

■ 論 文 ■

## 유비쿼터스 환경의 프로브 차량 정보를 활용한 표본 OD 전수화 (제주시 시범사업지역을 대상으로)

Expansion of Sample OD Based on Probe Vehicle Data in a Ubiquitous Environment

**정 소 영**  
(한국도로공사 도로교통연구원  
연구원)

**백 승 걸**  
(한국도로공사 도로교통연구원  
책임연구원)

**강 정 규**  
(한국도로공사 도로교통연구원  
연구위원)

### 목 차

- |  |  |
|--|--|
| <p>I. 서론</p> <p>II. 기존 문헌 검토</p> <p style="margin-left: 20px;">1. 이용정보에 따른 OD 추정 연구</p> <p style="margin-left: 20px;">2. 전수화 기법 연구</p> <p>III. 표본OD 산정</p> <p style="margin-left: 20px;">1. 개요</p> <p style="margin-left: 20px;">2. 교통존 및 네트워크 설정</p> <p style="margin-left: 20px;">3. 표본OD 산정</p> | <p>IV. 표본OD의 전수화</p> <p style="margin-left: 20px;">1. 전수화의 필요성</p> <p style="margin-left: 20px;">2. 전수화 과정</p> <p style="margin-left: 20px;">3. 표본OD 전수화</p> <p style="margin-left: 20px;">4. 시사점</p> <p>V. 결론 및 향후 연구과제</p> <p>참고문헌</p> |
|--|--|

Key Words : 표본OD, 전수화, 프로브 차량, 실시간OD, 정적OD  
Sample OD, Expansion, Probe Car, Real Time OD, Static OD

### 요 약

최근 교통·물류 분야에서도 유비쿼터스 환경의 정보수집체계 및 이를 응용한 서비스 개발의 필요성이 매년 높아지고 있다. 프로브 차량과 무선통신기술을 활용한 교통정보 수집체계는 그 대표적인 사례로 차량의 기종점 자료를 이용하여 시간대별 OD를 산정하는 것이 가능하다. 그러나 프로브 차량 정보를 활용하여 산정된 OD는 시간적·공간적으로 변동되는 표본OD이기 때문에 이를 정적OD로 전환하기 위해서는 수집정보를 집적하여 적정 표본율을 산정하고, 표본OD를 전수화하는 과정이 필요하다. 본 연구는 제주시를 대상으로 수집된 실제 데이터를 표본OD 산정 및 전수화 알고리즘에 적용하여 표본OD를 산정하고 이를 전수화하였다. 각 링크별 관측교통량과 배분교통량의 오차를 비교 검토한 결과 링크별 관측교통량과 배분교통량의 평균 오차율은 22.9%, 상·하위 10%의 이상 자료를 제거한 후의 평균 오차율은 17.6%로 각각 나타났다. 본 연구는 기존OD가 존재하지 않는 지역에서 프로브 차량의 경로정보를 활용하여 정적OD를 산정하였다는 점과 적정 오차를 내 수렴을 위한 적정 표본율을 제시하였다는 점에서 그 의의를 찾을 수 있다.

Information collection systems and applications in a ubiquitous environment has emerged as a leading issue in transportation and logistics. A productive application example is a traffic information collection system based on probe vehicles and wireless communication technology. Estimation of hourly OD pairs using probe OD data is a possible target. Since probe OD data consists of sample OD pairs, which vary over time and space, computation of sample rates of OD pairs and expansion of sample OD pairs into static OD pairs is required. In this paper, the authors proposed a method to estimate sample OD data with probe data in Jeju City and expand those into static OD data. Mean absolute percentage difference (MAPD) error between observed traffic volume and assigned traffic volume was about 22.9%. After removing abnormal data, MAPD error improved to 17.6%. Development of static OD estimation methods using probe vehicle data in a real environment is considered the main contribution of this paper.

## I. 서론

OD는 통행의 출발지 및 목적지에 대한 정보로서 경로선택, 통행배분 등 교통계획에서 가장 중요한 정보 중의 하나이다. OD 추정을 위하여 많은 기법들이 검토되었으며, 최근에는 실시간 교통관리와 교통정보 제공을 위한 동적OD 추정에 관한 연구들이 진행되고 있다.

기존의 OD 추정기법은 조사를 통한 OD 추정기법, 차량번호판 매칭을 통한 OD 추정기법, 관측교통량을 이용한 OD 추정기법 등 크게 세 가지로 구분할 수 있다. 조사를 통한 OD 추정기법은 행정동 등의 자료를 기반으로 하는 전통적인 기법으로 해당 지역에서 표본을 추출하여 표본OD를 구축한 후 이를 전수화하여 OD를 추정하는 것이 일반적이다. 그러나 토지이용과 인구가 급격하게 변화하는 지역에 대해서는 정확한 예측이 어려우며, 비용, 노동력, 시간이 많이 소요되고 표본 추출이나 전수화 과정에서 오차가 발생한다는 단점을 지닌다. 차량번호판 매칭을 통한 OD 추정기법은 차량 번호판에 대한 직접 샘플링을 이용하는 방법으로 각기 다른 교통존에서 관측된 번호판 자료에 대해 차량의 기점과 종점을 일치시켜 OD를 산정한다. 관측시간의 제약과 부정확한 차량 감지로 인한 관측오차가 존재한다. 관측교통량을 이용한 OD 추정기법은 링크교통량으로부터 통행배분과정을 역으로 거슬러 OD를 추정하는 방법이다. 이는 직접 관측이 불가능한 OD와 관측된 링크교통량과의 인과관계를 설정하여 OD를 추정하는 것으로 교통량 자동 수집기술을 이용하여 관측교통량을 용이하게 구할 수 있다는 장점을 지닌다.

최근에는 산업 전 분야에서 유비쿼터스 환경 구축을 위한 다양한 시도가 이루어지고 있으며, 특히 교통·물류 분야에서는 교통정보 수집체계 및 이를 응용한 서비스 개발의 필요성이 매년 증가하고 있다. 프로브 차량과 무선통신기술을 ITS에 접목한 교통정보 수집체계는 그 대표적인 사례이다. 프로브 차량을 통해 실시간 데이터 수집이 가능하므로 교통관리, 교통정보 제공 등에 활용될 수 있으며, 리더기를 통해 수집되는 프로브 차량 정보를 이용하여 차량의 기종점 확인이 가능하므로 시간대별 정적OD를 산정할 수 있다. 이에 반해 통행목적의 파악하기 어려우며, 사회경제지표를 이용한 장래 통행수요 예측이 어렵다는 단점도 지닌다. 또 네트워크를 기반으로 수집되는 정보이므로 기종점별로 일정 표본율을 유지

하기 어려우며, 자료 수집시간에 따라 표본율이 변동하게 된다.

