정의와 ②Big data기반의 지식 발견(Knowledge discovery) 접근법 그리고 ③교통 수요를 고려한 Network상에서 이동의 변화를 고려해야 할 것으로 판단됨

#### 다. 개선방향 설정

- 국내외 기술현황과 이론의 고찰과 한계, 그리고 시사점 분석결과, 기존의 단기 예측 (Short-term prediction)방법을 통해서는 본 연구의 최종목표인 익일 교통소통상태 예보를 위한 장래 상태를 수용할 만한 예측력을 가지기는 어려울 것으로 판단됨
- 이는 기존의 단기에측 방법은 시스템의 낮은 성능으로 이하여 단기적인 예측 전략에 적합하게 개발되었으며, 장기 교통예보를 위한 장래의 교통상태 예측을 수행하지 않았기 때문임. 즉, 개발모형의 목표와 용도가 다름
- 그러나 최근의 급속하게 발전된 Database처리기술과 연산능력, 그리고 검색엔진의 성능을 고려할 때, 이론적인 접근법은 Data기반의 접근법이 가장 합리적이라고 판단됨. 그 이유는 예보수준에서 장래 교통상태에 대한 불확실성을 극복하기 위한 적절한 방법론이 제시되고 있지 않기 때문임
- 따라서 기존의 단기에측모형은 본 연구에 적합하지 않으며, 앞서 설명한 한계 및 시사점에 근거하여 본 연구의 모형 개발방향을 설정하면 다음과 같음
  - 첫째, 중·장기 상태의 변화를 잘 설명하기 위해서는 Link기반의 공간 정의를 Network기반으로 정의해야 함
  - 둘째, Network의 소통상태 변화는 Link의 국부적인 변동보다는 순환성(Recurrency)과 경향성(Pattern), 그리고 내재된 알려지지 않은 Network의 상태 진화에 대한 교통 현상 (Traffic Behavior)인 속도와 수요(Probe 건수)의 변화 추세에 의하여 설명되어야 함
  - 셋째, Network의 소통 상태를 설명하기 위해서는 Network라는 유기체를 구성하는 개별 Link의 시간대별 변화 속성을 기반으로 Network 상태를 고려해야 함
  - 넷째, 인위적인 통계적 모델링 보다는 데이터 기반의 지식 발견(Knowledge discovery) 접근법이 본 과제의 현실적인 목표달성에 타당할 것으로 판단됨

## 제1절 교통예보를 위한 기초자료 구축

### 1. 구축 개요

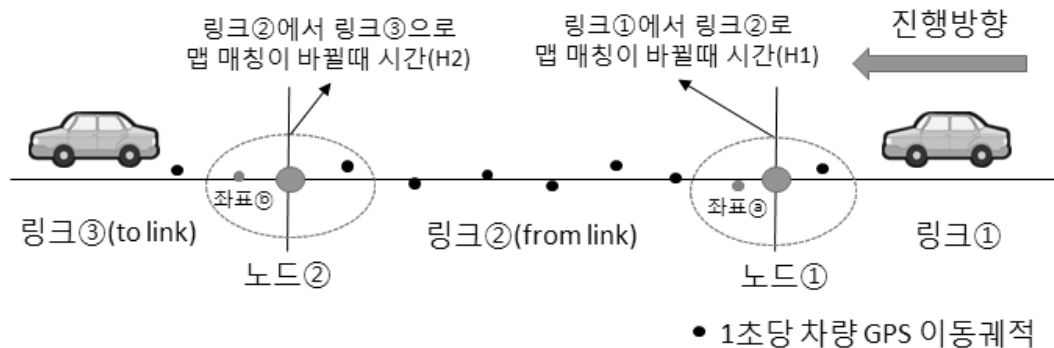
- 최근 IT기술의 급속한 발전에 따라 Car Navigation, 스마트폰 등 첨단기기의 보급이 보편화 되었으며, 이를 통하여 과거 수집하지 못했던 다양한 정보들을 수집할 수 있게 되었음. 즉, 개인 또는 개별 차량의 이동궤적을 확인할 수 있는 시/공간상의 연속적 동적 자료<sup>3)</sup>를 수집할 수 있음
- 이러한 자료는 과거 1세대와 2세대의 정적인 지점자료의 형태가 아닌 개인이 이동하면서 수집되는 동적인 3세대 자료라고 볼 수 있음
- 과거 2013년 한국교통연구원에서 수행한 “빅데이터를 이용한 교통정책 개발 및 활용성 증대 방안”에서는 3세대 자료는 차량의 연속적 이동을 통해 자동적으로 자료가 수집되기 때문에 링크와 링크 간의 시/공간적인 연결성을 통해서 링크별 통행속도, 이동경로 등 다양한 정보를 활용할 수 있는 측면에서 그 가치가 매우 높다고 설명하고 있음
- 따라서, 본 연구에서는 개별 차량의 이동궤적을 확인할 수 있는 Car Navigation 자료를 장래 교통예보를 위한 기초자료로서 구축하였으며, 이 자료는 국토교통부에서 제공하고 있는 ITS 표준노드링크 체계의 링크 단위로 개별 차량의 링크별 통행속도 정보를 수집하고 있음
- 여기서, 먼저 ITS 표준노드링크 기반의 개별 Car Navigation 자료 생성과정과 그 구축 결과를 제시하였으며, 기존 ITS 표준노드링크의 문제점을 보완하기 위하여 기 구축한 분석 맵의 링크 단위 기초자료 집계 방법과 그 결과를 제시하였음
- 마지막으로 구축된 Car Navigation 자료에 대한 기초자료 특성을 분석하여 해당 자료의 우수성을 제시하였으며, 본 장의 주요 내용은 아래와 같음
  - (1) ITS 표준노드링크 기반의 개별 Car Navigation 자료 구축
  - (2) 분석 맵 단위의 자료 집계
  - (3) 기초자료 특성분석

<sup>3)</sup> Dixon, M. P. and L. R. Rillet(2002), “Real-time OD estimation using automatic vehicle identification and traffic count data”, Journal of Computer-Aided Civil and Infrastructure Engineering, pp. 7-21.

## 2. ITS 표준노드링크 기반의 개별 Car Navigation 자료 구축

### 가. Car Navigation 자료 수집

- 본 연구에서 이용하고 있는 Car Navigation 자료는 개별 차량의 Navigation 단말기에서 수집되는 차량의 이동궤적 정보로서 국토교통부에서 제공하고 있는 ITS 표준노드링크 체계의 링크 단위로 개별 Probe 차량의 이동궤적을 맵 매칭하여 각 차량에 대한 링크별 Probe건수와 통행속도 정보를 수집하고 있음
- 여기서, 개별 차량의 이동궤적 정보로부터 링크별 Probe건수와 속도 정보로 데이터를 생성하는 과정은 <그림 2-1>과 같음



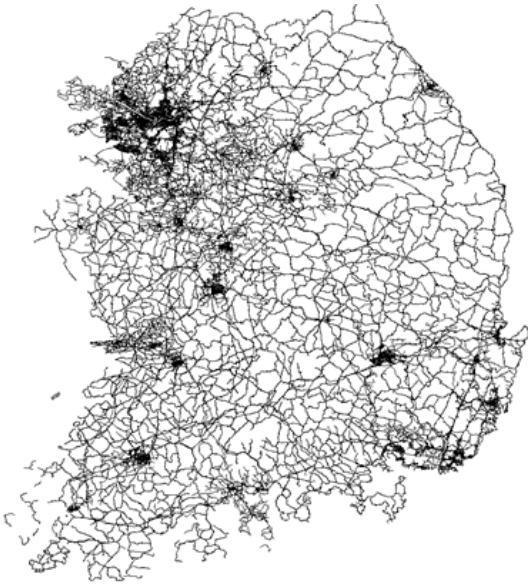
- 링크 ②의 Probe건수(건) = 링크 ②에 맵 매칭이 이루어지는 Probe 차량수
- 링크 ②의 속도(km/h) = (좌표㉑와 좌표㉒의 거리) / (H2 - H1)

<그림 2-17> 링크별 Probe건수와 속도 정보생성 개념

- 먼저, 개별 차량의 이동궤적 정보를 맵 매칭하기 위하여 이용되고 있는 ITS 표준노드링크에 대한 설명은 아래와 같음
  - ITS 표준노드링크는 교통정보를 수집 및 제공하기 위하여 교통체계 지능화 사업자가 구축하는 전자교통지도<sup>4)</sup>로서 민간에게 공개되고 있음
  - 노드(Node)는 도로법에 의한 도로로서 교차로, 분기점, 터널입구 등 교통흐름이 변경되는 지점을 전자적으로 형상화한 것을 말하며, 링크(Link)는 위의 노드와 노드를 도로선 형에 따라 연결하는 도로구간을 전자적으로 형상화한 것을 말함
  - 표준 노드/링크체계는 노드정보와 회전정보, 링크정보, 링크부가정보로 구성하며, 각각은

<sup>4)</sup> 국토해양부, “지능형교통체계 표준 노드·링크 구축관리지침”, 2008.

### 별도의 자료구조를 가짐



- 2014년 3월 국토교통부 배포자료 이용
- 총 링크개수 : 256,160개
- 총 노드개수 : 98,729개
- 표준 노드/링크 구성 체계
- 노드정보/회전정보/링크정보/링크부가정보
- 도로등급(ROAD\_RANK) : 8개 Type
- 고속국도(101)
- 도시고속국도(102)
- 일반국도(103)
- 특별/광역시도(104)
- 국가지원지방도(105)
- 지방도(106)
- 시/군도(107)
- 기타(108)

<그림 2-18> ITS 표준노드링크 자료 설명

- 본 연구의 Car Navigation 자료는 앞서 설명한 ITS 표준링크를 기반으로 일 단위의 수집시간 순으로 수집되며, 수집 자료의 저장 형태는 시/공간적 범위에 따라 아래와 같이 총 5가지 Column 형태로 수집일자별 텍스트로 저장됨
- 수집 Column(총 5개) : ①auth\_key/②link\_time/③from\_link/④to\_link/⑤speed
- 공간적 범위 : 전국
- 데이터 수집기간 : 2013년 10월 1일 ~ 2014년 1월 31일, 4개월치 이력자료

<표 2-8> Car Navigation 수집자료 테이블 정의서

Column	설명	Data Type	PK	FK
auth_key	단말기 ID	Datetime	-	-
link_time	수집 시간 (년/월/일/시/분/초)	Integer	-	-
from_link	진입 링크번호	Integer	-	-
to_link	진출 링크번호	Integer	-	-
speed	속도(km/h)	Integer	-	-



## ■ 2013년 10월 1일

1	4527234	2013-10-01	00:30:01	2220033204	2220226700	45
2	4527234	2013-10-01	00:30:01	2220226700	2220226600	48
3	4527234	2013-10-01	00:30:09	2220226600	2220226400	40
4	4527234	2013-10-01	00:31:54	2220226400	2220228000	56
5	4527234	2013-10-01	00:32:22	2220228000	2220227700	57
6	4527234	2013-10-01	00:38:54	2220018201	2210009900	90
7	4527234	2013-10-01	00:38:54	2210009900	2210010500	90
8	4527234	2013-10-01	00:39:03	2210010500	2210012000	82
9	4527234	2013-10-01	00:39:39	2210012000	2210012200	85
10	4527234	2013-10-01	00:40:21	2210012200	2210012600	87
11	4527234	2013-10-01	00:41:05	2210012600	2210013000	88
12	4527234	2013-10-01	00:41:55	2210013000	2210011100	83
13	4527234	2013-10-01	00:41:45	2210011100	2210040400	27
14	4527234	2013-10-01	00:41:45	2210040400	2210040500	36
15	4527234	2013-10-01	00:43:51	2210040500	1100012903	61
16	141669	2013-10-01	00:01:20	2220003701	2220003702	85
17	141669	2013-10-01	00:02:51	2220003702	2220003703	87
18	141669	2013-10-01	00:04:29	2220003703	2220003704	84
19	141669	2013-10-01	00:06:12	2220003704	2220004500	85
20	141669	2013-10-01	00:08:05	2220004500	2220005100	84
21	141669	2013-10-01	00:00:38	2220005100	2220005400	84
22	141669	2013-10-01	00:00:38	2220005400	2220005600	74
23	141669	2013-10-01	00:01:03	2220005600	2220006001	23
24	141669	2013-10-01	00:01:36	2220006001	2220006002	37
25	141669	2013-10-01	00:03:23	2220006002	2220006200	69
26	141669	2013-10-01	00:01:49	2220006200	2220006801	74
27	141669	2013-10-01	00:04:01	2220006801	2220006802	78
28	141669	2013-10-01	00:04:01	2220006802	2220006806	84
29	141669	2013-10-01	00:05:22	2220006806	2220008600	81
30	141669	2013-10-01	00:07:10	2220008600	2220008800	15
31	141669	2013-10-01	00:05:29	2220008800	2220009201	33
32	141669	2013-10-01	00:03:56	2220009201	2220009202	47
33	141669	2013-10-01	00:07:49	2220009202	2220012601	62
34	141669	2013-10-01	00:07:49	2220012601	2220012602	58
35	141669	2013-10-01	00:08:16	2220012602	2220012607	57
36	141669	2013-10-01	00:10:01	2220012608	2220012604	52

## ■ 2014년 1월 31일

1	2720812	2014-01-31	00:15:19	2850026304	2850121900	8
2	2720812	2014-01-31	00:15:19	2850121900	2850035401	50
3	2720812	2014-01-31	00:15:54	2850035401	2850539100	46
4	1897915	2014-01-31	00:00:37	2140191200	2140188300	75
5	1897915	2014-01-31	00:04:40	2140188300	2333264200	85
6	1897915	2014-01-31	00:04:40	2333264200	2333264100	87
7	1897915	2014-01-31	00:07:55	2333264100	2230077800	84
8	1897915	2014-01-31	00:07:55	2230077800	2230078900	59
9	1897915	2014-01-31	00:11:01	2230078900	2230078800	82
10	473418	2014-01-31	00:04:20	1320027200	1320019100	59
11	473418	2014-01-31	00:06:20	1320019100	1320020400	61
12	473418	2014-01-31	00:04:21	1320020400	1320020000	50
13	473418	2014-01-31	00:02:32	1320020000	1360027800	51
14	473418	2014-01-31	00:00:48	1360027800	1360027400	52
15	473418	2014-01-31	00:01:38	1360021600	1360021300	65
16	473418	2014-01-31	00:01:38	1360021300	1360021700	43
17	473418	2014-01-31	00:02:33	1360021700	1360022100	48
18	473418	2014-01-31	00:03:49	1360022100	1360022500	55
19	473418	2014-01-31	00:05:18	1360022500	1360022900	54
20	473418	2014-01-31	00:07:03	1360022900	1360023500	56
21	473418	2014-01-31	00:09:00	1360023500	1360028100	50
22	473418	2014-01-31	00:04:26	1360028100	1360028500	46
23	473418	2014-01-31	00:02:30	1430005900	1430006501	52
24	473418	2014-01-31	00:07:16	1430006501	1430006502	47
25	473418	2014-01-31	00:07:16	1430006502	1430008500	17
26	1627282	2014-01-31	00:36:56	1690003802	1690003801	52
27	1627282	2014-01-31	00:36:56	1690003801	1690003400	72
28	1627282	2014-01-31	00:38:55	1690011300	1690011300	82
29	1627282	2014-01-31	00:38:55	1690011300	1690002400	73
30	1627282	2014-01-31	00:39:16	1690002400	1690002202	48
31	1627282	2014-01-31	00:42:08	1690002202	1690002201	70
32	1627282	2014-01-31	00:42:08	1690002201	1690002000	99
33	1627282	2014-01-31	00:42:13	1690002000	1690001800	99
34	83007	2014-01-31	00:00:14	1160009900	1160009902	57
35	83007	2014-01-31	00:00:06	2100086000	2110018700	12
36	83007	2014-01-31	00:02:41	2100085900	1160008802	44

&lt;그림 2-19&gt; 수집자료의 텍스트 저장 형태

## 나. 수집자료 구축

- 본 연구의 Car Navigation 자료는 진입링크, 진출링크, 수집시간, 속도 등으로 구성되며, 수집된 자료는 1일 단위로 취합하여 Text 파일로 저장되고 있음
- 지난 2013년 한국교통연구원에서 수행한 “Car Navigation 자료를 이용한 교통혼잡지도 연구”에서는 1개월치(2013년 9월)의 Car Navigation 자료를 이용하였으나, 본 연구에서는 수집 대상(지역, 도로등급)을 점차적으로 확대하여 총 4개월치(2013년 10월 ~ 2014년 1월)의 Car Navigation 자료를 이용할 수 있게 되었음
- 먼저, ITS 표준노드링크 기반으로 수집된 Car Navigation 자료를 분석한 결과 아래와 같은 몇 가지의 문제점이 발생하였음
  - 2014년 3월 기준의 ITS 표준노드링크와 매칭한 결과, 일부 수집되지 않은 도로구간이 발생함

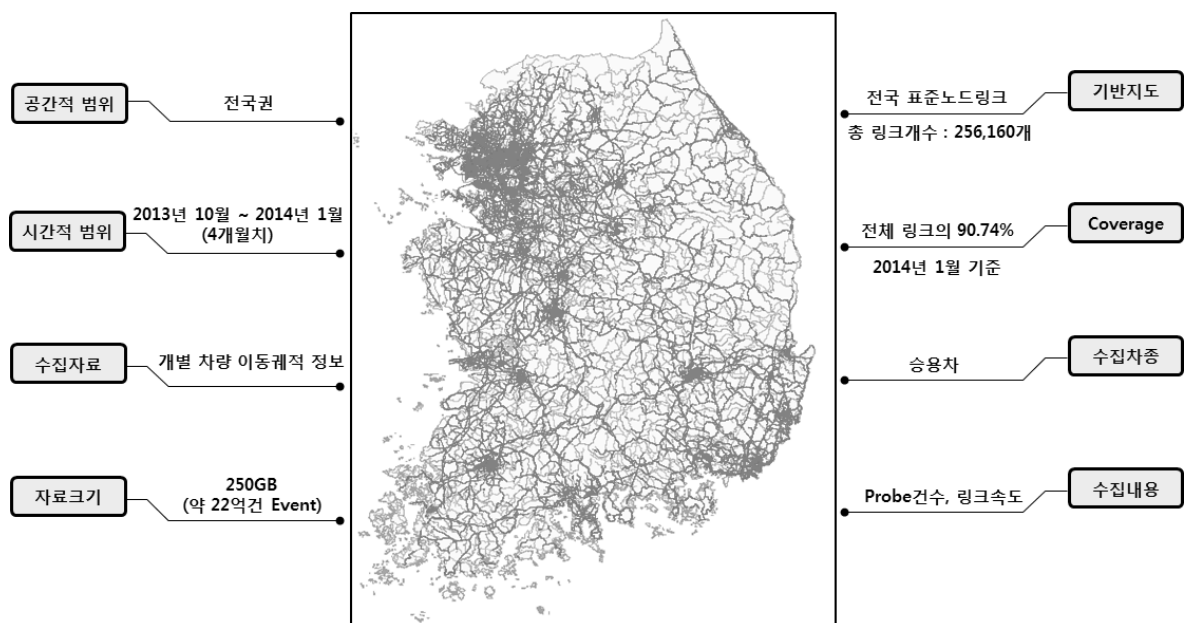
- 교차로 구간에서 차량의 우회전 또는 좌회전 시 진행 방향의 링크가 수집되지 않고 직진방향의 링크가 수집되는 경우가 발생함
- 수집시간 정보로 개별 차량의 주행 경로를 표출 시, 일부 차량의 주행 경로의 앞뒤가 엇갈리게 표출되며, 수집시간으로 데이터를 정렬하지 않고 데이터의 입력 순서로 정렬 시 차량의 주행경로가 정상적으로 표출됨
- 일부 Car Navigation 단말기 ID 중복이 발생함
- 이러한 문제점들을 보완하기 위하여 2014년 한국교통연구원에서 수행 중인 “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”에서는 ITS 표준노드링크 단위로 수집된 자료에 대한 보정을 아래와 같이 기 수행하였음
  - 경로 보정을 위한 ITS 표준노드링크 오류 검증 및 수정
    - 링크와 연결성이 없는 고립 노드, 노드 중복, 링크의 속성정보와 실제 노드와의 연결성 체크, 다른 링크와의 연결성이 없는 고립 링크 체크, 링크의 방향성 등에 대하여 ITS 표준노드링크 수정
  - 경로 탐색을 이용한 개별 Probe 경로 보정
    - 수집데이터는 단말기 별 데이터를 확인하여 이벤트 개수가 특정개수 미만(집계시간 단위 당 20건 미만)이면, 데이터를 필터링하여 불안정한 데이터 제거
    - 동일 단말기 내에서 이벤트 수집시간 간격이 30분 이상 발생 시에는 별도의 그룹으로 데이터 분류
    - 동일 단말기 내에서 이벤트의 누락 구간의 거리의 차가 최대 50Km이상 발생 시에는 별도의 그룹으로 데이터 분류
    - 별도의 그룹으로 분류된 데이터는 데이터 로딩 순서대로 경로탐색을 반복 수행하여 데이터의 연결성 검증 및 데이터 보정
    - 경로탐색 방법은 링크의 from\_node\_id(시작노드)와 다음 진행 링크의 to\_node\_id(종료노드)를 기준으로 경로 탐색하여 해당구간의 링크 검색
    - 검색한 링크 리스트 중에서 새로운 링크를 검색 시에는 이를 그룹 내의 링크로 추가하여 누락구간 부분을 보정
    - 검색한 링크 리스트 중에서 그룹 내의 링크가 해당하지 않으면 데이터 제거
- 수집자료에 대한 보정전, 후 자료의 분석 결과는 아래와 같으며, 2014년 1월 기준 데이터 Coverage가 당초 84.41%에서 보정 후 90.74%로 향상된 것으로 분석되었음
- 또한, 2014년 1월 기준 매칭 링크 당 1일 평균 이벤트 건수(건/링크)도 당초 46건/링크에서 보정 후 59건/링크로 향상된 것으로 분석되었음

&lt;표 2-9&gt; 수집자료에 대한 보정전, 후 데이터 분석 결과 비교

구분		2013년			2014년
		10월	11월	12월	1월
1일 평균 운행 대수(대)		184,709	180,452	162,959	143,397
보정 전	이벤트 건수(건)	440,395,473	445,596,480	401,343,793	310,960,262
	Link 매칭정보(전체 256,160개)	155,830	186,144	215,122	216,236
	Coverage	60.83%	72.09%	83.98%	84.41%
	차량 1대당 1일 평균 이벤트 건수(건/대)	77	82	79	70
	링크 당 1일 평균 이벤트 건수(건/링크)	91	80	60	46
보정 후	이벤트 건수(건)	609,807,948	618,102,595	550,804,284	422,274,073
	Link 매칭정보(전체 256,160개)	188,689	207,951	231,514	232,428
	Coverage	73.66%	81.18%	90.38%	90.74%
	차량 1대당 일 평균 이벤트 건수(건/대)	106	114	109	95
	링크 당 1일 평균 이벤트 건수(건/링크)	104	99	77	59

자료 : “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”, 한국교통연구원, 2014

- 따라서, 본 연구에서는 기존 연구에서 데이터 보정이 완료된 Car Navigation 자료를 이용하여, 해당 자료에 대한 전체 구축 현황은 아래와 같음

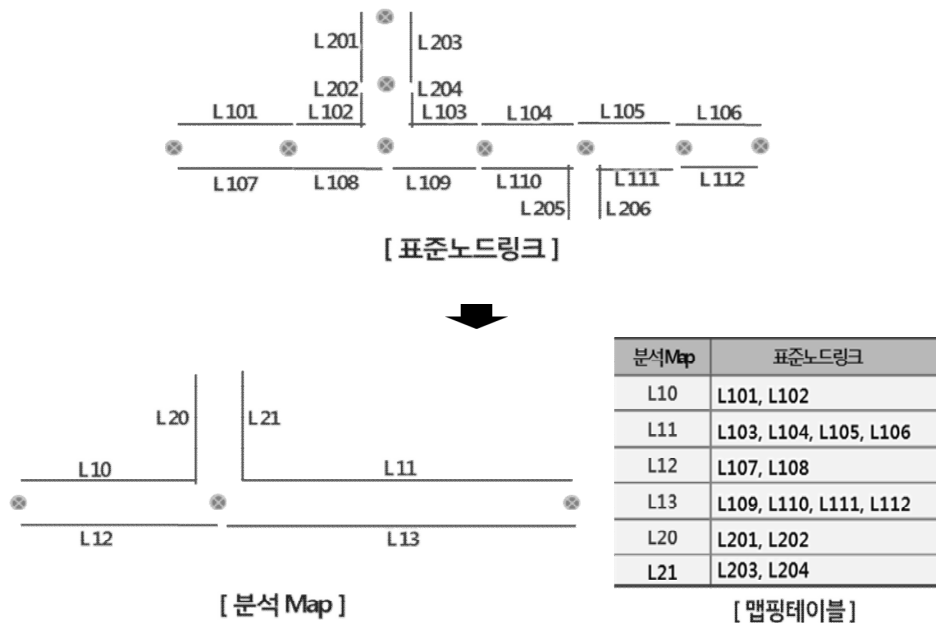


&lt;그림 2-20&gt; 본 연구의 Car Navigation 자료구축 현황

### 3. 분석맵 단위의 자료 집계

#### 가. 분석 맵 단위의 집계 구조

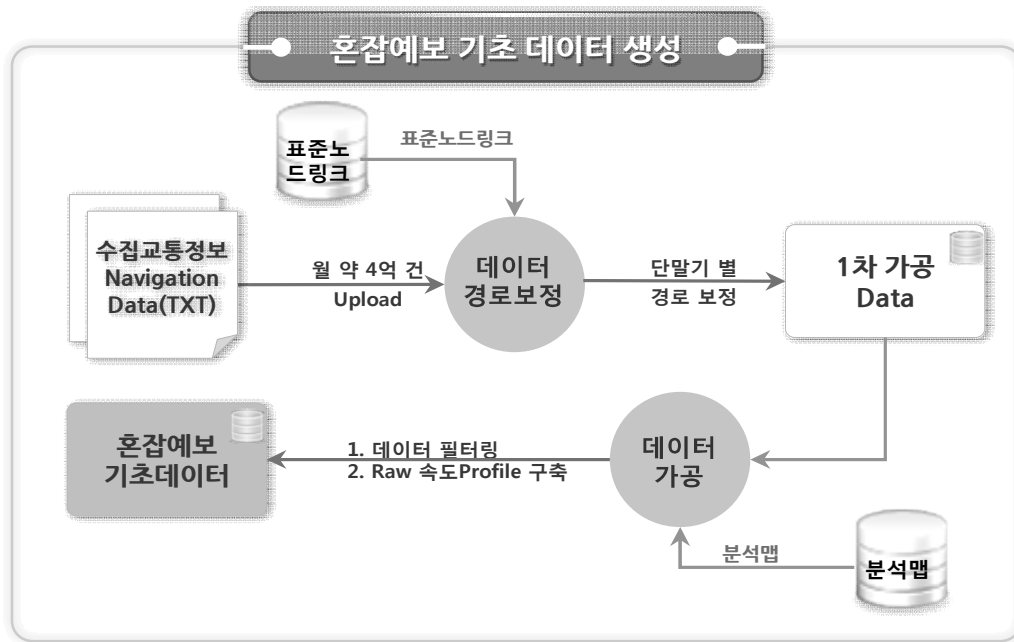
- ITS 표준노드링크에서는 동일한 구간이 교차로, 속성변환지점(속도, 차선수 등), 행정경계 등으로 인하여 상세하게 링크로 분할되어 있어 향후 예측정보 표출 시 동일 도로구간의 소통정보(정상 ↔ 혼잡)를 이용자들에게 효율적으로 제공하기 어려움
- 따라서, 동일 도로구간에 대한 소통정보를 이용자들이 모두 인지하기 쉽도록 동일 도로구간에 여러 개로 분할되어 있는 링크를 병합하여 장래 소통정보를 한눈에 파악 할 수 있도록 하는 분석 맵의 구축이 필요함
- 이러한 분석 맵은 2014년 한국교통연구원이 수행한 “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”에서 기 구축이 되어 있으며, 본 연구의 분석 맵은 기 구축되어 있는 전국 교통혼잡지도의 분석 맵을 이용하였음



<그림 2-21> 분석 맵의 구축 개념도

자료 : “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”, 한국교통연구원, 2014

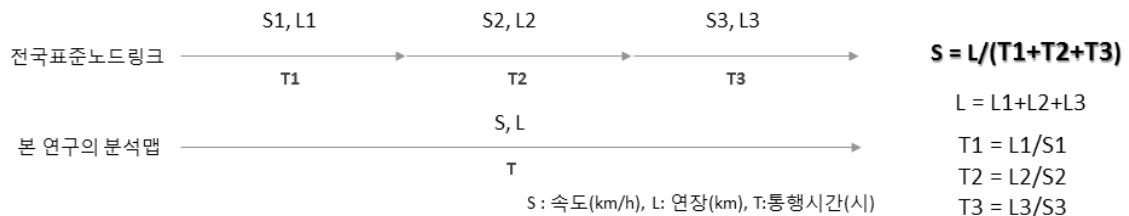
- 이러한 분석 맵 기준 Car Navigation 자료의 생성과정은 ①교통예보 Input 데이터 생성 선택, ②내비게이션 수집데이터 폴더 선택, ③실행, ④분석맵 기준의 개별 Car Navigation 자료 생성의 4단계에 걸쳐 수행하도록 개발하였음



<그림 2-22> 분석 맵 단위의 Car Navigation 자료 구축 프로세스

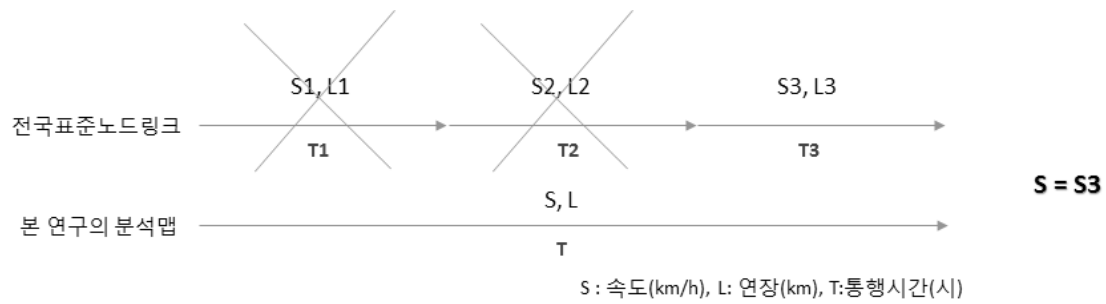
#### 나. 분석 맵 단위의 집계 방법

- 장래 구간 통행속도 예측을 위하여 이용되는 개별 Car Navigation 자료는 ITS 표준노드링크 기반으로 수집 및 저장되어 있어 본 연구의 분석 맵을 적용할 경우 해당 자료를 분석 맵의 링크 단위로 집계해야 함
- 이를 위하여 기존 전국 표준링크에서 분석 맵 단위의 링크로 병합되는 상황을 3가지의 CASE로 구분하고 각 CASE별 자료 집계 기준을 정립하여 적용하였음
- 여기서, 각 Case별 자료 집계 시 해당 링크별 수집시간에 따른 개별 차량 순으로 데이터를 집계함
  - CASE 1 : 기존 링크별 데이터가 모두 수집되어 있는 경우



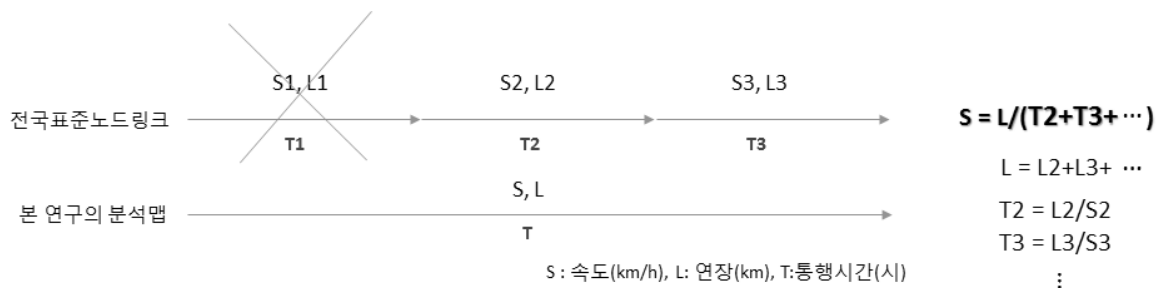
<그림 2-23> 분석 맵 단위의 자료 집계 기준(CASE 1)

- CASE 2 : 기존 링크별 데이터 중 하나만 수집되어 있는 경우



<그림 2-24> 분석 맵 단위의 자료 집계 기준(CASE 2)

- CASE 3 : 기존 링크별 데이터 중 두개 이상 수집되지 않은 경우



<그림 2-25> 분석 맵 단위의 자료 집계 기준(CASE 3)

다. 분석 맵 단위의 자료구축 현황

- 현재 전국 도로등급별 분석 맵 구축현황은 아래의 표와 같으며, 기존의 링크 총 254,896개에서 현재 총 53,541개로 즉, 링크 당 평균 약 4.8개를 병합하여 구축하였음

<표 2-10> 분석 맵의 도로등급 별 구축 현황

도로 등급	표준노드링크 개수 (전체 : 254,896개)	분석 맵의 링크 개수 (전체 : 53,541개)	비고
고속국도	6,296개	5,564개	- ITS 표준노드링크의 기타도로(108)은 단지 내 도로로 분석 맵 구축에서 제외하였음
도시고속국도	2,203개	1,375개	
일반국도	44,665개	12,563개	
국가지원지방도	9,963개	2,762개	
지방도	31,231개	6,315개	
특별·광역시도	35,548개	6,961개	
사군도	122,422개	1,8001개	

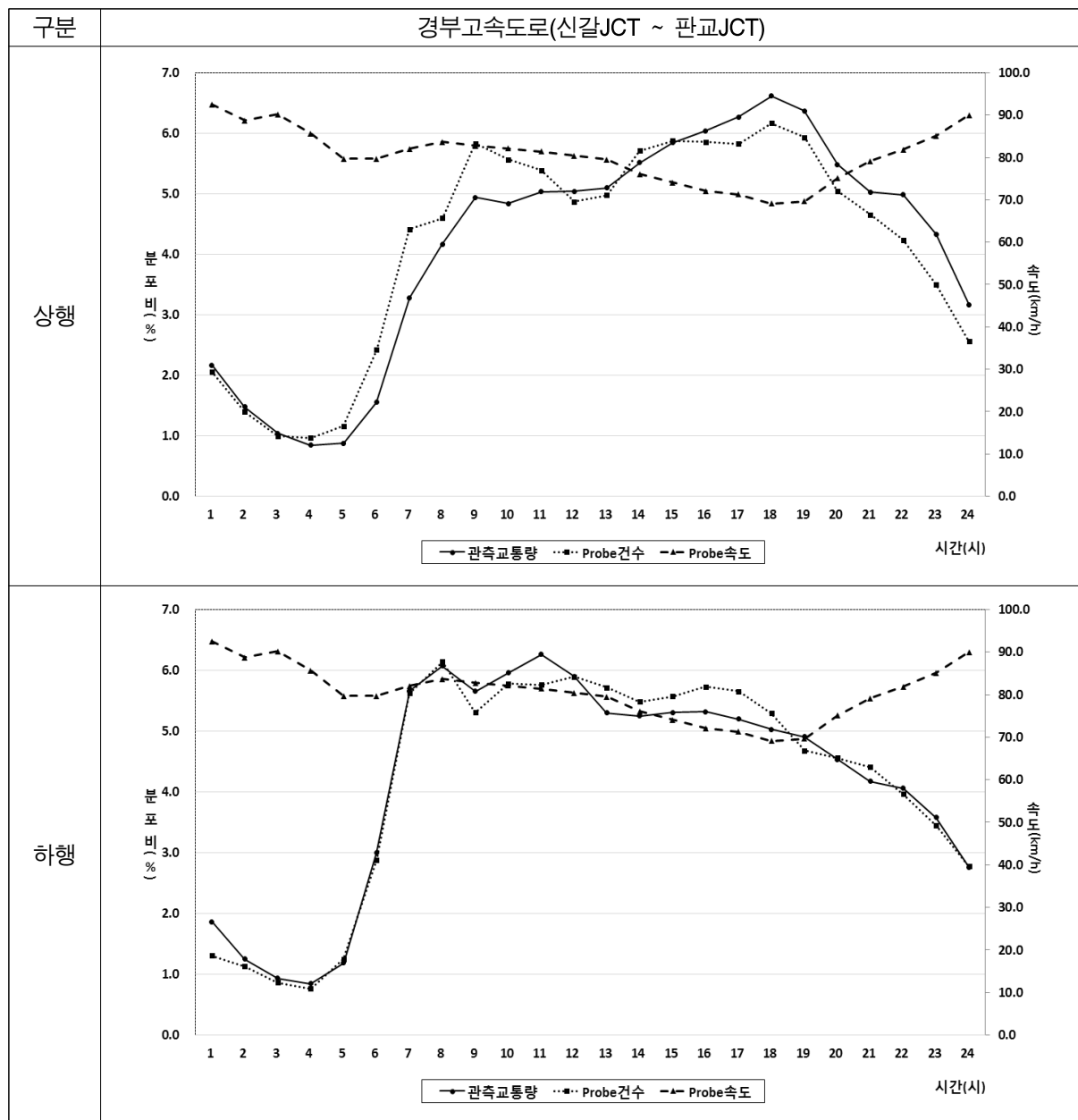
자료 : “Car Navigation 자료를 이용한 교통혼잡지도 검증 및 고도화 연구”, 한국교통연구원, 2014

#### 4. 기초자료 특성 분석

- 본 연구에서 구축한 Car Navigation 자료를 적용하기 전에 해당 자료에 대한 특성 분석이 선행되어야 하므로 장래 교통예보를 위한 기초자료인 Car Navigation 자료의 Probe 건수와 속도에 대한 특성을 분석하였음
- 즉, Car Navigation 자료를 이용하기 위해서는 해당 구간의 시간대별 Probe건수와 속도가 실제 시간대별 관측교통량 흐름을 반영하고 있어야 함
- 따라서 본 연구에서는 Car Navigation 자료의 Probe건수와 속도의 시간대별 분포가 실제 관측교통량의 시간대별 분포와 유사한 분포를 가지는 지에 대하여 구간별 관측교통량, Probe 건수, 속도의 분포 그래프를 통하여 분석하였음. 또한, 분석 구간은 주요 노선 중 연속류와 단속류 구간으로 구분하여 해당 자료의 특성을 분석하였음

##### 가. 연속류 구간

- 연속류 구간은 경부 고속도로의 신갈JCT ~ 판교JCT 구간에 대하여 방향별 평일 관측교통량 기준으로 해당 구간의 Probe건수와 속도 자료의 특성을 분석하였음
- 시간대별 관측교통량과 Probe건수의 총량적인 측면에서는 당연히 차이가 발생하기 때문에 이를 시간대별 분포비(%)로 살펴보면, 전체적으로 Probe건수의 시간대별 분포가 관측교통량과 유사하게 나타나는 것으로 분석되었음. 이는 Car Navigation 자료가 실제 현장의 교통흐름을 적절히 나타내고 있는 것으로 판단됨
- 시간대별 Probe속도의 경우는 교통량이 많지 않은 새벽 또는 심야 시간대의 Probe속도가 하루 중 가장 높다가 교통량이 점차 증가하면서 교통혼잡이 발생하여 Probe속도가 낮아지는 것을 볼 수 있음
- 또한, 관측교통량이 일정 수준 이상으로 유지되고 있는 상황(오전10시 ~ 저녁8시)에서는 Car Navigation의 Probe속도 분포도 일정한 범위 내에서 유지되고 있는 것으로 분석되었음
- 즉, 연속류의 실제 통행속도는 해당 구간의 교통수요의 변화와 기하구조 등과 같은 물리적인 제약에 의해 크게 영향을 받으므로 실제 해당 시간대에 그 구간을 주행하였던 Probe차량들의 통행속도 자료를 이용하는 것은 바람직한 것으로 판단됨



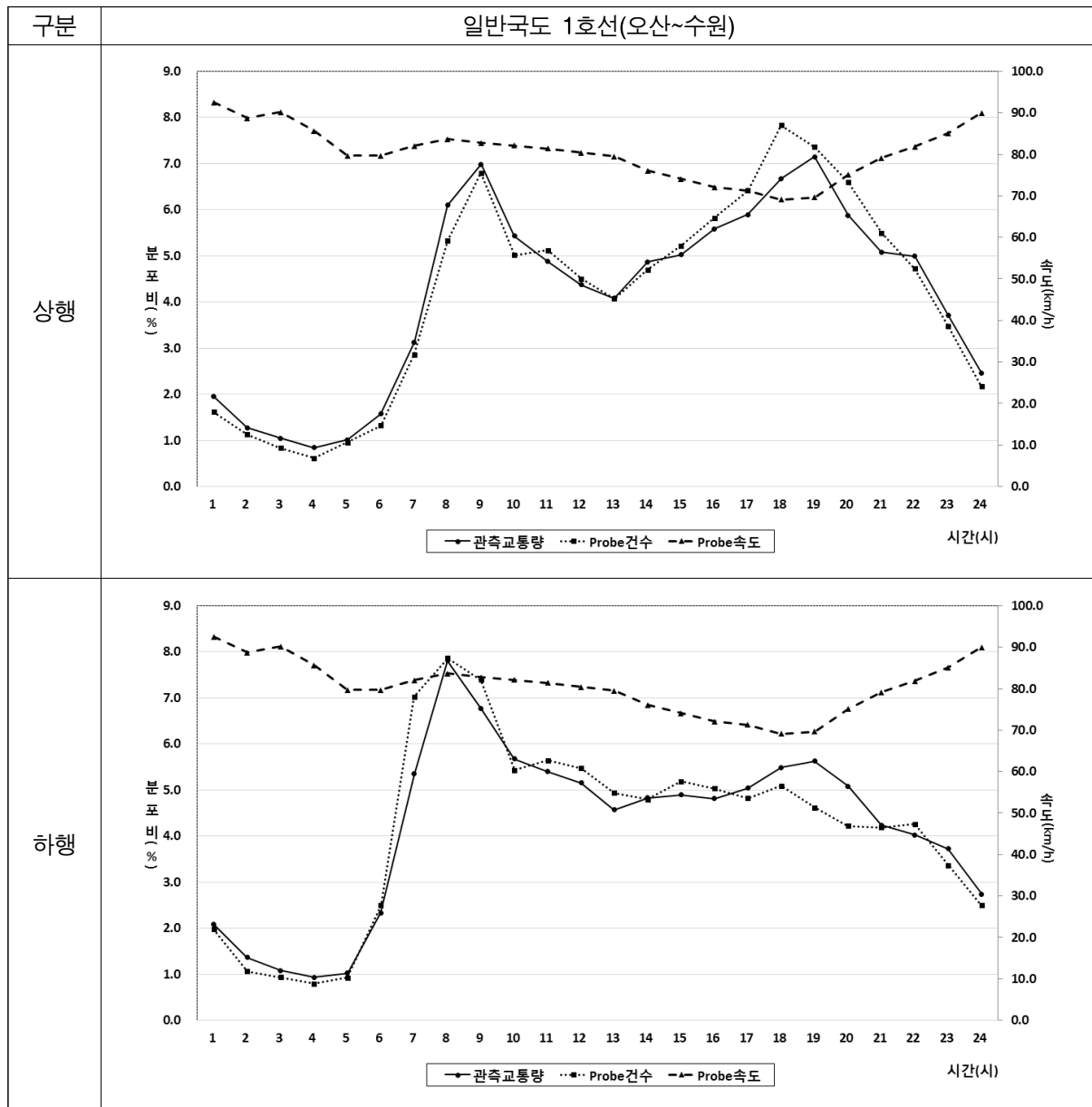
<그림 2-26> 연속류 구간의 기초자료 특성 분석

#### 나. 단속류 구간

- 단속류 구간은 도심지역을 통과하는 일반국도 1호선의 오산~수원 구간에 대하여 방향별 평일 관측교통량 기준으로 해당 구간의 Probe건수와 속도 자료의 특성을 분석하였음
- 먼저, 시간대별 관측교통량과 Probe건수 분포는 전체적으로 유사하게 나타나는 것으로 분석되었으며, 이는 Car Navigation 자료가 실제 현장의 교통흐름을 적절히 나타내고 있는 것으로 판단됨
- 시간대별 Probe속도의 경우는 교통량이 많지 않은 새벽 또는 심야 시간대의 Probe속도가 하