따라서 프로브 차량 정보를 활용하여 정적OD를 산정하기 위해서는 표본OD를 산정하고 시간적·공간적으로 변동되는 자료를 집적하여 전수화하는 과정이 필요하다. 백승걸 등(2007c)은 가상자료를 활용하여 전수화 알고리즘을 제시한 바 있다. 본 연구에서는 백승걸 등(2007c)이 제시한 알고리즘을 수정·보완한 후 제주시 대상의 실제 데이터를 활용한 표본OD 전수화기법을 제시하였다.

## II. 기존 문헌 검토

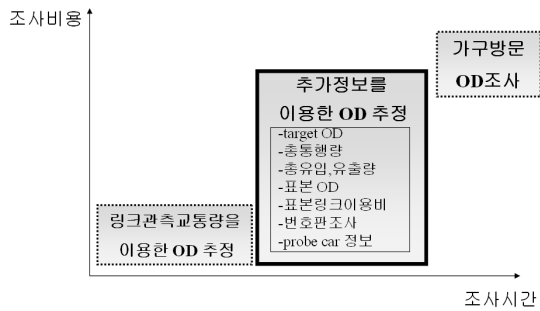
### 1. 이용정보에 따른 OD 추정 연구

관측교통량만을 이용한 OD 추정은 과소설정문제로 인해 유일한 해를 선정하기가 어려우며, 상이한 관측시간으로 인한 변동(variation)과 관측오차가 존재한다. 과소설정문제를 해결하기 위해 추가정보를 이용하거나 중력모형 또는 엔트로피 극대화 모형과 같이 모형의 구조를 제약하는 기법을 이용한다.

가구방문조사를 통해 구축하는 OD는 대규모 조사대상지역, 낮은 표본율, 조사 및 응답의 부정확성 등으로 오차나 신뢰도의 범위가 상대적으로 크다. 또한 대부분 정기적인 대규모 조사를 통해 구하는 것으로, 해당지역의 사회경제적 변수가 안정적이지 못한 곳에서는 신뢰성이 낮다. 따라서 추정해의 정확도를 높이기 위해서는 비용이 다소 소요되더라도 관측교통량과 target OD외에 "추가정보(additional information or data)"를 확보하여 OD를 추정하는 것이 필요하다. 이 경우 해의 정확도 증가정도와 조사비용의 증가수준이 추가로 정보를 확보하여야 하는지의 판단기준이 될 수 있다.

OD 통행 추정에 이용된 추가정보는 Tamin 등(1989)의 준별 사회경제적 변수, Cascetta 등(1984)의 표본OD, Spiess(1987)의 준별 유입·유출량, Kuwahara 등(1987)의 도로변 면적조사자료, Landau 등(1982)의 코든라인 조사자료 등이 있다. Watling(1994)은 차량번호판 조사를 통해 기종점을 파악하였으며, 김시곤 등(2005)은 휴대폰 기지국 정보를 이용하여 OD를 추정하였다.

표본조사자료를 이용한 OD 추정 연구들은 백승걸 등(2002)에 자세히 기술되어 있다.



〈그림 1〉 이용정보에 따른 OD 추정기법 구분 (백승걸 등, 2002)

## 2. 전수화 기법 연구

RFID와 관련된 연구로는 물류, 요금정책, 교통소통(통행시간 예측, 교통혼잡 등)에 관한 연구가 있으나, 교통계획 측면에서 RFID 리더기 수집정보를 활용하여 OD를 추정한 사례는 없는 것으로 조사되었다. 따라서 관측교통량 외의 추가정보를 활용하여 OD를 전수화한 사례를 중심으로 검토하였다.

한국교통연구원(2007)은 2005년 지역간 기종점통행량 조사를 기반으로 하여 전국 167개 교통존 간 기종점통행량 자료를 구축하고, 전국 지역간 승용차 여객 기종점통행량을 전수화하는 방법을 검토하였다. 승용차의 경우 지점별 노측면적 조사자료와 지역간 교통통계자료를 활용하여 모집단 대비 5%의 표본OD를 구축하였으며, 통과교통량이 배제된 교통존별 발생·도착량을 이용하여 전수화하였다. 전수화된 결과는 screen line에 따라 보정하였으나, 이와 같은 전통적인 기법은 전수조사가 어렵기 때문에 누락조사지점으로 인한 전수화 오차 보정이 필요하다.

이신해(2004)는 전수화지표와 전수화계수를 활용하여 기종점통행량을 구축하였다. 전수화지표는 표본자료로 모집단을 추정할 때 기준으로 사용하는 변수를 의미하는 것으로, 개인특성 및 통행특성에 대한 모집단으로 활용한 가구특성 변수를 전수화지표로 사용하였다. 전수화계수는 가구·가구원·통행 전수화계수로 분류할 수 있으며, 통행특성을 반영하는 통행 전수화계수의 경우 과소응답 보완조사, 통행목적별·수단별로 보정계수를 산출하여 보정하였다.

건설교통부(2004)는 GPS 위치정보를 활용하여 택

시의 통행경로정보를 수집하고, 이를 기반으로 OD를 산출하고 전수화하였다. 전수화는 표본 추출에 의한 결과로부터 모표본을 구하는 방식과 동일하게 산정하였으며, 택시 1대당 평균운행횟수에 총 운행차량대수를 곱하여 전수화하였다. 그러나 이는 단말기가 부착된 차량으로부터 직접 경로정보를 수집하는 방식으로 RFID 리더기와 같이 네트워크를 기반으로 설치되어 정보를 수집하는 방식과는 차이가 있다.

백승걸 등(2007c)은 본 연구의 선행연구로서 프로브 차량 정보를 기반으로 하는 전수화 알고리즘을 제시하였으나 실제 데이터가 아닌 가상 데이터를 활용한 알고리즘 검증에 중점을 두고 수행되었다.

## III. 표본OD 산정

### 1. 개요

본 연구는 “유비쿼터스 환경의 차세대 교통정보 수집 체계 개발 및 시범사업”에서 수집된 데이터를 기반으로 하여 수행되었으며, 관련 과제는 교통 분야에서 능동형 RFID 기술을 활용한 국내 최초의 시범사업이다. RFID는 도로 위를 주행하는 차량에 부착된 태그가 도로 위에 설치된 리더기와의 통신을 통하여 태그 속에 미리 입력된 정보(차량의 고유 ID)를 인식하여 읽어오는 원리로 구성되어 있으며, 수집된 정보를 활용하여 기·종점, 통행시간, 운행속도, 통행경로 등의 2차 정보 가공이 가능하다. 따라서 적정 위치에 리더기를 설치하여 도로 위를 주행하는 차량의 통행특성을 반영할 수 있도록 해야 한다.

RFID 리더기의 최적 설치지점에 대한 연구는 백승걸 등(2007a)에서 기수행한 바 있으며, 면(관측구역), 선(관측구간), 지점(관측지점)을 선정하는 과정을 반복검토하여 리더기 설치위치를 조정한다. 관측구역은 해당 지역의 상징성, 사회·경제지표, 실시간 교통관리 등을 검토하여 선정하며, 관측구간은 교통존, 교통현황, 도로 기하구조, 기존 ITS 관측구간 등을, 관측지점은 리더기의 성능, 자료수집 용이성, 현장설치 여건 등을 고려하여 선정한다.

대상구역인 제주시는 제주도를 남북으로 가르는 몇 개의 국도들과 해안선을 따라 순환하는 도로들로 연결되어 있어 전체적으로 정비된 격자형의 도로망을 갖추고 있으며, 총 19개의 행정동으로 구성되어 있다. 2005년 북

제주군이 제주시로 통합되면서 인구가 증가하여 2006년 기준 약 41만명으로 나타났으며, 자동차등록대수는 약 16만대로 나타났다.

〈그림 2〉는 RFID 리더기 시범설치 대상구역을 나타낸 것으로 ITS 관측정보를 이용한 실시간 교통관리가 가능하며, 주변에 제주시청, 은행과 같은 중심업무지역이 있어 통행량이 밀집되므로 설치대상구역으로 적합한 것으로 판단된다.

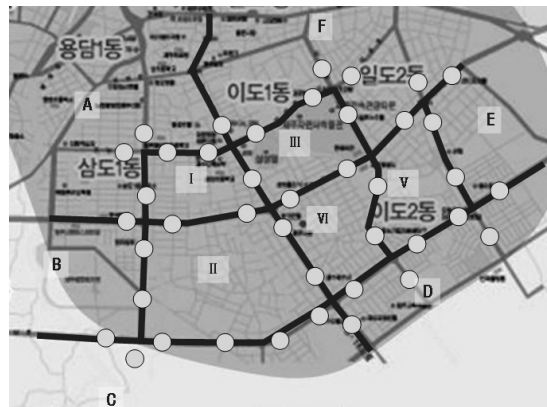
〈그림 2〉에서 제시된 시범설치 대상구역에 프로브 차량의 태그정보를 읽어서 중앙서버로 송신하는 리더기를 설치하고 약 2,000대의 차량에 태그를 부착하여 각 프로브 차량의 경로정보를 수집하였다. 90%인 1,800대가 승용차이며, 나머지 10%는 버스 및 제주도청 소속 차량 등으로 구성되어 있다. 2006년 말을 기준으로 한 제주시의 자동차등록대수 16만대와 비교했을 때 약 1.26% 정도로 태그 부착 표본율은 매우 낮은 편이다. 그러나 동일한 차량이 분석기간 동안 여러 번 분석대상 네트워크를 통행하게 되므로 분석대상 네트워크에서의 프로브 차량 이용 표본율은 이보다 높을 것으로 판단된다.



〈그림 2〉 대상구역

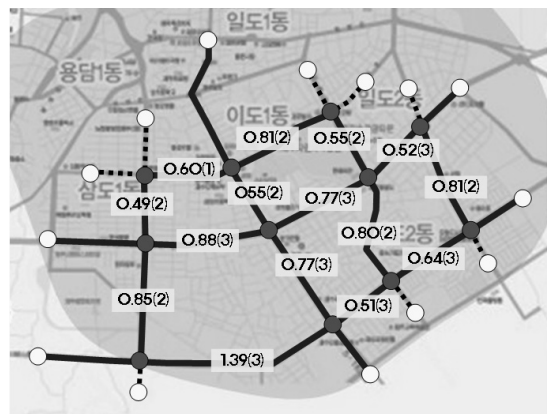
## 2. 교통존 및 네트워크 설정

리더기는 도로 네트워크를 따라 설치되고, 리더기를 통해 수집된 정보도 네트워크를 기반으로 하므로 교통존도 네트워크를 기반으로 하는 블록 단위로 설정할 필요가 있다. 본 연구에서는 〈그림 3〉과 같이 교통존을 설정하였으며, 내부 교통존 5개(I, II, III, IV, V)와 주요 간선축을 경계로 하는 외부 교통존 6개(A, B, C, D, E, F)로 구성된다.



〈그림 3〉 교통존 설정

〈그림 4〉는 대상구역의 네트워크 정보(연장 및 차로 수)를 나타낸 것으로 차로 수는 편도를 기준으로 작성하였다.

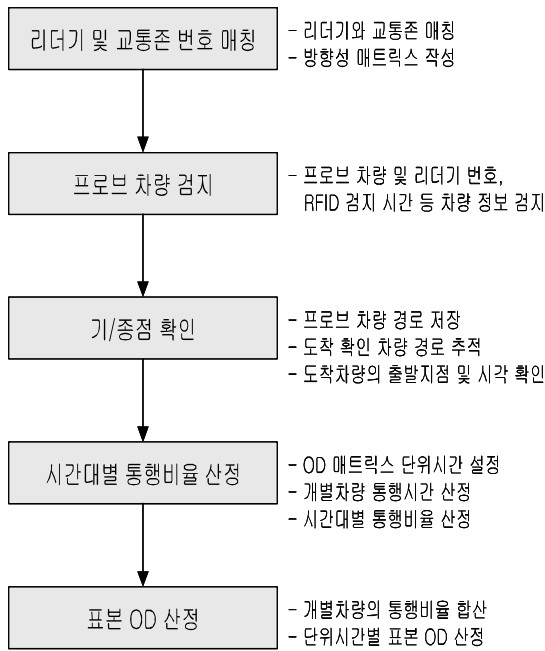


〈그림 4〉 네트워크 연장(km) 및 차로 수(편도 기준)

## 3. 표본OD 산정

프로브 차량 정보로부터 표본OD를 산정하는 과정은 〈그림 5〉와 같이 5단계로 구성되며, 자세한 내용은 백승걸 등(2007b)에 기술되어 있다.

현장조사를 통한 링크별 관측교통량은 목요일 오전 침두시(07~09시) 데이터를 활용하였다. 동일 시간의 프로브 차량 정보를 분석한 결과 전체 기종점 121쌍에 대한 표본OD 통행량은 141통행으로 나타나 표본 자료로 활용하기에 미흡한 것으로 나타났다. 따라서 프로브 차량 정보를 일주일 동안 수집, 총 4,529통행에 대한 표본OD를 산정하였다.



〈그림 5〉 표본OD 산정과정

〈표 1〉 시간대별·요일별 표본OD 통행량

요일	오전 침두시(07~09)	24시간
월	31	354
화	127	685
수	104	712
목	141	761
금	131	711
토	128	666
일	75	640
계	737	4,529

그러나 통행시간대에 따라 통행특성(기중점, 통행량, 통행비율 등)이 달라지기 때문에 통행시간이 일치하지 않는 자료들에 대해서는 동일한 통행특성을 지니는지에 대한 검증이 필요하다. 본 연구에서는 전체 통행량에 대한 각 기중점별 통행량의 비율을 산정하여 비교하였다. 목요일 오전 침두시 통행량 비율과 일주일의 통행량 비율을 비교하여 오차율을 산정한 결과(〈표 2〉 참조) 오차가 1% 이내인 기중점이 전체의 88.4%를 차지하는 것으로 나타났다. 이는 오전 출퇴근 침두시와 유사한 규모의 통행이 반복적으로 발생함을 의미하며, 제주시청을 포함하는 업무지구가 대상구역에 밀집해 있어 시간대에 따른 통행변화가 적었기 때문인 것으로 판단된다.