루 중 가장 높다가 교통량이 점차 증가하면서 교통혼잡이 발생하여 Probe속도가 낮아지는 것을 볼 수 있음. 즉, 단속류의 실제 통행속도는 해당 구간의 교통수요 변화 뿐만아니라 신호기, 기하구조 등과 같은 물리적인 제약에 의해 더 크게 영향을 받으므로 실제 해당 시간대에 그 구간을 주행하였던 Probe차량들의 통행속도 자료를 이용하는 것은 바람직한 것으로 판단됨



<그림 2-27> 단속류 구간의 기초자료 특성 분석

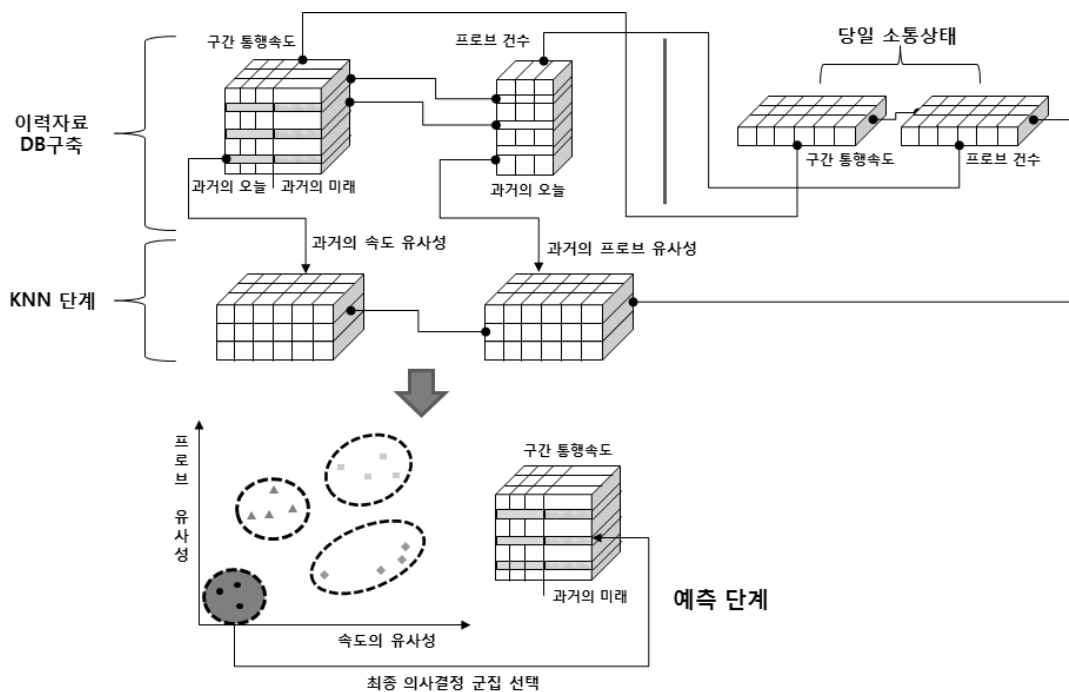
## 제4절 교통예보를 위한 소통상태 예측 핵심 방법론 개발

### 1. 개발개요

- 앞서 언급한 바와 같이 현대의 교통정보 수집원, 그리고 자료의 질과 양은 기존의 단일 ITS 시스템에서 수집/관리하던 정보와는 매우 다른 형태를 가짐
- Smart phone 내비게이션 App과 차량 내비게이션 단말기를 통하여 제공되는 교통소통정보는 IT의 기술 발달과 더불어 높은 정확도를 보유하고 있으며, 수집되는 정보의 양과 공간적 규모도 국내에서 가장 큰 교통 Big data라고 표현할 수밖에 없음. 이와 같이 가용자료의 양과 정확도가 급변하는 시점에서 장래 교통상태의 예측방법론도 이에 부합하게 변화해야 함
- 따라서, 본 연구에서는 기존 단기 예측기법의 한계를 극복함으로서 익일 이상의 보다 먼 장래 통행속도를 수용 가능한 오차의 범위 내에서 예측하고, 향후 교통예보 시스템에 탑재가 용이하고 교통류의 행태와 예측 모델링에 대한 전문지식이 없이도 운영이 가능한 데이터 기반의 예측방법론을 개발하고자 함
- 이러한 예측모형의 개발은 모형에 데이터를 맞추는 기존의 방식을 벗어나 모형이 데이터를 받아들이고 그 데이터를 이용하여 장래를 예측하는 자료 순응형 모형개발을 의미함
- 먼저, 본 연구에서 정의하는 빅데이터는 3장에서 설명한 개별 Car Navigation 단말기에서 수집되는 이동계적 자료(Probe건수, 속도)를 말하며, 교통예보는 당일 현재 시각 기준에서 익일의 링크별/시간대별 통행속도를 예측하여 각 예측 값에 해당하는 소통상태(정상↔혼잡)정보를 제공하는 것을 말함
- 본 연구의 장래 소통상태 예측 방법론은 아래와 같이 크게 2가지의 핵심 요소기술로 구분하고 각 요소기술별로 세부적인 개발모형들을 포함하고 있음
  - 핵심 요소기술① : Car Navigation 자료를 이용한 과거 이력자료 구축방법론 개발
    - 링크별/일별/시간대별 Probe건수와 속도 이력자료 DB 구축
    - ▶ Probe건수 산출 방법론 개발
    - ▶ 속도 유전자 지도 개발 및 구간 통행속도 산출 방법론 개발
  - 핵심 요소기술② : 구축된 이력자료를 이용한 장래 통행속도 예측방법론 개발
    - KNN 단계 : k-군집 추출모형 개발
    - 예측단계 : 최종 의사결정 및 통행속도예측 모형 개발

- 마지막으로 장래 교통예보를 위한 전체 개발 구조는 <그림 2-28>에서 보는 바와 같이 크게 4단계 걸쳐 개발을 수행하고자 함
  - Step 1 : Car Navigation 자료를 이용한 과거 이력자료 DB 구축(Probe건수, 속도)
    - 데이터 수집기간 : 2013년 10월 1일 ~ 2014년 1월 31일(총 4개월)
    - 대상범위 : 서울특별시 강남 일대
  - Step 2 : 장래 통행속도 예측을 위한 시·공간적 범위 설정
    - 시간적 범위 : 과거 이력자료 DB 중 예측하고자 하는 예측일(익일) 선택
    - 공간적 범위 : 과거 이력자료 DB 중 예측하고자 하는 분석 네트워크(링크번호) 선택
    - ▶ 당일의 소통 상태 : 선택된 예측일 하루 전 당일의 링크별/시간대별 소통상태 집합
    - ▶ 과거의 소통 상태 : 당일 기준보다 이전의 링크별/일별/시간대별 소통상태 집합

☞ 여기서 말하는 소통 상태는 Probe건수와 속도를 의미함
  - Step 3 : k-군집 추출모형 개발(KNN 단계)
    - 과거 소통상태의 집합 중 당일 소통 상태와 유사한 k개의 군집 추출
  - Step 4 : 최종 의사결정 및 통행속도예측 모형 개발(예측 단계)
    - 이전 단계의 k개의 군집 중 최종의사결정 군집 선택(유사성이 가장 높은 군집)
    - 장래 구간 통행속도 예측 : 각 구성요소들에 대한 익일의 링크별/시간대별 통행속도의 가중평균값

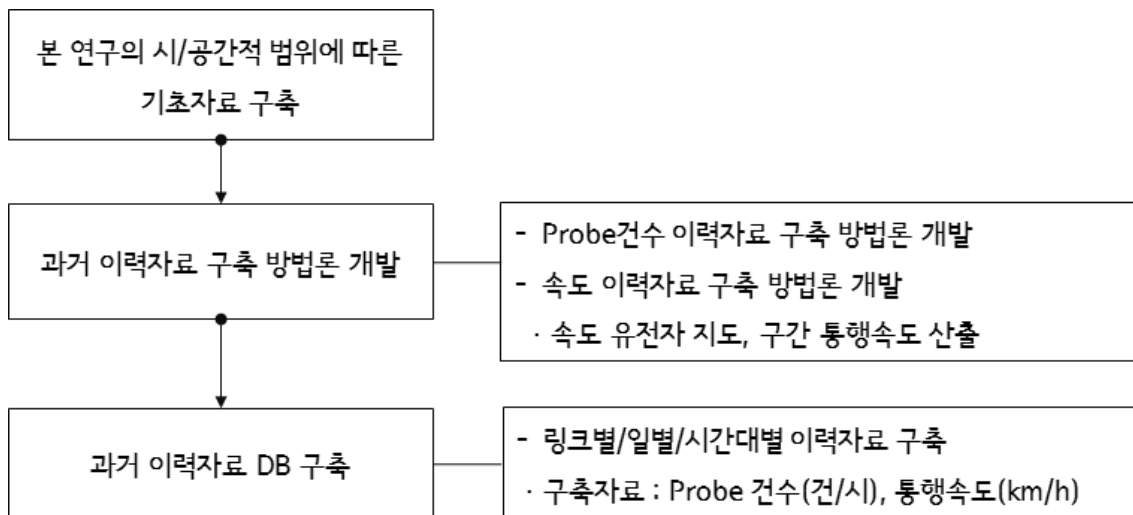


<그림 2-28> 장래 예측 방법론의 개발 구조

## 2. 과거 이력자료 구축방법론 개발

### 가. 이력자료 구축을 위한 개발 구조

- 본 연구는 기존의 확률적인(Stochastic) 상태를 가정하는 대부분의 단기 예측을 위한 통계기반 모형이 아닌 데이터 기반의 비모수(Non-parametric) 모형을 적용하며, 비모수 모형에서 최근린 이웃은 독립변수의 상태간 거리가 가장 가까운, 즉 유사성이 가장 높은 k개의 최근린 이웃(k-nearest neighbor)의 집합을 의미함
- 이러한 모형의 예측력은 모형에 적용되는 분석 데이터의 질과 양에 의해 크게 좌우되므로 무엇보다도 정확한 과거 이력자료 구축이 필요함
- 본 연구의 과거 이력자료는 분석 맵 단위의 Car Navigation 자료(건별 이벤트 자료)를 이용하여 Probe건수와 속도라는 두 가지 변수로 링크별/일별/시간대별로 이력자료 DB를 구축해야 함. 이를 위하여 장래 교통예보를 위한 Probe건수와 속도 이력자료 구축 방법론을 개발하고자 함
  - Probe건수(건/시) : 건별 이벤트 자료 → 링크별/일별/시간대별로 누적 합계한 Probe건수
  - 속도(km/h) : 건별 이벤트 자료 → 속도 유전자 지도를 이용한 링크별/일별/시간대별 가중 평균속도
- 본 연구의 과거 이력자료 DB를 구축하기 위한 전체 개발 구조는 다음과 같음



<그림 2-29> 과거 이력자료 구축을 위한 개발 구조

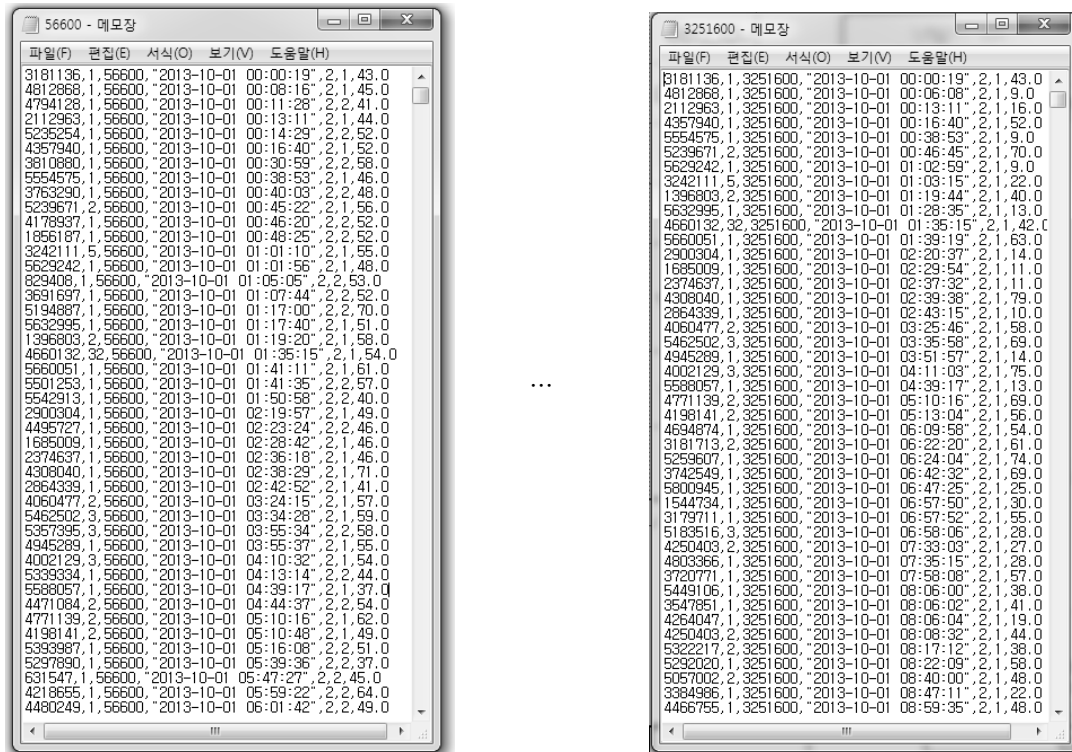
#### 나. 본 연구의 시/공간적 범위에 따른 기초자료 구축

- 본 연구는 향후 교통예보시스템 구축을 위하여 시스템 내 핵심 요소기술을 개발하는 기초 연구이므로 개발모형을 전국 단위로 평가하는 데에는 무리가 있음. 따라서, Car Navigation 자료의 수집기간을 고려하여 Probe샘플 수가 많으면서 연속류와 단속류가 혼재되어 다양한 교통류 흐름이 존재하는 서울시 강남 일대로 한정하여 공간적 범위를 정하였음
- 즉, 전국의 분석 맵 단위로 구축한 Car Navigation 자료에서 본 연구의 시/공간적 범위에 해당하는 자료만을 추출하여 본 연구의 과거 이력자료 DB구축을 위한 기초자료로 구축하였음
  - 데이터 수집기간 : 2013년 10월 1일 ~ 2014년 1월 31일, 4개월치 이력자료
  - 공간적 범위 : 서울특별시 강남 일대
  - 이용 자료 : Probe건수, 속도
- 해당 자료는 아래와 같이 총 7개의 Column으로 분석 맵의 링크별 수집시간 순(년/월/일/시/분/초)으로 구축하였음
  - 수집 Column(총 7개) : OBU\_ID/GROUP\_ID/from\_link/link time/week\_type/turn\_type/speed

<표 2-11> 분석 맵 기준의 Car Navigation 기초자료 정의서

Column	설명	Data Type	PK	FK
OBU_ID	Car Navigation 단말기 ID	Varchar	-	-
GROUP_ID	단말기별 그룹 ID	Varchar	-	-
from link	분석맵 기준 링크 ID	Varchar	-	-
link time	수집시간(년/월/일/시/분/초)	Varchar	-	-
week_type	요일구분 (0:일요일, 1:월요일, 2:화요일, 3:수요일, 4:목요일, 5:금요일, 6:토요일, 11:공휴일)	Integer	-	-
turn_type	이동류 구분 (0:좌회전, 1:직진, 2:우회전, 3:유턴)	Integer	-	-
speed	속도(km/h)	Double	-	-

- 위의 자료는 수집시간 순으로 링크별 텍스트 형식으로 저장하며, 해당 자료의 저장형태는 아래와 같음
  - 데이터 저장 규모 : 총 255개의 텍스트 파일(서울특별시 강남 일대, 총 225개 링크)
  - 연속류(77개) : 경부고속도로 30개, 올림픽대로 47개
  - 단속류(178개) : 기타 내부도로 178개



<그림 2-30> 분석 맵 단위의 텍스트 저장 형태

- 본 연구의 시/공간적 범위에 따른 기초자료 구축 현황은 아래의 표와 같으며, 링크개수는 기존의 링크 총 600개에서 현재 총 255개로 즉, 링크 당 평균 약 2.4개를 병합하여 구축되었음. 또한 링크당 1일 평균 이벤트 건수는 1,437건으로 충분한 샘플 수를 확보하였음

<표 2-12> 본 연구의 공간적 범위 내 기초자료 구축 현황

구분		링크 개수		링크당 1일 평균 이벤트 건수(건/일)	비고
		ITS 표준노드링크	분석 맵		
연속류	경부고속도로	32개	30개	1,893	링크 당 평균 약 2.4개 병합
	올림픽대로	62개	47개	1,798	
단속류	기타 내부도로	506개	178개	621	
합계		600개	255개	1,437	

#### 다. 과거 이력자료 DB 구축 방법론 개발

- 본 연구의 시/공간적 범위 내 분석 맵 단위의 건별 Car Navigation 자료를 이용하여 장래 통행속도 예측을 위한 입력자료 DB(Probe건수, 속도)를 구축해야 하며, 이를 위해서는 아래와 같이 해당 자료의 집계시간 길이와 최소 유효표본수를 설정해야 함
  - 집계시간 길이 : 1시간
  - 최소 유효 Probe 표본수 : 집계시간 길이 당 20건 이상
- 과거 이력자료 DB는 Probe건수와 속도의 두 가지 변수로 구축되며, Probe건수는 각 링크별로 수집되는 건별 Probe를 1시간 단위로 누적 합계하여 Probe건수 DB를 구축하고, 속도의 경우는 본 연구의 속도 유전자 지도와 구간 통행속도 산출 방법론을 적용하여 속도 DB를 구축해야 함

##### 1) Probe 이력자료 DB구축

- 교통류의 상태변화는 거시적 교통류 변수인 교통량( $q$ , 대/단위시간), 속도( $v$ , 단위거리/단위시간), 밀도( $\rho$ , 대/단위거리)의 관계로서 설명이 되며, 여기서 Probe건수(건/시)는 거시적 교통류 변수 중 교통량 변수의 역할을 수행하게 됨
- Probe건수의 이력자료 DB구축은 본 연구의 시/공간적 범위를 고려하여 각 링크별로 수집되는 건별 Probe를 1시간 단위로 누적 합계하여 링크별/일별/시간대별 Probe건수 자료로 구축함
  - 시간당 Probe건수(건/시) : 분석 맵 단위 링크별/일별/시간대별 Probe건수 누적 집계
    - 대상 링크개수 : 총 255개(서울특별시 강남 일대)
    - 구축일 개수 : 총 123개(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
    - 시간대 개수 : 총 24개(01~24시)
    - 전체 Probe건수 DB규모 : 752,756개( $255 \times 123 \times 24 = 752,760$ )
- 구축된 링크별/일별/시간대별 Probe건수의 DB구축 결과는 아래와 같음

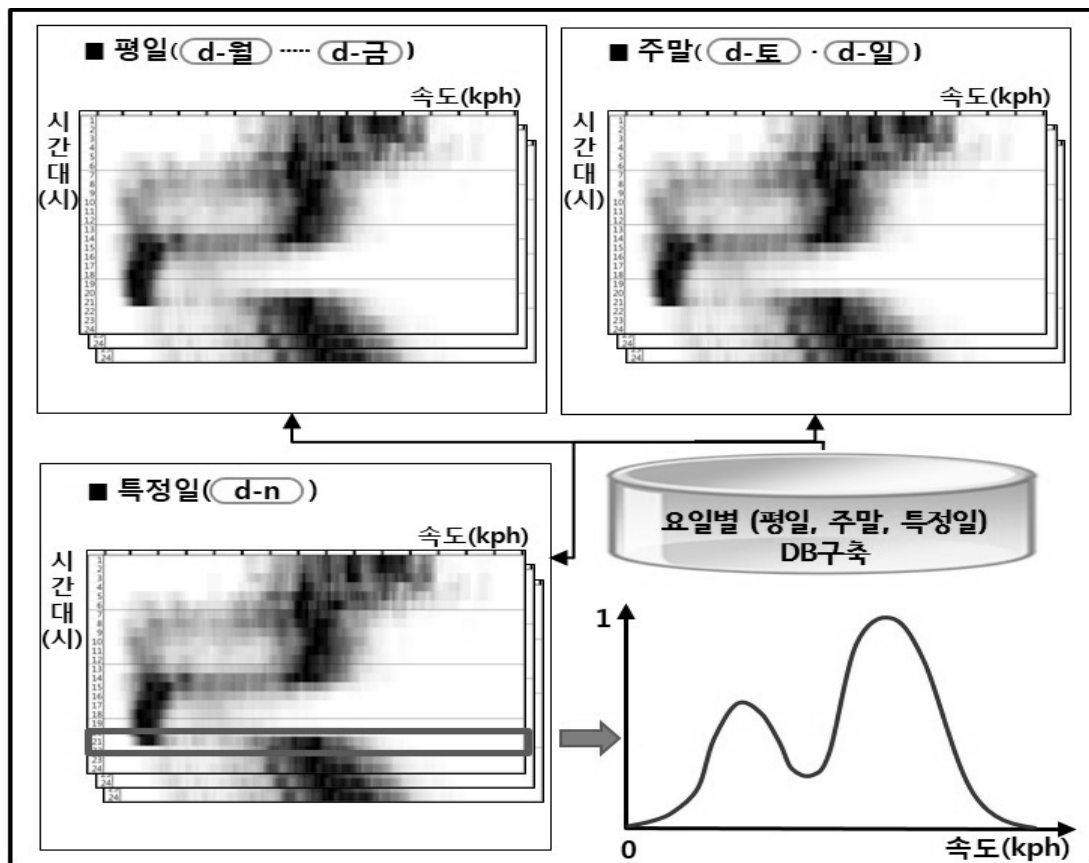
구분	Probe건수(건/시)				
	174400	174500	...	3251400	3251600
2013년10월1일 00~01시	24	21	...	22	25
2013년10월1일 01~02시	18	20	...	45	17
⋮	⋮	⋮	⋮	⋮	⋮
2014년1월30일 22~23시	24	45	...	34	21
2014년1월31일 23~24시	20	30	...	45	33

<그림 2-31> Probe건수의 DB구축 형태

## 2) 속도 이력자료 DB 구축

### ① 속도 유전자 지도 개발

- 본 연구의 속도 유전자 지도는 분석 맵 단위의 링크별로 수집되는 건별 Probe 차량의 속도자료를 이용하여 정수형태의 이산적인 속도 범위(1~150km/h)에서 링크별/요일별/시간대별 속도 profile<sup>5)</sup> 구축하고 각 해당 빈도 값에 따라 음영의 색깔로 구분하여 표출한 지도를 말함
- 이러한 속도 유전자 지도는 링크별/요일별/시간대별 속도 이력자료 DB를 구축하기 위하여 개발되는 구간 통행속도 산출방법에서 조화평균의 가중치로서 이용되며, Probe 건수가 낮은 구간의 통행속도를 산출할 때에도 적용이 가능함. 또한, 해당 링크 구간의 시간대별 속도상태의 변화와 더불어 장래 상태 변화의 방향성과 상태 정도에 대한 지도의 역할도 수행함

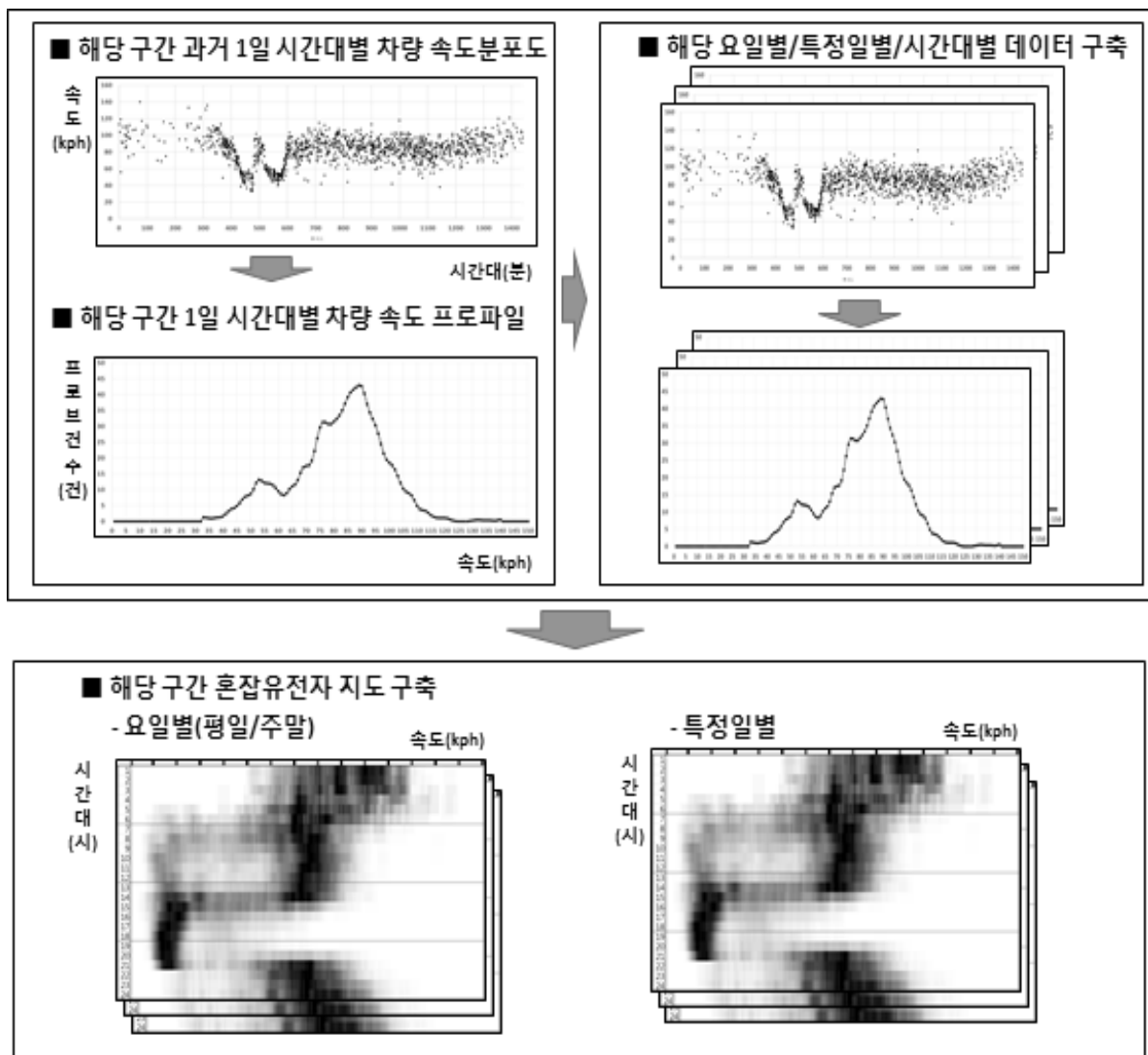


<그림 2-32> 속도 유전자 지도 개발 형태

<sup>5)</sup> 속도 profile은 1km/h의 계급구간으로 구성되는 속도 Histrogram임



- 먼저, 본 연구의 속도 유전자 지도 개발을 위하여 건별 Car Navigation 속도 자료를 이용하여 링크별/요일별/시간대별 속도 profile을 형태의 DB를 구축해야 하며, 이를 위한 단계별 구축 방법은 다음과 같음
  - Step 1 : 건별 Probe 속도의 시계열 자료 구축
  - Step 2 : 링크별/요일별/시간대별 속도 profile 구축
  - Step 3 : 속도 profile 이상치 제거
  - Step 4 : 속도 profile 정규화(0~1)
  - Step 5 : 링크별/요일별/시간대별 유전자 지도 DB 구축



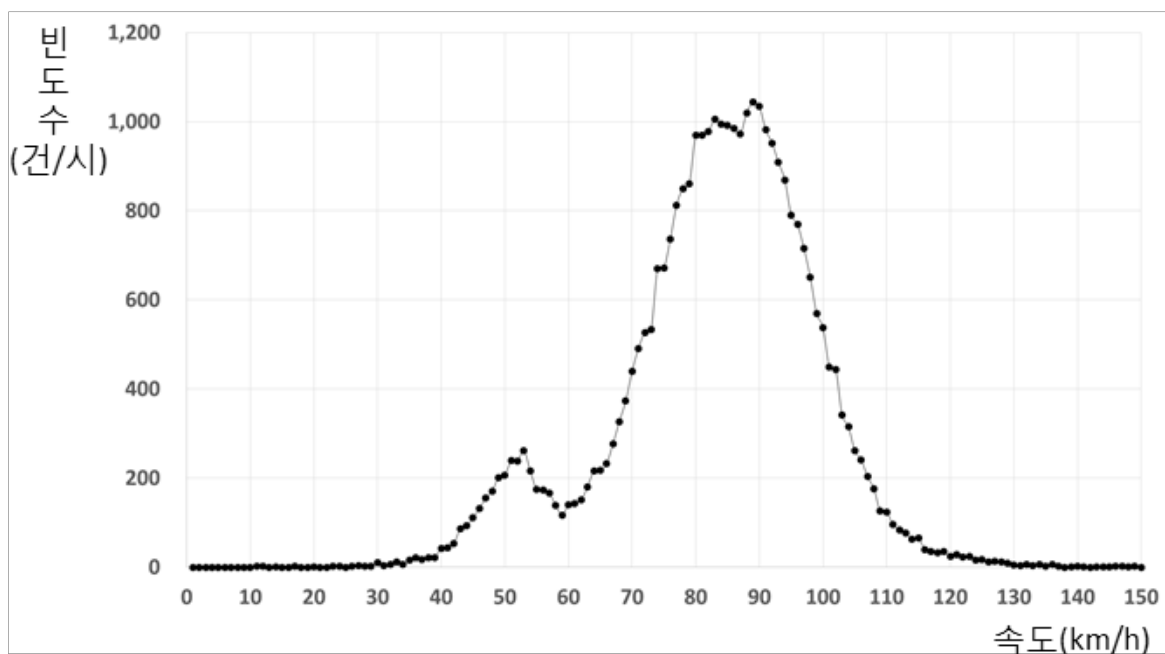
<그림 2-33> 속도 유전자 지도 개발 개념도

○ 건별 Probe 속도의 시계열 자료 구축

- 분석 맵의 링크 단위로 수집되는 건별 Probe 속도자료를 수집시간(④link time, 년/월/일/분/초) 순으로 정렬하여 링크별 시계열 자료를 구축함
  - 데이터 수집기간 : 2013년 10월 1일 ~ 2014년 1월 31일, 4개월치 이력자료
  - 공간적 범위 : 서울특별시 강남 일대
  - 집계 Column (총 7개)
  - OBU\_ID/GROUP\_ID/from\_link/link time/week\_type/turn\_type/speed

○ 링크별/요일별/시간대별 속도 profile 구축

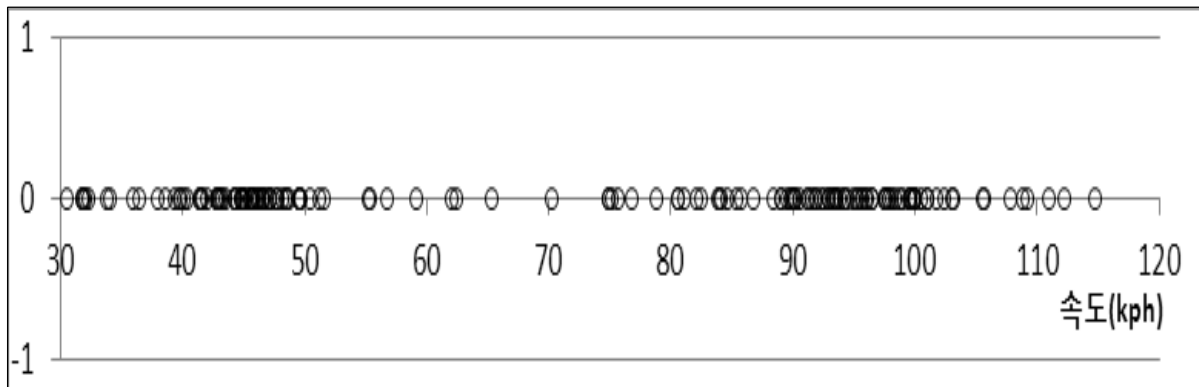
- 구축된 링크별 시계열 자료를 이용하여 요일별/시간대별로 정수형태의 이산적인 속도 범위 (1~150km/h)에 따라 누적 집계하여 profile를 구축함
  - 전체 링크개수 : 총 255개 (서울특별시 강남 일대)
  - 요일구분 개수 : 7개 (월요일~일요일)
  - 시간구분 개수 : 24개 (01~24시)
- 즉, 속도 profile은 링크별/요일별시간대별로 총 42,840개( $255 \times 7 \times 24 = 42,840$ )가 생성되며, 속도 profile의 형태는 아래와 같음



<그림 2-34> 초기 속도 profile의 형태

○ 속도 profile 이상치 제거

- 혼잡과 비 혼잡이 발생한 구간에 대한 수집된 개별 Car Navigation 차량의 속도 분포는 아래의 그림과 같음. ①교통류의 속도는 60kph를 전/후로 혼잡과 비 혼잡 교통류로 극명하게 양분되고 있으며, ②전형적인 쌍봉분포(Two-peak distribution)를 보이고 있음



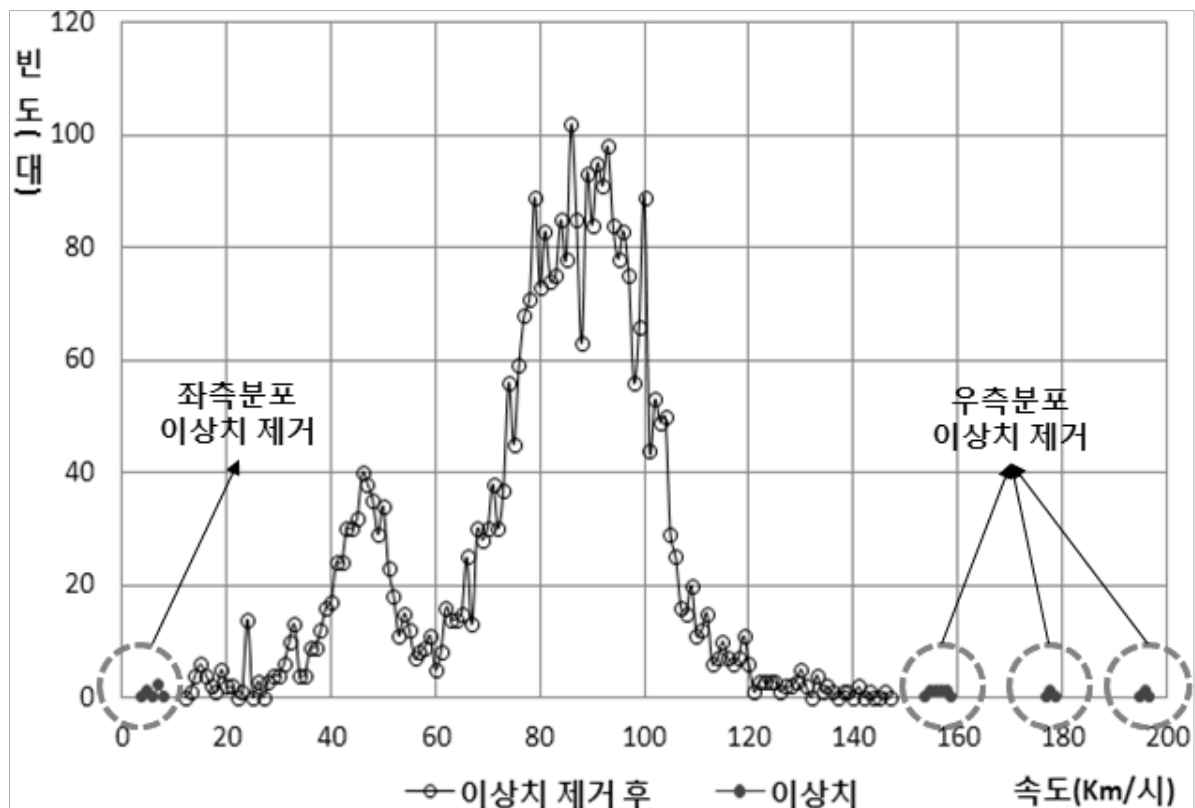
<그림 2-35> 개별 Car Navigation 차량의 속도 분포

- 앞서 설명한 개별 Car Navigation 차량의 속도 분포의 특성으로 인하여 기존의 정규분포(Normal distribution)를 가정한 통계적 접근 방법론으로 설명할 수 없는 한계를 가짐
- 따라서, 구축된 (링크별/요일별/시간대별) 속도 profile에 대한 이상치를 제거하기 위하여 기존의 통계적 이상치 제거 방법이 아닌 속도 profile의 사건이 발생한 도수 간 최인접 좌/우측 거리분포의 표준편차 최소화하는 이상치 제거 방법을 적용함
- 본 연구의 링크별/일별/시간대별 속도 profile 구축을 위한 단계별 이상치 제거 방법은 아래와 같음
  - 주어진 도로구간, 이동류, 시간조건(년, 월, 일, 시간대)에 대하여;
    - Step 1 : 초기 속도 Profile 분석
      - 1-1 : 1km/시의 계급구간으로 구성되는 속도 Profile 구축
      - 1-2 : 좌/우측 outlier 제거 비율값( $tv_{lt}$ ,  $tv_{rt}$ ,  $0.0 \leq tv_{lt}, tv_{rt} \leq 0.03$ )을 이용한 1차 이상치 제거
      - 1-3 : 쌍봉분포 또는 단일분포의 분석 및 첨두에서 속도값 결정( $v_{lp}$ ,  $v_{rp}$ )
    - Step 2 : 저속도 분포의 이상치 제거
      - 2-1 : Step 1에서 구축한 속도 Profile의 속도 0과 저속 분포에 대한 첨두시 속도( $v_{lp}$ )\*g 구간의 속도 Profile 구축, 여기서, g=황금비(0.618)
      - 2-2 : 쌍봉분포의 분석 및 첨두에서 좌측 outlier 제거 비율값( $tv_{lt}$ ) 결정

- Step 3 : 고속도 분포의 이상치 제거
  - 3-1 : Step 1에서 구축한 속도 Profile의 고속 분포에 대한 첨두시 속도( $v_{rp}$ )\*g와 최대값( $v_{max}$ )구간의 속도 Profile 구축
  - 3-2 : 쌍봉분포의 분석 및 첨두에서 좌측 outlier 제거 비율값( $tv_{rt}$ ) 결정
- Step 4 : 최종 속도 Profile 구축
  - 4-1 : 1km/시의 계급구간으로 구성되는 속도 Profile 구축
  - 4-2 : 갱신된 좌우측 outlier 제거 비율값( $tv_{lt}$ ,  $tv_{rt}$ )을 이용한 이상치 제거
  - 4-3 : 이상치 제거 후, 최종 속도 Profile 구축

<표 2-13> 본 연구의 이상치 제거 방법

구분	기존 방법	본 연구
이상치 제거 방법	- 이상치 < 평균-3 $\delta$ 또는 평균+3 $\delta$ < 이상치	- 속도 profile의 사건이 발생한 도수 간 최인접 좌/우측 거리분포의 표준편차 최소화하는 이상치 제거 방법 적용



<그림 2-36> 이상치 제거 전과 후의 속도 분포

○ 속도 profile 정규화(0~1)

- 이상치 제거 후의 속도 profile에서 Y축의 빈도수(건/시)를 0~1사이의 값으로 정규화하여 최종적인 (링크별/요일별/시간대별)속도 profile를 구축함

$$F_{l,d,h}^s = \frac{f_{l,d,h}^s}{\sum_{s=1}^{150} f_{l,d,h}^s}, \quad \sum_{s=1}^{150} f_{l,d,h}^s > 0$$

여기서,  $F_{l,d,h}^s$  : l링크별 d요일별 h시간대에 따른 s속도의 빈도비(0~1)

$f_{l,d,h}^s$  : l링크별 d요일별 h시간대에 따른 s속도의 빈도수(건/시)

○ 링크별/요일별/시간대별 유전자 지도 DB 구축

- 정규화된 속도 profile의 빈도비를 이용하여 링크별/요일별/시간대별 유전자 지도 DB를 구축하였으며, 그 결과는 아래와 같음
  - 링크개수 : 총 255개(서울특별시 강남 일대)
  - 요일구분 개수 : 7개(월요일~일요일)
  - 시간구분 개수 : 24개(01~24시)
  - 전체 유전자 지도 DB구축 규모 : 42,840개(255×7×24=42,840)

링크번호	요일	시간대	정규화된 속도 profile(0~1)						
			1km/h	2km/h	3km/h	...	148km/h	149km/h	150km/h
56600	월	00~01시	0.00	0.00	0.01	...	0.02	0.00	0.00
56600	월	01~02시	0.00	0.00	0.00	...	0.00	0.00	0.00
⋮	⋮	⋮	⋮	⋮	⋮	...	⋮	⋮	⋮
3319200	일	22~23시	0.00	0.00	0.00	...	0.01	0.01	0.00
3319200	일	23~24시	0.00	0.01	0.00	...	0.00	0.00	0.00

<그림 2-37> 유전자 지도 DB 구축 형태

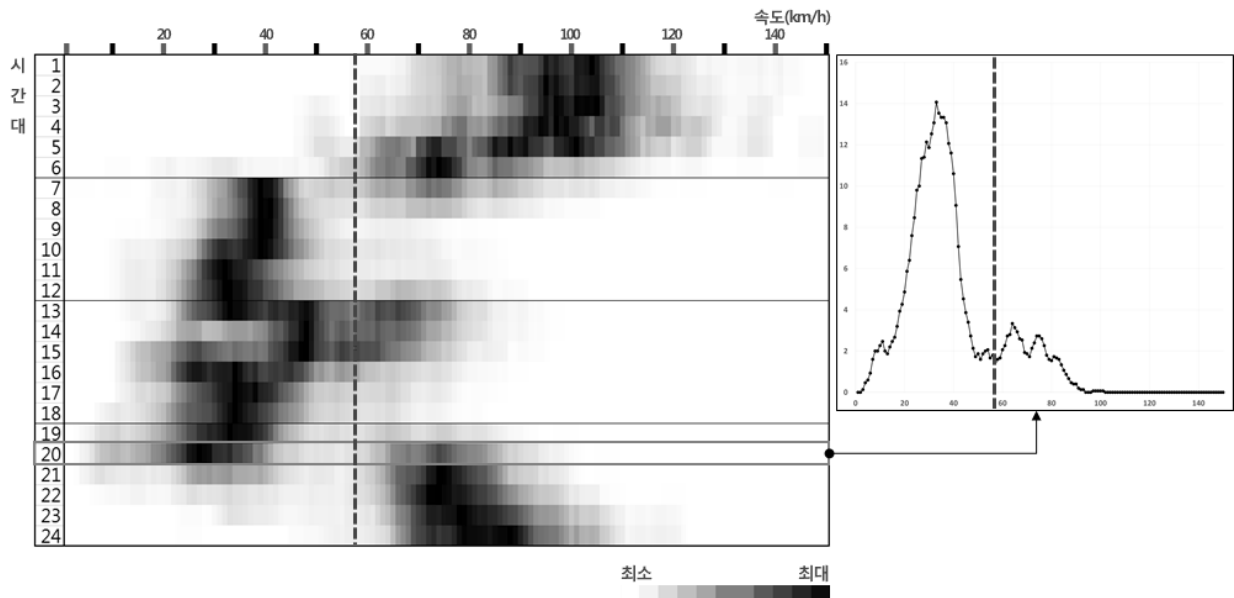
## ② 속도 유전자 지도 구축

- 본 연구에서는 해당 구간을 연속류와 단속류로 구분하여 링크별/요일별/시간대별로 속도 유전자 지도를 구축하였으며, 일부 구간에 대한 구축 결과를 아래와 같이 제시함
- 즉, 속도 유전자 지도에서 음영이 진할수록 해당 속도의 빈도수가 많다는 것을 의미이며, 그 반대의 경우는 해당 속도의 빈도수가 적다는 것을 의미함

### ○ 연속류 구간

#### - 경부고속도로(양재IC → 서초IC 구간)

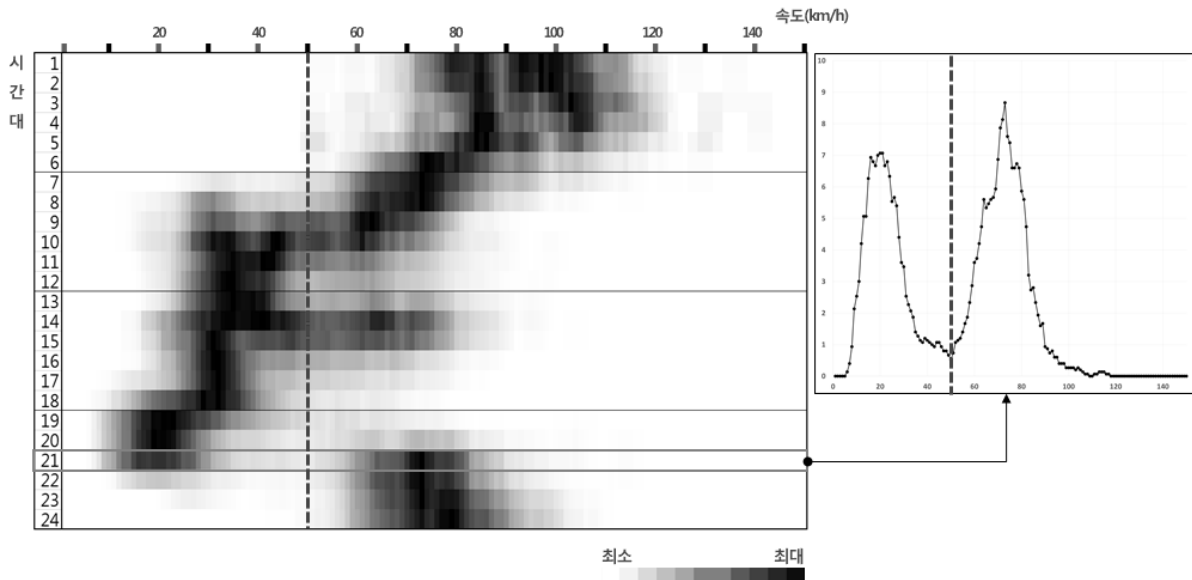
- 본 구간의 교통 혼잡은 출근시간 07시부터 시작하여 13~15시에 일부 해소되는 듯 하다가 그 이후로 퇴근시간 20시까지 지속되는 것을 알 수 있음
- 또한, 본 구간은 교통류 이론 측면에서 전형적인 양봉분포 형태를 가지며, 용량상태(불안정 교통류)를 구분하는 경계속도가 약 58km/h인 것으로 분석되었음



<그림 2-38> 경부고속도로(양재IC → 서초IC 구간) 속도 유전자 지도

#### - 올림픽대로(한남대교→동호대교 구간)

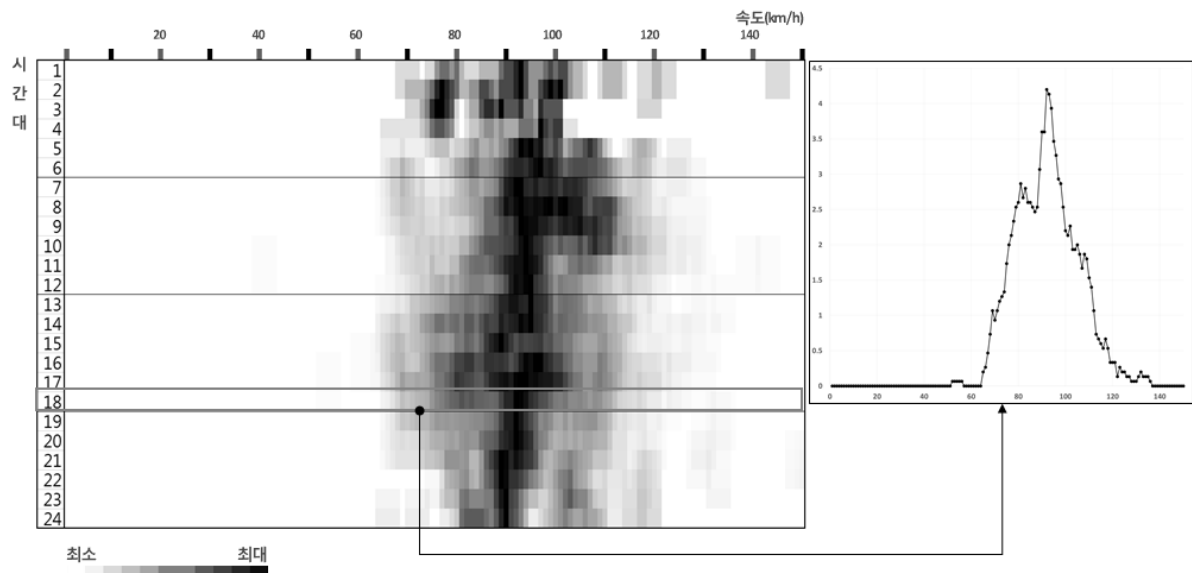
- 본 구간의 교통 혼잡은 출근시간대(07~09시) 보다는 오히려 퇴근시간대(18~21시)에 극심한 혼잡이 발생하며, 주간시간대(10~17시)에도 지속적으로 정적상태(Stationary condition)의 혼잡이 계속 발생하는 것으로 분석됨
- 또한, 본 구간은 교통류 이론 측면에서 전형적인 양봉분포 형태를 가지며, 용량상태를 구분하는 경계속도가 약 50km/h인 것으로 분석되었음



<그림 2-39> 올림픽대로(한남대교→동호대교 구간) 속도 유전자 지도

- 인천대교 고속도로(송도→인천공항 구간)

- 본 구간은 모든 시간대에서 혼잡이 발생하지 않는 자유교통류 상태의 구간이며, 시간대별 차량들의 속도 분포는 개별 운전자의 운전행태 따른 이질적인 변화라고 볼 수 있음
- 또한, 교통류 이론 측면에서 혼잡이 발생하지 않는 전형적인 일봉분포 형태를 가지며, 양봉분포와는 달리 용량상태를 구분하는 경계속도는 발생하지 않는 것으로 분석됨

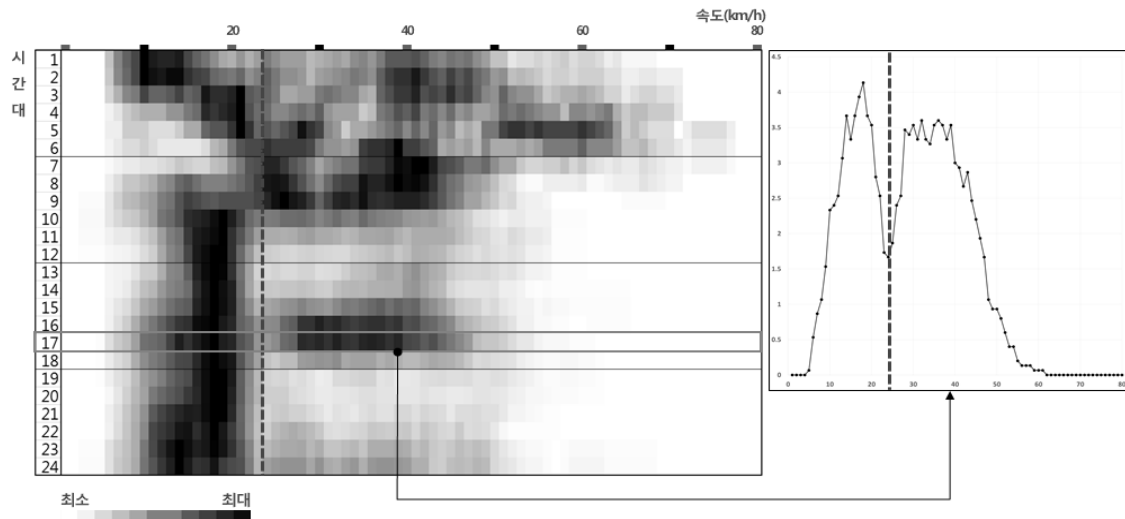


<그림 2-40> 인천대교 고속도로(송도→인천공항 구간) 속도 유전자 지도

○ 단속류 구간

- 강남대로(신논현역 → 강남역 구간)

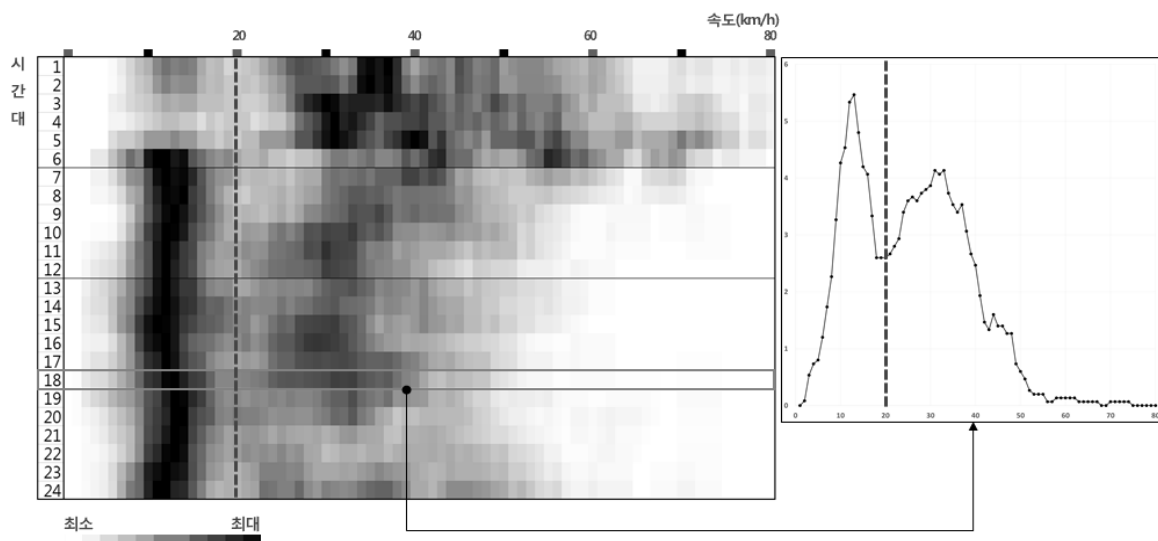
- 본 구간은 신호기에 의한 차량의 신호대기 길이에 따라 연속류에 비해 시간대별 속도분포의 변화가 심하며, 용량상태를 구분하는 경계속도는 약 23km/h인 것으로 분석됨



<그림 2-41> 강남대로(신논현역 → 강남역 구간) 속도 유전자 지도

- 종로(종로2가→종각역 구간)

- 본 구간은 신호기에 의한 차량의 신호대기 길이에 따라 연속류에 비해 시간대별 속도분포의 변화가 심하며, 용량상태를 구분하는 경계속도는 약 20km/h인 것으로 분석됨



<그림 2-42> 종로(종로2가→종각역 구간) 속도 유전자 지도

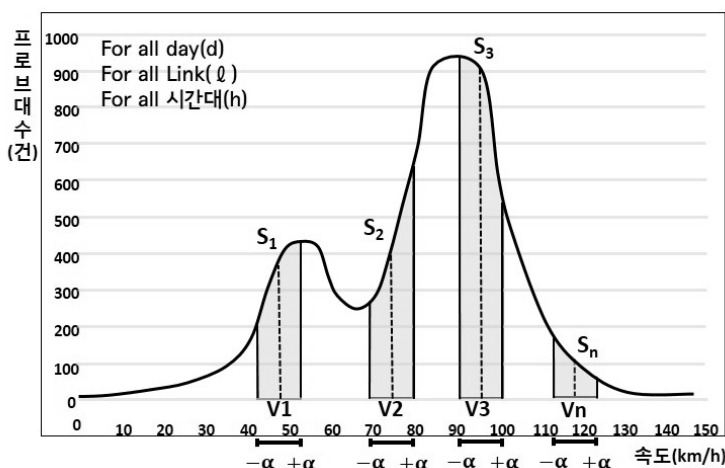


## 2) 구간 통행속도 산출 방법론 개발

- 본 구간 통행속도 산출 방법론은 다음과 같이 2개 핵심 요소기술로 개발하며, ①Probe 샘플이 최소샘플수 이상인 구간의 통행속도 산출방법과 ②Probe 샘플이 낮은 구간의 통행속도 산출방법으로 구성됨
- 또한, 구간 통행속도 산출방법은 속도 유전자 지도와 마찬가지로 단속류, 연속류 구간에 대하여 공통적으로 적용이 가능하도록 개발함

### ① Probe 샘플이 최소샘플수 이상인 구간의 통행속도 산출

- 구간 통행속도 산출방법은 건별 Car Navigation 속도 자료를 이용하여 링크별/일별/시간대별로 집계를 하기 위한 대표 속도 값을 산출하는 것이며, 이를 위하여 기 구축된 링크별/요일별/시간대별 속도 유전자 지도가 이용됨
- 즉, 속도 유전자 지도의 링크별/요일별/시간대별 속도 profile의 빈도비(%)가 가중 조화평균에서 가중치( $S_n$ )로 이용됨
- 따라서, Probe 샘플이 최소샘플수 이상(시간당 20건 이상)인 구간의 경우는 공통적으로 속도 유전자 지도의 Probe 빈도기반 가중 조화평균 값으로 산출하며, 해당 산출방법은 아래와 같음



- Probe 빈도기반 가중 조화평균 이용

$$V = \frac{\sum_{n=1}^n S_n}{\sum_{n=1}^n \frac{S_n}{V_n}}$$

여기서,  $S_n$ : n번째 속도분포 빈도(%)

$V_n$ : n번째 속도(kph)

<그림 2-43> Probe 샘플이 최소샘플수 이상인 구간의 속도 산출방법

## ② Probe 샘플이 낮은 이상구간의 통행속도 산출

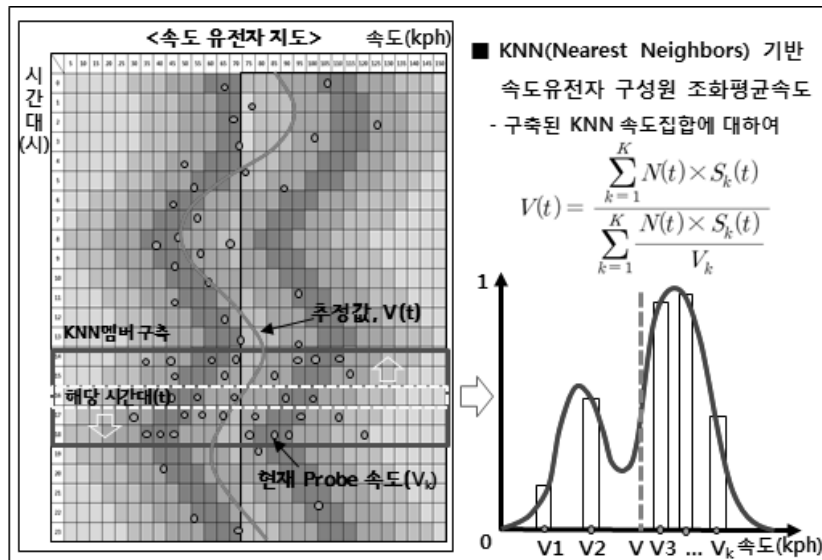
- Probe 샘플이 낮은 이상구간(시간당 20건 미만)의 경우, 속도 프로브를 이용한 구간속도 추정 시 낮은 프로브 샘플에 의한 대표 속도 산출 값이 편향될 우려가 있음
- 이를 해결하기 위하여 본 연구에서는 아래와 같이  $k$ 개의 최근린 이웃( $k$ -nearest neighbor, 이하 KNN)으로 KNN 멤버를 구축하고  $k$ 개의 속도 유전자 지도의 구성원에 대한 조화 평균속도를 이용하여 보다 현실적인 구간 통행속도를 추정하도록 함
  - KNN은 속도 유전자 구성원 조화평균속도를 산정하기 위한 멤버의 개수임
  - KNN의 산정은 시간대( $t$ )에서 전/후로 인접한 시간대로 시간대를 확장해 가면서 KNN의 개수(20개 이상) 만큼 속도 프로브 자료를 구축함
  - 최소 KNN 개수만큼의 속도 프로브 자료가 구축되면, 이를 이용하여 해당 시간대( $t$ )의 평균 구간속도는 다음과 같이 산정됨

$$\text{구축된 KNN 통행속도 집합에 대하여, } V(t) = \frac{\sum_{k=1}^K N(t) \times S_k(t)}{\sum_{k=1}^K \frac{N(t) \times S_k(t)}{V_k}}$$

여기서,  $K(t)$ : 해당 시간대( $t$ )의 총 수집 프로브 대수,

$S_k(t)$ : 해당 시간대( $t$ )의  $k$ -번째 프로브의 속도에 대한 구간 확률(0.0~1.0),

$V_k$ :  $k$ -번째 프로브의 속도,  $k \in K, k = \{1, 2, \dots, K\}$



<그림 2-44> Probe 샘플이 낮은 구간의 구간 통행속도 산출방법

## 라. 과거 이력자료 DB 구축 결과

- 본 연구의 과거 이력자료 DB는 건별 Car Navigation 자료(Probe건수, 속도)를 이용하여 속도 유전자 지도와 구간 통행속도 산출방법론을 적용한 후, 링크별/일별/시간대별 Probe건수와 속도의 자료로 구축하였음
- 과거 이력자료 DB의 규모는 시/공간적 범위와 Car Navigation 데이터 수집기간에 따라 각각 752,760개( $255 \times 123 \times 24 = 752,760$ )씩 Probe건수와 속도 자료가 구축되었음
  - 대상 링크개수 : 총 255개(서울특별시 강남 일대)
  - 구축일 개수 : 총 123개(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
  - 시간대 개수 : 총 24개(01~24시)
- 구축된 링크별/일별/시간대별 Probe건수와 속도의 DB구축 현황 및 저장 형태는 아래와 같음

&lt;표 2-14&gt; 과거 이력자료 구축 현황

구분		Probe건수(건/시)			속도(km/h)		
		데이터 개수	평균	표준편차	데이터개수	평균	표준편차
연속류	경부고속도로	88,560개	60.2	25.4	88,560개	64.2	24.0
	올림픽대로	138,744개	56.8	28.7	138,744개	65.7	19.2
단속류	기타 내부도로	525,456개	31.9	11.4	525,456개	41.8	11.8
전체		752,760개	49.6	21.8	752,760개	57.2	18.3

## - Probe건수(건/시) DB

구분	링크번호(총 255개)				
	174400	174500	...	3251400	3251600
2013년10월1일 00~01시	24	21	...	22	25
2013년10월1일 01~02시	18	20	...	45	17
⋮	⋮	⋮	⋮	⋮	⋮
2014년1월30일 22~23시	24	45	...	34	21
2014년1월31일 23~24시	20	30	...	45	33

## - 속도(km/h) DB

구분	링크번호(총 255개)				
	174400	174500	...	3251400	3251600
2013년10월1일 00~01시	120	109	...	99	94
2013년10월1일 01~02시	110	95	...	100	92
⋮	⋮	⋮	⋮	⋮	⋮
2014년1월30일 22~23시	100	110	...	108	98
2014년1월31일 23~24시	95	90	...	114	100

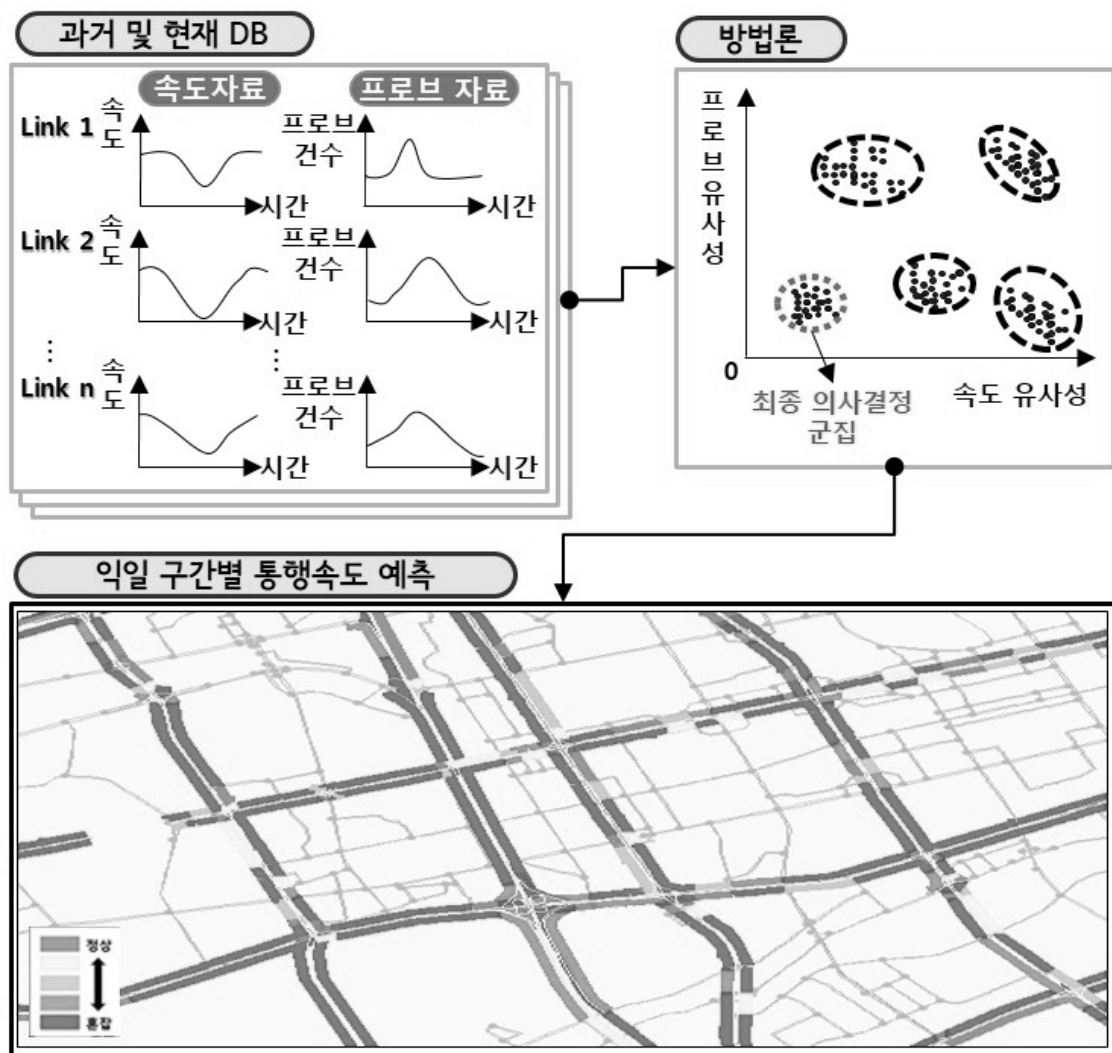
&lt;그림 2-45&gt; 과거 이력자료 DB 구축 저장 형태

### 3. 장래 통행속도 예측방법론 개발

#### 가. 예측방법론 개발 개요

- 본 연구는 기존의 단기 예측기법에 대한 한계를 극복함으로써 보다 먼 장래의 통행속도를 수용 가능한 오차의 범위 내에서 예측하고, 교통류의 행태와 예측 모델링에 대한 전문적인 지식이 없이도 운영이 가능한 데이터 기반의 예측방법론을 개발하고자 함
- 이러한 예측모형은 통계기반의 모형에 데이터를 맞추던 기존의 방식을 벗어나 개발모형이 데이터베이스(DB)를 탐색함으로써 선별적으로 데이터를 추출하며, 추출된 데이터를 지식발견(Knowledge discovery, 이하 KD)과 의사결정의 과정을 통하여 장래 통행속도를 예측하는 지식 탐색형 모형을 의미함
- 즉, 본 연구의 개발모형은 데이터 탐색을 기반으로 하는 예측 관련 선행 연구의 고찰을 통하여 연산 수행속도가 빠르면서 모형의 예측력이 우수한 비모수(non-parametric) 기반의 k-최근린 이웃 기법(k-nearest neighbor)을 선정하였음
- 이 기법의 파라미터인 k값에 대한 전역(Global) 최적화의 단점을 극복하기 위하여 모형의 수행 과정에서 국부(Local) 파라미터 최적화 과정을 통하여 기존의 전역 최적화의 문제점을 극복하고자 함. 또한, 국부(Local) k값의 최적화는 연산 속도가 빠르면서 효과적으로 군집을 구성할 수 있는 안정적인 군집화 통하여 달성하고자 함
- 먼저, 본 연구의 장래 통행속도 예측모형을 개발하기 위한 전제조건은 예측하고자 하는 분석 네트워크의 링크별 통행속도는 해당 네트워크의 교통량(즉, Probe 건수)과 통행속도의 관계에 의하여 정의됨
  - 장래 링크별 통행속도 =  $F_n\{\text{링크별 Probe 건수, 링크별 속도}\}$
- 위의 전제조건에 대한 상태벡터의 차원은 링크로 구성되는 네트워크로 정의되며, 각 상태벡터는 ①네트워크 구간별 프로브 대수와 ②네트워크 구간별 속도로 정의됨. 그리고 현행 상태벡터와 과거 이력 상태벡터 간의 유사성은 유클리디언 거리(Euclidean Distance)로 정의함
- 즉, Probe건수(네트워크 수요)의 유사성과 속도의 유사성을 동시에 고려하여 장래 네트워크의 링크별 통행속도를 예측하며, Probe건수와 속도의 유사성 높다는 것은 두 변수에 대한 현행 상태벡터와 과거 이력 상태벡터 간의 유클리디언 거리(Euclidean Distance)가 가깝다는 것을 의미함
  - 입력상태 벡터(Input State Vector) : 네트워크 상태벡터

- 속도 상태벡터 : 시간대 개수(24개)\*링크별 속도
- Probe 상태벡터 : 시간대 개수(24개)\*링크별 Probe건수
- 출력상태 벡터(Output State Vector)
- 속도 상태벡터 유사성
- Probe 상태벡터 유사성
- 익일의 링크별/시간대별 통행속도
- 또한, 네트워크 상태벡터는 연결로를 제외한 본선만을 이용하며, 이는 연결로에서 발생하는 국부적 Variation에 의한 영향을 배제함과 동시에 전체 네트워크의 전역(Global) 변화를 탐색하고자 함

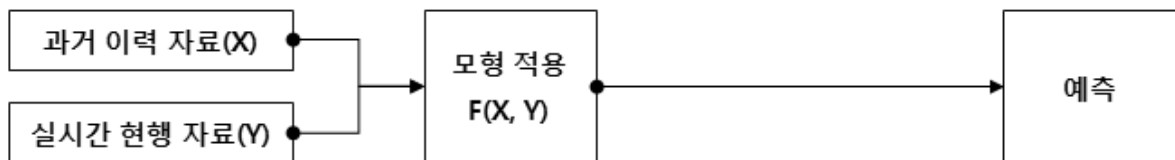


<그림 2-46> 교통예보를 위한 장래 통행속도 예측 개념도

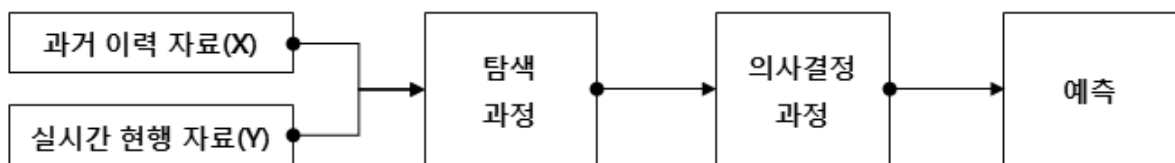
#### 나. 적용 예측기법의 이론적 고찰

- 본 연구의 개발모형은 입력 자료의 자료수집 환경, 자료의 질과 양을 고려하여 연산 수행속도가 빠르면서 모형의 예측력이 우수하고 데이터 탐색을 기반으로 하는 비모수(non-parametric) 기반의 k-최근린 이웃 기법(k-nearest neighbor, 이하 KNN)을 적용하였으며, 이 기법을 선택한 근거를 아래와 같이 선행 연구들에 대한 이론적 고찰을 통하여 제시하였음
- 먼저, 인위적인 파라미터 모델링의 한계를 극복하기 위한 대안 중 하나가 비모수 기반의 k-최근린 이웃 기법이며, 여기서 최근린 이웃은 현행 상태와 가장 유사한 k개의 최근린 이웃을 의미함. 최근린 이웃은 주어진 현 상태의 상태벡터와 유사한 과거의 상태벡터 그룹을 의미함 (Eubank, 1988)
- 비모수 기반의 KNN 모형은 모형 개발자에 의해 인위적으로 생성된 정보보다는 복잡한 정보의 대한 지식은 대량의 자료에 내재적으로 포함되어 있다고 가정함(Eubank, 1988)
- 모수기반의 모형과 비모수 기반의 KNN 모형에 대한 구조적 차이는 아래의 그림과 같음

##### < 모수기반 모형 >



##### < 비모수기반 KNN모형 >



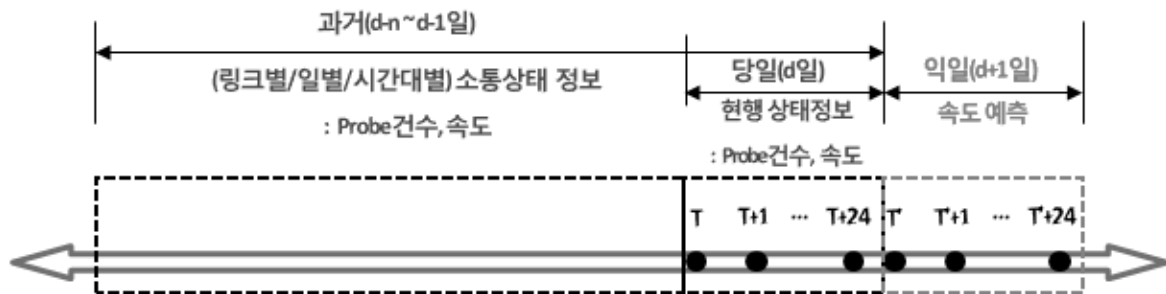
<그림 2-47> 모수기반 모형과 비모수 기반의 KNN모형 구조

- 모수기반의 회귀식은 설명변수에 의한 영향, 즉 파라미터가 장래에도 동일하게 종속변수에 영향을 미친다는 가정을 전제로 한다는 문제를 내포함(Oswald 등, 2000).
- 이에 비해 비모수 기반 모형은 현재 조건과 유사한 과거 관측치를 탐색하여 장래 상태를 추정하는데 적용이 용이함(Smith, 2000). 일반적인 회귀식 모형과 달리 비모수 기반 모형은 입력과 출력 자료간의 관계를 결정하는 것으로서 파라미터의 계산 없이 새로운 관측 자료를 쉽게 모형에 추가할 수 있으며, 자료의 통계적 구조에 대한 가정을 하지 않는 장점이 있음

- 비모수 기반 모형은 과거의 패턴을 바탕으로 장래의 상태를 결정하는 의사결정과정이라고 할 수 있으며, 예측시간대에 상태(state)와 유사한 과거의 상태 벡터로 구성된 의사결정집단인 군집을 지속적으로 탐색하는 동적 군집모형이라 할 수 있음. 이는 비모수 기반 모형이 예측시간대의 군집수를 정의하는 대신 현행(current) 입력상태와 유사한 과거 경우에 대한 군집을 정의하기 때문임
- KNN 모형의 군집(neighborhood)을 정의는 크게 최인접 개수  $k$ 를 정하는 KNN기법과 최인접 개수의  $\pm$ 범위를 설정하는 Kernel 이웃 기법이 있음. KNN 기법을 적용하기 위해서는 방대한 양의 이력자료(즉, 과거의 정보)가 요구되며, 과거 상태와 현재 상태를 이용한 장래 상태의 예측에 있어 다양한 분야에 이용되고 있음
- Robinson(1983), Mulhern와 Caprara(1994)는 혼재된 무질서 상태에서 비선형 시계열 예측 문제를 비모수 기반으로 모형화하였으며, Disbro와 Frame(1989), Mulhern와 Caprara(1994)의 연구는 무질서 상태에서 타 기법보다 비모수 기반의 모형이 장점을 갖는다고 언급하였음
- Karlsson와 Yakawitz(1987)은 비모수 기반으로 강우량을 예측하였으며, Mulhern와 Caprara(1994)는 시장반응(market response)의 예측에 비모수 기반 모형을 적용하였음
- 교통 분야에서 비모수 기반 모형은 교통량 또는 통행시간 예측에 주로 적용되었으며, Davis와 Nihan(1991), Smith et al.(2002), 그리고 Sun 등(2003)은 비모수 기반 모형을 단일시간대(Single interval) 교통량 예측에 적용하였으며, Qi와 Smith(2004)는 유고 지속시간을 예측하였음
- Chang et al.(2010.)은 BMS 이력자료와 실시간 통행시간 자료를 이용하여 버스 정류장간 다중시간대 출발지 기준 경로 통행시간을 예측하였으며, 우수한 결과를 도출하였음
- 비모수 기반의 KNN 모형의 단점으로  $k$ 개의 최근린 이웃을 구축하기 위한 탐색과정이 요구되기 때문에 연산 수행 속도가 느리다는 지적이 있었으나, 이러한 단점은 IT와 컴퓨터의 연산기술 첨단화로 극복되었고 실시간 적용이 가능한 수준에 이미 도달하였음(Smith and Oswald, 2003, Chang et al. 2010)

#### 다. 개발모형의 예측 문제 정의

- 장래 구간 통행속도를 예측하기 위하여 구축한 과거 이력자료 DB(Probe건수, 속도)의 시·공간적 범위는 아래와 같음
  - 시간적 범위 : 데이터 수집기간(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
    - 구축일 개수 : 총 123일
    - 시간대 개수 : 총 24시간대 (01~24시)
  - 공간적 범위 : 서울특별시 강남 일대 → 대상 링크개수(총 255개)
- 구축된 과거 이력자료 DB(Probe건수, 속도)에 따른 예측의 시·공간과 영역을 아래와 같이 정의하였음
  - 현행 소통상태 : 선택된 과거의 당일(d일)로서 링크별/시간대별 Probe건수와 속도 상태벡터
  - 과거 소통상태 : 선택된 과거의 당일(d일)보다 이전(d-n ~ d-1일)의 링크별/일별/시간대별 Probe건수와 속도 상태벡터
  - 예측 영역
    - 공간의 영역 : 과거 이력자료의 공간적 범위에서 예측을 위하여 선택된 네트워크(링크 집합)
    - 시간의 영역 : 선택된 과거의 당일(d일) 기준으로 익일(d+1)의 링크별/시간대별 속도 예측



<그림 2-48> 개발모형의 예측 영역

- 따라서, 본 연구의 개발모형에 대한 예측 문제는 아래의 식으로 정의함

Given  $X_p^s(d, t)$

Predict  $LV_s(d+1, t)$ ,  $\forall s$  with min. of uncertainty

여기서,  $X_p^s(d, t)$ : 링크집합( $S$ )와 입력변수( $p$ : probe건수와 속도)에 대한 당일( $d$ )의 시간대( $t$ )별 현행 상태벡터

$LV_s(d+1, t)$ : 링크집합( $S$ )에 대한 익일( $d+1$ )의 시간대( $t$ )별 통행속도(km/h)



## 라. 본 연구의 예측모형 구조

- 본 연구의 예측모형은 과거 이력자료 DB(Probe건수, 속도)를 이용하여 경험적 지식을 기반으로 하는 상향식(Bottom-up)체계 의사결정모형으로서 아래와 같이 ①KNN 단계, ②예측 단계로 구성됨

- ① KNN 단계는 과거 이력자료 DB(Probe건수, 속도)에서 현행 교통상태와 유사한 과거 교통상태를 탐색·추출하여 장래 교통상태를 설명하기 위한 의사결정집단인  $k$ 개의 최근린 이웃( $k$ -nearest neighbor, 이하 KNN)을 구성하는 일종의 정보추출 및 구축 모형으로 이용됨. 이상의 과정을 달성하기 위하여 ①입·출력 상태벡터(State vector), ②현행 상태와 과거 상태의 유사성(Similarity)을 설명하기 위한 상태거리(State metric), ③탐색된 유사성이 높은 과거 상태를 추출하고 KNN을 구축하기 위한 자료구조(Data structure)가 설정됨. 설정된 3개 구성요소는 KNN 정보추출 및 구축 모형을 구성하게 되며, 개발된 KNN구축 알고리즘으로 구현함
- ② 예측 단계는 이전 단계에서 구축된  $k$ 개의 군집 집단들 중 최종 의사결정 집단을 결정하고, 최종 의사결정 집단을 구성하는  $m$ 개 구성원( $m < k$ ,  $m > 1$ )의 장래 링크별 소통 상태를 결정하고 구간 통행속도를 예측하는 단계임. 해당 단계는 최종 의사결정 집단 선택 모형과 장래 구간통행속도 예측모형으로 구성되며, 최종의사결정 집단 선택모형은 예측 단계에서 선택된 군집들 중 현행 상태와 가장 유사한 군집부터 군집순위를 정하여 현행 상태와 가장 유사한 군집 집단을 선택함. 장래 구간통행속도 예측모형은 최종적으로 선택된 군집 집단의  $m$ 개의 구성원에 해당하는 과거 상태와 KNN단계에서 산정한 개별 구성원의 상태거리를 기반으로 하는 유사성을 이용하여 장래 통행속도를 예측함

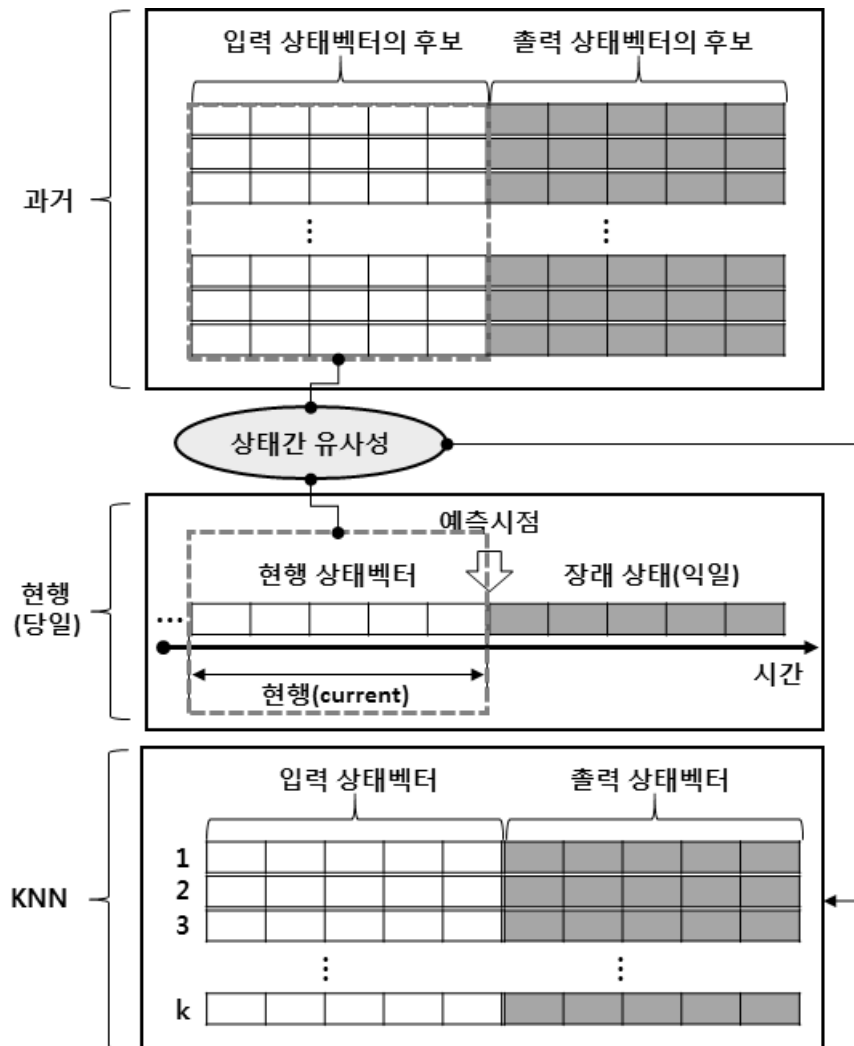


<그림 2-49> 본 연구의 예측모형 구조

## 마. KNN 단계

### 1) 상태벡터의 정의

- KNN 구축을 위한 상태벡터는 현행 상태벡터, 입력 상태벡터, 그리고 출력 상태벡터로 구성되며, 현행 상태벡터는 예측시점 바로 이전의 상태의 시계열을 의미함
- 여기서, KNN은  $k$ 개의 입력 상태벡터와 출력 상태벡터로 구성되며, KNN의 입력 상태벡터와 출력 상태벡터는 과거의 이력자료에 포함된 모든 입력 상태벡터의 후보와 출력상태 벡터의 후보 중 현행 상태벡터와 유사한  $k$ 개의 선택된 후보임
- 본 연구의 KNN 구축을 위한 상태벡터의 개념은 아래의 그림과 같음



<그림 2-50> KNN 구축을 위한 상태벡터의 개념

- KNN단계는 KNN모형을 기반으로 현행 상태벡터와 유사한 과거의  $k$ 개의 상태벡터를 탐색하여  $k$ 개의 이웃을 추출하여 자료 구축을 수행하는 단계로서 먼저 KNN을 구축하기 위해서는 상태벡터가 정의되어야 함
- 상태벡터는 현재 상태를 설명하기 위한 현행 상태벡터( $X_c$ ), 현행 상태의 탐색 대상이 되는 과거 상태벡터( $X_k$ ) 그리고 이웃으로 선정된 과거 상태벡터( $X_k$ )의 현재 상태의 미래에 해당하는 이력상태인 출력상태 벡터( $O_k$ )로 정의함
- 본 연구의 독립변수는 통행속도(kph)와 Probe건수(건/시)로 구성되며, 독립변수를  $p$ 로 정의함.  $p = \{\text{구간 통행속도, Probe건수}\} = \{v, b\}$ 으로 구성되는 집합으로  $X_{c,p}^s$ 로 정의함

#### ① 현행 상태벡터 정의

- 구간 통행속도( $v$ )의 현행 상태벡터( $X_{c,v}$ )는 입력변수와 주어진 현행 시간대의 개수( $l$ ), 그리고 구간의 개수( $s$ )를 고려하면 상태벡터의 차원은  $l \times s$ 이며, 이를 수식화 하면 아래와 같음

$$X_{c,v}^s = \begin{bmatrix} x_{c,v}^1(t) & , x_{c,v}^2(t) & , x_{c,v}^3(t) & , ..., x_{c,v}^s(t) \\ x_{c,v}^1(t-1) & , x_{c,v}^2(t-1) & , x_{c,v}^3(t-1) & , ..., x_{c,v}^s(t-1) \\ \dots & \dots & \dots & \dots & \dots \\ x_{c,v}^1(t-l) & , x_{c,v}^2(t-l) & , x_{c,v}^3(t-l) & , ..., x_{c,v}^s(t-l) \end{bmatrix}$$

- Probe건수의 현행 상태벡터( $X_{c,b}$ )는 구간속도 현행상태 벡터와 동일한 차원이며, 이를 수식화 하면 아래와 같음

$$X_{c,b}^s = \begin{bmatrix} x_{c,b}^1(t) & , x_{c,b}^2(t) & , x_{c,b}^3(t) & , ..., x_{c,b}^s(t) \\ x_{c,b}^1(t-1) & , x_{c,b}^2(t-1) & , x_{c,b}^3(t-1) & , ..., x_{c,b}^s(t-1) \\ \dots & \dots & \dots & \dots & \dots \\ x_{c,b}^1(t-l) & , x_{c,b}^2(t-l) & , x_{c,b}^3(t-l) & , ..., x_{c,b}^s(t-l) \end{bmatrix}$$

#### ② 과거 상태벡터 정의

- $k$ 번째 과거 구간 통행속도( $v$ )의 과거 상태벡터( $X_{h,v,k}$ )는 현행 구간속도 상태벡터와 동일 차원에서 아래의 수식과 같음

$$X_{h,v,k}^s = \begin{bmatrix} x_{h,v,k}^1(t) & , x_{h,v,k}^2(t) & , x_{h,v,k}^3(t) & , ..., x_{h,v,k}^s(t) \\ x_{h,v,k}^1(t-1) & , x_{h,v,k}^2(t-1) & , x_{h,v,k}^3(t-1) & , ..., x_{h,v,k}^s(t-1) \\ \dots & \dots & \dots & \dots & \dots \\ x_{h,v,k}^1(t-l) & , x_{h,v,k}^2(t-l) & , x_{h,v,k}^3(t-l) & , ..., x_{h,v,k}^s(t-l) \end{bmatrix}$$