〈표 2〉 통행시간 변화에 따른 OD통행량 비율 오차 검토

오차율	기중점 수	구성비
5% ~ 4%	1	0.008
4% ~ 3%	3	0.025
3% ~ 2%	4	0.033
2% ~ 1%	6	0.050
1% 이하	107	0.884
계	121	1.000

통행시간에 따른 기중점 통행량 비율의 변화가 적으므로 데이터 확보를 위하여 일주일의 데이터를 활용하여 표본OD를 산정하였다(〈표 3〉 참조).

〈표 3〉 표본OD

기중점 \ 종점	I	II	III	IV	V	A
I	246	10	7	23	8	107
II	9	190	6	42	9	6
III	49	15	225	7	15	28
IV	9	52	1	179	15	3
V	4	99	10	42	177	3
A	72	22	15	14	5	363
B	28	9	5	15	15	11
C	9	10	5	11	36	11
D	12	21	21	32	32	4
E	1	5	1	4	2	2
F	48	9	44	16	13	40
합계	487	442	340	385	327	578

기중점 \ 종점	B	C	D	E	F	합계
I	9	22	13	2	17	464
II	40	10	9	2	11	334
III	7	2	9	2	53	412
IV	26	15	29	3	21	353
V	144	5	48	5	34	571
A	7	13	8	1	9	529
B	105	2	5	3	4	202
C	7	226	243	13	21	592
D	9	3	267	23	32	456
E	14	4	5	60	10	108
F	11	6	12	5	304	508
합계	379	308	648	119	516	4,529

## IV. 표본OD의 전수화

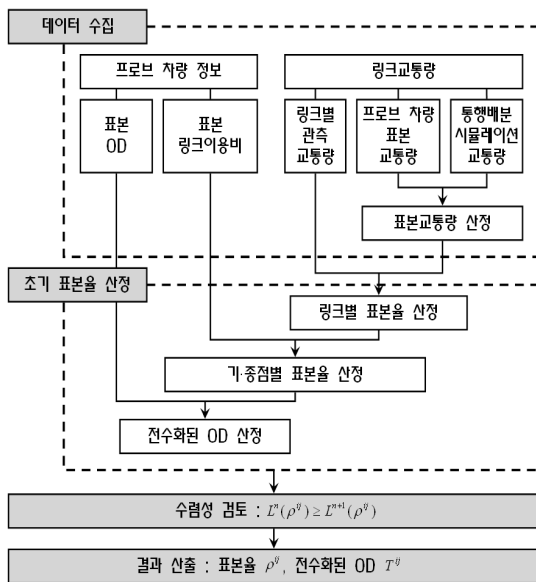
### 1. 전수화의 필요성

본 연구에서는 프로브 차량의 경로 정보를 활용하여 표본OD를 산정하였다. 그러나 이는 태그를 부착한 일부

차량에 대한 데이터이기 때문에 전통적인 OD 조사처럼 각 기종점별로 일정 표본율을 산정하기 어려우며, 리더기가 미설치된 구간이나 설치구간 외부로 프로브 차량이 통행하는 경우 표본을 확보하기 어려워진다. 또 실시간 데이터이기 때문에 통행시간에 따라 표본율이 변동하게 된다. 따라서 프로브 차량 정보를 활용하여 적정OD를 산정하기 위해서는 변동하는 표본율을 대표하는 적정 표본율을 산정하고, 이에 따라 표본OD를 전수화하는 과정이 필요하다. 본 연구에서는 앞서 제시한 문제점들을 해결하기 위해 프로브 차량 정보 수집기간을 일주일로 설정하여 누락되는 데이터를 최소화하였으며, 외부 교통존을 설정하여 미설치구간에 대한 통행량을 반영할 수 있도록 하였다.

### 2. 전수화 과정

본 연구에서는 특정 링크를 통행한 프로브 차량으로부터 수집된 표본교통량과 관측교통량을 이용하여 각 링크별 표본율을 산정하고, 이를 표본링크이용비를 이용, 기종점별 표본율로 전환하여 표본OD를 전수화하였다. 프로브 차량 표본OD의 전수화 과정은 <그림 6>과 같이 데이터 수집, 초기 표본율 산정, 수렴성 검토, 결과 산출로 분류할 수 있으며, 구체적인 내용은 표본OD의 전수화에서 기술하였다.



<그림 6> 프로브 차량 표본OD의 전수화 과정

### 3. 표본OD의 전수화

#### 1) 데이터 수집

##### (1) 프로브 차량 정보

기종점별로 산정되는 표본OD는 <표 3>에서 제시하였으며, 표본링크이용비는 식(1)을 이용하여 산정하였다. 표본링크이용비는 링크에 대한 정보와 기종점에 대한 정보를 모두 가지고 있기 때문에 링크별 표본율을 기종점별 표본율로 전환하는 데 활용된다.

$$\hat{R}_a^{ij} = \frac{\hat{x}_a^{ij}}{\sum_a \hat{x}_a^{ij}} \quad (1)$$

여기서,  $\hat{R}_a^{ij}$  :  $ij$  통행량의 링크  $a$  이용비

$\hat{x}_a^{ij}$  :  $ij$ 별 링크  $a$ 의 프로브 차량 통행량

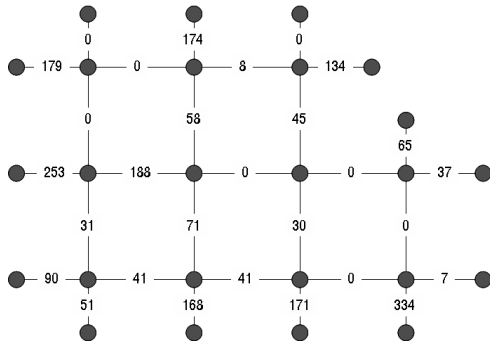
##### (2) 링크교통량

링크별 관측교통량과 프로브 차량 표본교통량은 <그림 7>과 같다. 통행시간에 따른 통행특성의 변화가 적으므로 프로브 차량 표본교통량도 표본OD와 동일하게 일주일의 데이터를 합산하여 산정하였으며, 관측교통량은 목요일 오전 침두시의 데이터를 활용하였다.