- Probe건수의  $k$ 번째 과거 상태벡터( $X_{h,b,k}$ )는 현행 상태벡터( $X_{c,b}$ )와 동일 차원이며, 아래의 수식과 같음

$$X_{h,b,k}^s = \begin{bmatrix} x_{h,b,k}^1(t) & , x_{h,b,k}^2(t) & , x_{h,b,k}^3(t) & , ..., x_{h,b,k}^s(t) \\ x_{h,b,k}^1(t-1) & , x_{h,b,k}^2(t-1) & , x_{h,b,k}^3(t-1) & , ..., x_{h,b,k}^s(t-1) \\ \dots & \dots & \dots & \dots & \dots \\ x_{h,b,k}^1(t-l) & , x_{h,b,k}^2(t-l) & , x_{h,b,k}^3(t-l) & , ..., x_{h,b,k}^s(t-l) \end{bmatrix}$$

### ③ 출력 상태벡터 정의

- 앞서 현행 상태벡터와 과거 상태벡터가 정의되면 출력 상태벡터를 정의해야 하며, 출력상태 벡터는 구간별 통행속도 출력상태 벡터로 정의됨
- 구간별 통행속도의 출력상태 벡터( $O_{h,v,k}^s$ )는 시간대( $t+1$ )에서 시간대( $t+m$ )까지의 모든 링크에 대하여 통행속도를 출력하게 되며,  $k$ 개로 구성되는 구간별 통행속도의 출력상태 벡터는  $m \times l$ 의 차원에서 아래의 수식과 같음

$$O_{h,v,k}^s = \begin{bmatrix} x_{h,v,k}^1(t) & , x_{h,v,k}^2(t) & , x_{h,v,k}^3(t) & , ..., x_{h,v,k}^s(t) \\ x_{h,v,k}^1(t+1) & , x_{h,v,k}^2(t+1) & , x_{h,v,k}^3(t+1) & , ..., x_{h,v,k}^s(t+1) \\ \dots & \dots & \dots & \dots & \dots \\ x_{h,v,k}^1(t+m) & , x_{h,v,k}^2(t+m) & , x_{h,v,k}^3(t+m) & , ..., x_{h,v,k}^s(t+m) \end{bmatrix}$$

### 2) 상태간 거리 정의

- 현행 상태벡터( $X_{c,p}^s$ )와 유사한 과거(historical) 상태벡터( $X_{h,p,k}^s$ )를 탐색하여 출력 상태벡터( $O_h$ )로 구성된 군집을 구성하기 위해서는 탐색 알고리즘을 이용하게 되며, 탐색의 과정에서 일반적으로 현행 상태벡터( $X_{c,p}^s$ )와 과거 상태벡터( $X_{h,p,k}^s$ )간의 상태간 거리( $L_m$ )를 이용함
- 상태간 거리( $L_m$ )은  $m=1$ 인 경우 맨하탄 거리(Manhattan distance,  $MD$ ),  $m=2$ 이면 유클리

드 거리(Euclidean distance,  $ED$ )로 알려져 있음

- 본 연구에서는 상태벡터간 거리가 증가할수록 상태거리가 증가하는 유클리드 거리를 적용하도록 함. 즉, 유클리드 거리는 맨하탄 거리와 같이 상태벡터 간의 절대거리가 같은 경우라도 상태벡터의 요소간 거리를 고려하기 때문에 유사한 상태간의 관계를 설명할 수 있음

$$L_m = (\sum_i^d |x_i - y_i|^m)^{1/m}, \quad d \geq i, \quad m = 2$$

- 또한, 본 연구에서는 모형의 구조화와 효과적인 연산과정을 설명하기 위하여 행렬의 곱( $\cdot$ )에 대한 연산을 아래와 같이 정의하며, 이는 실제 알고리즘의 구현에 있어 매우 효율적임

$$A \cdot B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} a \times 1 & b \times 2 \\ c \times 3 & d \times 4 \end{pmatrix}$$

### 3) KNN의 DB구조

- KNN 군집구축 알고리즘을 수행하기 위해서는 입·출력 상태벡터를 저장하기 위한 공간이 필요하며, 저장 공간은  $k=[1, 2, 3, \dots, K, \dots, k]$  이웃에 대하여 각각 ①입력자료 상태벡터 즉, 추출된 현행 상태벡터와 유사한 과거 상태벡터, ②출력 상태벡터 그리고 ③  $\mathbb{R}^p$ 에서 대표 거리로 구성됨
- KNN DB공간은 KNN 군집구축 알고리즘의 탐색결과 저장 공간이면서 동시에 Clustering 단계의 입력 자료로 이용됨

$i$	[입력상태 벡터]	[출력 상태벡터]
1	$[X_{h,v,1}, X_{h,b,1}]$	$[O_{h,v,1}^s], [ED_1]$
2	$[X_{h,v,2}, X_{h,b,2}]$	$[O_{h,v,2}^s], [ED_2]$
...	....	....
K	$[X_{h,v,K}, X_{h,b,K}]$	$[O_{h,v,K}^s], [ED_K]$
...	....	....
k	$[X_{h,v,k}, X_{h,b,k}]$	$[O_{h,v,k}^s], [ED_k]$

<그림 2-51> KNN 군집구축 알고리즘의 DB구조

#### 4) KNN 군집 구축 알고리즘

- (현행, 과거)상태벡터, 상태거리, 그리고 KNN DB가 정의되면, KNN 군집 구축 알고리즘을 이용하여 KNN DB를 구축하게 됨
- 본 연구의 KNN 군집 구축 알고리즘은 다음과 같음

Member=0

For each  $[X_{h,v,j}, X_{h,b,j}]$ ,  $j \in N$ ,  $N$ : 이력자료의 개수

Calculate  $ED_j$  Between  $[X_{h,v,j}, X_{h,b,j}]$  and  $[X_{c,v}, X_{c,b}]$

IF Member  $< k$

Update  $[X_{h,v,j}, X_{h,b,j}], [O_{h,v,j}^s], [ED_j]$

Member = Member + 1

Else

If  $ED_j < ED_{\max}$  then

(여기서,  $ED_{\max} = \max\{ED_1, ED_2, \dots, ED_i, \dots, ED_k\}$

Delete  $[X_{h,v,i}, X_{h,b,i}], [O_{h,v,i}^s], [ED_i]$  from KNN DB

(여기서,  $ED_i$ 는  $ED_{\max}$ 에 해당하는 자료임)

Update  $[X_{h,v,j}, X_{h,b,j}], [O_{h,v,k}^s], [ED_j]$

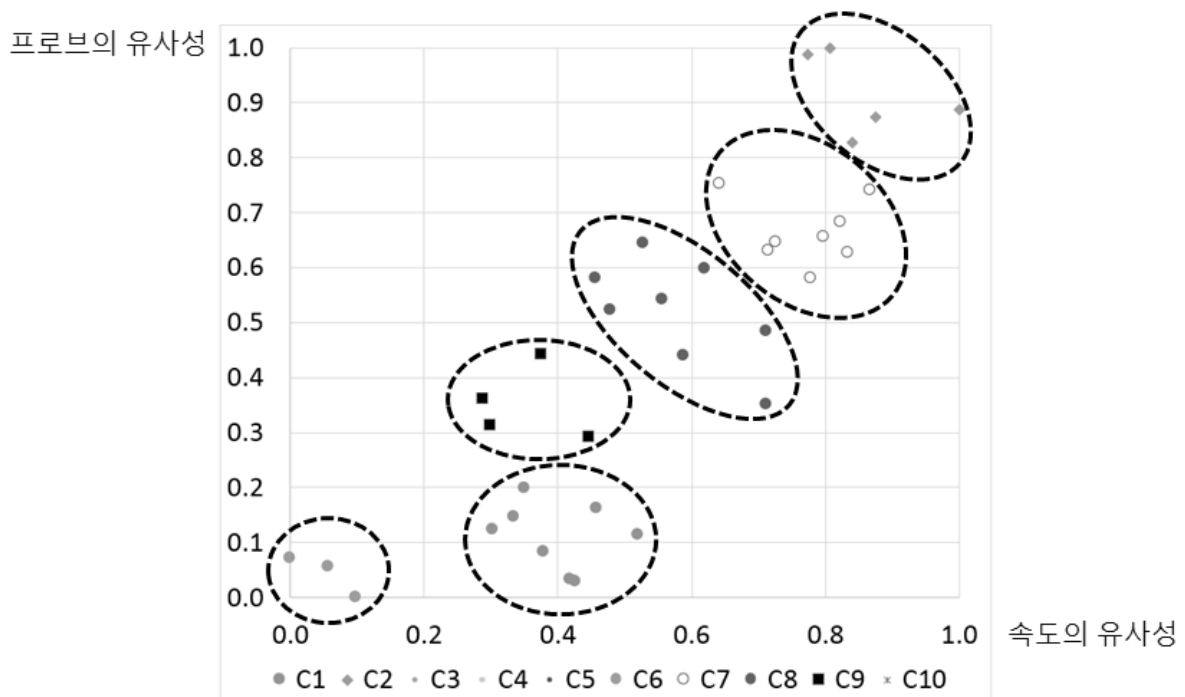
Search  $ED_{\max}$  in KNN DB

<그림 2-52> KNN 군집 구축 알고리즘

## 바. 예측 단계

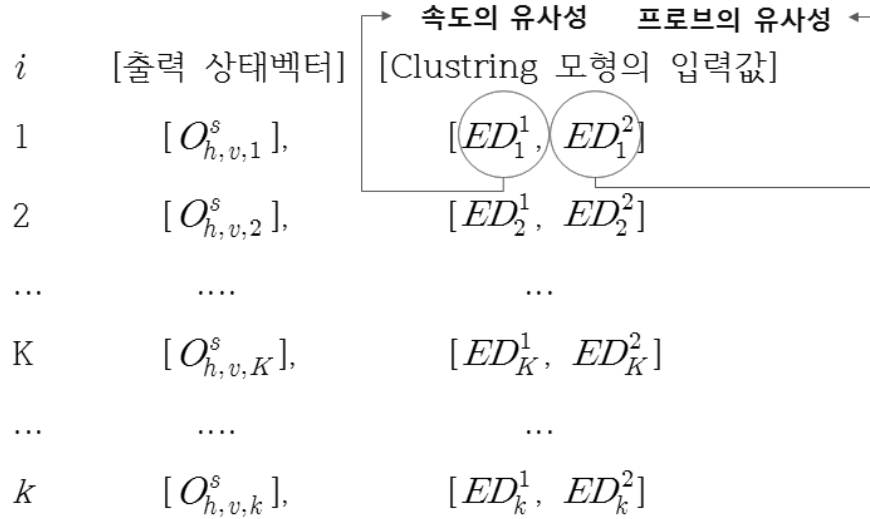
### 1) 입/출력 자료의 정의

- 예측 단계에서는 앞서 KNN 모형에서 구축한 KNN DB의 출력값 즉, 1차 의사결정 집단을 입력값으로 하여 군집화의 과정을 통해 다수의 의사결정 그룹을 결정함으로써 장래의 불확실성을 줄여주는 역할을 수행함
- 즉, 두 독립변수(Probe건수, 속도)의 상태간 거리( $L_m$ )를 0~1로 정규화하여 산출한 <그림 2-27>과 같은 2차원 그래프{x축: 속도의 유사성( $ED_1$ ), y축: Probe의 유사성( $ED_2$ )}에서 각 군집의 중심점과 원점에서 거리가 가장 가까운(유사성이 가장 높은) 최종의사결정 군집(C6)을 선택하기 위한 역할을 수행하게 됨



<그림 2-53> 두 변수의 상태간 거리에 따른 의사결정 그룹 군집화

- 따라서, 예측 단계의 입력값은 KNN단계의 출력값을 이용하게 되며, 즉, 의사결정 그룹의 군집화를 수행하기 위한 입력값은 아래의 그림과 같음



<그림 2-54> 예측 단계의 입력값

- 예측 단계의 출력값은 <그림 2-28>과 같이 입력값  $k$ 개의 이웃에 대한 출력 상태벡터와 두 변수의 상태간 거리에 따른 군집의 개수 및 각 군집의 중심점에 대한 정보를 이용하여 최종의 사결정을 위한 우선 순위를 정함
- 즉,  $k$ 개의 이웃 중 군집의 중심점( $c_n$ )은 최종의사결정 그룹을 결정하는데 이용되며, KNN 단계의 출력값인 상태거리( $d_k$ )는 장래 교통상태의 예측에 있어 상태 결정에 대한 가중치로 이용됨



$i$	출력 상태벡터	상태거리	군집( $N$ )	중심좌표( $c_n$ )
1	$O_{v,1}$	$d_1$	1	$c_1$
2	$O_{v,2}$	$d_2$		
3	$O_{v,3}$	$d_3$		
4	$O_{v,4}$	$d_4$	2	$c_2$
5	$O_{v,5}$	$d_5$		
6	$O_{v,6}$	$d_6$		
7	$O_{v,7}$	$d_7$		
...	...	...	...	...
$k-2$	$O_{v,k-2}$	$d_{k-2}$	n	$c_n$
$k-1$	$O_{v,k-1}$	$d_{k-1}$		
$k$	$O_{v,k}$	$d_k$		

&lt;그림 2-55&gt; 예측 단계의 출력값

## 2) 최종의사결정 및 장래 통행속도 예측

- 예측 단계는 <그림 2-28>과 같이  $k$ 개의 출력값을 이용하여 의사결정 우선순위에 의하여 의사결정 집단을 결정하고, KNN 단계의 출력값과 결정된 의사결정 우선순위를 이용하여 최종의사결정 군집을 선택함
- 즉, 최종 의사결정 방법론은 유사한 과거의 경험에 기초해서 유사하거나 새로운 문제를 해결하는 사례기반 추론(Case-based reasoning, 이하 CBR)에 기초하며, <그림 2-29>과 같이 구축된  $k$ 개의 군집 중 현행상태와 비교하여 속도와 Probe의 상태간 거리가 가장 가까운(유사성이 가장 높은) 군집을 최종의사결정 군집으로 선택함

군집(N)	원점과의 거리(D)	우선순위	비고
1	D1	1	여기서, 우선순위는 원점과의 거리가 가장 가까운 순서임  최종의사결정 군집 선택
2	D2	2	
...	...	...	
N	D3	N	

<그림 2-56> 최종 의사결정 군집 선택

- 선택된 최종 의사결정 군집 내 각 구성원들의 익일(d+1일)에 대한 링크별/시간대별 통행속도 자료를 이용하여 속도와 Probe 유사성 그래프에서의 원점과 선정된 Member와의 거리의 역수로 가중치를 부여하여 통행속도( $LS^s$ )를 산정함
- 이는 현행의 교통류 상태와 가장 유사한 상태에 장래 상태에 대한 의사결정 권한을 더 부여하는 방법으로 직접 평균을 이용하는 것 보다 좋은 예측력을 보였음
- 원점 상태간 거리가 가까운 member에게 의사 결정권을 더 부여

$$LS^s = \frac{\sum_{i=1}^m (O_v^s(i) \times UD_{c,z}(i))}{\sum_{i=1}^m (1/UD_{c,z}(i))}, \quad UD_{c,z}(i) > 0$$

여기서,  $UD_{c,z}(i)$  : i번째 의사 결정자의 원점과의 상태간 거리

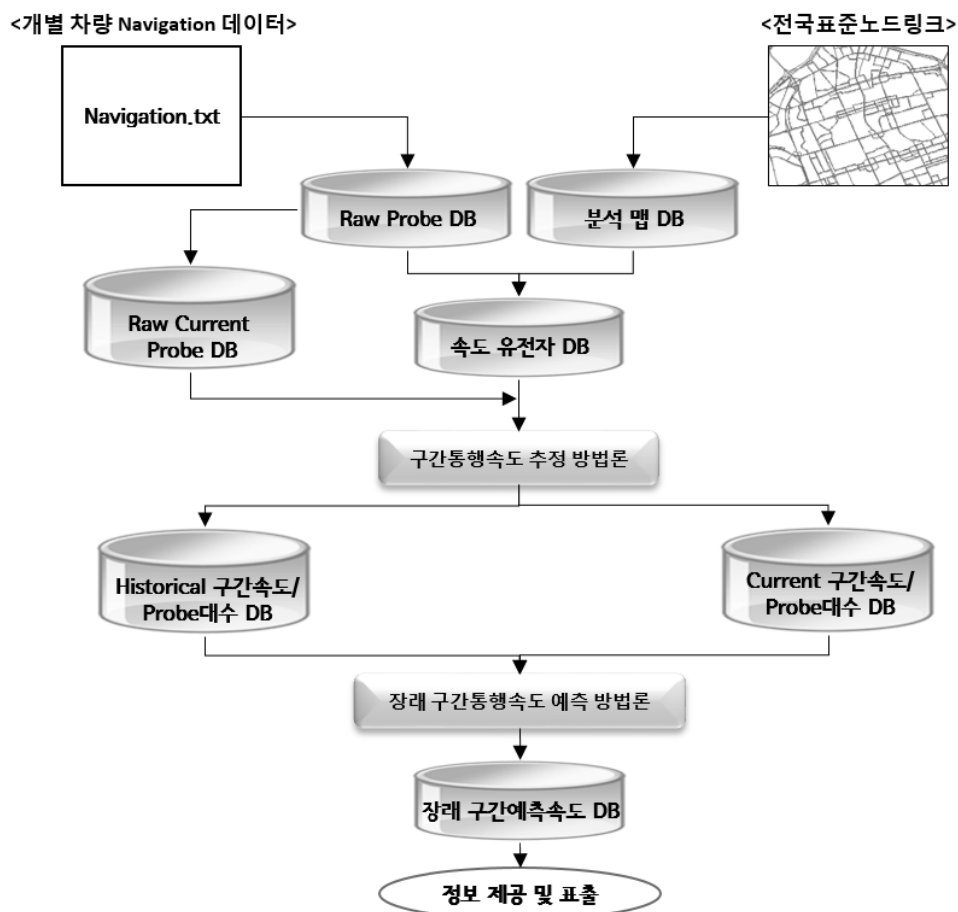
$O_v^s(i)$  : 모든 링크에 대한 i번째 의사 결정자(입력자료)에 대한 익일의 모든 링크의 통행속도

## 제5절 교통예보를 위한 시스템 구축 방안

### 1. 개념적 DB설계

#### 가. 시스템 DB 구조 설계

- 본 연구의 DB 구조는 먼저 개별 차량의 Navigation에서 추출되는 이동계적 데이터(속도, Probe대수)를 Raw Probe DB로 구축하고, 입력 및 결과 정보를 표출하기 위한 분석 Map DB를 구축함
- 위의 Raw Probe DB와 분석 Map DB를 이용하여 구축하는 속도 유전자 DB를 기반으로 Historical DB(이력자료 DB)와 Current DB(현재자료 DB)를 구축하고, 이 두 DB를 이용하여 장래 구간별 통행속도를 예측하고 그 결과를 DB로 구축함



<그림 2-57> DB구조 설계 개념도



#### 다. DB 테이블 정의서 구축

- 본 연구의 DB구축 위한 각 데이터의 테이블 정의서는 아래와 같음
  - 개별 Navigation 차량 수집 데이터(기본 데이터)
  - 실시간 및 단말기 수집 교통 정보 텍스트 파일

<표 2-15> 개별 Navigation 차량 수집 데이터 테이블(기본 데이터)

Column	설명	Data Type	PK	FK
auth_key	단말기 ID	Datetime		
link_time	수집 시간	Integer		
from_link	진입 링크	Integer		
to_link	진출 링크	Integer		
speed	속도	Integer		

- 표준노드링크 DB 테이블
- Navigation 차량 수집 데이터의 기준 네트워크

<표 2-16> 표준노드링크의 노드 테이블

No	컬럼명	Type	설명	코드	코드정보
1	NODE_ID	Integer	노드 ID	-	-
2	NODE_TYPE	Char	노드유형	101 102 103 104 108	도로교차점 도로시·종점 속성변화점 도로시설물 IC및JC
3	NODE_NAME	Char	교차로 명칭	-	-
4	X	Double	X 좌표		
5	Y	Double	Y 좌표		
6	SIDO_ID	Integer	시도 ID		
7	SIGUN_ID	Integer	시군 ID		
8	SIGUNGU_ID	Integer	시군구 ID		
9	EMD_ID	Integer	읍면동 ID		

&lt;표 2-17&gt; 표준노드링크의 링크 테이블

No	컬럼명	Type	설명	코드	코드정보
1	LINK_ID	Integer	링크 ID	-	-
2	F_NODE	Integer	시작노드	-	-
3	T_NODE	Integer	종료노드	-	-
4	ROAD_RANK	Char	도로등급	101 102 103 104 105 106 107	고속국도 도시고속국도 일반국도 특별·광역시도 국가지원지방도 지방도 시·군도
5	ROAD_TYPE	Char	도로유형	000 001 002 003 004	일반도로 고가차도 지하차도 교량 터널
6	ROAD_NO	Char	도로번호	-	-
7	NAME	Char	도로명	-	-
8	LENGTH	Integer	도로길이	-	-
9	CONNECT	Char	연결로 코드	000 101 102 103 104 105 106 107	연결로 아님 고속국도 연결로 도시고속국도 연결로 일반국도 연결로 특별·광역시도 연결로 국가지원지방도 연결로 지방도 연결로 시·군도 연결로
10	LANES	Integer	차선수	-	-
11	MAX_SPD	Integer	최고속도	-	-
12	LENGTH	Double	링크길이(km)	-	-
13	EMD_ID	Integer	읍면동 ID	-	-

## - 분석 Map DB 테이블

&lt;표 2-18&gt; 분석 Map의 링크 테이블

No	컬럼명	Type	설명	코드	코드정보
1	K_LINK_ID	Integer	분석 Map 링크 ID	-	-
2	F_NODE	Integer	시작노드	-	-
3	T_NODE	Integer	종료노드	-	-
4	ROAD_RANK	Char	도로등급	101 102 103 104 105 106 107	고속국도 도시고속국도 일반국도 특별·광역시도 국가지원지방도 지방도 시·군도
5	ROAD_TYPE	Char	도로유형	000 001 002 003 004	일반도로 고가차도 지하차도 교량 터널
6	ROAD_NO	Char	도로번호	-	-
7	ROAD_NAME	Char	도로명	-	-
8	LENGTH	Integer	도로길이	-	-
9	CONNECT	Char	연결로 코드	000 101 102 103 104 105 106 107	연결로 아님 고속국도 연결로 도시고속국도 연결로 일반국도 연결로 특별·광역시도 연결로 국가지원지방도 연결로 지방도 연결로 시·군도 연결로
10	DATA_INFO	Char	데이터 이력	0 1	신규 수정

&lt;표 2-19&gt; 분석 Map의 노드 테이블

No	컬럼명	Type	설명	코드	코드정보
1	NODE_ID	Integer	표준노드링크의 노드 ID	-	-
2	NODE_TYPE	Char	노드유형	101 102 103 104 108	도로교차점 도로시·종점 속성변화점 도로시설물 IC및JC
3	NODE_NAME	Char	교차로 명칭	-	-

&lt;표 2-20&gt; 분석 Map의 링크 맵핑 테이블

No	컬럼명	Type	설명	코드	코드정보
1	K_LINK_ID	Integer	분석 Map 링크 ID	-	-
2	LINK_ID	Integer	표준노드링크의 링크 ID	-	-
3	SEQ	Integer	맵핑 순서	-	-

- Raw Probe DB(Probe 대수 포함)

&lt;표 2-21&gt; Raw Probe DB 테이블

Column	설명	Data Type	PK	FK
year	수집 년	Varchar	○	
month	수집 월	Varchar	○	
day	수집 일	Varchar	○	
form_link_id	진입 링크 ID	Integer	○	○
time	시간	Integer	○	
week_time	구분(주중/주말/특송)	Integer	○	
to_link	진출 링크 ID	Integer		
direction	이동류(1:좌회전, 2:직진, 3:우회전)	Integer		
vehicle_count	probe 대수	Double		
spd_0km_h	속도 0km 빈도수	Double		
~	~ (1단위로 0부터 150km까지 컬럼 생성)	Double		
spd_150km_h	속도 150km 이상 빈도수	Double		



## - 속도 유전자 지도 DB 테이블

&lt;표 2-22&gt; 속도 유전자 지도 DB 테이블

Column	설명	Data Type	PK	FK
year	수집 년	Varchar	○	
month	수집 월	Varchar	○	
day	수집 일	Varchar	○	
form_link_id	진입 링크 ID	Integer	○	○
time	시간	Integer	○	
week_time	구분(주중/주말/특송)	Integer	○	
to_link	진출 링크 ID	Integer		
direction	이동류(1:좌회전, 2:직진, 3:우회전)	Integer		
p_spd_0	속도 0km 정규화 비율(0~1)	Double		
~	~ (1단위로 0부터 150km까지 컬럼 생성)	Double		
p_spd_150	속도 150km 이상 정규화 비율(0~1)	Double		

## - Historical 구간속도/Prob대수 DB 테이블

&lt;표 2-23&gt; Historical 구간속도/Prob대수 DB 테이블

Column	설명	Data Type	PK	FK
year	수집 년	Varchar	○	
month	수집 월	Varchar	○	
day	수집 일	Varchar	○	
form_link_id	진입 링크 ID	Integer	○	○
time	시간	Integer	○	
week_time	구분(주중/주말/특송)	Integer	○	
to_link	진출 링크 ID	Integer		
direction	이동류(1:좌회전, 2:직진, 3:우회전)	Integer		
h_probe	probe 대수(건)	Double		
h_spd	가중조화평균속도(km/h)	Double		

- Current 구간속도/Prob대수 DB 테이블

<표 2-24> Current 구간속도/Prob대수 DB 테이블

Column	설명	Data Type	PK	FK
year	수집 년	Varchar	○	
month	수집 월	Varchar	○	
day	수집 일	Varchar	○	
form_link_id	진입 링크 ID	Integer	○	○
time	시간	Integer	○	
week_time	구분(주중/주말/특송)	Integer	○	
to_link	진출 링크 ID	Integer		
direction	이동류(1:좌회전, 2:직진, 3:우회전)	Integer		
c_probe	probe 대수(건)	Double		
c_spd	가중조화평균속도(km/h)	Double		

- 장래 구간 예측속도 DB 테이블

<표 2-25> 장래 구간 예측속도 DB 테이블

Column	설명	Data Type	PK	FK
year	수집 년	Varchar	○	
month	수집 월	Varchar	○	
day	수집 일	Varchar	○	
time	시간	Integer	○	
week_time	구분(주중/주말/특송)	Integer	○	
direction	이동류(1:좌회전, 2:직진, 3:우회전)	Integer		
f_probe	probe 대수(건)	Double		
f_spd	속도(km/h)	Double		

## 2. 개발 방법론의 시스템 구축 방안

### 가. 개발 방법론의 개선방향

- 본 연구에서 제안한 개발 방법론은 이력자료를 기반으로 일일의 장래 소통상태의 예측에 적용하기 위함임. 이를 위하여 본 연구에서 제안한 모형은 (ITS 예측분야의 관점에서)상태의 시간적 진화(Temporal evolution)에 대한 전반적인 경향을 과거 상태와 현행상태의 유사성을 기반으로 분석함으로써 장래 상태를 예측함
- 따라서 향후 모형의 고도화를 위해서는 ①빅데이터에 검색에 소요되는 모형의 연산시간의 단축이 요구되며, ②계절적 특성과 여가 특성을 보이는 Network에 대한 이력자료의 시간 종속성에 대한 연구가 추가되어야 할 것으로 판단됨. ③그리고 이력자료에 포함된 시간적으로 종속적이지 않지만 특정 계절에 포함된 행사에 대한 추가 연구가 필요할 것으로 판단됨. ④가까운 시간대의 단기에측과 중/장기 예측사이에 발생하는 예측구간에 대한 예측에 대한 연구가 필요함
- 첫째, 빅데이터의 검색에 소요되는 연산시간은 향후 구축되는 시스템의 성능에 맞게 조정되어야 할 것으로 판단됨. 현재 빅데이터 기반의 연산체계는 매우 높은 검색능력과 연산성능을 가지고 있어 본 연구에서 제안한 기술의 적용에는 무리가 없을 것으로 판단됨. 추가적으로 대용량 데이터 기반의 Knowledge discovery 기법을 이용한 ITS 예측분야에 있어 검색에 의한 연산저하는 극복되었음(Smith and Oswald, 2003; Chang et al., 2010)
- 둘째, 이력자료와 현행자료간의 시간종속성은 검색자료의 양과 더불어 실제 장래 상태의 예측에 영향을 크게 미칠 수 있는 변수임. 그러나 이는 수년간의 방대한 양의 이력자료가 구축되어야 가능한 분야임. 따라서 향후 시스템 구축이후 이에 대한 추가적인 연구가 필요함
- 셋째, 행사(지방축제, 장날, 단풍, 스키시즌, 하계휴가 등)와 자연재해(폭설, 폭염, 태풍 등)와 같이 특정일에 시간적으로 종속적이지 않지만 Season에 종속성이 강한 변수에 대한 개선방향이 필요하며, 이는 향후 수년간의 방대한 양의 이력자료가 구축되어야 가능한 분야임
- 넷째, 가까운 시간대의 단기에측과 중/장기 예측사이에 발생하는 예측구간은 현재의 기술로 예측의 정확도를 담보할 수 없는 상태임. 따라서 향후 본 연구에서 개발된 기술을 활용한 추가적인 연구를 통하여 Short-term ⇨ Mid-term ⇨ Long-term에 대한 연속적인 예측이 가능하도록 해야 할 것으로 판단됨

#### 나. 개발 방법론을 적용하기 위한 시스템 요구사항

- 본 연구에서 제안한 기술은 기존의 ITS에서 적용해 오면 현행자료기반의 단기 예측기법에서 벗어나 방대한 양의 Raw data와 1/2차 가공된 DB를 관리/저장해야 함. 따라서 고성능의 검색 및 연산능력을 가진 시스템의 구축이 필요할 것을 판단됨
- 따라서 본 연구에서 개발되는 기술을 적용하기 시스템 요구사항으로서 ①본 연구의 장래예측 기술에 요구되는 방대한 양의 DB 크기, ②정보제공 시스템의 정보제공 요구시간에 적합하게 예측을 수행할 수 있는 예측시스템의 성능, ③다중 접속자에 의한 정보제공시스템의 처리능력 등 다양한 시스템 요구사항을 분석하고 제시하도록 함
- 시스템 요구사항을 정의하고 세부 요구사항을 제시하는 다음의 5개 사항을 고려하여 수행하도록 함
  - 첫째, 개념적 시스템 설계를 통하여 실제 구축되는 시스템에서 요구되는 요구사항을 정의하고, HW/SW 그리고 운영인력에 대한 요구사항을 정의하도록 함
  - 둘째, 개발 모형의 실제 운영이 가능한 자료의 수집범위와 크기에 대한 정의를 통하여 지속적인 자료의 수집이 가능하도록 요구사항을 제시하도록 함
  - 셋째, Big data 기반의 ITS 시스템 전문가 집단의 검토를 통하여 실제 시스템의 구축, 개발 모형의 탑재 시 DB의 구조, DB검색과 추출을 위한 데이터 추론기술 등 전반적인 시스템의 HW와 SW의 최적화에 대한 요구사항을 제시하도록 함
  - 넷째, ITS 표준노드링크의 링크개수는 약 250,000로서 본 연구에서 개발된 방법론은 국내에서 가장 큰 Network 사이즈와 개수 그리고 데이터의 크기를 처리해야 함. 이는 용이하게 Big data를 처리할 수 있는 기존의 슈퍼컴퓨터보다 매우 뛰어난 성능을 보유한 실시간 Hadoop을 도입하고 이를 지원할 수 있는 시스템을 설계해야 함을 의미함
  - 다섯째, 전국을 대상으로 장래 혼잡예보 정보를 다양한 요구수준을 가진 국민에게 제공하기 위해서는 정보제공을 위한 소통정보 제공용 맵과 효율적인 정보제공 콘텐츠가 필요함. 따라서 생성된 장래정보를 효과적으로 제공하기 위한 서비스 방안을 도출함과 더불어 서비스 콘텐츠 개발에 필요한 요구사항을 제시하도록 함

## 제6절 사례분석 및 평가

### 1. 평가 개요

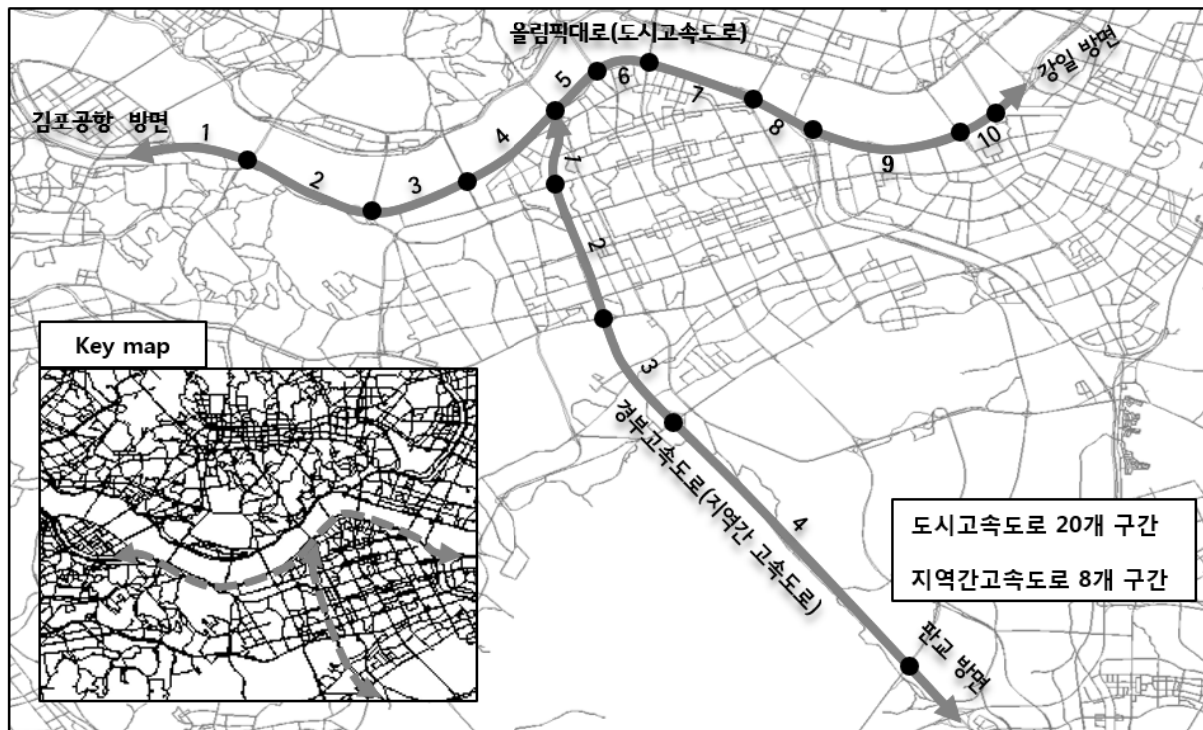
- 본 장에서는 앞서 구축한 링크별/일별/시간대별 과거 이력자료 DB(Probe건수, 속도)를 이용하여 개발된 장래 통행속도 예측모형의 성능 평가를 수행하고자 함
- 성능 평가의 수행과정은 ①사례대상 구간 선정, ②사례대상 구간의 분석DB 구축, ③분석 자료의 특성, ④평가지표 및 분석시나리오 설정, ⑤방법론 적용 및 평가 그리고 소결로 구성됨
- 먼저, 사례대상 구간 선정은 연속류와 단속류가 혼재되어 있는 서울시 강남구 일대의 전체 도로축을 대상구간으로 선정하였음
  - 연속류 구간 : 경부 고속도로(한남대교 ~ 판교JC), 올림픽 대로(원효대교 ~ 올림픽대교)
  - 단속류 구간 : 강남 일대 내부 14개 도로축
- 분석DB 구축은 본 연구의 사례대상 구간에 해당하는 분석 맵의 링크 기준으로 링크별/일별/시간대별 Probe건수와 속도 자료를 구축하였음
  - 대상링크 개수 : 총 255개
  - 구축일 개수 : 총 123개(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
  - 시간대 개수 : 총 24개(01~24시)
- 분석 자료의 특성에서는 링크별/일별/시간대별로 구축되어 있는 Probe건수와 속도 자료에 대한 Historical 데이터 분석을 수행하여 각 자료에 대한 시계열적 Variation을 검토하였음
- 사례분석 및 평가를 위한 지표는 백분율 절대평균오차(MAPE, %: Mean absolute percentage error)와 적중률(%)을 적용하였으며, 분석 시나리오는 예측하고자 하는 일자에 따라 경부고속도로, 올림픽대로, 기타 내부도로로 분석 네트워크로 구분하여 설정하였음. 여기서, 본 예측모형의 평가를 위한 비교모형으로는 현재 ITS시스템에서 많이 이용되고 있는 이력자료 평균(Historical average) 기법을 적용하였음
- 방법론 적용 및 평가에서는 설정된 평가지표를 이용하여 분석 시나리오별로 관측치와의 오차 분석을 수행하여 이를 종합적으로 평가하였음. 마지막으로 소결에서 본 개발모형의 목표 달성과 적용의 측면을 앞서 평가한 결과들을 토대로 설명하였음

## 2. 사례대상 구간 선정

- 본 연구의 개발 방법론(㉔과거 이력자료 DB 구축방법론 개발, ㉕구축된 이력자료와 현행자료를 이용한 (예보수준의)장래 통행속도 예측기술 개발)에 대한 적용 및 평가를 위하여 연속류와 단속류가 혼재되어 있는 서울시 강남구 일대의 전체 도로축을 대상구간으로 선정하였음

### 가. 연속류 구간

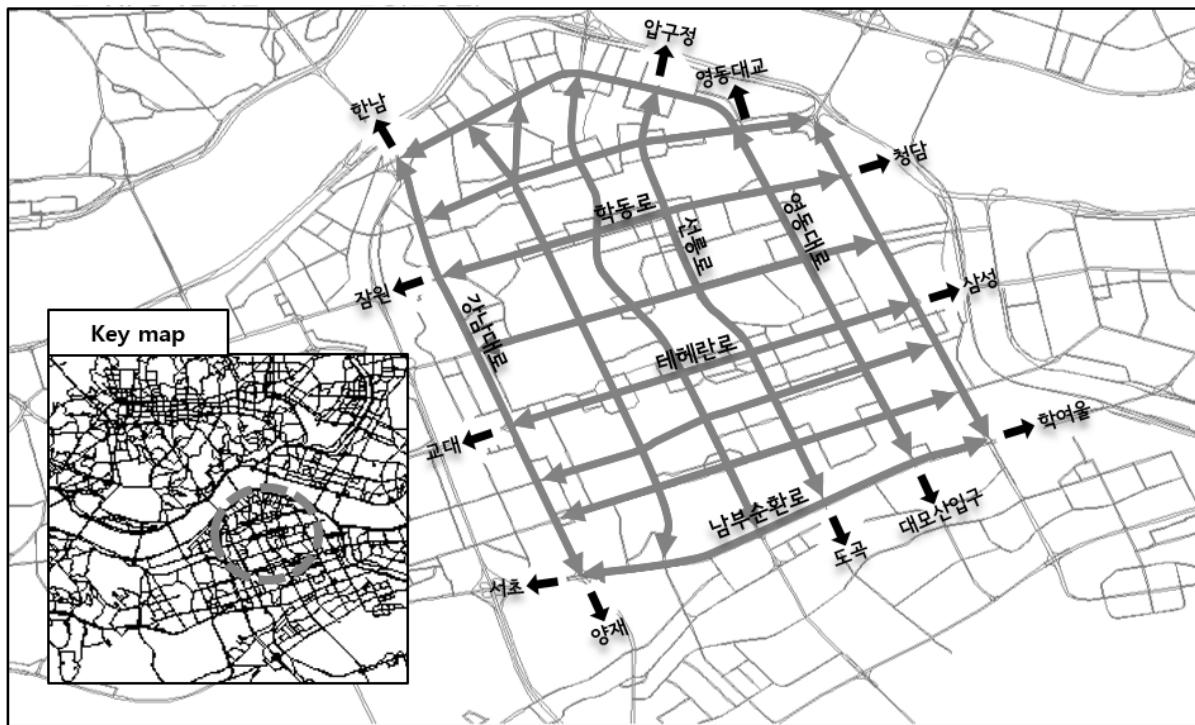
- 본 연구의 연속류에 대한 사례대상 구간은 서울시를 동서와 남북으로 이어주는 주요 교통축인 올림픽대로와 경부고속도로 일부 구간으로 선정하였으며, 해당 위치는 아래의 그림과 같음
  - 경부고속도로(고속도로, 양방향) : 한남대교 ~ 판교JC 구간
  - 올림픽대로(도시고속도로, 양방향) : 원효대교 ~ 올림픽대교 구간



<그림 2-59> 사례분석 대상구간 위치도(연속류)

### 나. 단속류 구간

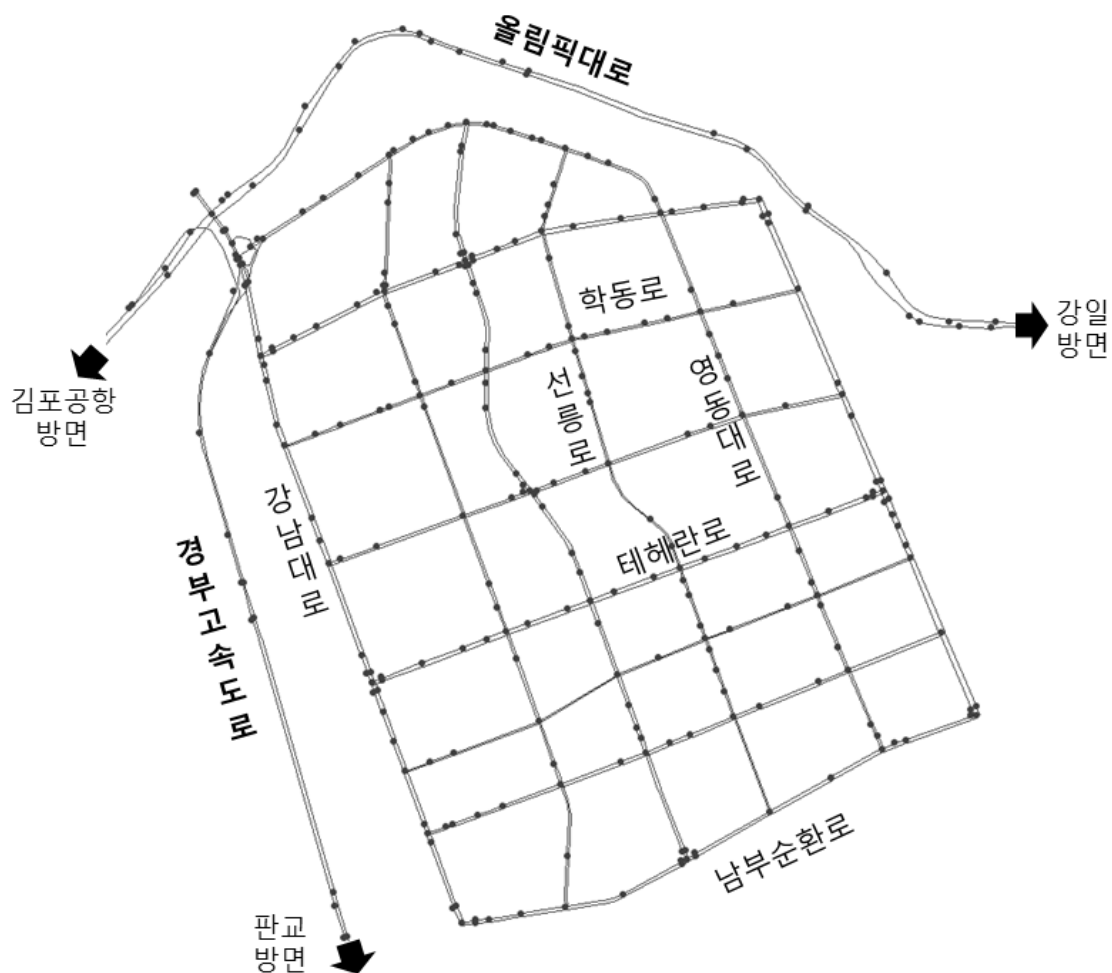
- 본 연구의 단속류에 대한 사례대상 구간은 격자형 가로망 형태를 띠는 서울시 강남 일대 내부 도로구간으로 설정하였으며, 해당 위치는 아래의 그림과 같음
- 남북축 : 강남대로, 선릉로, 영동대로 등 6개 교통축
- 동서축 : 학동로, 테헤란로, 남부순환로 등 8개 교통축



<그림 2-60> 사례분석 대상구간 위치도(단속류)

### 3. 사례대상 구간에 대한 분석 DB구축

- 기 구축된 ITS 표준노드링크 기반의 분석 맵에서 사례대상 구간에 해당하는 노드와 링크들을 추출하여 아래와 같이 본 연구의 사례분석을 위한 분석 네트워크를 구축하였음
- 사례분석 대상구간 : 총 255개 링크(서울특별시 강남 일대)
- 연속류(77개 링크) : 경부고속도로 30개 링크, 올림픽대로 47개 링크
- 단속류(178개 링크) : 기타 내부도로 178개 링크



<그림 2-61> 사례분석 대상구간의 분석 네트워크 구축

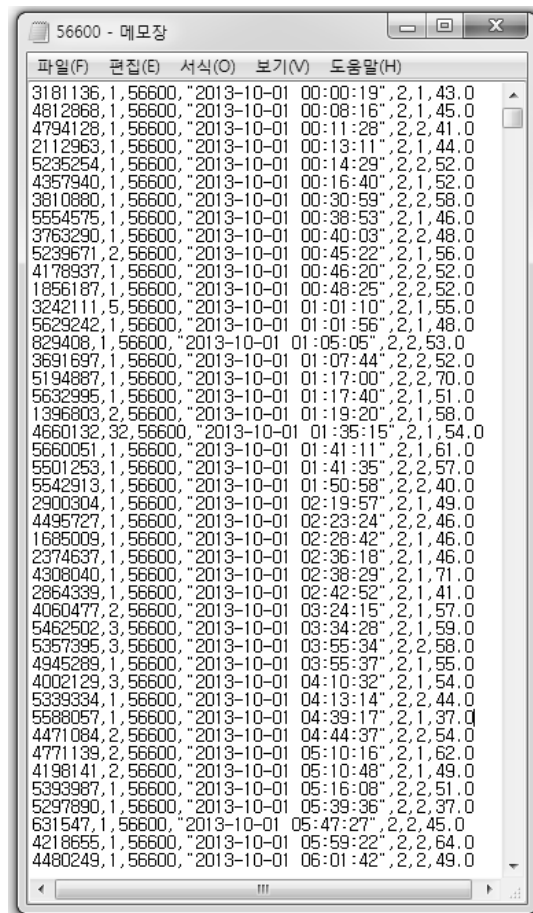
- ITS 표준노드링크 기반으로 수집된 개별 Car Navigation 데이터를 아래와 같이 총 7개의 Column으로 분석 맵의 링크별 수집시간 순(년/월/일/시/분/초)으로 자료를 집계하여 텍스트 형식으로 구축하였음



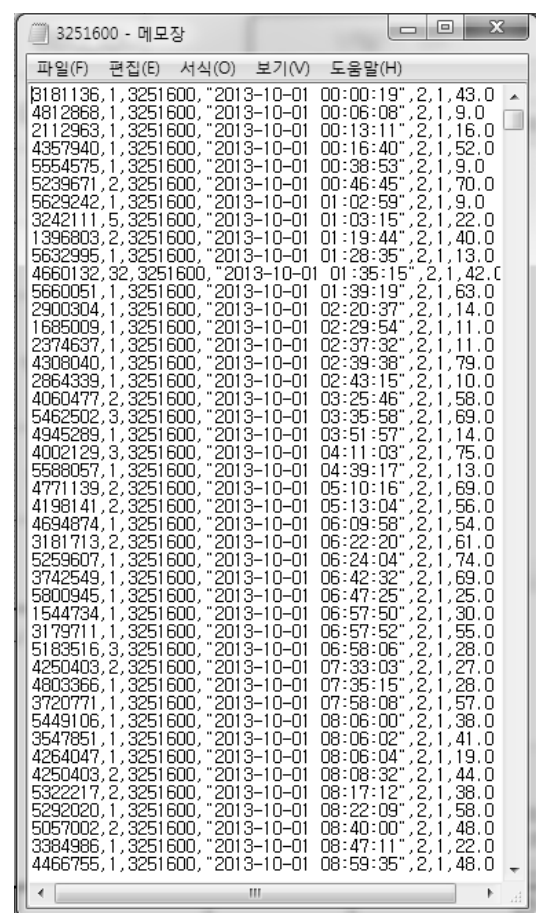
&lt;표 2-26&gt; 분석 맵 기준의 Car Navigation 데이터 정의서 및 저장 형식

Column	설명	Data Type	PK	FK
OBU_ID	Cae Navigation단말기 ID	Varchar	-	-
GROUP_ID	단말기별 그룹 ID	Varchar	-	-
from link	분석맵 기준 링크 ID	Varchar	-	-
link time	수집시간(월/일/시/분/초)	Varchar	-	-
week_type	요일구분 (0:일요일, 1:월요일, 2:화요일, 3:수요일, 4:목요일, 5:금요일, 6:토요일, 11:공휴일)	Integer	-	-
turn_type	이동류 구분 (0:좌회전, 1:직진, 2:우회전, 3:유턴)	Integer	-	-
speed	속도(km/h)	Double	-	-

데이터 저장 형식(텍스트 파일) - 총 255개 파일



...



- 위의 분석 맵 링크 기준으로 수집시간(월/일/시/분/초) 순으로 구축되어 있는 건별 Car Navigation 데이터를 링크별/일별/시간대별로 집계하여 사례분석을 위한 분석 DB(Probe건수, 속도)로 구축함
  - 구축 자료 : Probe건수(건/시), 속도(km/h)
  - 대상링크 개수 : 총 255개
  - 구축일 개수 : 총 123개(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치)
  - 시간대 개수 : 총 24개(01~24시)
  - 분석DB 규모 : 각 변수별 752,760개( $255 \times 123 \times 24 = 752,760$ ) → 총 1,505,520개
- 본 연구의 사례분석을 위한 분석 DB(Probe건수, 속도) 구축 결과는 아래와 같음
  - 링크별/일별/시간대별 Probe건수(건/시) 분석DB 구축 결과

구분	시간대	연속류(77개 링크)				단속류(178개 링크)	
		경부고속도로 (30개 링크)		도시고속도로 (47개 링크)		기타 내부도로	
		174400	...	1380700	...	56600	...
2013년10월1 일	00~01시	24	...	45	...	15	...
2013년10월1 일	01~02시	15	...	33	...	17	...
⋮	⋮	⋮	...	⋮	...	⋮	...
2014년1월30 일	22~23시	32	...	45	...	21	...
2014년1월31 일	23~24시	42	...	44	...	18	...

- 링크별/일별/시간대별 속도(km/h) 분석DB 구축 결과

구분	시간대	연속류(77개 링크)				단속류(178개 링크)	
		경부고속도로 (30개 링크)		도시고속도로 (47개 링크)		기타 내부도로	
		174400	...	1380700	...	56600	...
2013년10월1 일	00~01시	120	...	109	...	99	...
2013년10월1 일	01~02시	110	...	95	...	100	...
⋮	⋮	⋮	...	⋮	...	⋮	...
2014년1월30 일	22~23시	100	...	110	...	108	...
2014년1월31 일	23~24시	95	...	90	...	114	...

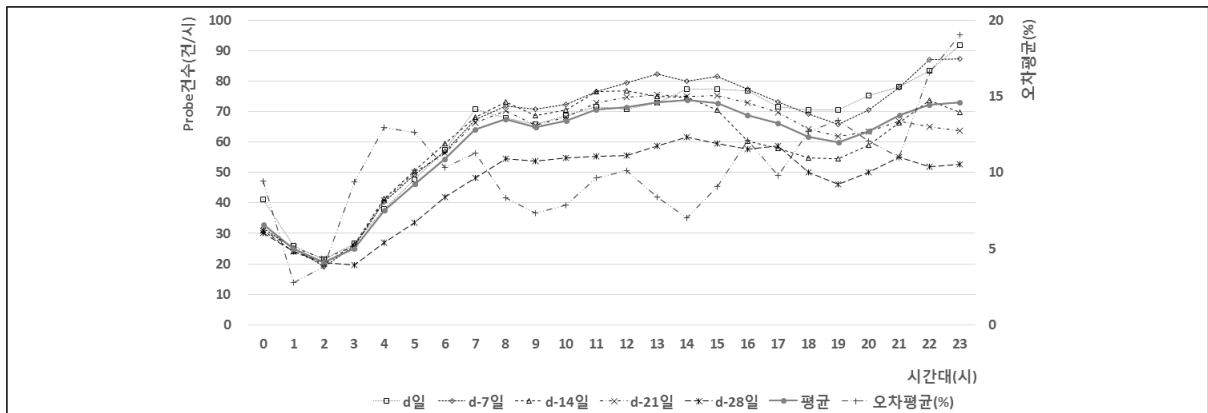
#### 4. 분석 자료의 특성

- 장래 통행속도 예측모형의 성능을 평가하기 위하여 이용되는 분석 자료는 ITS 표준노드링크 기반의 분석 맵 단위 링크별/일별/시간대별 Probe건수와 속도 자료임
- 본 연구의 사례대상 구간 중 연속류(경부고속도로, 올림픽도로)와 단속류(강남대로)로 구분하여 각 주요 구간에 대한 분석 자료의 historical 특성 분석을 수행하였음
- 즉, 본 연구의 시/공간적 범위를 고려하여 최근의 분석일(d일, 14년 1월 20일)을 정한 후, 해당 일과 같은 요일이면서 일주일 간격으로 과거 1개월치(d-7일~d-28일)의 이력자료에 대한 특성을 검토하였음

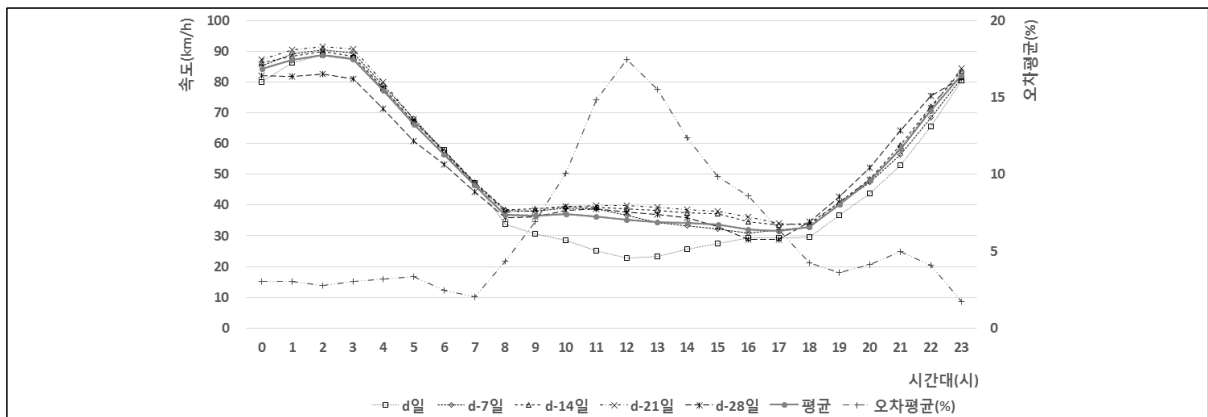
##### 가. 연속류 구간

###### 1) 경부고속도로(서초IC→양재IC)

- 연속류의 경부고속도로(서초IC→양재IC) 구간에 대한 historical 특성 분석을 수행한 결과, 먼저 Probe건수의 경우는 평균 Probe건수 분포를 기준으로 일별/시간대별 Probe건수 분포의 변화량이 평균 오차 범위 2.8~19.0% 내에서 전체적으로 비슷한 패턴을 보이는 것으로 분석되었음
- 즉, 시간대별 Probe건수 분포를 살펴보면, 새벽시간에는 전체 교통량의 감소로 인하여 Probe건수도 감소하며, 출근시간이 가까워질수록 점차 교통량이 증가함에 따라 Probe건수도 증가하고 그 이후로 일정 수준의 Probe건수를 유지하는 것으로 분석됨
- 오히려 출/퇴근 시간에는 교통 혼잡으로 인하여 Probe건수가 감소하는 경향을 보이고 있음. 따라서 일별/시간대별 Probe건수의 분포는 실제 교통량의 분포와 유사한 것으로 판단되며, 해당 Probe건수를 교통량과 동일한 거시적 교통변수로 이용이 가능할 것으로 판단됨
- 다음으로 속도자료의 경우, 새벽시간대(00~08시)와 야간시간대(18~24시)에는 과거 이력자료의 속도 분포는 큰 변화가 없는 것으로 분석되었으며, 그 외 시간대(08~17시)에서는 평균 오차가 약 18%까지 증가하여 해당 일별 이력자료에 대한 속도 변화가 다소 큰 것으로 분석됨
- 이를 경우, 새벽시간대(00~08시)와 야간시간대(18~24시)에서는 기존의 과거 이력자료 평균방법과 본 연구방법의 예측결과의 차이가 크지 않을 것으로 판단되며, 그 외 시간대(08~17시)에는 해당 일별 속도 분포의 변화가 크므로 기존의 방법보다 현행 기준으로 가장 유사한 과거 이력자료의 패턴을 찾는 본 연구방법이 더 우수한 예측력을 보일 것으로 판단됨



<그림 2-62> Probe 이력자료 특성분석(경부 고속도로, 서초IC→양재IC)

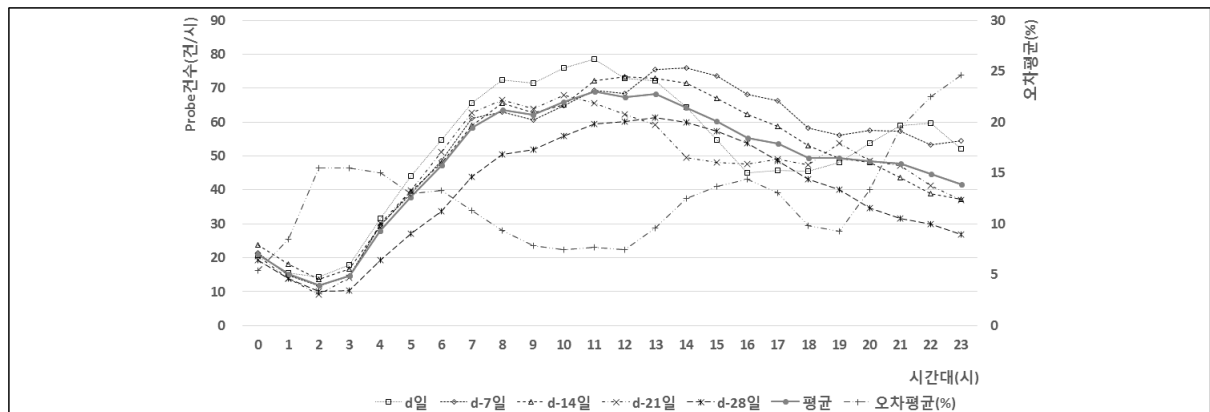


<그림 2-63> 속도 이력자료 특성분석(경부 고속도로, 서초IC→양재IC)

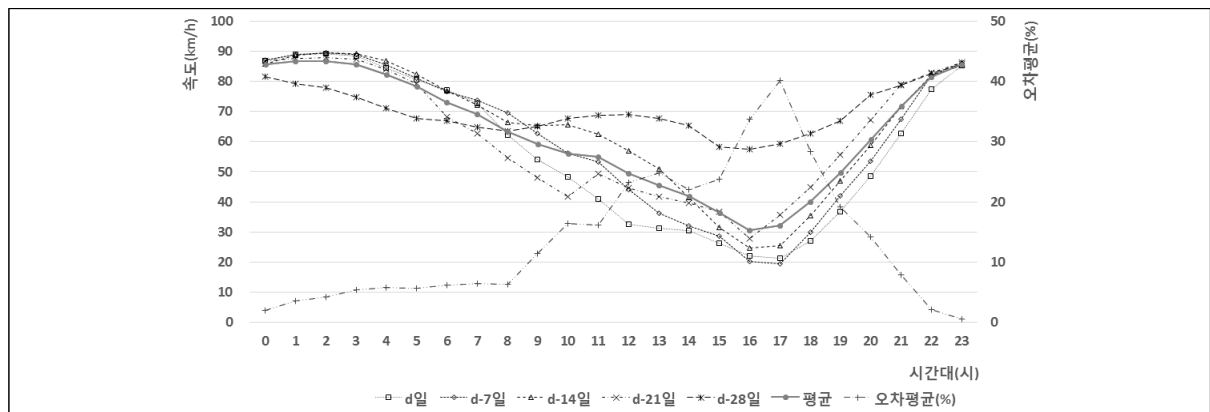
## 2) 올림픽대로(반포대교→한남대교)

- 연속류의 올림픽대로(반포대교→한남대교) 구간에 대한 historical 특성 분석을 수행한 결과, 먼저 Probe건수의 경우는 평균 5.4~24.6%의 오차 범위 내에서 일별/시간대별 분포 변화량이 전체적으로 비슷한 패턴을 보이는 것으로 분석되었음
- 즉, 시간대별 Probe건수 분포를 살펴보면, 새벽시간에는 전체 교통량의 감소로 인하여 Probe건수도 감소하며, 출근시간이 가까워질수록 점차 교통량이 증가함에 따라 Probe건수도 증가하다가 낮 12시 이후로는 교통 혼잡으로 인하여 Probe건수가 감소하는 경향을 보이고 있음
- 따라서, 일별/시간대별 Probe건수의 분포는 실제 교통량의 분포와 유사한 것으로 판단되며, 해당 Probe건수를 교통량과 동일한 거시적 교통변수로 이용이 가능할 것으로 판단됨

- 다음으로 속도자료의 경우, 새벽시간대(00~08시)를 제외하고는 전체적으로 과거 이력자료의 속도 변화가 심한 것으로 분석되었으며, 특히 오후 17시경에는 평균 오차가 40%까지 증가하여 과거 이력자료의 속도 변화가 가장 큰 것으로 분석됨
- 이를 경우, 새벽시간대(00~08시)를 제외하고 그 외 시간대(08~24시)에는 해당 일별 속도 분포의 변화가 크므로 기존의 방법보다 현행 기준으로 가장 유사한 과거 이력자료의 패턴을 찾는 본 연구방법이 더 우수한 예측력을 보일 것으로 판단됨



<그림 2-64> Probe 이력자료 특성분석(올림픽대로, 반포대교→한남대교)



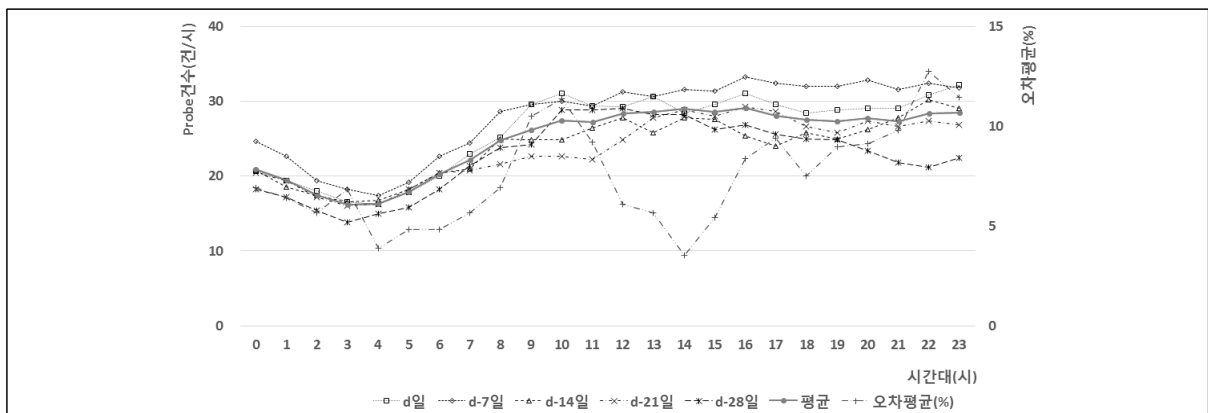
<그림 2-65> 속도 이력자료 특성분석(경부 고속도로, 반포대교→한남대교)

#### 나. 단속류 구간(강남대로, 신논현역 → 강남역)

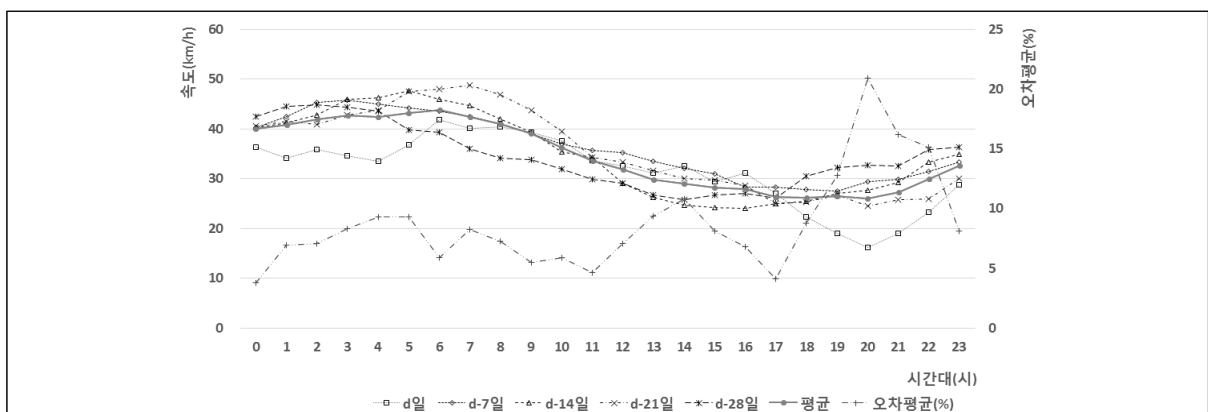
- 단속류의 올림픽대로(반포대교→한남대교) 구간에 대한 분석 결과는 앞서 연속류의 구간보다 일별/시간대별 분포 변화량이 전체적으로 좀 더 비슷한 패턴을 보이는 것으로 분석되었음
- 시간대별 Probe건수 분포를 살펴보면, 새벽시간에는 전체 교통량의 감소로 인하여 Probe건수가 감소하며, 출근시간이 가까워질수록 점차 교통량이 증가함에 따라 Probe건수도 증가하고

그 이후로 일정 수준의 Probe건수를 유지하는 것으로 분석됨

- 즉, 일별/시간대별 Probe건수의 분포는 실제 교통량의 분포와 유사한 것으로 판단되며, 해당 Probe건수를 교통량과 동일한 거시적 교통변수로 이용이 가능할 것으로 판단됨
- 다음으로 속도자료의 경우, 새벽 00시부터 오후 17시까지는 평균 오차 3.8~10.8%의 범위 내에서 과거 이력자료의 속도 분포는 대체적으로 변화가 없는 것으로 분석되었으며, 오후 17시 이후에서는 평균 오차가 최대 20.9%까지 증가하여 해당 일별 이력자료에 대한 속도 변화가 커지는 것으로 분석됨
- 이를 경우, 새벽 00시부터 오후 17시까지는 기존의 과거 이력자료 평균방법과 본 연구방법의 예측결과의 차이가 크지 않을 것으로 판단되며, 오후 17시 이후에서는 해당 일별 속도 분포의 변화가 크므로 기존의 방법보다 현행 기준으로 가장 유사한 과거 이력자료의 패턴을 찾는 본 연구방법이 더 우수한 예측력을 보일 것으로 판단됨



<그림 2-66> Probe 이력자료 특성분석(강남대로, 신논현역→강남역)



<그림 2-67> 속도 이력자료 특성분석(강남대로, 신논현역→강남역)

## 5. 평가지표 및 분석 시나리오 설정

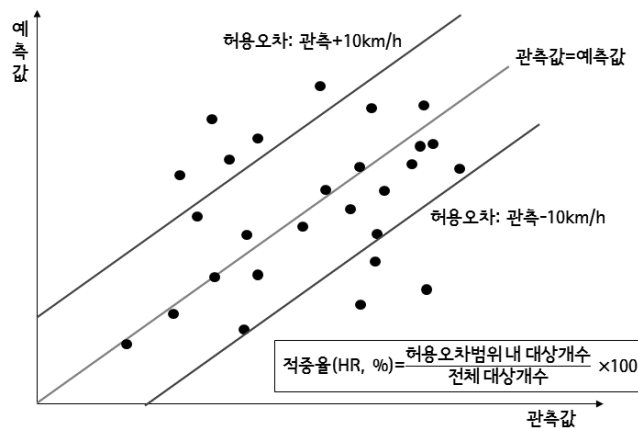
### 가. 평가지표 설정

- 실제 통행속도와 예측모형에 의해 추정된 통행속도 간의 정확성을 평가하기 위한 방법으로는 관측치와 예측치 간의 상대적 오차를 이용하는 방법임. 이는 관측치와 예측치 간의 오차를 관측치로 나누어서 백분율로 계산하기 때문에 서로 다른 단위의 예측치를 이용할 때 예측 방법의 정확성을 평가할 수 있음
- 예측치의 상대적 오차를 측정하는 방법으로 주로 백분율 평균오차와 백분율 절대평균오차(MAPE, %: Mean absolute percentage error)를 측정하는 방법이 있음. 이 중에서 백분율 평균오차는 추정 오차들이 서로 상쇄되기 때문에 예측분석에 있어 불합리한 결과를 초래할 수 있음

$$MAPE(\%) = \frac{\sum_{i=1}^n (|x_i - \hat{x}_i|/x_i) \times 100}{n}$$

여기서,  $x_i$ : 관측값,  $\hat{x}_i$ : 예측값,  $n$ : 전체 자료의 개수

- 적중률(%)은 전체 예측값 추출 대상 중 관측값과 비교하여 허용오차 범위 내( $\pm 10\text{km/h}$ )에 속하는 대상에 대한 적중 비율(%)을 말함



<그림 2-68> 본 연구의 적중률(%)

- 따라서, 본 연구에서는 예측모형의 평가임을 고려하여 MAPE(%)와 적중률(%)을 이용하여 평가를 수행하도록 함

## 나. 분석 시나리오 설정

- 본 연구에서 개발한 예측모형을 평가하기 위하여 먼저 평가 비교대상이 있어야 하며, 다양한 소통 상태에 따른 분석 시나리오를 설정해야 함
- 먼저, 평가 비교대상은 기존 ITS시스템에서 많이 이용되고 있는 과거평균(Historical average) 기법을 적용하며, 이는 해당 예측일(d+1)의 당일(d일) 기준보다 이전 날짜들의 과거 속도자료 중 동일한 요일의 지난 1개월간의 시간대별 평균 속도를 산출하는 것을 말함
- 분석 시나리오 설정 시, 과거 이력자료 DB(Probe건수, 속도)의 시/공간적 구축 범위를 고려하여 시간적으로는 예측일을 구정 연휴를 포함하지 않은 가장 최근의 기간으로 설정하고, 공간적으로는 연속류와 단속류로 구분하여 총 3개의 분석 네트워크로 설정하였음
  - 시간적 범위 : 총 123일(2013년 10월 1일 ~ 2014년 1월 31일, 4개월치), 24개 시간대
    - 장래 예측일 설정 : 총 5일, 2014년 1월 20일(월) ~ 2014년 1월 24일(금)
    - 여기서, 주말의 경우는 과거 이력자료에서 해당 주말의 소통 상태와 유사한 패턴을 찾기 위하여 필요로 하는 샘플 수 문제로 본 연구의 장래 예측일 설정 범위에서 제외하였으며, 향후 1년 이상의 충분한 데이터가 확보 될 경우에 가능할 것으로 판단됨
  - 공간적 범위 : 총 255개 링크(서울특별시 강남 일대)
    - 연속류 분석 네트워크 설정 : 경부고속도로(총 30개 링크), 올림픽대로(총 47개 링크)
    - 단속류 분석 네트워크 설정 : 기타 내부도로(총 178개 링크)
- 위의 평가 비교대상과 시/공간적 분석의 범위에 따라 아래의 표와 같이 분석 시나리오를 설정하였으며, 설정된 분석 시나리오별로 적용방법에 따라 장래 링크별/시간대별 통행속도를 예측하여 해당 링크의 관측값과 비교·분석을 수행하였음

<표 2-27> 분석 시나리오 설정

구분		분석 시나리오	연속류				단속류	
			경부고속도로		올림픽대로		기타 내부도로	
			기존	본 연구	기존	본 연구	기존	본 연구
장래 예측일자	2014년1월20일(월)	시나리오1(S1)	S1-1		S1-2		S1-3	
	2014년1월21일(화)	시나리오2(S2)	S2-1		S2-2		S2-3	
	2014년1월22일(수)	시나리오3(S3)	S3-1		S3-2		S3-3	
	2014년1월23일(목)	시나리오4(S4)	S4-1		S4-2		S4-3	
	2014년1월24일(금)	시나리오5(S5)	S6-1		S6-2		S6-3	

주 : 기준은 과거평균(Historical average) 기법이며, 본 연구는 본 연구의 개발 모형을 의미함.



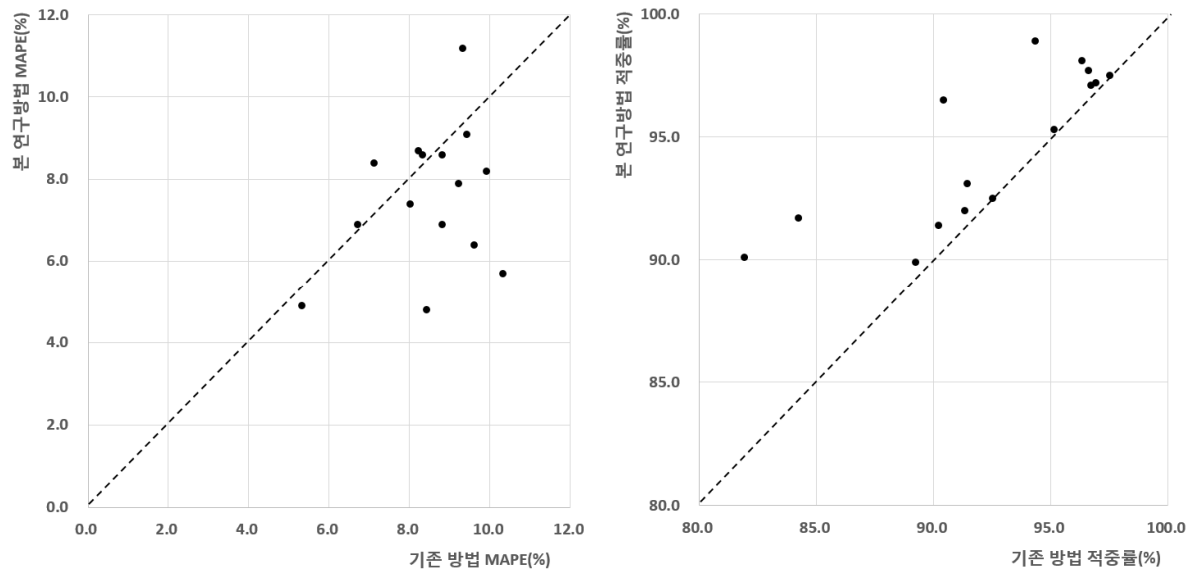
## 6. 방법론 적용 및 평가

### 가. 종합평가 결과

- 본 연구에서는 앞서 설명한 장래 예측일과 분석 네트워크 설정에 따라 총 5가지의 분석 시나리오를 설정하였으며, 각 시나리오별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 그 결과, 전체 분석 시나리오와 모든 네트워크에서 본 연구의 예측모형이 기존 과거평균기법보다 MAPE는 0.6~1.4%, 적중률은 0.7~4.0%의 예측력을 향상시키는 것으로 분석되었음. 이는 앞서 4절에서 설명한 바와 같이 과거 속도자료의 패턴이 대체로 일정한 비첨두시간대(00~06시)와 혼잡 시간대(07~09시, 18~20시)를 제외한 나머지 속도의 변화가 큰 시간대에서 본 모형의 예측력이 높은 것으로 판단됨
  - MAPE 1.4%의 향상은 기존 방법 보다 본 연구의 예측모형 오차가 평균 1.4% 감소했다는 의미이며, 적중률 4.0%의 향상은 기존 방법 보다 본 연구의 예측모형이 허용오차 범위( $\pm 10\text{km/h}$ )에 속하는 표본이 86개(24시간 $\times$ 255개 링크 $\times$ 4%)가 더 증가된 것을 의미함
- 시나리오 3~5의 경우, 일부 분석 네트워크에서 기존 과거평균기법의 예측력이 본 연구의 예측모형과 비슷하거나 오히려 다소 향상된 결과가 산출된 것은 해당 분석 네트워크의 과거 일별/시간대별 속도자료 분포의 Variation이 크지 않고 전체 시간대에서 일정한 패턴을 유지하는 데에서 기인된 것으로 판단됨
- 위의 종합 평가결과에 대하여 먼저, 과거 속도자료의 일별/시간대별 패턴변화가 크지 않고 일정한 경우(일별/시간대별 패턴의 주기성이 있는 경우)에는 기존 방법과 본 연구방법의 예측결과에 대한 차이는 크지 않으며, 그 반대로 과거 속도자료의 일별/시간대별 패턴변화가 크고 일정하지 않은 경우(일별/시간대별 패턴의 주기성이 없는 경우)에는 기존 방법보다 본 연구방법의 예측력이 더 우수한 것으로 분석되었음
- 이는 기존의 과거평균방법이 본 연구방법에 비해 시계열적으로 변화하는 통행속도의 차이에 민감하게 반응하지 못하며, 또한, 기존의 개별 링크가 아닌 네트워크의 차원에서 장래 상태를 설명하기 때문에 장래 교통상태의 불확실성을 어느 정도 감소시킬 수 있었기 때문으로 판단됨

&lt;표 2-28&gt; 종합평가 결과

구분	평가지표	연속류 네트워크						단속류 네트워크		
		경부고속도로			올림픽대로			기타 내부도로		
		기존	본연구	증감	기존	본연구	증감	기존	본연구	증감
시나리오1	MAPE(%)	10.3	5.7	-4.6	5.3	4.9	-0.4	8.8	8.6	-0.2
	적중률(%)	90.4	96.5	6.1	95.1	95.3	0.2	96.9	97.2	0.3
시나리오2	MAPE(%)	8.8	6.9	-1.9	9.6	6.4	-3.2	9.9	8.2	-1.7
	적중률(%)	84.2	91.7	7.5	81.9	90.1	8.2	96.3	98.1	1.8
시나리오3	MAPE(%)	8.4	4.8	-3.6	8.0	7.4	-0.6	8.2	8.7	0.5
	적중률(%)	94.3	98.9	4.6	91.3	92.0	0.7	97.5	97.5	-
시나리오4	MAPE(%)	9.3	11.2	1.9	8.3	8.6	0.3	9.4	9.1	-0.3
	적중률(%)	92.5	92.5	-	90.2	91.4	1.2	96.7	97.1	0.4
시나리오5	MAPE(%)	7.1	8.4	1.3	6.7	6.9	0.2	9.2	7.9	-1.3
	적중률(%)	91.4	93.1	1.7	89.2	89.9	0.7	96.6	97.7	1.1
전체평균	MAPE(%)	8.8	7.4	-1.4	7.6	6.8	-0.7	9.1	8.5	-0.6
	적중률(%)	90.6	94.5	4.0	89.5	91.7	2.2	96.8	97.5	0.7



&lt;그림 2-69&gt; 기존 방법과의 MAPE(%)와 적중률(%) 비교

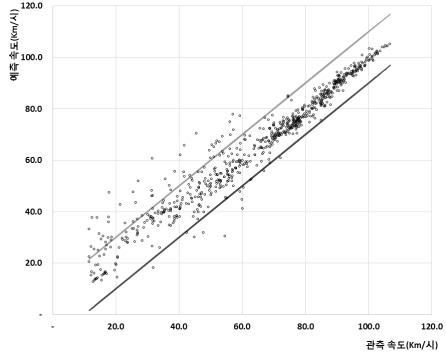
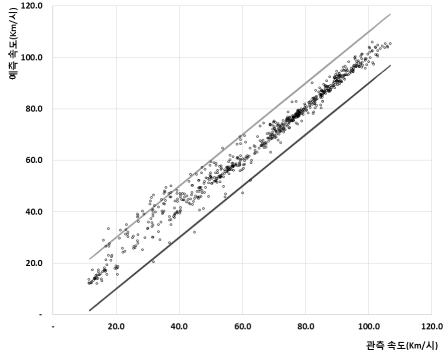
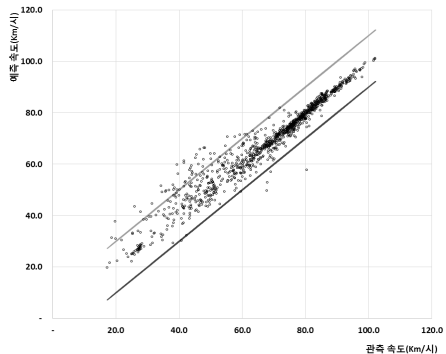
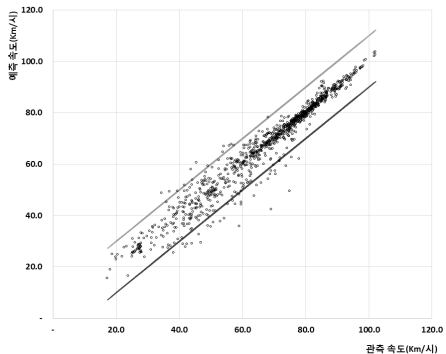
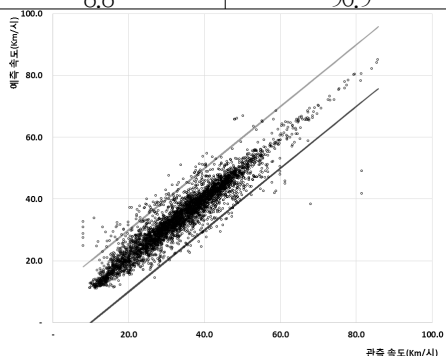
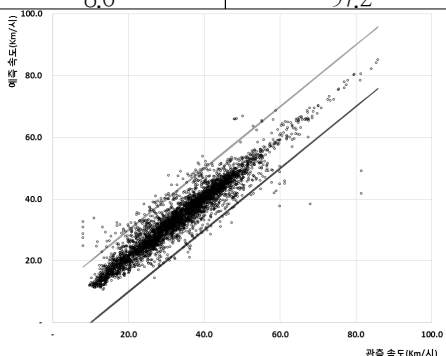
## 나. 시나리오별 평가 결과

### 1) 시나리오 1

#### ① 분석 네트워크별 평가 결과

- 본 연구의 사례대상 구간 중 연속류 네트워크는 경부고속도로(30개 링크)와 올림픽대로(47개 링크)로 구성되며, 단속류 네트워크는 연속류 네트워크를 제외한 나머지 기타 내부도로(178개 링크)로 구성됨
- 여기서, 시나리오 1의 경우는 2014년 1월 20일(월)을 장래 예측일로 설정하고 위의 분석 네트워크별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 경부고속도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었으며, 이는 시간대별 통행속도가 높은 비첨두시간 때보다 혼잡이 발생하는 첨두시간대에서 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석됨. 즉, 비첨두시간대에서 혼잡이 발생하는 첨두시간대로 넘어갈 때, 과거 속도자료의 일별/시간대별 패턴 변화가 큰 것에서 기인된 것으로 판단됨
  - MAPE(%) : 기존의 과거평균기법 10.3% → 본 연구의 예측모형 5.7%, MAPE 4.6% 감소
  - 적중률(%) : 기존의 과거평균기법 90.4% → 본 연구의 예측모형 96.5%, 적중률 6.1% 증가
- 올림픽대로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었으며, 이는 위의 경부고속도로의 경우와 동일함
  - MAPE(%) : 기존의 과거평균기법 5.3% → 본 연구의 예측모형 4.9%, MAPE 0.4% 감소
  - 적중률(%) : 기존의 과거평균기법 95.1% → 본 연구의 예측모형 95.3%, 적중률 0.2% 증가
- 기타 내부도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었음. 이는 전체 시간대에서 걸쳐 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석되었으나, 위의 연속류의 경우 보다 개선효과는 비교적 낮은 편임. 즉, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 대체적으로 크지 않으므로 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 8.8% → 본 연구의 예측모형 8.6%, MAPE 0.2% 감소
  - 적중률(%) : 기존의 과거평균기법 96.9% → 본 연구의 예측모형 97.2%, 적중률 0.3% 증가

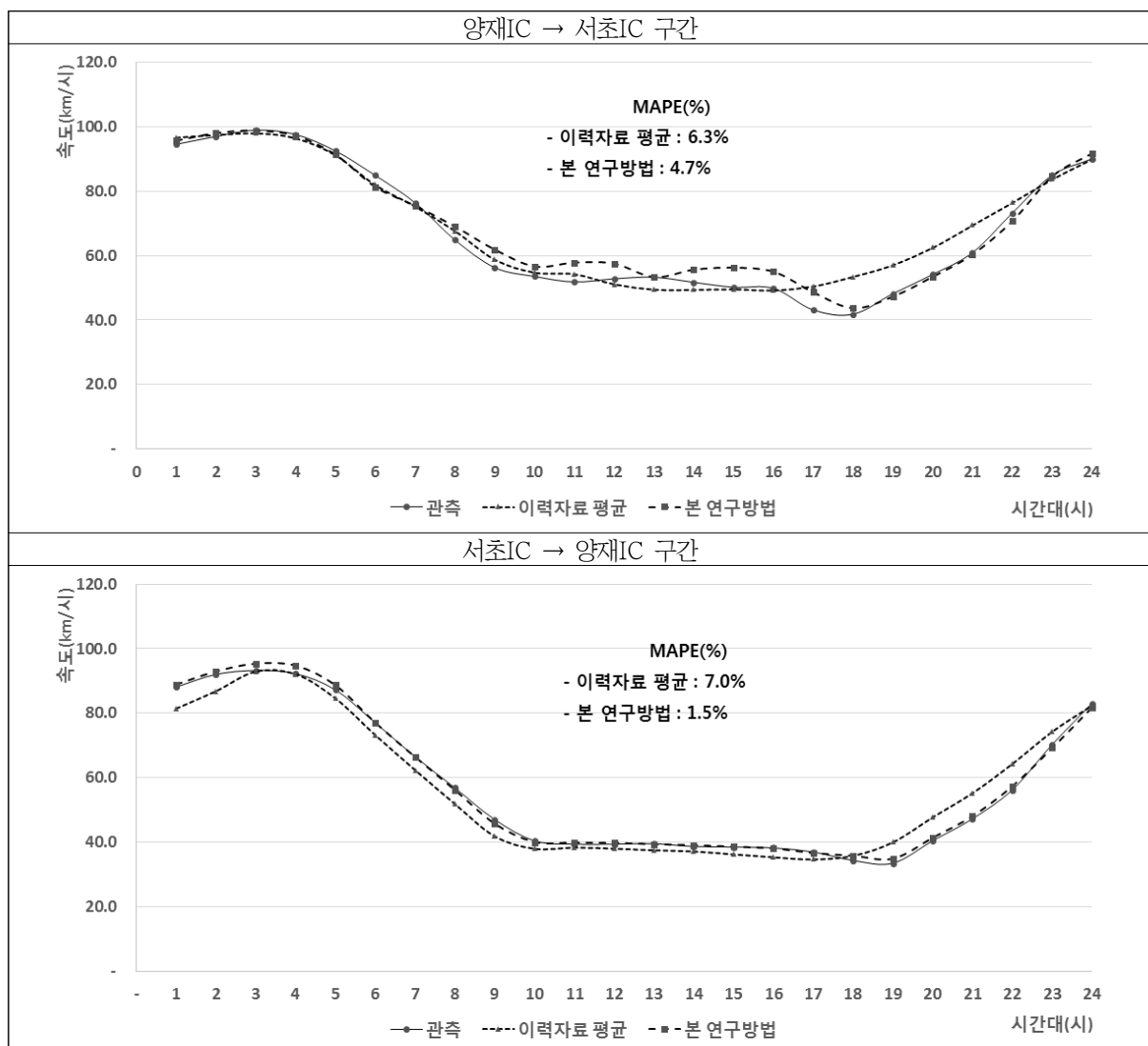
&lt;표 2-29&gt; 전체 분석 네트워크의 평가 결과(시나리오 1)