구분	교통량																																																																		
관측교통량 (오전 침두시)	<table border="1"> <tr> <td>●</td><td>2,050</td><td>●</td><td>2,036</td><td>●</td><td>1,719</td></tr> <tr> <td>●</td><td>1,315</td><td>●</td><td>1,446</td><td>●</td><td>2,279</td><td>●</td><td>1,859</td></tr> <tr> <td>●</td><td>2,277</td><td>●</td><td>2,592</td><td>●</td><td>2,265</td><td>●</td><td>3,184</td></tr> <tr> <td>●</td><td>3,886</td><td>●</td><td>5,117</td><td>●</td><td>4,532</td><td>●</td><td>5,008</td><td>●</td><td>4,223</td></tr> <tr> <td>●</td><td>2,769</td><td>●</td><td>3,817</td><td>●</td><td>2,084</td><td>●</td><td>2,663</td></tr> <tr> <td>●</td><td>7,133</td><td>●</td><td>8,176</td><td>●</td><td>6,129</td><td>●</td><td>5,546</td><td>●</td><td>5,778</td></tr> <tr> <td>●</td><td>1,742</td><td>●</td><td>3,570</td><td>●</td><td>1,997</td><td>●</td><td>2,022</td></tr> <tr> <td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td></tr> </table>	●	2,050	●	2,036	●	1,719	●	1,315	●	1,446	●	2,279	●	1,859	●	2,277	●	2,592	●	2,265	●	3,184	●	3,886	●	5,117	●	4,532	●	5,008	●	4,223	●	2,769	●	3,817	●	2,084	●	2,663	●	7,133	●	8,176	●	6,129	●	5,546	●	5,778	●	1,742	●	3,570	●	1,997	●	2,022	●	●	●	●	●	●	●	●
	●	2,050	●	2,036	●	1,719																																																													
	●	1,315	●	1,446	●	2,279	●	1,859																																																											
	●	2,277	●	2,592	●	2,265	●	3,184																																																											
	●	3,886	●	5,117	●	4,532	●	5,008	●	4,223																																																									
●	2,769	●	3,817	●	2,084	●	2,663																																																												
●	7,133	●	8,176	●	6,129	●	5,546	●	5,778																																																										
●	1,742	●	3,570	●	1,997	●	2,022																																																												
●	●	●	●	●	●	●	●																																																												
프로브 차량 표본교통량 (일주일)	<table border="1"> <tr> <td>●</td><td>349</td><td>●</td><td>229</td><td>●</td><td>175</td></tr> <tr> <td>●</td><td>123</td><td>●</td><td>299</td><td>●</td><td>276</td><td>●</td><td>145</td></tr> <tr> <td>●</td><td>325</td><td>●</td><td>322</td><td>●</td><td>156</td><td>●</td><td>172</td></tr> <tr> <td>●</td><td>227</td><td>●</td><td>591</td><td>●</td><td>681</td><td>●</td><td>272</td><td>●</td><td>24</td></tr> <tr> <td>●</td><td>62</td><td>●</td><td>189</td><td>●</td><td>376</td><td>●</td><td>300</td></tr> <tr> <td>●</td><td>645</td><td>●</td><td>1,038</td><td>●</td><td>825</td><td>●</td><td>717</td><td>●</td><td>123</td></tr> <tr> <td>●</td><td>1</td><td>●</td><td>305</td><td>●</td><td>133</td><td>●</td><td>134</td></tr> <tr> <td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td><td>●</td></tr> </table>	●	349	●	229	●	175	●	123	●	299	●	276	●	145	●	325	●	322	●	156	●	172	●	227	●	591	●	681	●	272	●	24	●	62	●	189	●	376	●	300	●	645	●	1,038	●	825	●	717	●	123	●	1	●	305	●	133	●	134	●	●	●	●	●	●	●	●
	●	349	●	229	●	175																																																													
	●	123	●	299	●	276	●	145																																																											
	●	325	●	322	●	156	●	172																																																											
	●	227	●	591	●	681	●	272	●	24																																																									
●	62	●	189	●	376	●	300																																																												
●	645	●	1,038	●	825	●	717	●	123																																																										
●	1	●	305	●	133	●	134																																																												
●	●	●	●	●	●	●	●																																																												

<그림 7> 관측교통량 및 표본교통량

그러나 <그림 7>에서와 같이 관측교통량은 모든 링크에 통행량이 고르게 분포되어 있는 반면 프로브 차량의 표본교통량은 링크별 통행량의 편차가 크게 나타나 이에 대한 보완이 필요하다. 따라서 본 연구에서는 <표 3>에 산정된 표본OD에 대하여 통행배분 시뮬레이션을 수행하였으며, 이를 프로브 차량의 표본교통량에 보완하였다 (<그림 8> 참조).



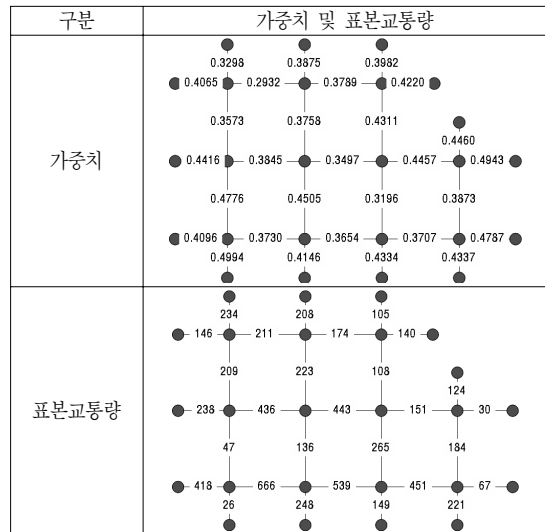
<그림 8> 통행배분 시뮬레이션 교통량

표본교통량은 앞서 산정된 프로브 차량 표본교통량과 통행배분 시뮬레이션 교통량에 가중치를 부여하여 산정된다. 가중치는 표본율(관측교통량에 대한 프로브 차량 표본교통량 비율)이 높을수록 표본교통량의 비율을 높이고 표본율이 낮을수록 통행배분 시뮬레이션 교통량의 비율을 높일 수 있도록 설정하였다. 단 통행배분 시뮬레이션은 현장 데이터보다 가중치가 높아지지 않도록 최대값을 0.5로 제한하였다. 표본교통량은 식(2)와 같이 산정된다.

$$\hat{x}_a = \omega \cdot \hat{x}_a^{sim} + (1-\omega) \cdot \hat{x}_a^{probe} \quad (2)$$

- 여기서,  $\hat{x}_a$  : 링크  $a$ 의 표본교통량
- $\omega$  : 가중치 ( $= \max\{0, 0.5 - \hat{\rho}_a\}$ )
- $\hat{\rho}_a$  : 표본율 ( $= \frac{\hat{x}_a^{probe}}{x_a}$ )
- $\hat{x}_a^{sim}$  : 통행배분 시뮬레이션 교통량
- $\hat{x}_a^{probe}$  : 프로브 차량 표본교통량
- $\bar{x}_a$  : 링크  $a$ 의 관측교통량

<그림 9>는 식(2)에 의해 산정된 가중치와 표본교통량을 나타낸 것이다.