구분	적용 방법론			
	이력자료 평균		본 연구 방법	
	MAPE(%)	적중률(%)	MAPE(%)	적중률(%)
경부고속도로 (30개 링크)	10.3	90.4	5.7	96.5
				
울림픽대로 (47개 링크)	5.3	95.1	4.9	95.3
				
기타 내부도로 (178개 링크)	8.8	96.9	8.6	97.2
				

## ② 주요 링크구간의 평가 결과

### ○ 연속류 구간(경부고속도로, 양재IC~서초IC)

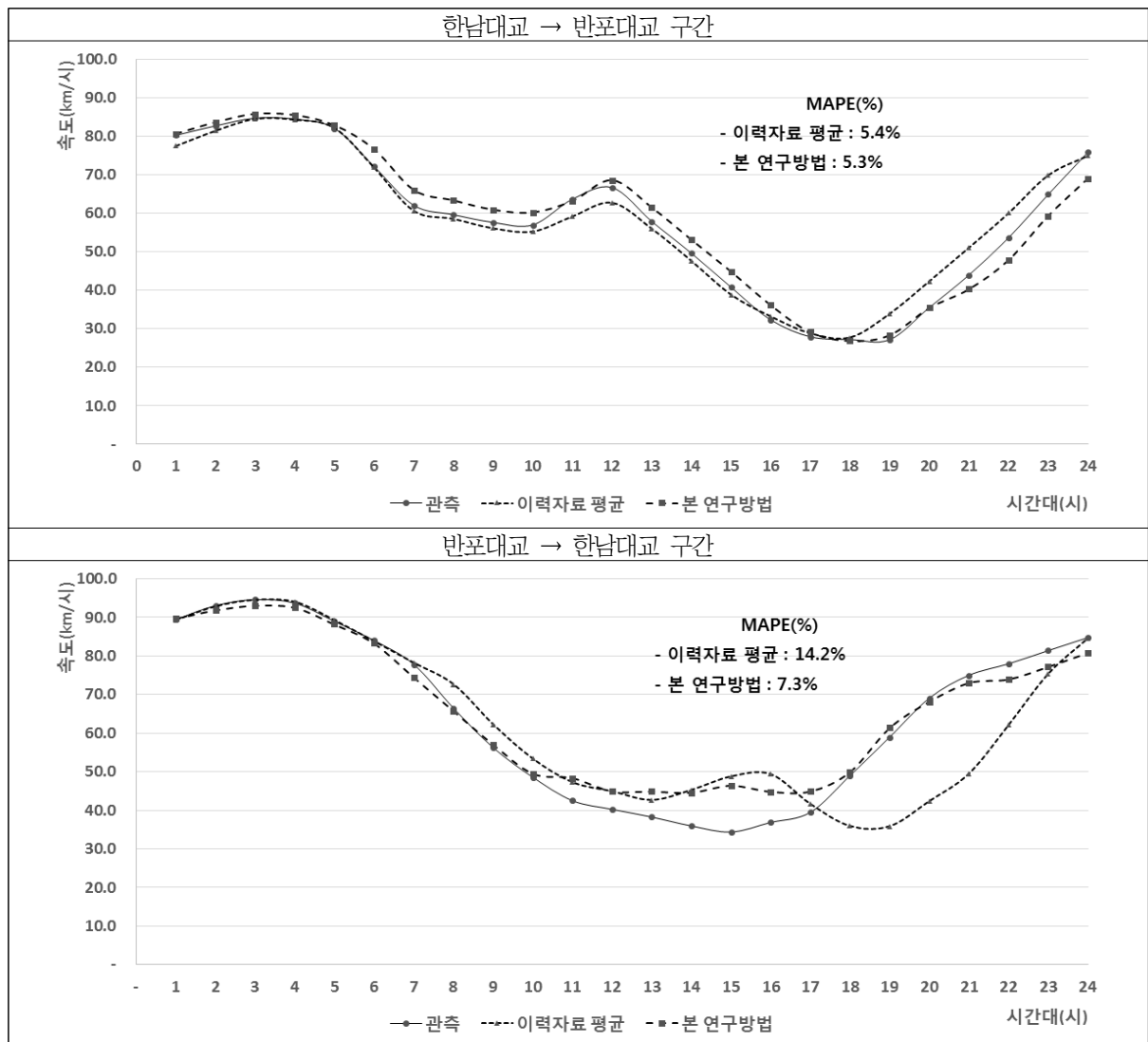
- 시나리오 1의 연속류 구간(경부고속도로, 양재IC~서초IC)은 방향별로 모두 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 이는 시간대별 통행속도가 높은 비첨두시간 때보다 퇴근 혼잡이 시작되는 17~18시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석됨. 즉, 혼잡이 시작되는 시간대에서 과거 속도 자료의 일별/시간대별 패턴 변화가 크므로 기존 방법의 오차가 더 증가하기 때문임



<그림 2-70> 시나리오 1의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC)

○ 연속류 구간(올림픽대로, 한남대교~반포대교)

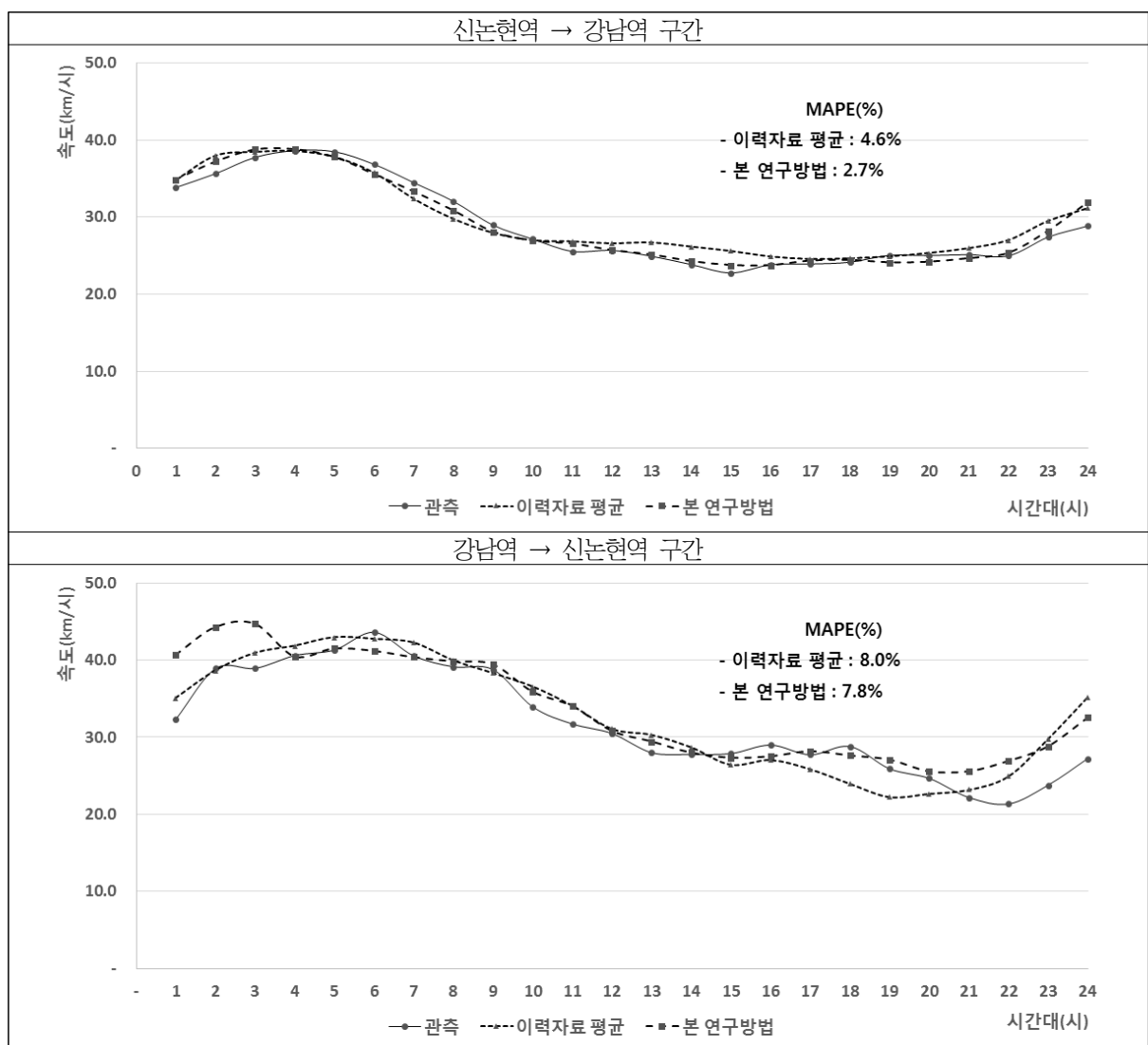
- 시나리오 1의 연속류 구간(올림픽대로, 한남대교~반포대교)은 한남대교→반포대교 방향의 경우는 기존 방법에 비해 예측결과의 차이가 크지 않으나, 반대 방향의 경우는 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 즉, 한남대교→반포대교 방향에서는 과거 속도자료의 일별/시간대별 패턴 변화가 일정하여 두 적용방법의 차이가 없는 것으로 분석되었으며, 그 반대 방향은 시간대별 통행속도가 높은 비첨두시간 때보다 퇴근 혼잡이 시작되는 17시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석됨



<그림 2-71> 시나리오 1의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교)

○ 단속류 구간(강남대로, 신논현역~강남역)

- 시나리오 1의 단속류 구간(강남대로, 신논현역~강남역)은 방향별로 모두 기존 과거평균기법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 이는 전체 시간대에서 걸쳐 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석되었으나, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 대체적으로 크지 않으므로 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨



<그림 2-72> 시나리오 1의 단속류 구간 평가결과(강남대로, 신논현역~강남역)

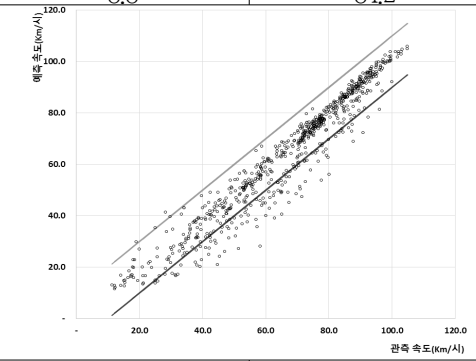
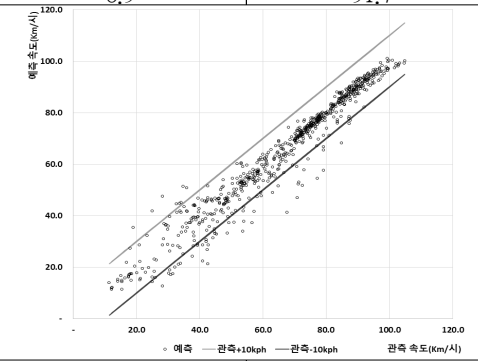
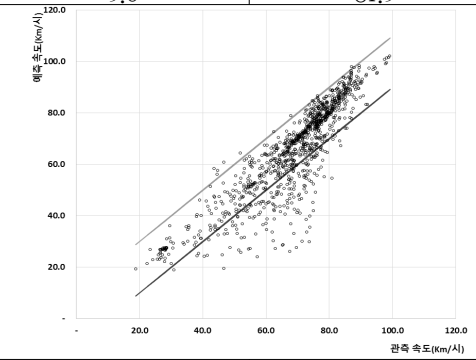
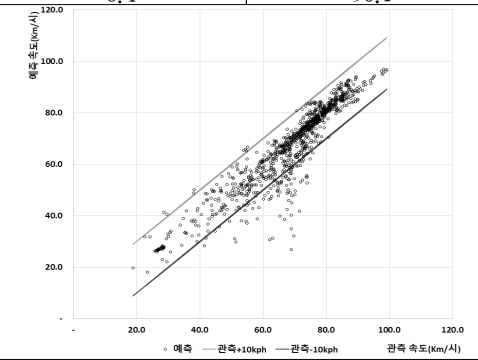
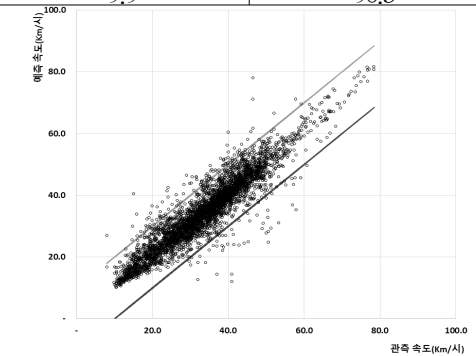
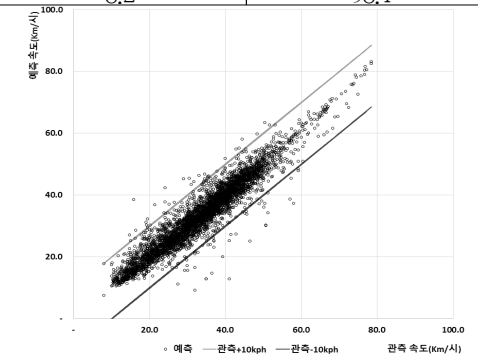
## 2) 시나리오 2

### ① 분석 네트워크별 평가 결과

- 본 연구의 사례대상 구간 중 연속류 네트워크는 경부고속도로(30개 링크)와 올림픽대로(47개 링크)로 구성되며, 단속류 네트워크는 연속류 네트워크를 제외한 나머지 기타 내부도로(178개 링크)로 구성됨
- 여기서, 시나리오 2의 경우는 2014년 1월 21일(화)을 장래 예측일로 설정하고 위의 분석 네트워크별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 경부고속도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었으며, 이는 시간대별 통행속도가 높은 비첨두시간 때보다 혼잡이 발생하는 첨두시간대에서 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석됨. 즉, 비첨두시간대에서 혼잡이 발생하는 첨두시간대로 넘어갈 때, 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 기존 방법의 오차가 더 증가하기 때문임
  - MAPE(%) : 기존의 과거평균기법 8.8% → 본 연구의 예측모형 6.9%, MAPE 1.9% 감소
  - 적중률(%) : 기존의 과거평균기법 84.2% → 본 연구의 예측모형 91.7%, 적중률 7.5% 증가
- 올림픽대로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었으며, 이는 위의 경부고속도로의 경우와 동일함
  - MAPE(%) : 기존의 과거평균기법 9.6% → 본 연구의 예측모형 6.4%, MAPE 3.2% 감소
  - 적중률(%) : 기존의 과거평균기법 81.9% → 본 연구의 예측모형 90.1%, 적중률 8.2% 증가
- 기타 내부도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었음. 이는 전체 시간대에서 걸쳐 본 연구방법의 예측력이 높은 것으로 분석되었으나, 위의 연속류의 경우 보다 개선효과는 비교적 낮은 편임. 즉, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 대체적으로 크지 않기 때문에 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 9.9% → 본 연구의 예측모형 8.2%, MAPE 1.7% 감소
  - 적중률(%) : 기존의 과거평균기법 96.3% → 본 연구의 예측모형 98.1%, 적중률 1.8% 증가



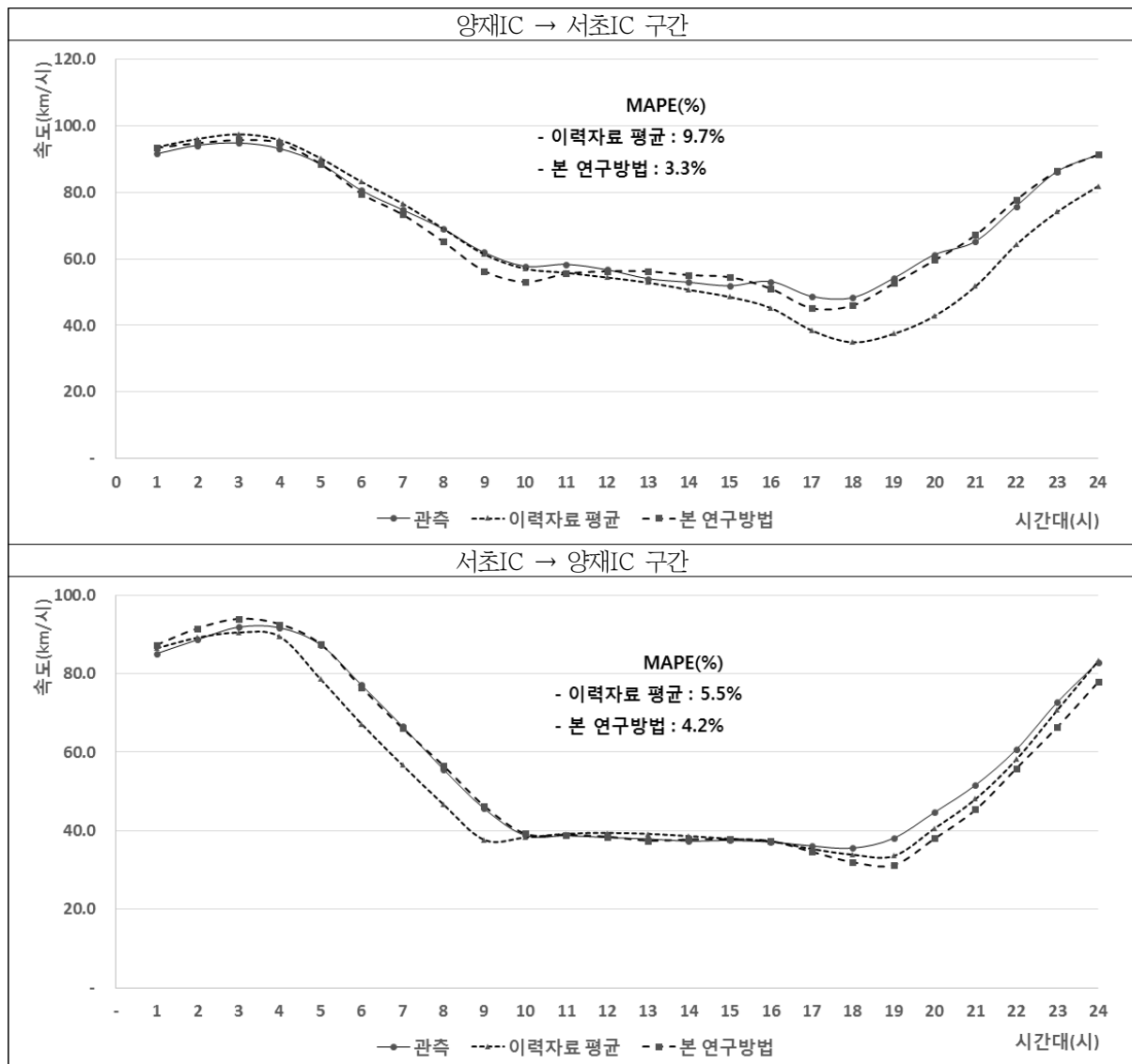
&lt;표 2-30&gt; 전체 분석 네트워크의 평가 결과(시나리오 2)

구분	적용 방법론			
	이력자료 평균		본 연구 방법	
	MAPE(%)	적중률(%)	MAPE(%)	적중률(%)
경부고속도로 (30개 링크)	8.8	84.2	6.9	91.7
				
울림픽대로 (47개 링크)	9.6	81.9	6.4	90.1
				
기타 내부도로 (178개 링크)	9.9	96.3	8.2	98.1
				

## ② 주요 링크구간의 평가 결과

### ○ 연속류 구간(경부고속도로, 양재IC~서초IC)

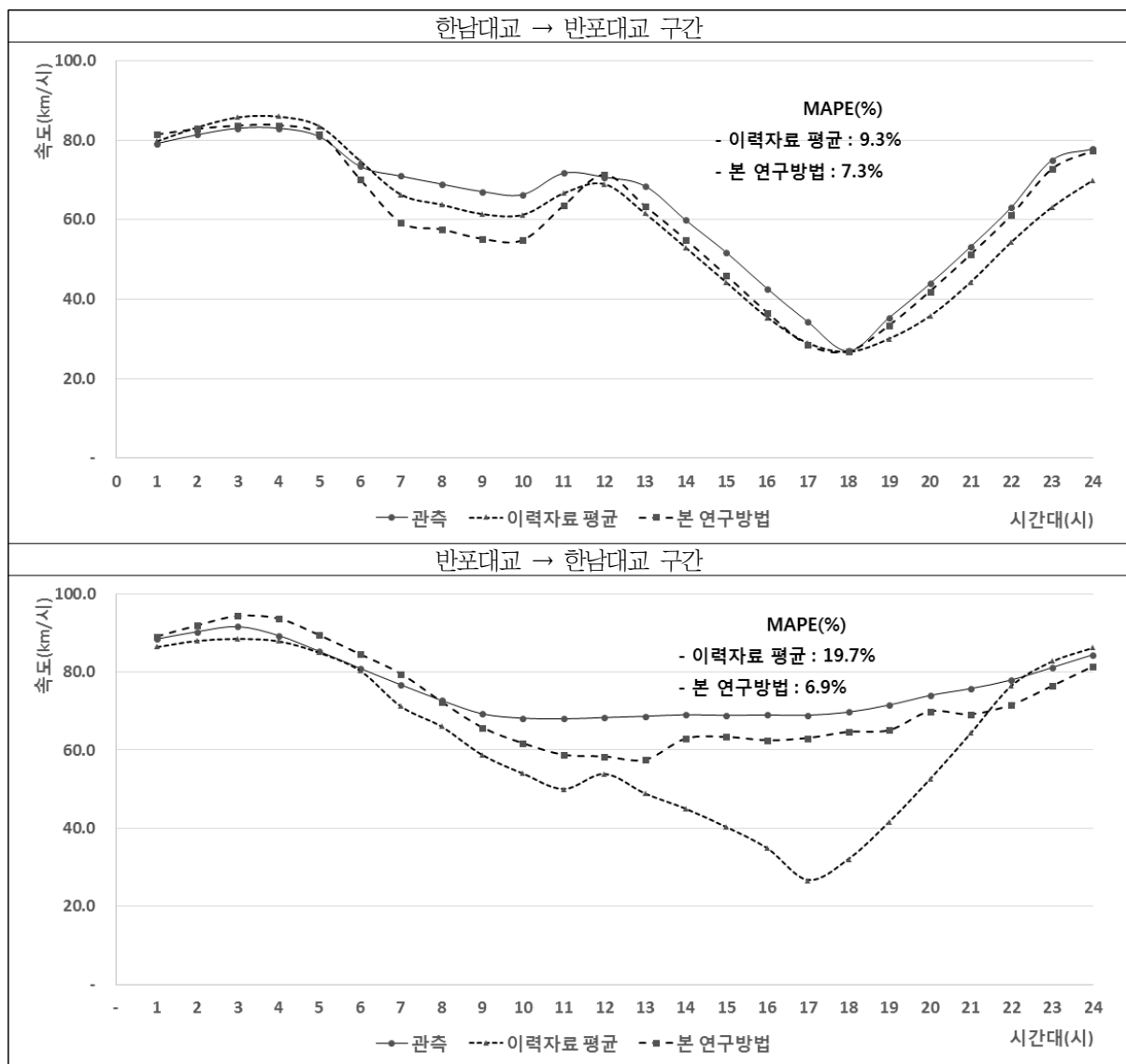
- 시나리오 2의 연속류 구간(경부고속도로, 양재IC~서초IC)은 방향별로 모두 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 이는 시간대별 통행속도가 높은 비첨두시간 때보다 출/퇴근 혼잡이 발생하는 06~10시와 18시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석됨. 즉, 혼잡이 발생하는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문임



<그림 2-73> 시나리오 2의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC)

○ 연속류 구간(올림픽대로, 한남대교~반포대교)

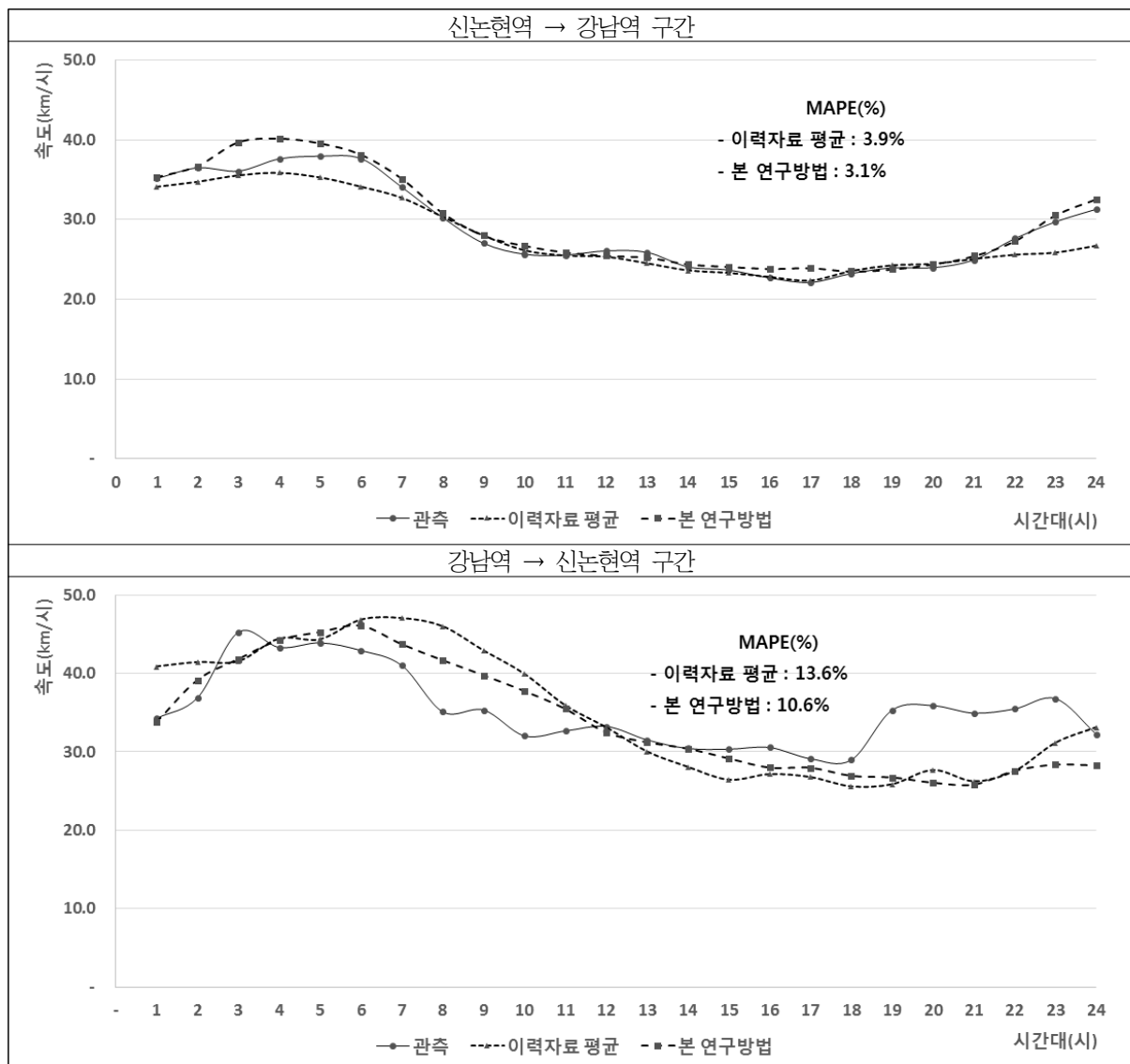
- 시나리오 2의 연속류 구간(올림픽대로, 한남대교~반포대교)은 방향별로 모두 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 즉, 한남대교→반포대교 구간에서는 시간대별 통행속도가 높은 비첨두시간 때보다 퇴근 혼잡이 시작되는 17시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석되었으며, 그 반대 방향은 출근 혼잡이 시작되는 08시 이후부터 전체적으로 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨. 이는 혼잡이 발생하는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 기존 방법의 오차가 더 증가하기 때문임



<그림 2-74> 시나리오 2의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교)

○ 단속류 구간(강남대로, 신논현역~강남역)

- 시나리오 2의 단속류 구간(강남대로, 신논현역~강남역)은 방향별로 모두 기존 과거평균기법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 이는 전체 시간대에서 걸쳐 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석되었으나, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 분포의 변화가 크지 않고 대체적으로 일정한 패턴을 가지므로 연속류 구간에 비해 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨



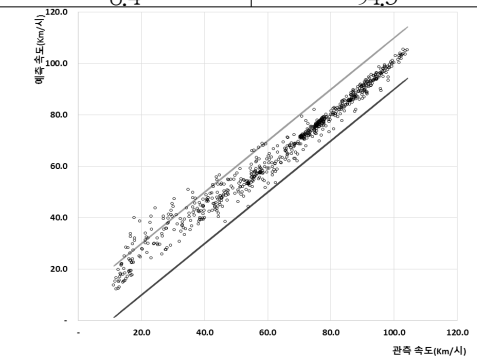
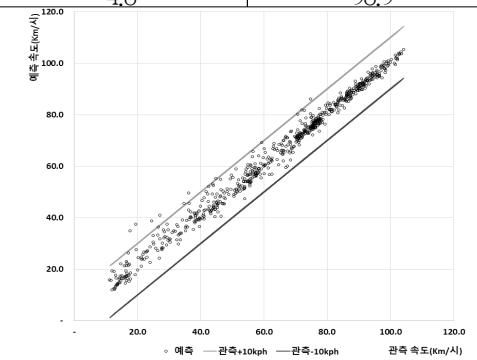
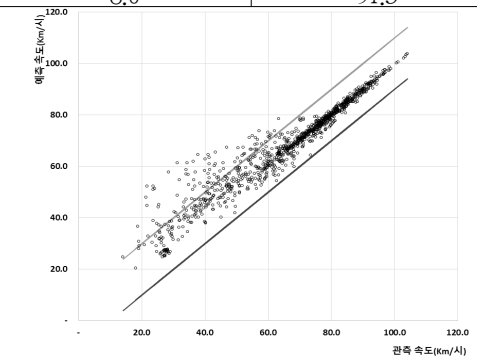
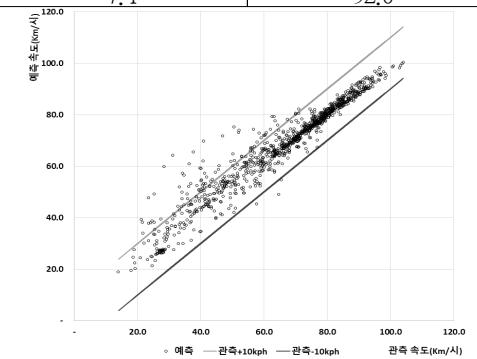
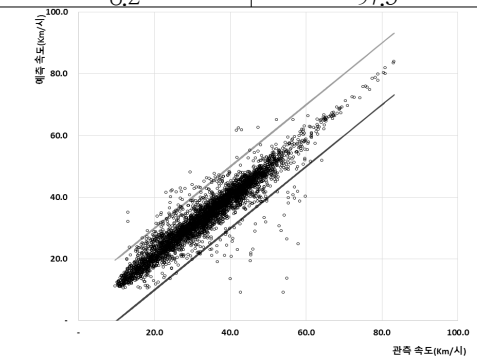
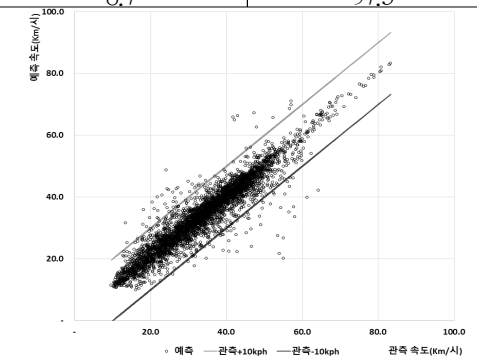
<그림 2-75> 시나리오 2의 단속류 구간 평가결과(강남대로, 신논현역~강남역)

### 3) 시나리오 3

#### ① 분석 네트워크별 평가 결과

- 본 연구의 사례대상 구간 중 연속류 네트워크는 경부고속도로(30개 링크)와 올림픽대로(47개 링크)로 구성되며, 단속류 네트워크는 연속류 네트워크를 제외한 나머지 기타 내부도로(178개 링크)로 구성됨
- 여기서, 시나리오 3의 경우는 2014년 1월 22일(수)을 장래 예측일로 설정하고 위의 분석 네트워크별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 경부고속도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었으며, 이는 시간대별 통행속도가 높은 비첨두시간 때보다 통행속도가 낮은 혼잡한 시간대에서 본 연구방법의 예측력이 좀 더 높은 것으로 분석됨. 즉, 비첨두시간대 보다 혼잡이 발생하는 첨두시간대의 경우에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문임
  - MAPE(%) : 기존의 과거평균기법 8.4% → 본 연구의 예측모형 4.8%, MAPE 3.6% 감소
  - 적중률(%) : 기존의 과거평균기법 94.3% → 본 연구의 예측모형 98.9%, 적중률 4.6% 증가
- 올림픽대로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었으나, 위의 경부고속도로 보다 개선 효과는 비교적 낮게 분석되었음. 이는 경부고속도로의 분석 네트워크 보다 올림픽대로의 경우가 전체적으로 과거 속도자료의 일별/시간대별 분포 변화가 크지 않고 비교적 일정한 패턴을 가지는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 8.0% → 본 연구의 예측모형 7.4%, MAPE 0.6% 감소
  - 적중률(%) : 기존의 과거평균기법 91.3% → 본 연구의 예측모형 92.0%, 적중률 0.7% 증가
- 기타 내부도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)는 다소 증가하나 적중률(%)은 비슷한 것으로 분석되었음. 이는 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 비교적 크지 않으므로 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 8.2% → 본 연구의 예측모형 8.7%, MAPE 0.5% 증가
  - 적중률(%) : 기존의 과거평균기법 97.5% → 본 연구의 예측모형 97.5%, 적중률 변화 없음

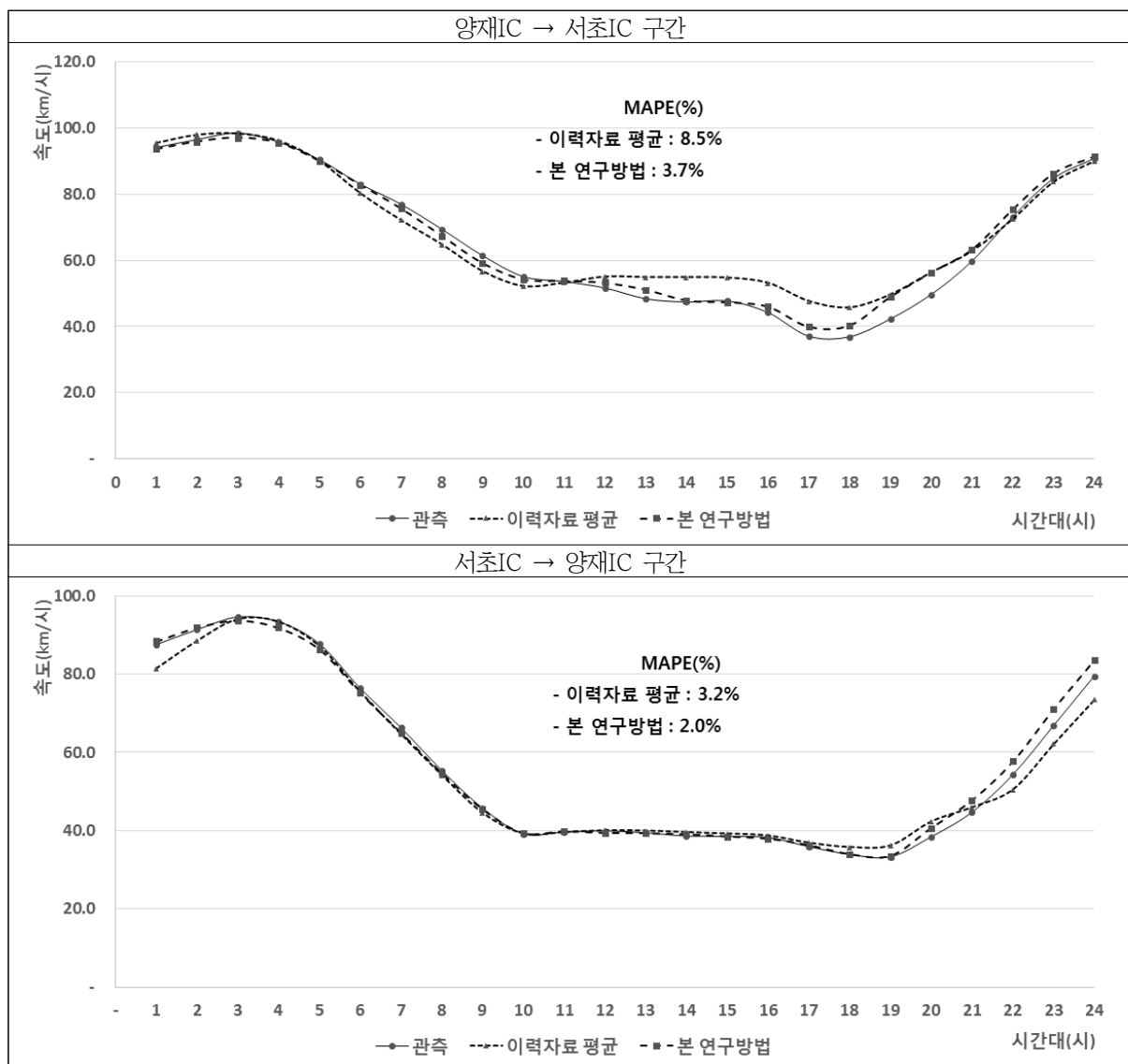
&lt;표 2-31&gt; 전체 분석 네트워크의 평가 결과(시나리오 3)

구분	적용 방법론			
	이력자료 평균		본 연구 방법	
	MAPE(%)	적중률(%)	MAPE(%)	적중률(%)
경부고속도로 (30개 링크)	8.4	94.3	4.8	98.9
				
올림픽대로 (47개 링크)	8.0	91.3	7.4	92.0
				
기타 내부도로 (178개 링크)	8.2	97.5	8.7	97.5
				

## ② 주요 링크구간의 평가 결과

## ○ 연속류 구간(경부고속도로, 양재IC~서초IC)

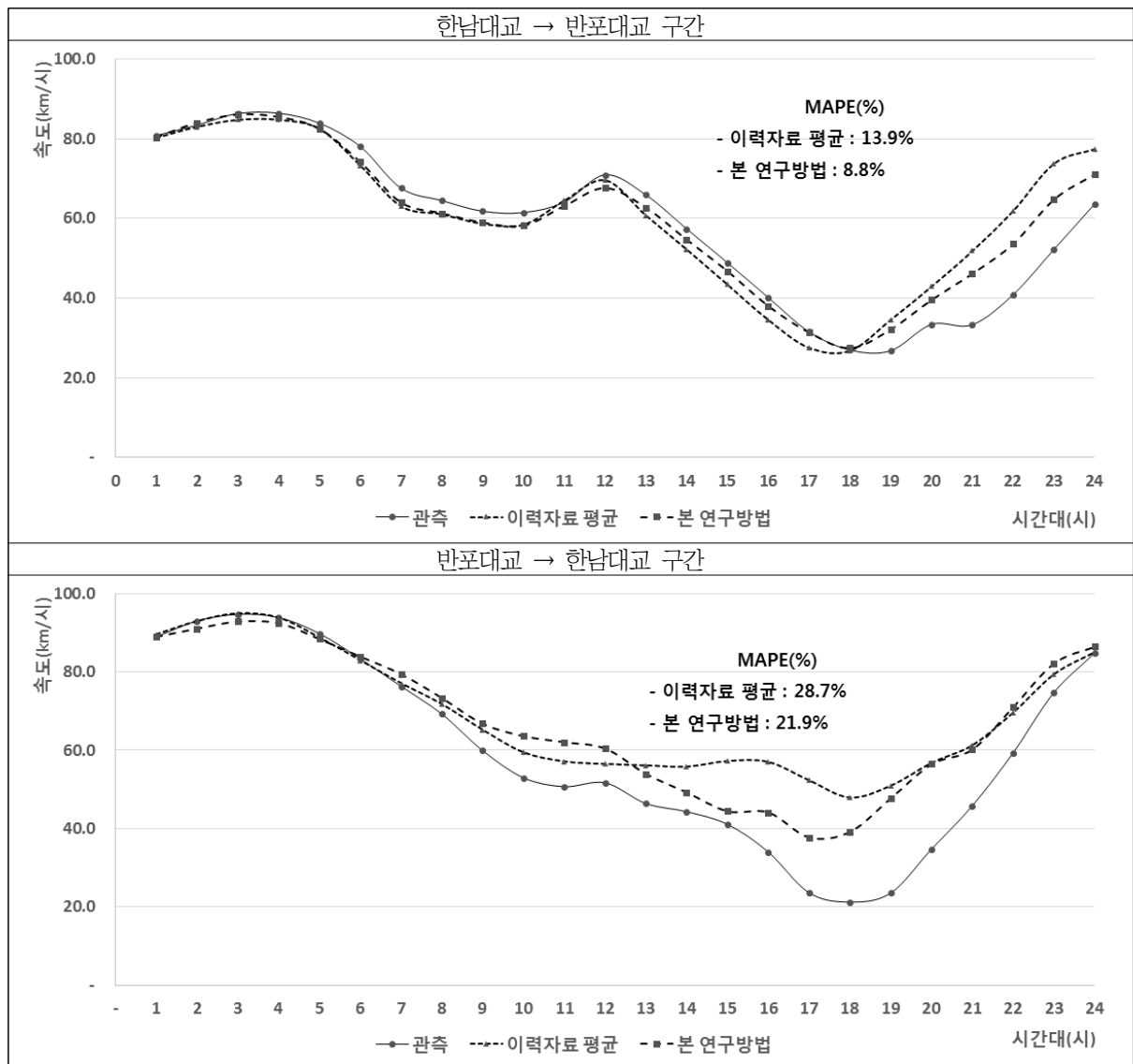
- 시나리오 3의 연속류 구간(경부고속도로, 양재IC~서초IC)은 방향별로 모두 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨. 즉, 시간대별 통행속도가 높은 비첨두시간대가 아닌 교통량 증가로 인하여 혼잡이 발생하는 출/퇴근 시간 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석됨. 이는 혼잡이 발생하는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문임



&lt;그림 2-76&gt; 시나리오 3의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC)

○ 연속류 구간(올림픽대로, 한남대교~반포대교)

- 시나리오 3의 연속류 구간(올림픽대로, 한남대교~반포대교)은 방향별로 모두 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨.
- 즉, 한남대교→반포대교 구간에서는 시간대별 통행속도가 높은 비첨두시간 때보다 퇴근 혼잡이 시작되는 17시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석되었으며, 그 반대 방향은 교통량이 비교적 많은 주간 시간대 13시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석됨. 이는 혼잡이 발생하는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문임