<그림 9> 가중치 및 표본교통량

## 2) 초기 표본을 산정

<그림 9>에서 산정된 표본교통량과 <그림 7>의 관측교통량을 이용하여 링크별 표본율  $\rho_a$ 를 계산한다.

$$\rho_a = \hat{x}_a / \bar{x}_a \quad (3)$$

- 여기서,  $\rho_a$  : 링크  $a$ 의 표본율
- $\hat{x}_a$  : 링크  $a$ 의 표본교통량
- $\bar{x}_a$  : 링크  $a$ 의 관측교통량

식(3)에서 산정된 링크별 표본율은 표본링크이용비를 이용하여 기중점별 표본율로 전환한다. 식(4)에서 제시된 표본율은 초기 표본율로 수렴성을 검토하는 과정에서 반복조정된다.

$$\rho^{ij} = \frac{\sum_a \rho_a \times \hat{R}_a^{ij}}{\sum_a \hat{R}_a^{ij}} = \frac{\rho_1 \times \hat{R}_1^{ij} + \rho_2 \times \hat{R}_2^{ij} + \dots}{\hat{R}_1^{ij} + \hat{R}_2^{ij} + \dots} \quad (4)$$

- 여기서,  $\rho^{ij}$  : 기중점  $ij$ 의 표본율
- $\rho_a$  : 링크  $a$ 의 표본율
- $\hat{R}_a^{ij}$  :  $ij$  통행량의 링크  $a$  이용비

기중점별 표본율이 산정되면 식(5)와 같이 전수화된 OD를 산정한다.

$$T^{ij} = \frac{\hat{t}^{ij}}{\rho^{ij}} \tag{5}$$

여기서,  $T^{ij}$  : 전수화된 OD  
 $\hat{t}^{ij}$  : 표본OD  
 $\rho^{ij}$  : 기중점  $ij$  간 표본을

3) 수렴성 검토

수렴성 검토를 위한 목적함수는 기중점별 표본율에 따른 관측교통량과 배분교통량(표본교통량에 표본율을 적용한 교통량)의 오차율로 설정하였다. 관측교통량과 배분교통량의 편차는 합산할 때 서로 상쇄될 우려가 있어 제곱오차를 사용하였다. 목적함수에 대한 구체적인 내용은 백승걸 등(2007c)에 기술되어 있다.

$$L(\rho^{ij}) = \sum_a \left[ \frac{x_a - \bar{x}_a}{x_a} \right]^2 \tag{6}$$

여기서,  $L(\rho^{ij})$  :  $\rho^{ij}$ 에 따른 오차율의 총합  
 $\rho^{ij}$  : 기중점  $ij$  간 표본율  
 $x_a$  : 링크  $a$ 의 배분교통량  
 $\bar{x}_a$  : 링크  $a$ 의 관측교통량

수렴성은  $n$ 번째 표본율과  $n+1$ 번째 표본율에 대하여 각각 오차율을 산정하여 비교하며, 최소 오차율이 나올 때까지 반복한다. 백승걸 등(2007c)에서 제시한 알고리즘은 가상데이터에 random 계수를 적용하여 검증하였기 때문에 목표함수인 오차율의 목표값을 가정하고 목표값에 도달할 때까지 알고리즘을 반복수행하도록 설정하였으나, 본 연구는 <그림 7>에서와 같이 통행량의 편차가 크게 발생하여 특정 목표값을 산정하기 어려우므로 오차율이 최소값이 될 때까지 반복하는 것으로 알고리즘을 수정하였다.  $n+1$ 번째 표본율은 Newton-Raphson method를 이용하여 산정하였다.

$$\rho_{n+1}^{ij} = \rho_n^{ij} - \frac{L(\rho_n^{ij})}{L'(\rho_n^{ij})} \tag{7}$$

여기서,  $\rho_{n+1}^{ij}$  :  $n+1$ 번째 표본율  
 $\rho_n^{ij}$  :  $n$ 번째 표본율

$L(\rho_n^{ij})$  :  $n$ 번째 오차율  
 $L'(\rho^{ij})$  :  $L(\rho^{ij})$ 의 편미분

목적함수인  $L(\rho^{ij})$ 의 편미분은 식(8)과 같다.

$$L'(\rho^{ij}) = 2 \sum_a \left[ \frac{x_a - \bar{x}_a}{x_a} \right] \times x_a' \tag{8}$$

여기서,  $L'(\rho^{ij})$  :  $L(\rho^{ij})$ 의 편미분  
 $x_a$  : 링크  $a$ 의 배분교통량  
 $\left( = \sum_{ij} \frac{\hat{t}^{ij} \cdot \hat{R}_a^{ij}}{\rho^{ij}} \right)$   
 $\bar{x}_a$  : 링크  $a$ 의 관측교통량  
 $x_a'$  :  $x_a$ 의 편미분  $\left( = -\sum_{ij} \frac{\hat{t}^{ij} \times \hat{R}_a^{ij}}{(\rho^{ij})^2} \right)$   
 $\hat{t}^{ij}$  : 표본OD  
 $\hat{R}_a^{ij}$  :  $ij$  통행량의 링크  $a$  이용비  
 $\rho^{ij}$  : 기중점  $ij$ 의 표본율

식(7)과 식(8)에 의해서  $n+1$ 번째 표본율은 식(9)와 같이 산정된다.

$$\rho_{n+1}^{ij} = \rho_n^{ij} - \frac{\sum_a \left[ \frac{x_a - \bar{x}_a}{x_a} \right]^2}{2 \sum_a \left[ \frac{x_a - \bar{x}_a}{x_a} \right] \times x_a'} \tag{9}$$

여기서,  $\rho_{n+1}^{ij}$  :  $n+1$ 번째 표본율  
 $\rho_n^{ij}$  :  $n$ 번째 표본율  
 $x_a$  : 링크  $a$ 의 배분교통량  
 $\bar{x}_a$  : 링크  $a$ 의 관측교통량  
 $x_a'$  :  $x_a$ 의 편미분

식(9)와 같이  $n+1$ 번째 표본율이 산정되면 식(6)에 의해 오차율을 검토하고 검토 결과에 따라 전수화된 OD를 재산정한다.

4) 사례분석 결과 및 검증

최소 오차율  $\rho^{ij}$ 에 따른 표본OD의 전수화 결과는 <표 4>와 같다.