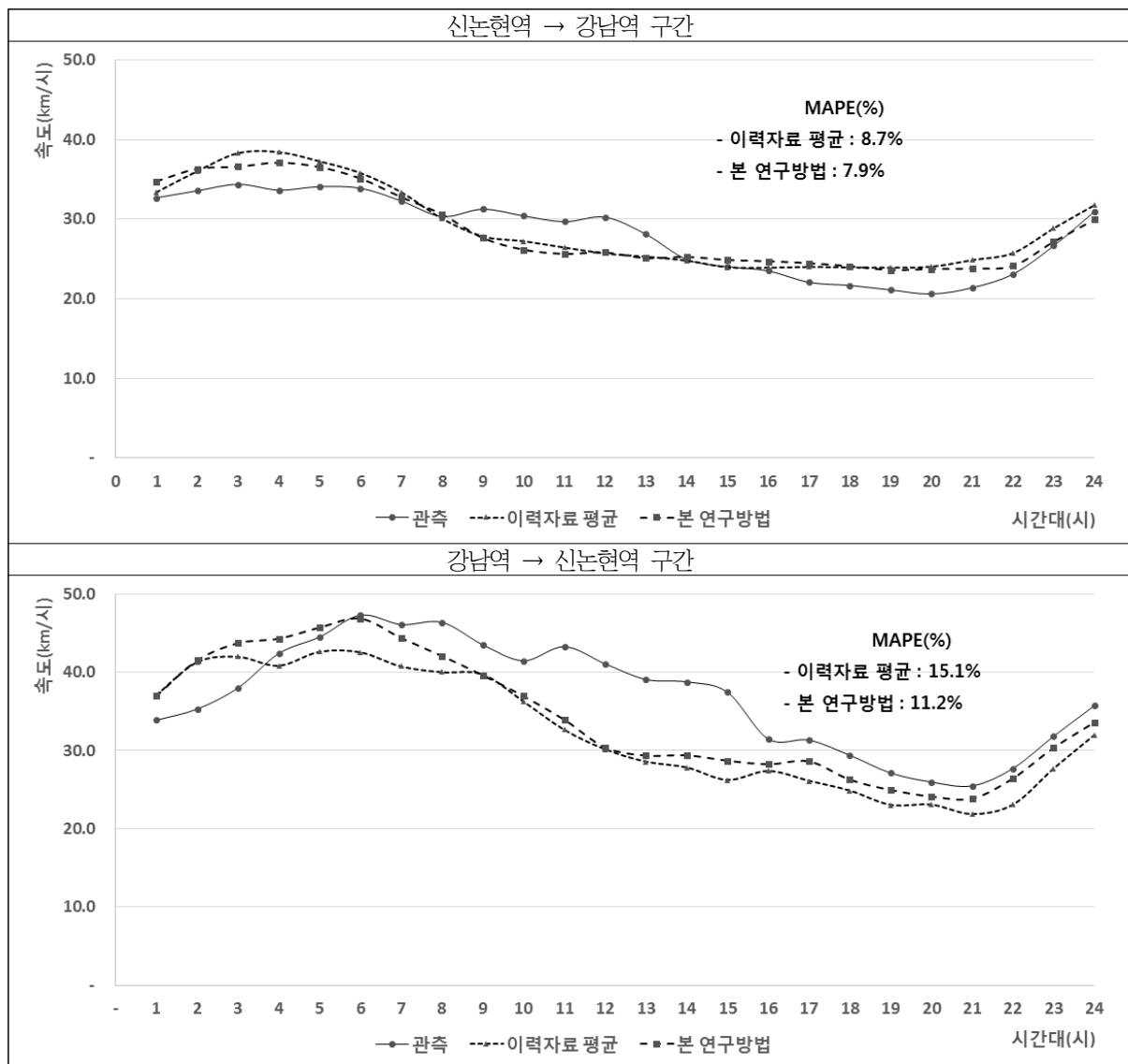


<그림 2-77> 시나리오 3의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교)



○ 단속류 구간(강남대로, 신논현역~강남역)

- 시나리오 3의 단속류 구간(강남대로, 신논현역~강남역)은 방향별로 모두 기존 과거평균기법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 이는 전체 시간대에서 걸쳐 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석되었으나, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 비교적 크지 않으므로 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨



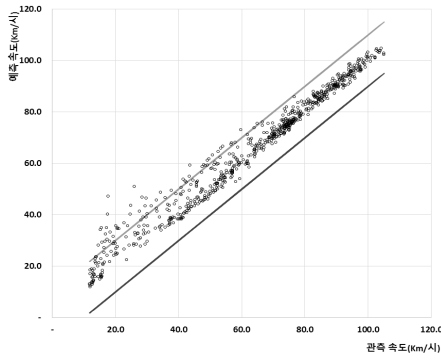
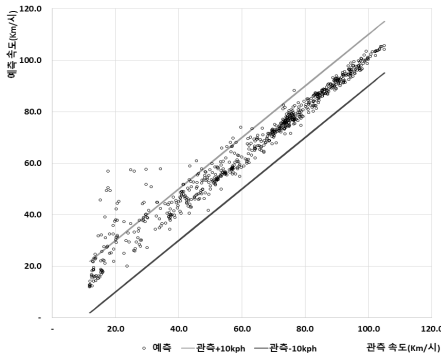
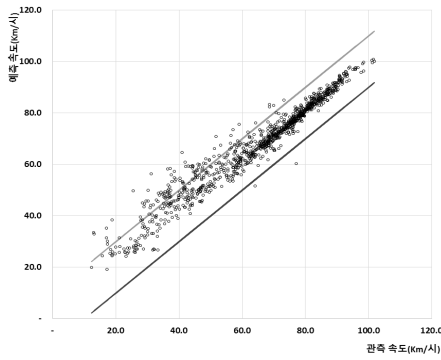
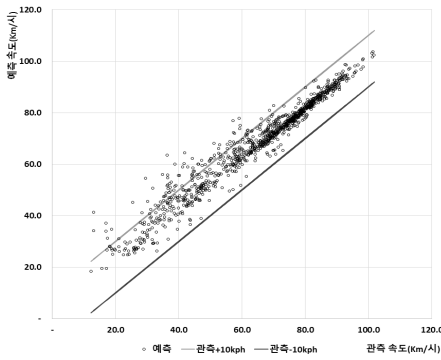
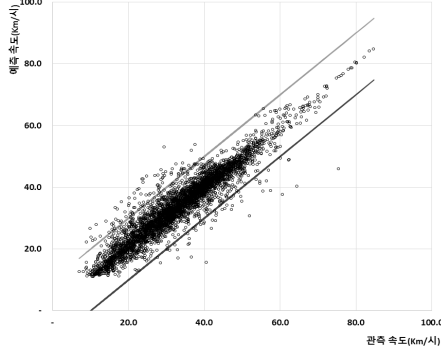
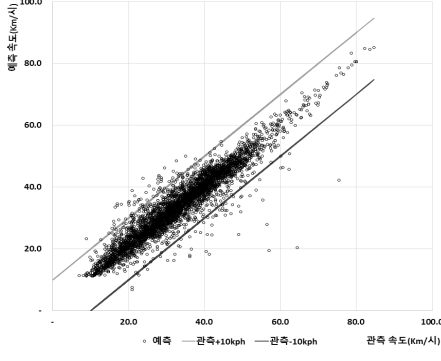
<그림 2-78> 시나리오 3의 단속류 구간 평가결과(강남대로, 신논현역~강남역)

#### 4) 시나리오 4

##### ① 분석 네트워크별 평가 결과

- 본 연구의 사례대상 구간 중 연속류 네트워크는 경부고속도로(30개 링크)와 올림픽대로(47개 링크)로 구성되며, 단속류 네트워크는 연속류 네트워크를 제외한 나머지 기타 내부도로(178개 링크)로 구성됨
- 여기서, 시나리오 4의 경우는 2014년 1월 23일(목)을 장래 예측일로 설정하고 위의 분석 네트워크별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 경부고속도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)는 다소 증가하나 적중률(%)은 비슷한 것으로 분석되었음. 이는 앞서 시나리오 1~3의 경우와는 다르게 해당 과거 속도자료의 일별/시간대별 패턴의 변화가 크지 않은 것을 의미하며, 이에 따라 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 9.3% → 본 연구의 예측모형 11.2%, MAPE 1.9% 증가
  - 적중률(%) : 기존의 과거평균기법 92.5% → 본 연구의 예측모형 92.5%, 적중률 변화 없음
- 올림픽대로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)는 다소 증가하나 적중률(%)은 좀 더 향상하는 것으로 분석되었으며, 이는 위의 경부고속도로의 경우와 동일함
  - MAPE(%) : 기존의 과거평균기법 8.3% → 본 연구의 예측모형 8.6%, MAPE 0.2% 증가
  - 적중률(%) : 기존의 과거평균기법 90.2% → 본 연구의 예측모형 91.4%, 적중률 1.2% 증가
- 기타 내부도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었음. 이는 전체 시간대에서 걸쳐 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석되었으나, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 대체적으로 크지 않으므로 기존 방법의 예측 결과와의 큰 차이는 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 9.4% → 본 연구의 예측모형 9.1%, MAPE 0.3% 감소
  - 적중률(%) : 기존의 과거평균기법 96.7% → 본 연구의 예측모형 97.1%, 적중률 0.4% 증가

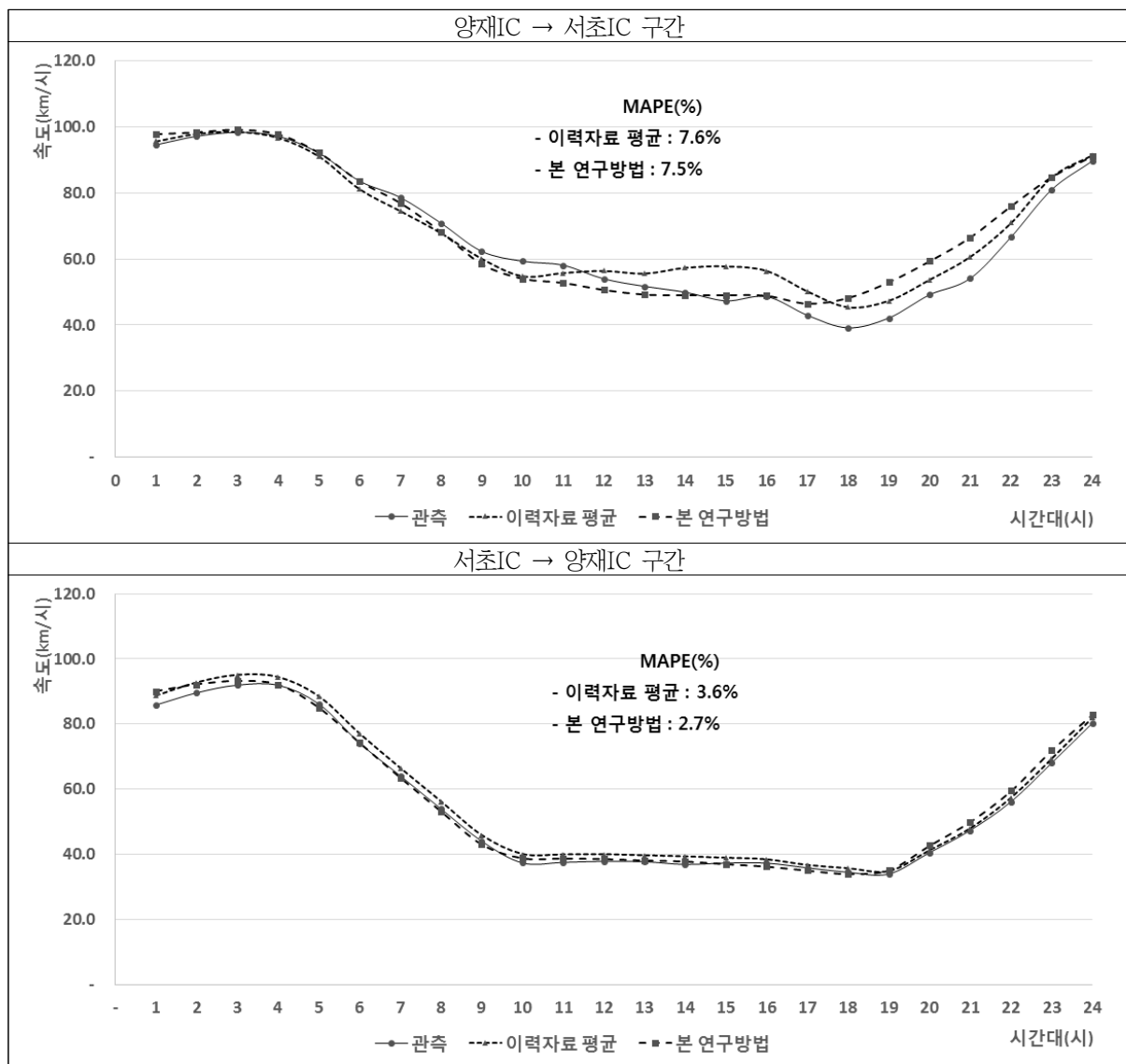
&lt;표 2-32&gt; 전체 분석 네트워크의 평가 결과(시나리오 4)

구분	적용 방법론			
	이력자료 평균		본 연구 방법	
	MAPE(%)	적중률(%)	MAPE(%)	적중률(%)
경부고속도로 (30개 링크)	9.3	92.5	11.2	92.5
				
울림픽대로 (47개 링크)	8.3	90.2	8.6	91.4
				
기타 내부도로 (178개 링크)	9.4	96.7	9.1	97.1
				

## ② 주요 링크구간의 평가 결과

### ○ 연속류 구간(경부고속도로, 양재IC~서초IC)

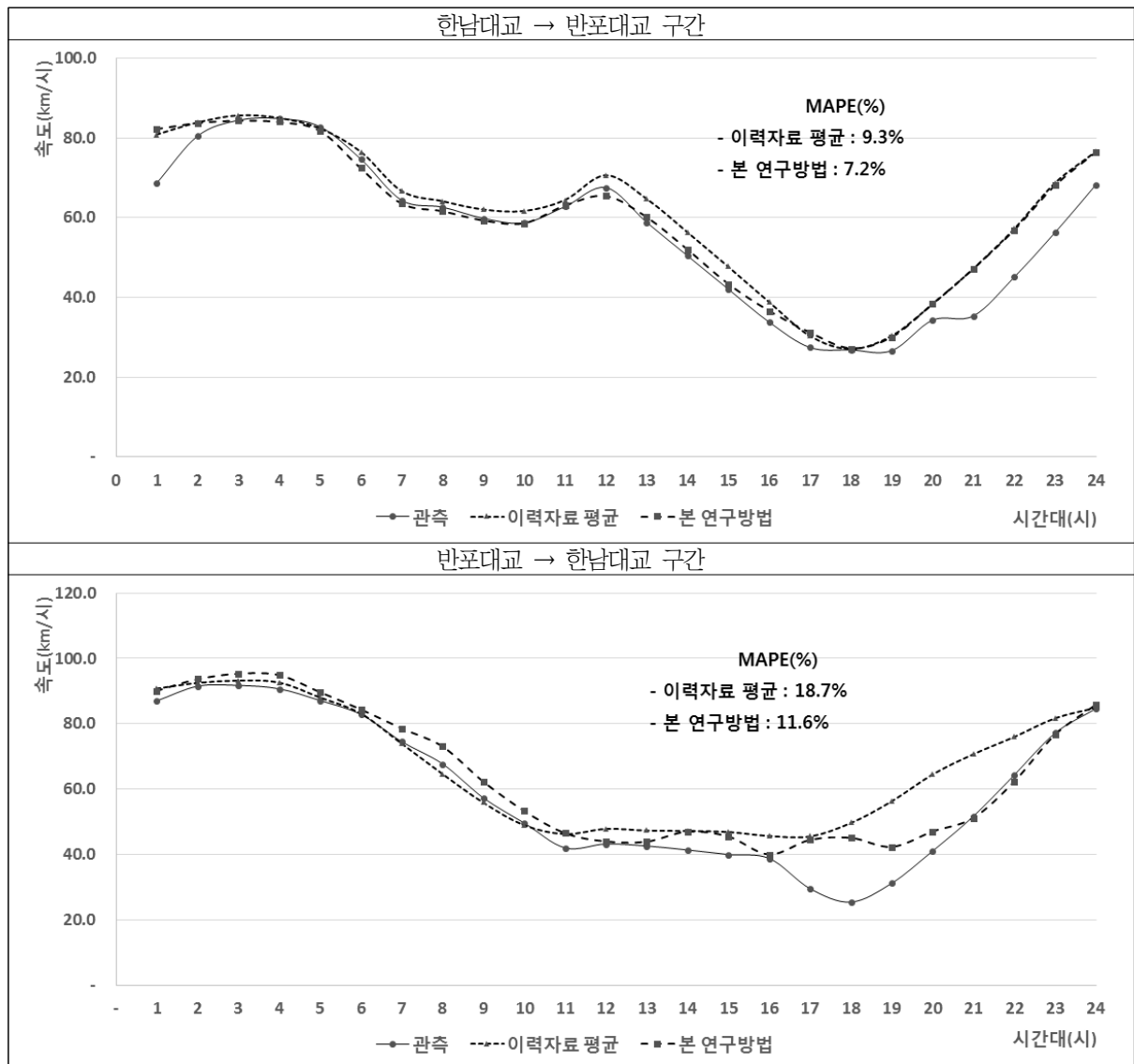
- 시나리오 4의 연속류 구간(경부고속도로, 양재IC~서초IC)은 방향별로 모두 기존 방법에 비해 예측결과의 차이가 크지 않음
- 즉, 모든 방향에서 과거 속도자료의 일별/시간대별 패턴 변화가 일정하여 두 적용 방법의 차이가 크지 않은 것으로 분석되었음



<그림 2-79> 시나리오 4의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC)

○ 연속류 구간(올림픽대로, 한남대교~반포대교)

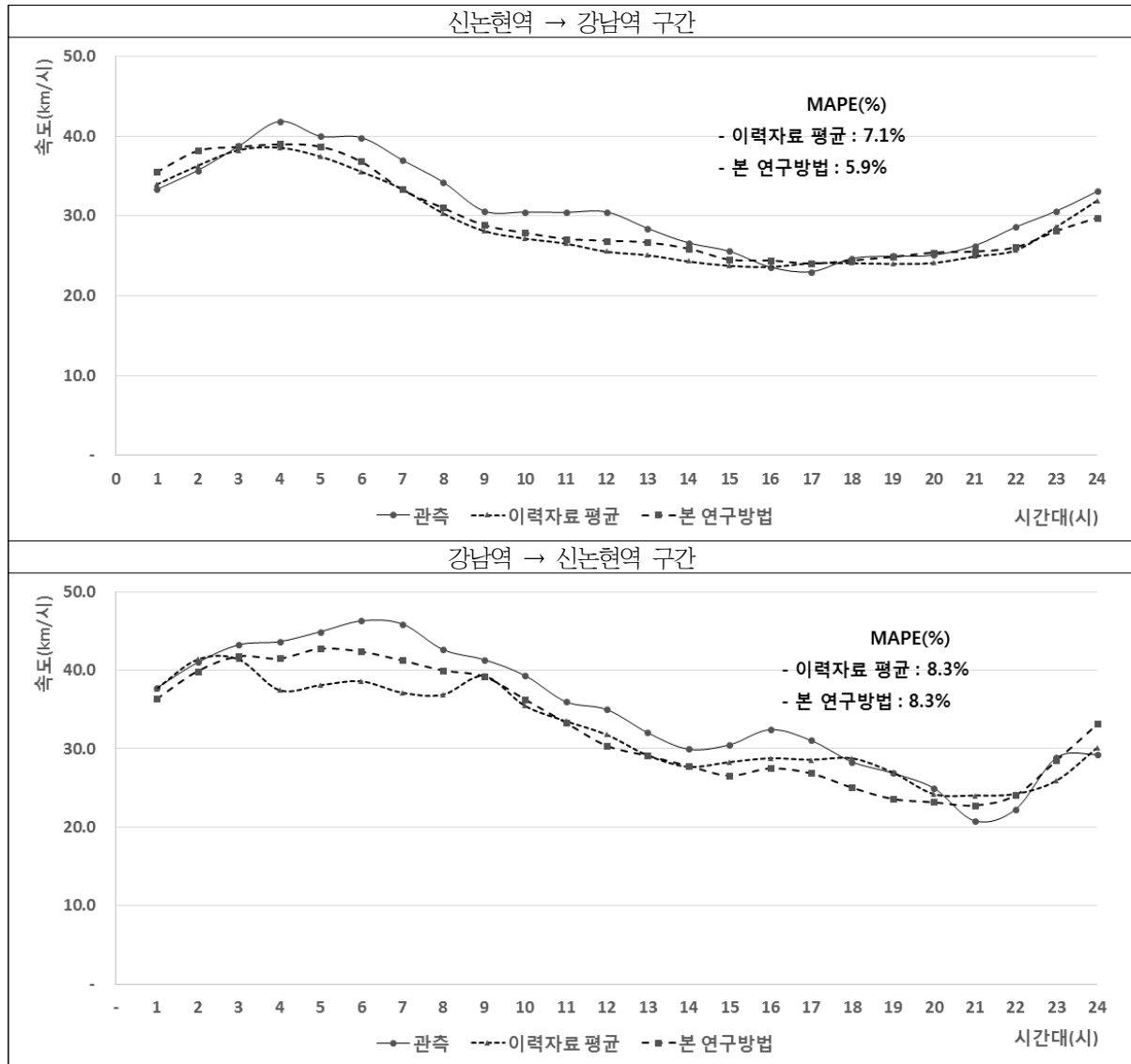
- 시나리오 4의 연속류 구간(올림픽대로, 한남대교~반포대교)은 방향별로 모두 기존 과거평균 기법 보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 즉, 한남대교→반포대교 구간에서는 과거 속도자료의 일별/시간대별 패턴 변화가 일정하여 두 적용방법의 차이가 크지 않은 것으로 분석되었으며, 그 반대 방향은 시간대별 통행속도가 높은 비첨두시간 때가 아닌 퇴근시간 혼잡이 다가오는 16시 이후부터 21시까지 본 연구방법의 예측력이 더 높은 것으로 분석됨. 이는 혼잡이 발생하는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문임



<그림 2-80> 시나리오 4의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교)

○ 단속류 구간(강남대로, 신논현역~강남역)

- 시나리오 4의 단속류 구간(강남대로, 신논현역~강남역)은 방향별로 모두 기존 방법에 비해 예측결과와 차이가 크지 않음
- 이는 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 비교적 크지 않으므로 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨



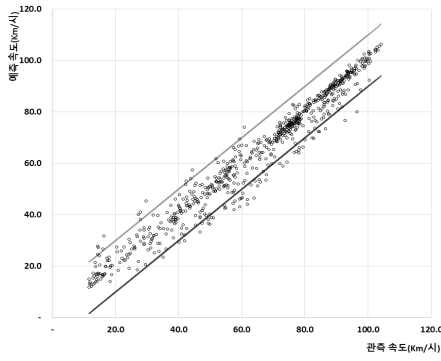
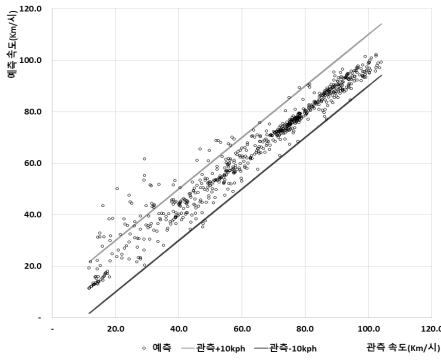
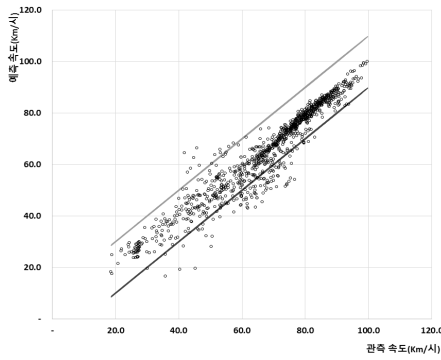
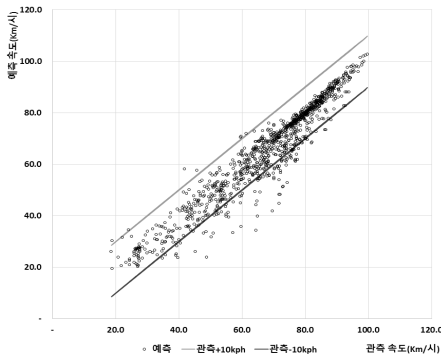
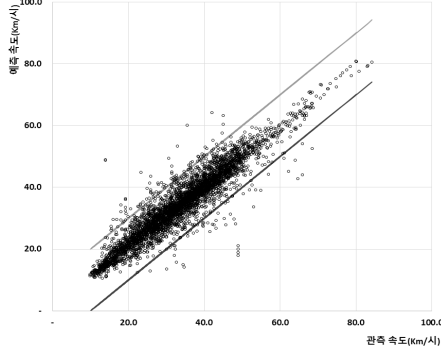
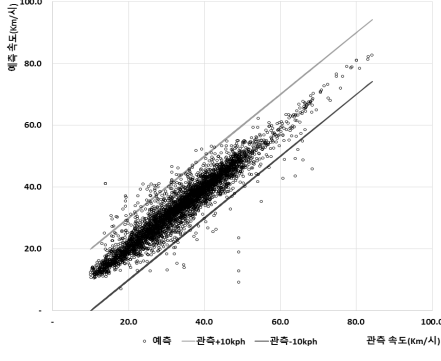
<그림 2-81> 시나리오 4의 단속류 구간 평가결과(강남대로, 신논현역~강남역)

## 5) 시나리오 5

### ① 분석 네트워크별 평가 결과

- 본 연구의 사례대상 구간 중 연속류 네트워크는 경부고속도로(30개 링크)와 올림픽대로(47개 링크)로 구성되며, 단속류 네트워크는 연속류 네트워크를 제외한 나머지 기타 내부도로(178개 링크)로 구성됨
- 여기서, 시나리오 5의 경우는 2014년 1월 24일(금)을 장래 예측일로 설정하고 위의 분석 네트워크별로 기존 과거평균(Historical average)기법과 본 연구의 예측모형을 적용하여 장래 통행속도 예측결과를 비교·분석하였음
- 경부고속도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)는 다소 증가하나 적중률(%)은 오히려 향상하는 것으로 분석되었음. 앞서 시나리오 1~3의 경우와는 다르게 해당 과거 속도자료의 일별/시간대별 패턴의 변화가 크지 않은 것을 의미하며, 이에 따라 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 7.1% → 본 연구의 예측모형 8.4%, MAPE 1.3% 증가
  - 적중률(%) : 기존의 과거평균기법 91.4% → 본 연구의 예측모형 93.1%, 적중률 1.7% 증가
- 올림픽대로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)는 다소 증가하나 적중률(%)은 오히려 향상하는 것으로 분석되었으며, 이는 위의 경부고속도로의 경우와 동일함
  - MAPE(%) : 기존의 과거평균기법 6.7% → 본 연구의 예측모형 6.9%, MAPE 0.2% 증가
  - 적중률(%) : 기존의 과거평균기법 89.2% → 본 연구의 예측모형 89.9%, 적중률 0.7% 증가
- 기타 내부도로의 경우, 본 연구의 예측모형이 기존의 과거평균기법 보다 MAPE(%)와 적중률(%) 모두 향상되는 것으로 분석되었음. 이는 전체 시간대에서 걸쳐 본 연구의 예측모형이 좀 더 예측력이 높은 것으로 분석되었으나, 도시부 단속류의 경우는 신호기에 의한 제약으로 인해 과거 속도자료의 일별/시간대별 패턴 변화가 대체적으로 크지 않으므로 기존 방법의 예측 결과와의 큰 차이는 없는 것으로 분석됨
  - MAPE(%) : 기존의 과거평균기법 9.2% → 본 연구의 예측모형 7.9%, MAPE 1.3% 감소
  - 적중률(%) : 기존의 과거평균기법 96.6% → 본 연구의 예측모형 97.7%, 적중률 1.1% 증가

&lt;표 2-33&gt; 전체 분석 네트워크의 평가 결과(시나리오 5)

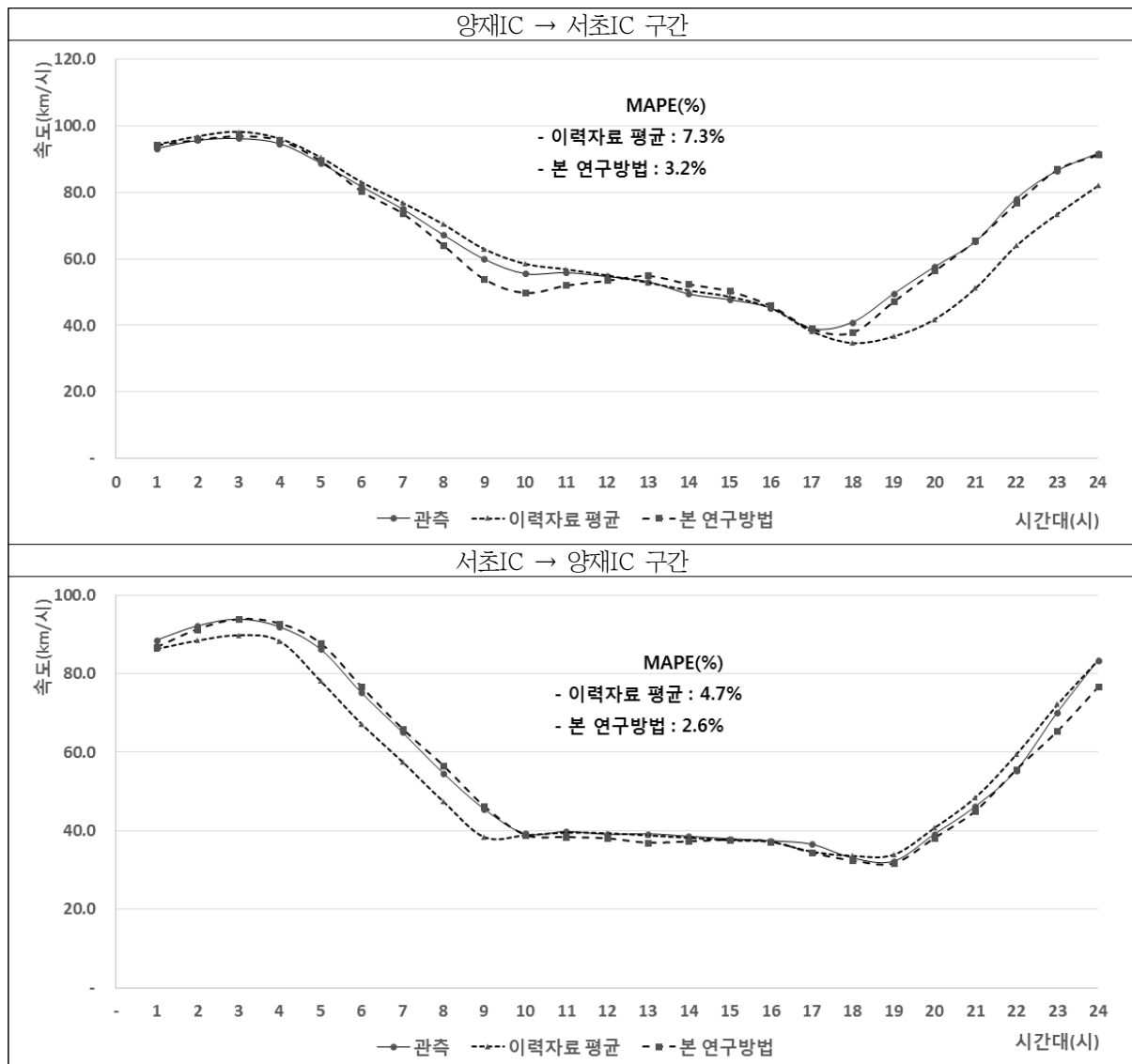
구분	적용 방법론			
	이력자료 평균		본 연구 방법	
	MAPE(%)	적중률(%)	MAPE(%)	적중률(%)
경부고속도로 (30개 링크)	7.1	91.4	8.4	93.1
				
울림픽대로 (47개 링크)	6.7	89.2	6.9	89.9
				
기타 내부도로 (178개 링크)	9.2	96.6	7.9	97.7
				



## ② 주요 링크구간의 평가 결과

### ○ 연속류 구간(경부고속도로, 양재IC~서초IC)

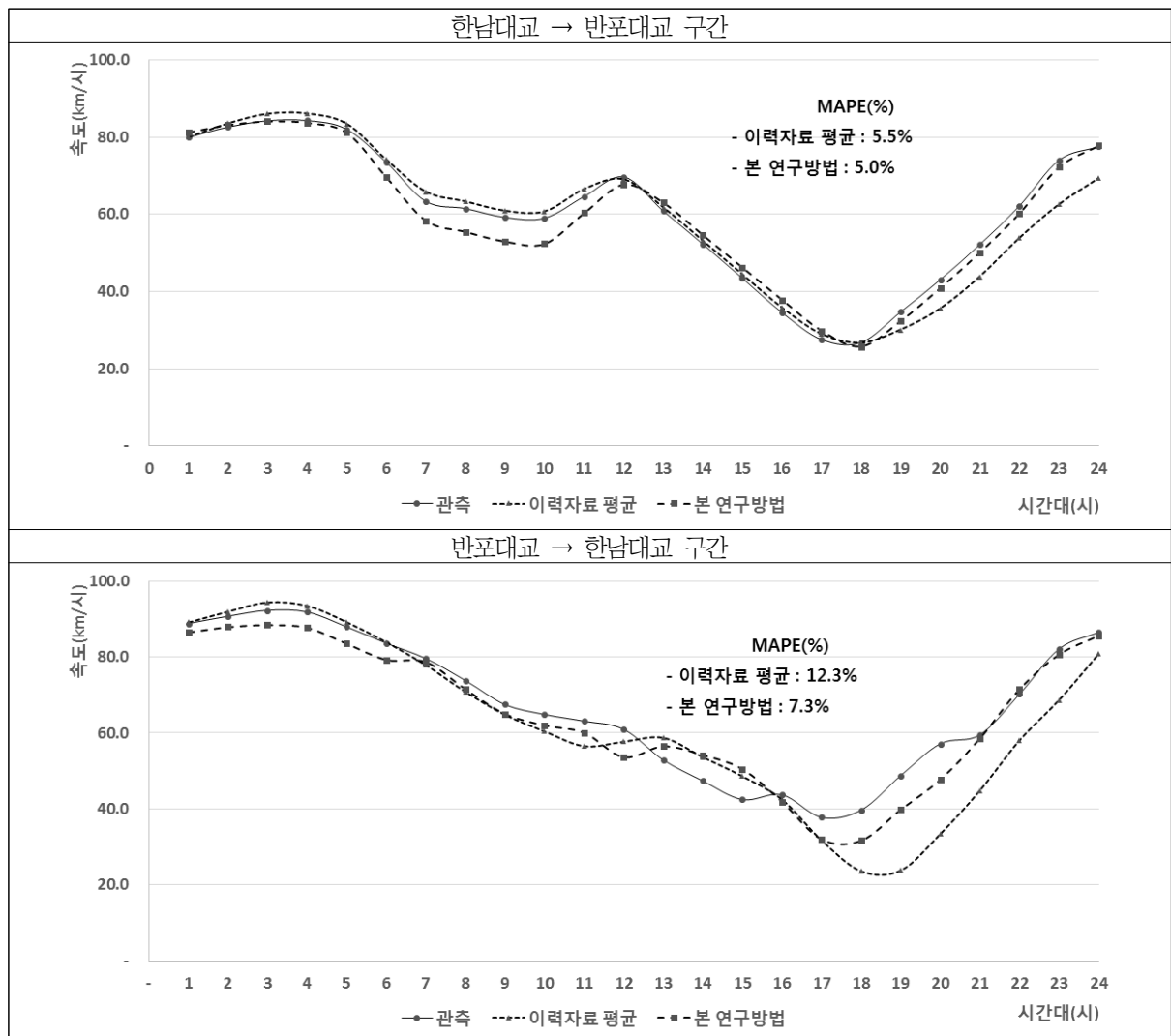
- 시나리오 5의 연속류 구간(경부고속도로, 양재IC~서초IC)은 방향별로 모두 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 즉, 시간대별 통행속도가 높은 비첨두시간 때가 아닌 출/퇴근 혼잡이 시작되는 시간대에서 본 연구방법의 예측력이 더 높은 것으로 분석됨. 이는 혼잡이 시작되는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문임



<그림 2-82> 시나리오 5의 연속류 구간 평가결과(경부고속도로, 양재IC~서초IC)

○ 연속류 구간(올림픽대로, 한남대교~반포대교)

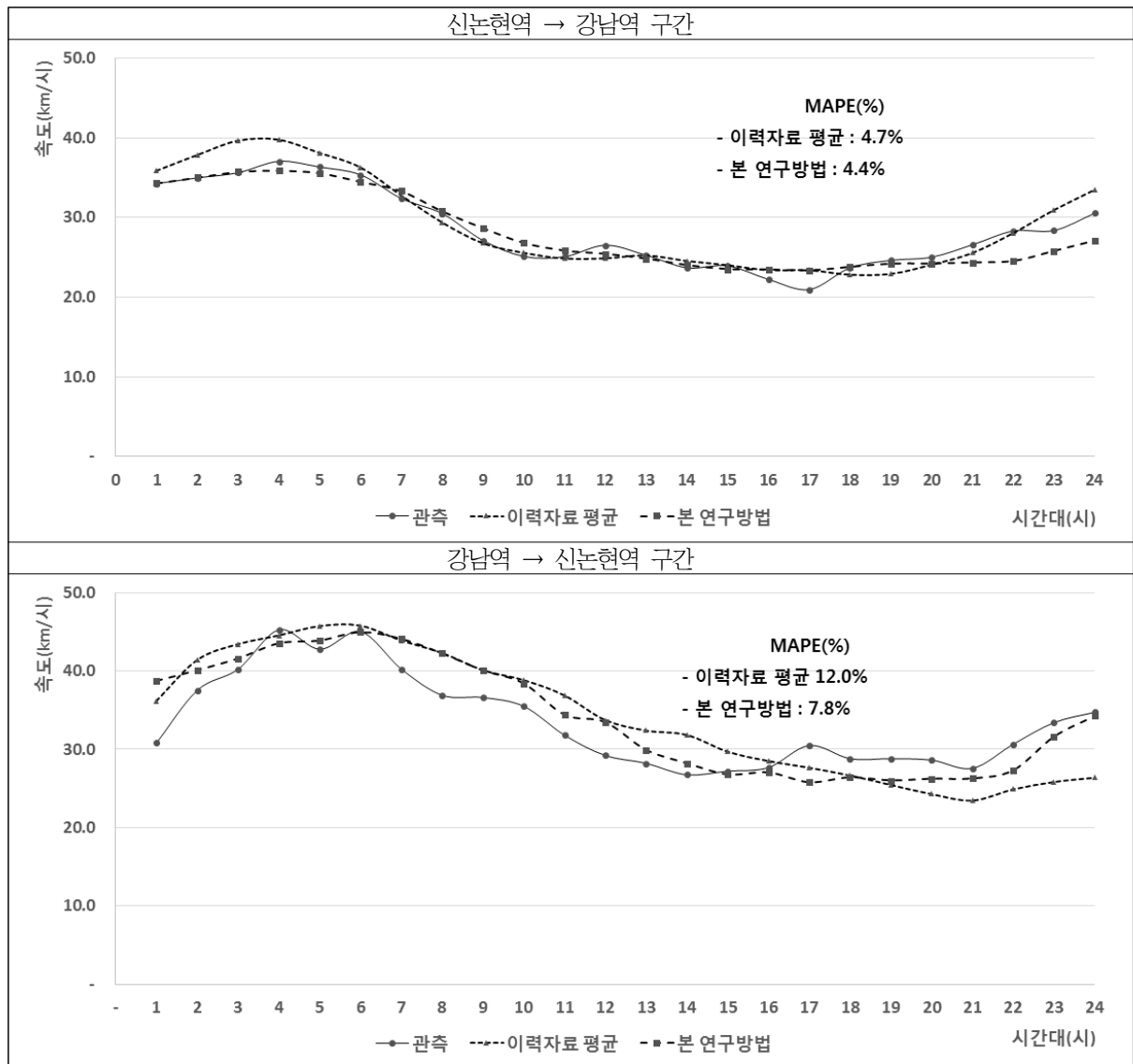
- 시나리오 5의 연속류 구간(올림픽대로, 한남대교~반포대교)은 한남대교→반포대교 방향의 경우는 기존 방법과의 예측결과 차이가 크지 않으나, 반대 방향의 경우는 기존 방법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 즉, 한남대교→반포대교 구간에서는 과거 속도자료의 일별/시간대별 패턴 변화가 일정하여 두 적용방법의 차이가 없는 것으로 분석되었으며, 그 반대 방향은 퇴근 혼잡이 시작되는 16시 이후부터 본 연구방법의 예측력이 더 높은 것으로 분석됨. 이는 혼잡이 시작되는 시간대에서 과거 속도자료의 일별/시간대별 패턴 변화가 크므로 이에 따라 기존 방법의 오차가 더 증가하기 때문이다



<그림 2-83> 시나리오 5의 연속류 구간 평가결과(올림픽대로, 한남대교~반포대교)

○ 단속류 구간(강남대로, 신논현역~강남역)

- 시나리오 5의 단속류 구간(강남대로, 신논현역~강남역)은 방향별로 모두 기존 과거평균기법보다 본 연구방법의 예측력이 더 높은 것으로 분석됨
- 특히 강남역 → 신논현역 방향의 전체 시간대에서 걸쳐 본 연구방법의 예측력이 높은 것으로 분석되었으나, 그 반대 방향의 경우는 과거 속도자료의 일별/시간대별 패턴 변화가 대체적으로 크지 않으므로 기존 방법의 예측 결과와 큰 차이가 없는 것으로 분석됨



<그림 2-84> 시나리오 5의 단속류 구간 평가결과(강남대로, 신논현역~강남역)

## 제7절 결론 및 향후 연구

### 1. 결론

- 본 연구는 개별 Car Navigation의 차량이동궤적 데이터를 이용하여 장래 교통예보를 위한 핵심 요소기술(①과거 이력자료 DB구축 방법론 개발, ②장래 통행속도 예측방법론 개발)을 개발하였으며, 개발모형의 평가 결과는 보편적으로 이용되고 있는 기존의 이력자료 평균(Historical average)방법 보다 전체적으로 우수한 예측력을 보였음
- 또한, 본 연구의 핵심 요소기술 개선방향 및 실용화를 위하여 향후 개념적인 시스템 구축방안도 제시하였음
- 본 연구의 빅데이터(Car Navigation의 차량이동궤적 자료) 기반 교통예보(익일 링크별/시간대별 통행속도 예측)를 위한 핵심 요소기술 개발에 대한 결론을 요약하면 아래와 같음
  - 첫째, 기존의 단기예측(일중 소통상태 예측)의 예측영역을 중장기 예측영역(익일 소통상태 예측)으로 확장하였으며, 이를 위하여 통계기반의 인위적 모델링에 의존하지 않고 과거 데이터를 탐색하여 내재된 정보를 도출하는 데이터 탐색 기반의 예측모형을 적용하였음
  - 둘째, Network의 소통상태 변화는 Link의 국부적인 변동보다는 네트워크 차원에서의 순환성(Recurrency)과 경향성(Pattern), 그리고 내재된 알려지지 않은 Network의 상태 진화에 대한 교통 현상(Traffic Behavior)인 속도와 수요(Probe 건수)의 변화 추세에 의하여 영향을 받는다는 것을 알 수가 있었음
  - 셋째, 모형의 정확도가 요구수준을 만족하면 모형의 연산 수행속도와 더불어 운영상의 용이성을 고려해야 함. 개발 모형의 예측력이 뛰어나더라도 시스템 측면에서 연산 수행시간이 추정정보를 활용할 수 있는 시간적 범위를 초과하거나, 운영측면에서 지속적으로 모형의 파라미터를 조정해야 한다면 실제 모형의 적용성은 현저히 떨어지게 됨. 따라서 본 연구에서는 파라미터의 조정이 시스템적으로 간편하면서 모형의 연산속도가 빠른 모형으로 개발하였음
  - 넷째, 현재 보다 방대한 양의 Car Navigation 자료를 이용할 수 있을 경우, 익일 이상의 주간 단위(1주일)의 교통소통 상태까지도 예측할 수 있는 가능성을 제시하였음
  - 다섯째, 본 연구에서 개발한 예측모형은 향후 교통예보 시스템 탑재를 목표로 하기 때문에 입력 자료의 요구조건, 파라미터의 정산, 예측모형 적용 등에 대하여 관련 전문가가 아닌 일반 현장 요원도 시스템을 운영함에 있어 큰 어려움이 없도록 개발하였음

## 2. 향후 연구

- 본 연구에서는 지난 4개월 간(2013년 10월 1일 ~ 2014년 1월 31일)의 Car Navigation 자료만을 이용하여 장래 구간 통행속도 예측모형을 개발하였음. 즉, 구축된 과거 입력 자료의 양이 많지 않으므로 예측 모형의 k군집 추출 단계에서 충분한 표본 수를 확보하지 못하였음
- 향후 개발 모형의 예측력을 더욱 높이기 위해서는 다 년간 축적된 Car Navigation 자료의 이용이 필요하며, 이를 위하여 실시간 Car Navigation 이동궤적 정보를 수집 및 저장할 수 있는 교통예보시스템 개발이 요구되어짐
- 즉, 개발모형에 대한 장래 소통상태 예측기법 및 현행 상태를 재구성하기 위한 다양한 형태의 접근방법을 적용하여 모형을 고도화 할 수 있을 것으로 판단됨
- 본 연구의 교통예보는 익일의 교통 소통상태만을 예측하지만, 향후 입력 자료가 충분히 확보된다면 익일 이상의 주간 교통예보까지도 예측이 가능할 것으로 판단됨
- 마지막으로 익일 이상의 중장기 교통 소통 상태를 예측할 때 발생하는 오차는 특히 교통사고, 공사 등과 같은 이벤트 상황 즉, 비 반복적 행태의 소통상태가 발생 시 이를 고려하여 장래 소통 상태를 예측하는 연구가 지속적으로 이루어져야 된다고 봄



## **제3장 교통정보DB 통합관리 시스템 설계**

---

**제1절 추진배경 및 경위**

**제2절 교통정보 공유현황**

**제3절 교통정보 통합관리 시스템 구축 방안**

**제4절 지자체 교통정보 현황**





## 제3장 교통정보DB 통합관리 시스템 설계

### 제1절 추진배경 및 경위

#### 1. 추진배경

- 국회 등에서 KTDB를 활용한 예측 교통량과 관측 교통량간의 오차가 발생함에 따라 KTDB의 신뢰성 제고 필요성 지적
- 국가교통통계 생성을 위한 교통정보DB의 연계·공유가 필요
- 이에 공공기관·지자체 등에서 수집하는 교통량·속도 정보 등을 연계·공유하여 교통정보 활용을 극대화하는 교통정보 통합관리 방안 마련('14.9.30), 단계적 추진
  - 1단계 국토부·공공기관 교통정보, 2단계 지자체 교통정보, 3단계 민간 교통정보 통합관리
- 공공기관 교통정보 연계를 현행 한국교통연구원·한국도로공사에서 한국건설기술연구원·교통안전공단까지 추가 확대·운영
  - 국토부·한국교통연구원(KTDB센터), 한국도로공사(국가교통정보센터) 교통정보DB 협의회 구성·운영 중('14.7)

#### 2. 추진경위

- 2014. 6. 국가교통DB센터와 국가교통정보센터 연계 협력 회의
  - 양 센터간 인터넷 전용선 설치 등 수집정보 연계·공유 방안 논의
- 2014. 7. 교통정보 연계·공유를 위한 교통정보DB 협의회 구성
  - 국토부(교통정책조정과, 첨단도로환경과), 한국교통연구원(국가교통DB센터), 한국도로공사(국가교통정보센터)로 구성
- 2014. 8. 제1차 교통정보DB 협의회 회의 개최
- 2014. 9. 국가교통DB 신뢰도 제고를 위한 교통정보 통합관리 방안 수립
- 2014.11. 교통정보DB 협의회 확대를 위한 실무협의회 회의 개최
  - 국토부(도로운영과·신교통개발과·교통안전복지과), 한국건설기술연구원(도로교통량정보화사업단), 교통안전공단(정보운영처, 교통안전처) 참석

- 2014.12. 제2차 교통정보DB 협의회 회의 개최
  - 공공기관 교통정보 연계·공유 추진현황 및 부문별 향후계획 제시

## 제2절 교통정보 공유현황

### 1. 주요기관 DB 종류 및 형태

#### 가. 국토교통부(국가교통정보센터)

- 고속국도·국도·지방도·시·군도 등 교통시설에 대한 지능형교통체계 구축·운영시 필요한 기본적인 교통정보를 교환·연계중에 있음
  - 기본교통정보는 아래와 같이 9개로 구성되어 있으며, 이들 중 KTDB 신뢰도 제고를 위한 관측교통량 분석 및 교통분석 효율성 제고 등을 위해서는 ‘교통소통 정보’의 연계 필요
  - 실제로 ‘교통소통정보’를 중심으로 제공되고 있으며, 이는 KTDB와 기 연계 중에 있음

<표 3-1> 국가교통정보센터 기본교통정보 분류 및 항목구성

ID	정보명	정보주기	정보세항목
101	교통소통정보	상시교환	- 속도, 교통량, 밀도, 통행시간, 대기길이, 점유율
102	교통통제정보	이벤트 발생시	- 위치, 통제 유형, 대상, 시간
103	돌발상황 발생정보	유고상황 발생시	- 위치, 시각, 사상자수, 피해정도
104	돌발상황정보	유고상황 발생시	- 관리기관, 상황유형, 대상유형, 조치상태, 갱신상태
105	도로상태정보	요청시	- 노면상태, 이용가능여부, 강우/강설수위, 표면온도
106	기상정보	요청시	- 기온, 날씨, 확률, 가시거리, 풍속, 풍향, 습도, 기압, 일출일몰시간
107	도로관리정보	정적정보	- 위치, 관할구역, 도로유형, 도로명, 길이, 포장유형, 운영조건, 중앙분리형태, 차선수, 노면폭
108	프로브정보	상시교환	- 차량종류, 검지시간, 통행시간, 검지위치
109	차량검지정보	상시교환	- 검지위치, 속도, 교통량, 점유율, 대기길이

자료 : 건설교통부, 「기본교통정보교환 기술기준 적용방법 해설서」, 2007, p.3.