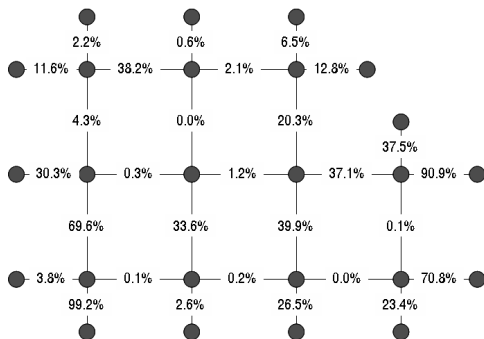
〈표 4〉 전수화된 OD

종점 기점	I	II	III	IV	V	A
I	1,660	82	51	162	63	755
II	69	1,555	44	302	74	44
III	350	115	1,678	51	147	204
IV	72	364	9	1,192	131	22
V	30	753	75	320	1,335	23
A	426	156	104	97	37	2,252
B	191	75	39	116	127	79
C	64	88	38	91	266	83
D	101	162	169	265	227	29
E	9	44	8	36	21	14
F	392	72	388	117	115	305
sum	3,364	3,466	2,603	2,749	2,543	3,810

종점 기점	B	C	D	E	F	sum
I	74	186	103	16	122	3,274
II	340	90	78	18	84	2,698
III	55	16	74	21	400	3,111
IV	196	109	207	27	173	2,502
V	1,144	41	382	50	283	4,436
A	52	99	60	8	60	3,351
B	1,007	20	43	27	31	1,755
C	60	1,854	1,951	105	165	4,765
D	72	23	2,053	180	255	3,536
E	124	37	46	546	87	972
F	92	49	95	45	2,616	4,286
sum	3,216	2,524	5,092	1,043	4,276	34,686

기존OD가 존재하지 않기 때문에 전수화 결과에 대한 검증은 표본율을 적용한 링크별 배분교통량과 관측교통량과의 오차율의 총합(목적함수)으로 검증하였다. 〈그림 10〉은 링크별 오차율을 나타내는 것으로 평균 22.9%로 나타났으며, 전체 29개 링크 중 상·하위 10%를 제외한 경우 오차율 평균은 17.6%로 나타났다. 〈표 5〉는 링크별 오차율을 구성비로 나타낸 것으로 일반적인 수요분석 오차범위 30% 이내인 링크가 전체의 65.5%로 나타났다.



〈그림 10〉 배분교통량과 관측교통량과의 오차율

〈표 5〉 배분교통량과 관측교통량과의 오차율

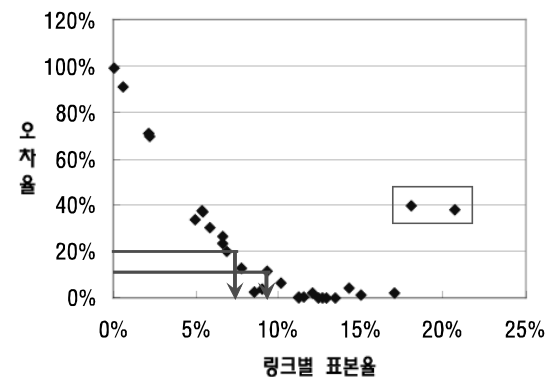
오차율	링크 수	구성비
1% 미만	7	24.14%
1% ~ 5%	6	20.69%
5% ~ 10%	1	3.45%
10% ~ 20%	2	6.90%
20% ~ 30%	3	10.34%
30% ~ 50%	6	20.69%
50% 이상	4	13.79%
계	29	100.00%

#### 4. 시사점

본 연구는 실제 프로브 차량 데이터를 활용한 정적 OD 산정에 관한 기초적 연구로 다음과 같은 시사점을 지닌다.

첫째, 기존OD가 존재하지 않는 지역에서 수집정보를 활용하여 정적OD를 산정하였다는 점이다. 기존의 링크 교통량을 이용한 OD 추정방법은 기존OD가 존재하는 지역에서만 OD의 수정·보완이 가능하였으나, 본 연구에서는 기존OD가 존재하지 않아도 프로브 차량을 통해 수집된 정보만을 활용하여 OD를 산정할 수 있다. 또 현장에서 수집되는 정보를 최대한 활용하기 위하여 블록 단위의 교통준을 설정하거나, 데이터 부족 문제를 보완하기 위하여 통행배분 시뮬레이션을 활용하는 방안을 제시하였다는 점에서도 그 의미를 찾을 수 있을 것이다.

둘째, 오차의 발생원인 검토 과정에서도 주요 시사점을 찾을 수 있다. 배분교통량과 관측교통량의 오차율은 평균 22.9%로 나타났으며, 주요 원인은 데이터 부족과 네트워크 특성으로 분류할 수 있다. 〈그림 11〉은 링크별 표본율과 오차율의 상관관계를 나타낸 것으로 표본율이

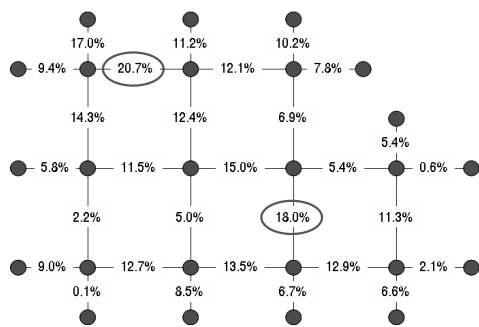


〈그림 11〉 오차율과 표본율의 상관관계

클수록 오차율이 감소함을 알 수 있다. 오차율을 20% 이내로 수렴하기 위해서는 약 7%의 표본율이, 10% 이내로 수렴하기 위해서는 약 9%의 표본율이 요구되는 것으로 나타났다.

표본율이 높음에도 불구하고 오차율이 크게 나타나는 경우(〈그림 11〉의 □로 표시된 부분)는 네트워크의 특성에서 그 원인을 찾을 수 있는데, 오차율이 크게 나타난 2개의 링크를(〈그림 12〉에 표시하였다. 표시된 구간은 양방향 2차로 또는 2차로/4차로가 혼재되어 있는 구간으로 다음과 같은 이유로 오차율이 높은 것으로 판단된다. 첫째, 교차로를 중심으로 리더기를 설치하였기 때문에 리더기가 설치된 링크를 따라 통행하지 않고 가로질러 이동하거나, 교통혼간 이면도로로 리더기가 미설치된 지점을 이동하는 차량의 경우 데이터가 누락될 우려가 있다. 리더기에서 수집된 정보만으로는 이러한 데이터 누락을 고려하기 어려우며, 특히 차로수가 적은 구간의 경우 교통혼간 프로브 차량의 통행이 더욱 빈번할 것으로 예상된다.

둘째, 본 연구에서 산정된 표본OD는 백승걸 등(2007b)에서 제시한 방법을 활용하였으며, 교통혼에서 유입·유출하는 차량은 우회전 진입·진출을 가정하였다. 때문에 차로 수가 적은 구간에서 발생하는 비보호 좌회전은 차량통행특성(기·종점)을 정확히 반영하기 어렵다. 따라서 오차율을 감소시키기 위해서는 네트워크 정보를 고려하여 최적 리더기 설치 위치를 선정하는 것이 중요하다고 할 수 있다.



〈그림 12〉 관측교통량 대비 표본교통량 표본율

주) ○ 표시된 링크는 오차율이 높은 링크

### V. 결론 및 향후 연구과제

OD는 교통계획 측면에서 필수적인 교통정보로 최근에는 교통관리, 교통정보 제공을 위한 동적OD 추정에

대한 연구들이 진행되고 있다. 프로브 차량을 통해 수집된 교통정보는 기종점 확인이 가능한 실시간 데이터로 이를 활용하여 정적OD를 산정하기 위해서는 시간적·공간적으로 변동되는 자료를 집적하여 표본OD를 산정하고 전수화하는 과정이 필요하다.

본 연구에서는 제주시를 대상으로 수집된 실제 데이터를 기존 연구에서 제시된 표본OD 산정 및 전수화 알고리즘에 적용하여 표본OD를 산정하고 전수화하였다. 제주시 일부지역을 대상으로 일주일의 데이터를 누적하여 활용하였으며, 링크별 통행량 편차를 감소시키기 위하여 통행배분 시뮬레이션을 활용하여 표본교통량을 보완하였다. 표본율은 표본교통량과 관측교통량을 기반으로 산정되는 초기 표본율을 배분교통량과 관측교통량의 편차를 최소화하기는 알고리즘을 반복 수행하여 수정하였으며, 산정된 표본율에 따라 표본OD를 전수화하였다. 전수화 결과 링크별 배분교통량과 관측교통량의 평균 오차율은 22.9%로 나타났으며, 상·하위 10%의 이상 자료를 제거한 후의 평균 오차율은 17.6%로 나타났다.

본 연구는 다음과 같은 2가지 시사점을 지닌다. 첫째, 기존OD가 존재하지 않는 지역에서 실시간으로 수집되는 프로브 차량 정보를 활용하여 정적OD를 산정하였다는 점이다. 기존의 OD 추정기법들은 기존OD가 있는 상황에서 추가 정보를 활용하여 정적OD를 산정하였으나, 본 연구에서는 현장에서 수집되는 정보와 통행배분 시뮬레이션 분석 결과를 활용하여 OD를 산정하는 기법을 제시하였다. 둘째, 오차율과 표본율의 상관관계를 검토하여 적정 오차율 내에서 수렴하기 위한 적정 표본율을 제시하였다는 점이다. 관련 자료가 미흡한 실정에서 OD를 산정하기 위한 적정 규모의 표본을 선정하는 일은 매우 어려운 일이다. 본 연구에서는 이를 위한 적정 표본율을 오차율 20% 이내로 수렴하기 위해서는 7% 이상으로, 오차율 10% 이내로 수렴하기 위해서는 9% 이상으로 설정할 것을 제안하고 있다.

그러나 데이터를 확보하기 위하여 설정된 블록 단위의 교통혼이 기존OD의 교통혼 설정방법과 다르기 때문에 기존OD가 존재하는 경우 이를 보완할 수 있는 방법에 대한 검토가 부족하다는 점과 차로 수가 양방향 4차로 이하인 도로의 경우 오차가 상대적으로 크게 발생된다는 한계점도 지닌다. 특히 기존OD가 존재하는 경우 이를 시간적·공간적으로 결합하여 수정·보완하는 방법에 대해서는 향후 연구과제로 수행되어야 할 것이며, 상대적으로 오차를 감소시킬 수 있는 리더기의 최적 설치위치

와 기존OD와 표본OD, 표본OD와 시뮬레이션 OD 등 데이터 보완에 따른 가중치 부여 방법에 대해서도 향후 연구과제로 제안한다.

**참고문헌**

1. 백승걸·정소영(2007a), "RFID 리더기 최적 설치지점 선정 연구", 제55회 학술발표회 논문집, 대한교통학회, pp.419~428.
2. 백승걸·정소영·김범진(2007b), "RFID 프로브 차량을 이용한 정적 표본 OD 추정", 대한토목학회 정기학술대회, 대한토목학회.
3. 정소영·백승걸(2007c), "프로브 차량 표본 OD의 전수화기법 연구", 제57회 학술발표회 논문집, 대한교통학회, pp.137~145.
4. 한국교통연구원(2007), "전국 지역간 여객 기종점통행량 자료의 전수화".
5. 한국교통연구원(2006), "기종점 통행량 신뢰성 향상을 위한 기초연구".
6. 김시곤·유병석·강승필(2005), "휴대폰 기지국 정보를 이용한 O/D 추정기법 연구", 대한교통학회지, 제23권 제1호, 대한교통학회, pp.93~102.
7. 건설교통부(2004), "LBS를 기반으로 한 정고속도 측정 및 O/D 산출기법 개발".
8. 손기민(2004), "Mobile 위치정보를 이용한 동적OD 생성방안 기초연구", 서울시정개발연구원.
9. 백승걸·김현명·신동호(2002), "표본링크이용비를 추가정보로 이용한 OD 행렬 추정", 대한교통학회지, 제

- 20권 제4호, 대한교통학회, pp.83~93.
10. 백승걸(2001), "유전알고리즘을 이용한 링크관측교통량으로부터의 기종점 통행행렬 추정", 서울대학교 대학원 박사학위논문.
11. 백승걸·임강원·이승재(1998), "기종점 추정을 위한 최적관측위치로서의 존 코든라인", 대한국토 및 도시계획학회지, 제33권 제6호.
12. Watling D. P.(1994), "Maximum likelihood Estimation of OD matrices from a partial registration plate number", Transpn. Res. 28B, pp.289~314.
13. Tamin O. Z. and Willumsen G.(1989), "Transport demand model estimation from traffic counts", Transpn. Vol.16, pp.3~2.
14. Cascetta E.(1984), "Estimation of trip matrices from traffic counts and survey data : generalized least squares estimator", Transpn. Res. 18B, pp.289~299.
15. Kuwahara M, Sullivan E. C.(1987), "Estimating OD matrices from roadside survey data", Transpn. Res. 21B, pp.233~248.
16. Spiess H.(1987), "A Maximum likelihood model for estimating O-D matrices", Transpn. Res. 21B, pp.395~412.
17. Landau U., Hauer E., and Geva I.(1982), "Estimation of cross-cordon OD flows from cordon studies", Transportation Research Record, 891, pp.5~10.

✉ 주 작 성 자 : 정소영  
 ✉ 교 신 저 자 : 백승걸  
 ✉ 논문투고일 : 2008. 3. 21  
 ✉ 논문심사일 : 2008. 5. 24 (1차)  
                   2008. 6. 10 (2차)  
 ✉ 심사판정일 : 2008. 6. 10  
 ✉ 반론접수기한 : 2008. 12. 31  
 ✉ 3인 익명 심사필  
 ✉ 1인 abstract 교정필