LINK_ID	ORGANCODE	GENERATEDATE	RECEIVEDATE	SPEEDRATE	VOLUMERATE	DENSITYRATE	TRAVELTIME	DELAYTIME	VEHICLEQUEUE	OCCUPANCYPERCENT	UPDATEDATE	MAKE_FLAG
1	2600021800	22093700	14/08/02 11:16:00	14/08/02 11:15:43	64	6	1	626	0	0	16/14/08/02 11:22:45	0
2	2600022500	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	2	1	657	0	0	3/14/08/02 11:22:45	0
3	2600022600	22093700	14/08/02 11:16:00	14/08/02 11:15:43	60	6	1	716	0	0	7/14/08/02 11:22:45	0
4	2600023300	22093700	14/08/02 11:16:00	14/08/02 11:15:43	64	6	1	626	0	0	16/14/08/02 11:22:45	0
5	2600023400	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	3	1	612	0	0	11/14/08/02 11:22:45	0
6	2600024700	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	2	1	657	0	0	3/14/08/02 11:22:45	0
7	2600024800	22093700	14/08/02 11:16:00	14/08/02 11:15:43	60	6	1	716	0	0	7/14/08/02 11:22:45	0
8	2600025700	22093700	14/08/02 11:16:00	14/08/02 11:15:43	67	4	1	300	0	0	4/14/08/02 11:22:45	0
9	2600025800	22093700	14/08/02 11:16:00	14/08/02 11:15:43	68	4	1	300	0	0	3/14/08/02 11:22:46	0
10	2600025900	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	3	1	612	0	0	11/14/08/02 11:22:46	0
11	2600026000	22093700	14/08/02 11:16:00	14/08/02 11:15:43	64	6	1	626	0	0	16/14/08/02 11:22:46	0
12	2600027500	22093700	14/08/02 11:16:00	14/08/02 11:15:43	64	6	1	626	0	0	16/14/08/02 11:22:46	0
13	2600027600	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	3	1	612	0	0	11/14/08/02 11:22:46	0
14	2600027900	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	3	1	612	0	0	11/14/08/02 11:22:46	0
15	2600028000	22093700	14/08/02 11:16:00	14/08/02 11:15:43	64	6	1	626	0	0	16/14/08/02 11:22:46	0
16	2600028100	22093700	14/08/02 11:16:00	14/08/02 11:15:43	65	3	1	612	0	0	11/14/08/02 11:22:46	0
17	2600028200	22093700	14/08/02 11:16:00	14/08/02 11:15:43	64	6	1	626	0	0	16/14/08/02 11:22:46	0
18	2600031300	22093700	14/08/02 11:16:00	14/08/02 11:15:43	66	3	1	401	0	0	4/14/08/02 11:22:46	0

&lt;그림 3-1&gt; 국가교통정보센터 기본교통정보 중 ‘교통소통정보’ 데이터 예시

## 나. 한국건설기술연구원(교통량정보시스템)

- 도로법 제87조 (도로에 관한 조사)에 따라 일반국도의 통과 교통량 자료 등을 정확하게 파악하기 위해 매년 상시조사와 수시조사를 수행하고, 조사된 자료를 통계 처리하고 도로교통량 통계연보를 작성하고 있음

&lt;표 3-2&gt; 도로교통량 조사 지점 수 및 구간 수 (2012년 기준)

도로 구분	수시조사 (1일)		상시조사 (365일)	
	수	수*	상	상*
고속국도	482	-	16	-
일반국도	843	171	433	127
국가지원지방도	339	-	-	-
지방도	1,144	-	-	-
계	2,808	171	449	127

- 국도를 대상으로 상시조사 및 수시조사 지점을 직접 조사
  - 상시조사지점 중 일반국도 ITS 구축구간에 대하여 자료 연계를 추진하여, 지방 국토관리청에 자료를 제공 중임

&lt;표 3-3&gt; 국토관리청별 연계현황 (2012년 현재)

(단위 : 지점)

구분	합계	서울청	원주청	대전청	익산청	부산청
연계지점	62	13	5	15	14	15

- 14년 11월 교통정보 통합관리방안 협의 이후 상시조사 지점 전체에 대한 자료 연계 요청 및 실시
- 고속국도 및 지방도의 교통량조사 결과를 취합 및 통계 처리
  - 고속국도와 국가지원지방도, 지방도의 경우는 한국도로공사와 지방자치단체에서 매년 3월부터 11월까지 화, 수, 목요일에 조사

#### 다. 교통안전공단 DTG(운행기록분석시스템)

- 교통안전법에 근거하여 전체 사업용 차량에 대하여 디지털운행기록계 장착이 의무화 되었으며, 교통안전공단에서 통합 시스템을 구축하여 주행궤적 자료를 수집중에 있음(매10초 간격)
  - 주요정보: 일련번호, 차량종류( 시내버스 등), 시각, 누적 주행거리, 속도, X·Y좌표
  - 수집대상 차량종류 : 시내버스, 시외버스, 전세버스, 법인택시, 개인택시, 일반화물

SEQ_NO	CAR_TP_NM	DTTM	ACCM_DIST	SPD	GPS_X	GPS_Y
00000001	시내버스	2014110113045300	39177	21	126.738718	37.537298
00000001	시내버스	2014110113050000	39177	0	126.738748	37.537611
00000001	시내버스	2014110113051000	39177	0	126.738746	37.53766
00000001	시내버스	2014110113052000	39177	0	126.738746	37.537676
00000001	시내버스	2014110113053000	39177	0	126.738748	37.537681
00000001	시내버스	2014110113054000	39177	0	126.738743	37.537676
00000001	시내버스	2014110113055000	39177	0	126.738741	37.537678
00000001	시내버스	2014110113060000	39177	0	126.738741	37.537681
00000001	시내버스	2014110113061000	39177	29	126.738798	37.537835
00000001	시내버스	2014110113062000	39177	41	126.738863	37.538758
00000001	시내버스	2014110113063000	39178	31	126.738885	37.539608
00000001	시내버스	2014110113064000	39178	10	126.738983	37.53993
00000001	시내버스	2014110113065000	39178	43	126.7391	37.540471
00000001	시내버스	2014110113070000	39178	42	126.739308	37.541705
:	:	:	:	:	:	:

&lt;그림 3-2&gt; DGT 샘플 자료

## 라. 교통안전공단 TAGO(국가대중교통정보센터)

- 버스, 철도, 항공, 해운에 대하여 노선 및 운행정보 등에 관한 DB를 구축
  - 고속·시외버스, 철도, 지하철, 항공, 해운부문은 노선운행정보까지, 시내버스의 경우 실시간 위치정보까지 구축 및 제공중에 있으며, 대중교통네트워크 구축 등 교통수요분석 고도화를 위하여 해당 데이터의 연계 필요

&lt;표 3-4&gt; 국가대중교통정보 데이터 현황

서비스명	상세 서비스	주요내용
열차정보	출·도착지 기반 열차정보	- 출발역, 출발시간, 도착역, 도착시간, 차량종류, 운임
	차량종류	- 차량코드, 차량종류
	시도별 기차역 목록	- 도시코드, 도시명, 역코드, 역명
	도시코드 목록	- 도시코드, 도시명
지하철정보	키워드 기반 지하철역 목록	- 지하철역코드, 지하철역명
	지하철역출구별 버스노선 목록	- 지하철역코드, 지하철역명, 출구번호, 버스번호
	지하철역출구별 주변시설 목록	- 지하철역코드, 지하철역명, 출구번호, 주요시설명
	지하철역별 시간표 목록	- 지하철역코드, 지하철역명, 도착시간, 출발시간, 종착역코드, 종착역명, 호선명, 상·하행 구분
국내항공 운행정보	항공운행정보 목록	- 출발공항명, 출발시간, 도착공항명, 도착시간, 항공편코드, 항공사명, 일반석운임, 비즈니스석운임
	공항 목록	- 공항코드, 공항명
	항공사 목록	- 항공사코드, 항공사명
국내선박 운행정보	선박운행정보 목록	-
	여객선터미널 목록	- 터미널코드, 터미널명, 주소, 전화번호
	선박종류 목록	- 선박코드, 선박종류
시외버스 정보	시외버스등급 목록	- 등급코드, 등급명
	시외버스터미널 목록	- 터미널코드, 터미널명
	출·도착지 기반 시외버스정보	- 출발터미널코드, 출발터미널명, 출발시간, 도착터미널코드, 도착터미널명, 도착시간, 운임, 버스등급, 노선ID
고속버스 정보	출·도착지 기반 고속버스정보	- 출발터미널코드, 출발터미널명, 출발시간, 도착터미널코드, 도착터미널명, 도착시간, 운임, 버스등급, 노선ID
	고속버스등급 목록	- 등급코드, 등급명
	고속버스터미널 목록	- 터미널코드, 터미널명
	도시코드 목록	- 도시코드, 도시명
도착정보 조회	정류소별 도착예정정보 목록	- 정류장코드, 정류장명, 노선코드, 노선명, 도착예정시간, 노선종류, 차량종류
	정류소별 특정노선버스 도착예정정보 목록	- 정류장코드, 정류장명, 노선코드, 노선명, 도착예정시간, 노선종류, 차량종류
	도시코드 목록	- 도시코드, 도시명
	노선번호목록	- 노선코드, 노선명, 노선종류, 출발정류장, 종착정류장, 첫차시간, 막차시간
버스노선 정보조회	노선별 경유정류소 목록	- 노선코드, 정차순번, 정류장코드, 정류장명
	노선정보항목	- 노선코드, 노선명, 노선종류, 출발정류장, 종착정류장, 첫차시간, 막차시간
	도시코드 목록	- 도시코드, 도시명

자료 : 국가대중교통정보센터(TAGO) 오픈API

주 : 오픈API를 통하여 구득할 수 있는 정보에 한하여 작성하였음

## &lt;표계속&gt;

서비스명	상세 서비스	주요내용
버스정류소 정보조회	정류소번호목록	- 정류장코드, 정류장명
	좌표기반 근접정류소 목록	- 위도, 경도, 정류장코드, 정류장명
	도시코드 목록	- 도시코드, 도시명
버스위치 정보조회	노선별 버스위치 목록	- 노선명, 정류장순번, 정류장코드, 정류장명, 위도, 경도
	노선별 특정정류소 접근버스 위치정보	-
	도시코드 목록	- 도시코드, 도시명

자료 : 국가대중교통정보센터(TAGO) 오픈API

주 : 오픈API를 통하여 구득할 수 있는 정보에 한하여 작성하였음

### 마. 기타기관

- 한국철도공사, 지방도시철도공사, 전국고속버스운송사업조합, 한국공항공사, 인천국제공항공사, 한국해운조합, 해양수산부 등 교통관련 기관의 운행현황 및 수송실적 통계

## 2. KTDB 자료 수집 현황 및 공유·연계 방안

### 가. KTDB 자료 수집 현황

#### 1) 자료수집 목적

- O/D 구축 및 검증, 코든/스크린라인 검정, 교통네트워크 구축, 장래 O/D 및 네트워크 구축 등 교통분석용 데이터 구축 및 결과검증 등 분석자료의 신뢰도 제고
- 여객·화물부문 국가교통통계 생성 및 DB구축
- 자료수집 절차 : 공문발송 → 기관 답변 및 자료 수집(2주~2개월 가량 소요)

&lt;표 3-5&gt; KTDB 신뢰도 제고 및 통계DB 구축을 위한 각 기관별 자료 수집 현황

수집목적	자료명	기관	요청 주기	자료 주요내용
O/D 구축 및 검증	일반철도 출발·도착지 간 여객·화물 수송실적	한국철도공사	매분기	- 여객: 주운행선, 열차종류, 출발역, 도착역, 인원 - 화물: 주운행선, 품목분류, 출발역, 도착역, 수송톤수
	지하철 출발·도착지 간 수송실적	한국철도공사, 지방도시철도공사	매분기	- 출발역코드, 출발역명, 도착역코드, 도착역명, 인원, 거리
	고속버스 터미널 간 여객 수송실적	고속버스조합	매분기	- 노선, 거리, 버스등급, 총 운행횟수, 일평균 회수, 총 인원, 일 평균 인원, 공급 좌석수, 이용률, 인-km
	공항 간 항공 여객·화물 수송실적	한국공항공사, 인천국제공항공사	매분기	- 국내선: 연월, 출발공항코드, 도착공항코드, 운항횟수, 여객수, 여객 인-km, 화물톤수, 우편물톤수, 화물 톤-km, 우편물 톤-km - 국제선: 연월, 노선, 국가, 도시, 운항횟수, 여객수
	해운 여객·화물 수송실적	한국해운조합 (국내)	매분기	- 계획회수, 결항회수, 여객수송능력, 수송인원, 인-km, 차량대수, 화물톤수, 화물 톤-km, 여객선수입
		해양수산부(국제)	매분기	- 노선, 여객수(입항, 출항), 여객 인-km, 차량대수, 일반화물 톤수, 컨테이너 수송실적
코드/스크린 라인 검증	고속도로 TCS 자료	한국도로공사	매년	- 평일/주말 요금소간 통행량
	교통량 자료	한국건설 기술연구원	매년	- 도로위계별 지점별 교통량, 교통량 조사지점
교통네트워 크 구축	도로/철도 시설 현황	지자체 한국철도공사	매년	- 도로 : 준공도로 현황, 차선수, 연장 등 - 철도 : 운행시각표, 역간거리, 운영현황 등
	대중교통 노선 현황	고속버스조합, 지자체 등	매년	- 버스노선번호, 노선유형, 시·종점, 운행횟수, 순번별 정차정류장 등
장래 O/D, 네트워크 구축	토지이용계획	지자체, LH공사, 지역별 도시공사	매년	- 택지개발사업, 주택건설사업, 도시개발사업, 산업단지개발사업, 도시재정비촉진사업, 주거환경개선사업, 주택재개발사업, 주택재건축사업
	교통시설계획	지자체, 한국철도시설공단 , 지방도시철도공사	매년	- 개발계획 연장, 준공연도, 노선도, 역간거리, 운영계획

#### 나. 자료 공유·연계 방안

- 노선 및 운행정보 등 동적자료: 교통량·속도 및 노선별 등 자료용량이 방대하고 실시간성이 강한 자료의 경우 KTDB와 실시간 연계를 통한 통합관리 필요
  - 국토교통부(국가교통정보센터), 교통안전공단(운행기록분석시스템), 교통안전공단 TAGO(국가대중교통정보센터) 등

- 경우에 따라서는 매일 수집된 자료를 저장하였다가 한번씩 전송하는 방법도 가능
- 수송 실적 등 정적자료: 자료용량이 비교적 작으며, 일년에 몇 번 변하지 않는 자료에 대해서는 현재와 같은 수집체계를 유지하며, 자료공유를 통한 통합관리 필요
  - 기관별 수송실적
  - 지자체 교통량 수시조사
  - 도로/철도 시설현황
  - 장래 계획 등

### 3. 자료 공유·연계에 따른 기대효과

#### 가. 국토교통부(국가교통정보센터) 및 교통안전공단 TAGO(국가대중교통정보센터)

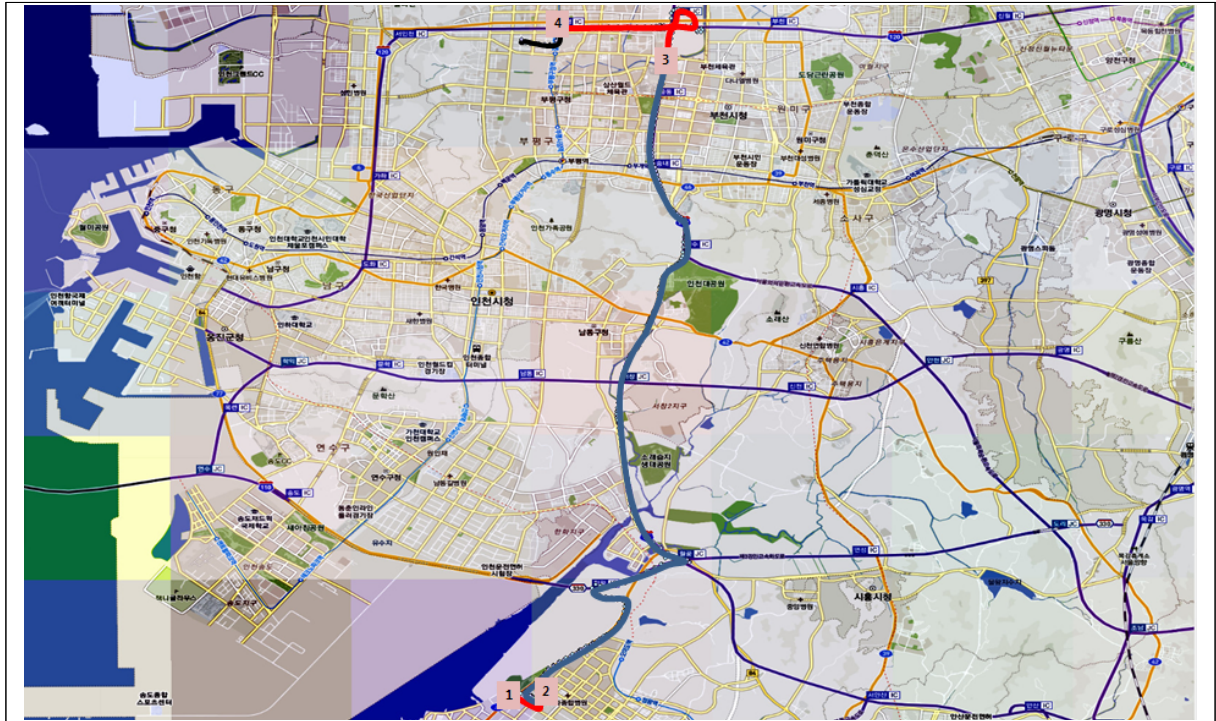
- 교통량, 속도, 수송실적 등의 교통정보를 활용하여 O/D 총량 검증 및 네트워크 성능 검증 등 KTDB 신뢰도 제고
  - 기준년도 O/D의 정확한 추정시 활용
  - 혼잡의 주 척도 제공(V/C, 교통량, 속도)
  - 지자체별 교통네트워크 성능에 따른 맞춤형 유지관리
- 교통정보 연계확보를 통해 조사비용 절감 및 세밀화된 대중교통 정보를 이용한 대중교통 수요분석의 효율성 및 정확성 증대
  - TAGO 시스템의 대중교통 노선 정보(운행 정보 포함)에 교통카드 이용실적 등을 결합하여 대중교통의 수요 분석시 입력 자료 및 검증자료로 활용
    - \* 대중교통 O/D 및 네트워크, 대중교통 성능 지표
  - 교통카드를 활용하는 지역을 대상으로 대중교통 재차인원 산출 가능



- 내비게이션 자료를 이용한 교통망에 차량궤적정보를 통합하여 보다 현실적인 통행패턴 분석 가능
  - 내비게이션 자료 기반 O/D estimation : 차량궤적정보의 기종점과 링크 교통량간 관계 분석을 통해 O/D estimation의 정확도 향상
  - 도로 통행시간 신뢰도 분석 : 차량궤적정보의 경우 장기간 동안 반복 관측된 자료 확보가 가능하므로 경로나 기종점의 통행시간의 안정성, 즉 신뢰도 평가가 가능
- 대국민 교통정보 서비스 및 정책 의사결정 지원 콘텐츠 개발
  - 수집된 교통DB의 표출로만 제한되던 활용 범위를 교통DB의 가공·분석을 통해 다양한 교통DB를 확대·재생산
  - 이를 통해 다양한 교통서비스 콘텐츠 개발 및 교통SOC의 효율적 운영 분석 등 다양한 교통 분야로 확대함으로써 교통DB의 활용도 제고 및 경쟁력 강화

#### 나. 교통안전공단 DTG(운행기록분석시스템)

- 여객부분 수송실적 산출 (택시 및 전세버스)
  - 통행당 평균재차인원('13년도 자동차이용실태조사 결과 활용)을 이용하여 인-km 단위의 수송실적 산출
  - '14년 11월 인천지역 수송실적 산출 예
    - ☞ (통행당 평균 재차인원: 24.4인) × (통행당 평균 통행거리: 58.6 km) × (평균 통행횟수: 4.8회) × (등록대수/표본수: 2223/32) × 30일 = 14,303,404인-km
  - 참고: '13년 11월 인천지역 전세버스 수송실적 추정량 = 146,833,000인-km
    - ☞ 차이 이유: 표본 추출 과정없이 표본 자료만을 이용하여 추정하여 자료의 신뢰성 저하
  - 문제점
    - i) 표본 추출의 랜덤성이 떨어져 추정 수송실적의 대표성 확보가 어려움
    - ii) 주어진 자료로는 전세버스의 차종 구분 (중형: 16인승 이상 35인승 이하, 대형 : 36인승 이상)이 어려움
- 화물부분 O/D 정확성 검증 및 물류지도 작성
  - 운행궤적 확보
    - 다음 <그림 3>은 일일 화물자동차의 운행궤적을 통행으로 구분한 것임
    - ※ 차량별로 DTTM의 시간차가 5분 이상 일 경우 하나의 통행으로 간주



<그림 3-3> 차량번호 No.5011(일반화물)의 11월 8일 운행궤적 및 통행구분 예시

○ 문제점 : 낮은 자료 수집율

- 추정된 자료의 신뢰성 확보가 어려움

- 현재 전체 영업용 차량의 약 8% 정도만 수집되고 있어 분석 자료의 대표성을 확보할 수 있는 방안 모색 필요
- 전수 자료일 경우에는 표본추출의 과정이 필요 없으나, 현재까지 주어진 DTG 자료는 여타의 표본조사에서 사용하는 표본을 보이나, 표본 추출시 랜덤성이 확보가 안되어 추정된 값의 신뢰성이 떨어짐

- 화물 O/D의 경우 개별차량의 경로 파악은 가능하나 지역간 O/D의 통행 배정에서 보이는 주요 경로를 찾기에는 표본율이 떨어짐

## 제3절 교통정보 통합관리 시스템 구축 방안

### 1. KTDB 시스템 현황 및 계획

#### 가. 시스템 현황

- 기존 KTDB 장비에 추가하여 2014년 외부 시스템 연계를 위한 신규 장비(서버, 네트워크 장비 및 DB 소프트웨어) 도입

<표 3-6 > KTDB 서버 현황

연번	이름	용도	비고
1	KTDB 웹서버	KTDB 메인 홈페이지 및 GIS, 혼잡지도 등 홈페이지	
2	KTDB DB서버	KTDB 분석 자료 저장	
3	KTDB 문서 변환	발간문, 보고서 등의 문서 변환 및 홈페이지 제공	
4	KTDB 이메일 서버	KTDB 회원에게 뉴스레터 송신	
5	분석용 서버		신규 구매
6	통신용 서버		신규 구매
7	DB 서버		신규 구매

<표 3-7> 보안 및 네트워크 장비 현황

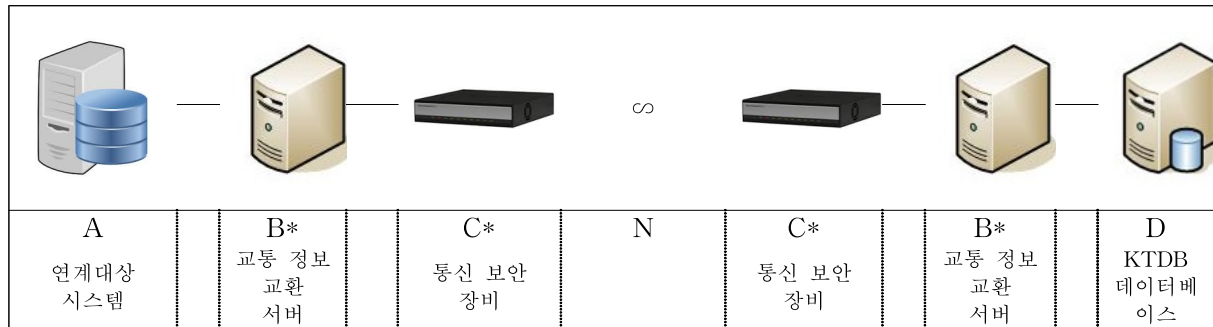
연번	이름	용도	비고
1	VPN	국가교통정보센터와 통신하기 위한 보안 장비	신규 구매
2	방화벽	내부 서버 보호	
3	웹 방화벽	내부 서버 보호	
4	네트워크 스위치	서버 통신 연결	

<표 3-8> DB 소프트웨어 현황

연번	이름	용도	비고
1	MS-SQL	KTDB 홈페이지 등 제공	
2	ORACLE	국가교통정보센터에서 수신된 자료 관리	신규 구매

## 2. 교통정보 통합관리 시스템 구축

### 가. 일반적인 데이터 연계 시스템



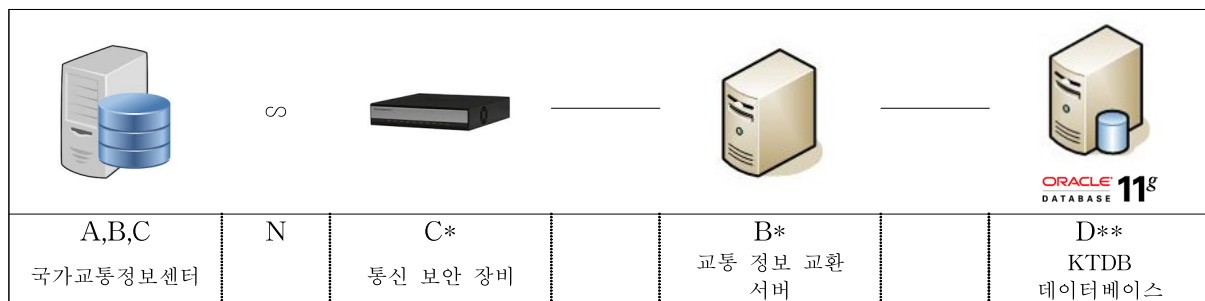
\*: 신규 도입 장비

<그림 3-4> 시스템 구성 현황

- 데이터 연계 시스템
  - 일반적으로 <그림 3-4>와 같이 구성
  - 연계 대상 시스템(A)에 정보교환용 서버(B) 및 통신보안 장비(C)를 추가
  - KTDB에 이에 대응하는 정보교환용 서버(B) 및 통신보안 장비(C)를 추가
- 데이터 교환 전용 회선 (N)
  - 현재 KTDB 웹사이트 회선을 이용하여 데이터 수신 중
  - 데이터 연동 기관수와 송·수신 데이터 정보량에 따라 회선 속도 결정
- 통신 보안 장비(VPN) (C)
  - 데이터 연동의 보안 유지를 위해 연계 기관과 본원 각각 1대씩 필요
  - 송·수신 데이터 정보량에 따라 보안 장비의 성능 결정
- 데이터 송·수신 서버 (B)
  - 데이터 연동기관에 송신 서버 및 본원에 수신 서버 각각 설치
  - 송·수신 데이터 정보량에 따라 서버의 성능 결정
- 데이터베이스 저장 서버 (D)
  - 수신된 교통데이터를 저장하기 위한 충분한 저장 공간 필요
  - 저장을 위한 공간 외에 데이터 백업을 위해 별도의 저장 장치 필요
  - 국가교통정보센터의 ITS 데이터는 하루에 1GB씩 저장 중

- DBMS(Database Management System) (D)
  - DB 서버에서 실제로 데이터를 저장하는 소프트웨어
  - 사용자 수와 서버 CPU 사양에 따라 가격 결정
- 현안 이슈
  - 장비 및 시스템 연계의 경우 대상 시스템에 따라 개별적인 설계 및 협력이 필요
  - 소요 비용 및 구축 시간에 대한 절대적인 고려가 필요

#### 나. 국가교통정보센터 시스템 연계 및 구축



\*: 2014년 도입 완료

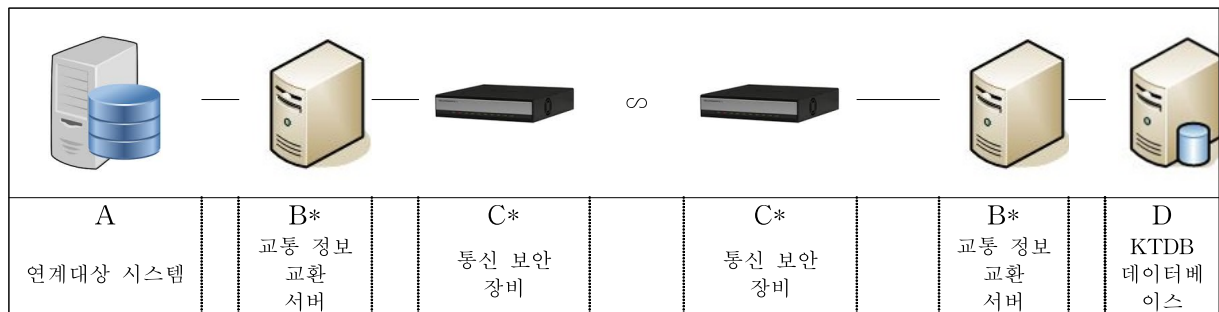
\*\*: 하드웨어는 기존 장비 재활용. DB SW는 2014년 도입 완료

<그림 3-5> 시스템 구성 현황

- 국가교통정보센터 서버와의 연계
  - 기존 국가교통정보센터에서는 기존 시스템 활용 (외부에 데이터를 제공하는 기능을 활용)
    - KTDB에 정보교환용 서버(B) 및 통신보안 장비(C)만을 추가
  - 교통 정보 교환 서버(B)
    - 교통 정보 교환 기술 기준에 따라 A와 D 간의 중계
    - 기존 장비 재활용
    - 국가교통정보센터에서 개발한 데이터 송·수신 SW 사용
  - 통신 보안 장비(C)
    - 2014년 신규 구매 후 설치
    - VPN(Virtual Private Network) 장비를 이용한 전용 통신
    - 데이터를 암호화하여 1:1 통신
  - KTDB 데이터베이스(D)
    - 서버는 기존장비 재활용, DBMS는 2014년 신규 구매 후 설치

- 국가교통정보센터의 교통소통정보 데이터의 저장
- 1일 수신량 : 약 1GB
- 국가교통정보센터 활용 시스템 구축 및 운영
  - 공유 현황
    - 국가교통정보센터 → 국가교통DB센터 : ‘교통소통정보’ 수신
    - 국가교통DB센터 → 국가교통정보센터 : 없음
  - 추진 경과
    - 2014. 6. 17 : 데이터 연동을 위한 실무진 회의
    - 2014. 6. 18 ~ 23 : 데이터 연동 테스트 서버, 데이터베이스 구축
    - 2014. 6. 24 ~ 7. 5 : 국가교통정보센터와 통신 연결 테스트
    - 2014. 7. 5 ~ 2014. 12. 4 : 국가교통정보센터로부터 데이터 수신
    - 2014. 12. 5 ~ 현재 : 기관 이전에 따른 데이터 전송 중단  
15년 2월부터 연동을 위한 실무진 협의 준비중
  - 현안 이슈
    - 국가교통정보센터에 저장되는 DB의 정보가 아니라 타기관에 제공하는 실시간 소통정보를 제공하고 있음
    - 기존 실시간 교통소통정보의 경우 저장 목적이 아니라 실시간 확인 및 표출 위주이기 때문에 간헐적으로 연결이 끊기고, 데이터가 유실됨
    - 국가교통정보센터와 실시간으로 데이터를 받아오는 방식 대신에 일정기간 동안 누적된 데이터를 받아오는 방식으로 전환 협의가 필요
    - 데이터 송·수신 절차 개선 후 데이터 연결 재설정 예정
    - 건기연의 상시 교통량 조사자료가 국토관리청을 통해 국가교통정보센터로 실시간 전송 중(14년 12월부터) : 이에 대한 확인 및 검증이 필요

#### 다. 교통안전공단 TAGO 시스템 연계 및 구축

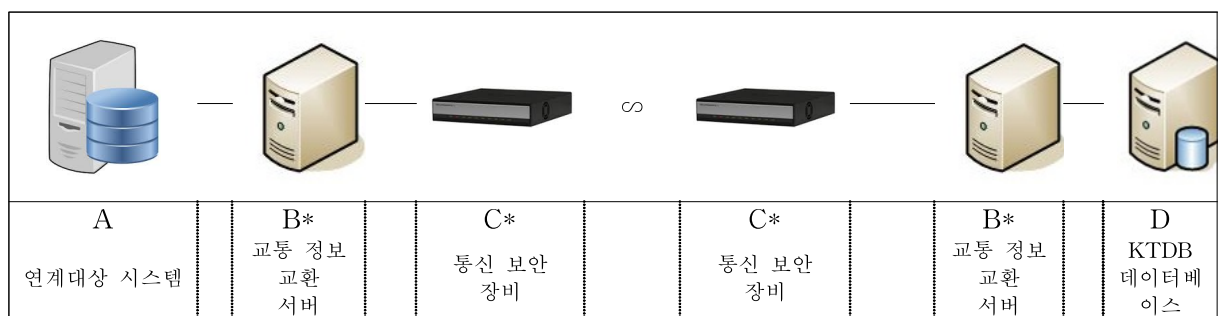


\*: 신규 도입 장비

<그림 3-6> 시스템 구성 현황

- 시스템 연계를 위한 실무진 협의 예정
  - 2015년 1월16일 계획
- 기존 연계 시스템 활용 예정
  - 교통안전공단 TAGO 시스템의 경우 이미 네이버, 다음, SK TMap 등에 실시간 정보를 제공 중임
  - 기존 시스템을 최대한 활용하여 연계할 계획임

#### 라. 교통안전공단 DTG



\*: 신규 도입 장비

<그림 3-7> 시스템 구성 현황

- 시스템 연계를 위한 실무진 협의 예정

#### 마. 기타 교통 유관기관

- 추후 협의 필요

## 제4절 지자체 교통정보 현황

### 1. 지자체 교통정보 수집 및 관리현황

- 지자체 교통정보 수집은 크게 ITS를 통한 상시조사와 지자체 자체 예산으로 실시되는 수시조사로 구분됨
- 상시조사로 도심 내 도로의 교통정보가 실시간으로 수집되어 지자체 교통정보정보센터 등에서 관리하고 현재 국가교통정보센터에 연계 및 저장되고 있음
- 수시조사는 도심 내 도로와 지방도 및 국지도를 대상으로 지자체가 매년 일년에 한번 조사하고 있으며 각각의 연계 및 관리체계는 상이함
  - 도심 내 도로는 지자체가 수집한 교통정보를 자체적으로 보유 및 관리하고 있으며, 연계체계는 없음
  - 지방도 및 국지도는 지자체가 직접 수집하나 한국건설기술연구원에 연계되고 관리하고 있음
- 현재 인구 30만 이상(33개 도시)에 대한 지자체 교통정보 수집 및 연계현황은 <표 6>과 같음
  - 특광역시
    - 서울시는 ITS를 통하여 교통량, 속도는 수집하여 TOPIS와 국가교통정보센터에 연계하고 있는 반면 현장조사를 실시하고 있지 않음
    - 반면 대부분의 광역시는 교통량, 속도정보가 ITS 및 현장조사를 통하여 수집되어 지자체 교통정보센터와 국가교통정보센터에 연계하고 있음
    - 단, 현장조사 교통정보 중 지방도 및 국지도는 한국건설기술연구원에서 연계되나 도심 내 도로는 연계되고 있지 않음
  - 그 외 지자체
    - ITS를 통해 수집되는 지자체의 교통정보는 교통량, 속도정보 모두 수집되어 지자체 교통정보센터와 국가교통정보센터에 연계됨
    - 현장조사를 통한 교통정보는 대부분 교통량이 대상되며, 이중 지방도 및 국지도를 제외한 나머지 도심 내 도로는 미연계됨



&lt;표 3-9&gt; 인구 30만 이상 지자체 교통정보 수집 및 연계현황

인구	도시		교통량				속도			
			ITS		현장조사 <sup>1)</sup>		ITS		현장조사	
			수집 여부	관리 및 연계	수집 여부	관리 및 연계	수집 여부	관리 및 연계	수집 여부	관리 및 연계
1,000만 이상	서울특별시		○		x		○		x	
100만 이상	울산광역시		○		○		○		○	
	대전광역시		○		○		○		○	
	부산광역시		x		○		○		○	
	대구광역시		x		○		○		○	
	광주광역시		○		○		○		○	
	인천광역시		○		○		○		x	
	경기도	수원시	○		x		x		x	
	경남	창원시	x		x	해당 지자체 관리,	x		x	
	경기도	고양시	○		○		○		x	
50~ 100만	경기도	성남시	○	지자체 교통정보 센터 및 국가교통 정보센터	x	지방도 및 국지도: 한국건 설기술 연구원 연계  , 도심내 도로 : 미연계	○	지자체 교통정 보센터 및 국가교 통정보 센터	x	해당 지자체 관리,  모두 미연계
		안양시	○		x		○		x	
		용인시	x		○		○		x	
		안산시	○		x		○		x	
		남양주시	x		○		○		x	
		화성시	○		○		○		x	
		부천시	○		○		○		x	
	충북	청주시	x		○		x		x	
	충남	천안시	○		○		○		x	
	전북	전주시	x		○		x		x	
	경북	포항시	x		x		x		x	
	경남	김해시	x		○		x		x	
30~ 50만	경기도	의정부시	x		○		x		x	
		광명시	○		x		○		x	
		평택시	x		x		x		x	
		시흥시	○		x		○		x	
		파주시	x		x		x		x	
		김포시	x		x		○		x	
	강원도	원주시	x		x		x		x	
	전북	익산시	x		x		x		x	
	경북	구미시	x		x		x		x	
	경남	진주시	x		○		x		x	
	제주도	제주시	○		○		○		x	

1) 현장조사 : 지자체 자체적으로 조사하는 도심내 도로 또는 지방도 및 국지도에서 실시하는 수시조사를 의미

## 2. 지자체 교통정보 공유방안

- 지자체 ITS로 실시간으로 수집되는 도심 내 도로의 교통정보는 현재 국가교통정보센터와의 연계되고 있으므로 연계공유 가능(교통정보통합관리방안 1단계에서 완료)
- 수시조사를 통하여 수집하는 지방도 및 국지도의 교통정보는 한국건설기술연구원에 연계되어 관리되고 있으므로 이와 연계하여 공유가능
- 반면 도심 내 도로의 교통정보는 수시조사를 통하여 지자체가 수집 및 관리하지만 연계·제공은 이루어지지 않음. 이에 이를 공유할 수 있는 별도의 시스템이 필요함
- 수시조사를 통하여 수집되는 교통정보의 통합관리함으로서 지자체 교통정보 수집지점 관리, 중복조사 방지, 도심 내 도로의 다양한 교통정책분석 가능 등의 기대효과를 제시할 수 있음

## 3. 지자체 교통정보 활용방안

- 지자체 교통정보를 연계·공유시, 이를 활용하여 교통분석 및 정책 등에 다양한 활용이 가능
  - 개별사업에서 이루어지고 있는 교통량, 속도의 중복조사 방지 및 예산절감 도모
  - 지자체 교통정보 통합관리로 교통정보 수집현황 파악용이, 시간 및 예산절감
  - O/D 구축 및 검증에 활용
  - 기존 연계되지 못했던 지자체 도심내 도로의 도로유형별 차량주행거리 산정과 교통량 예측
  - 교통혼잡비용, 차량운행비용 등의 교통비용 산정
  - 교통정책 의사결정(혼잡도로개선사업, 혼잡요금, 특별교통대책 등)활용