

■ 論 文 ■

표본링크이용비를 추가정보로 이용한 OD 행렬 추정

OD matrix estimation using link use proportion sample data as additional information

백 승 걸

김 현 명

신 동 호

(한국도로공사 도로연구소 책임연구원) ((주)하이콘엔지니어링 선임연구원)

(원광대학교 도시공학과 부교수)

목 차

- I. 서론
 - II. 기존연구 고찰
 - 1. 추가정보를 이용한 OD 행렬 추정
 - 2. 표본조사자료를 이용한 OD추정연구
 - III. 모형 및 알고리즘
 - 1. 표본링크이용비를 이용한 OD추정모형 개발
 - 2. 해도출 알고리즘
 - IV. 사례분석
 - 1. 사례 네트워크 및 입력자료
 - 2. 분석결과
 - V. 결론
- 참고문헌

Key Words : OD 행렬 추정, 추가정보, 표본링크이용비, 비균형 통행배정, 도로변 면접조사

요 약

조사비용이 비교적 소요되더라도 추정해의 정확도를 높이기 위해 링크교통량과 target OD외에 추가정보를 확보하여 OD추정을 하는 연구들이 시도되고 있다. 그러나 추가정보를 이용한 기존 OD추정기법은 대부분 추가정보의 특성 또는 장점을 유지하는 목적함수 구성 및 분석을 수행하지 못하여, 확보된 정보를 효율적으로 이용하지 못하는 문제를 가지고 있다. 본 연구의 목적은 관측교통량과 target OD외에 비용효율적인 추가정보를 이용하여 가능해의 범위를 좁힘으로써 추정OD의 정확도를 증진하는 것으로, 이를 위해 표본링크이용비(sample link use proportion)를 추가정보로 이용하였다. 즉 OD통행량과 링크교통량과의 관계를 target OD의 통행배분에서 구하지 않고, 도로변 면접조사에서 확보가능하며 신뢰성 높은 정보인 표본링크이용비를 이용하여 구하였다. 이에따라 본 연구에서는 경로기반 비균형 통행배분개념 하에서 링크교통량 보존법칙을 고려할 필요가 없는 OD추정 해도출 알고리즘을 제시하였다.

시험네트워크에 대한 사례분석결과, 표본링크이용비는 추가정보의 정확도가 낮은 경우에도 효율적으로 OD추정력을 향상시킬 수 있었다. 그리고 표본링크이용비를 이용한 OD 추정기법은 target OD 오차나 관측교통량 오차에 크게 영향을 받지 않아, 링크교통량이나 OD 행렬이 변화된 곳에서도 비교적 안정적인 OD 추정이 가능하였다. 또한 표본링크이용비를 추가정보로 이용할 경우에 이용정보간의 정밀도문제가 발생하기 때문에 다른 이용정보의 정밀도 수준을 고려하여 추가정보의 조사수준을 설정하여야 하며, 관측교통량을 기본정보로 하는 추정기법은 링크교통량을 일정수준까지는 관측하여야 추가정보의 자료활용성을 높일 수 있다는 점을 제시하였다. 마지막으로 링크상의 추가정보는 최적조사위치문제를 고려하여야 하며, 특히 정보의 정밀도 측면에서 볼 때 링크교통량의 최적관측위치문제보다 표본링크이용비의 최적조사위치문제가 추정력 향상에 더 중요한 영향을 미칠 수 있는 것으로 파악되었다.

I. 서론

1920년에서 1950년 대까지 미국에서 도로변 면접 조사(roadside survey)는 대도시의 cordon line들을 가로지르는 외부(external) 통행의 패턴을 추정하고, 교량이나 터널, 우회로와 같은 중요지점에 대한 국지적 통행패턴을 추정하기 위해 많이 이용되었다. 그러나 이러한 기법들을 통해서서는 전체 OD 통행량 패턴을 나타내기 어렵기 때문에, 1950년대부터는 종합적인 가정방문조사(home survey)가 많이 발달하였다. 그러나 가정방문조사는 많은 노력과 비용, 그리고 시간이 요구된다. 따라서, 이러한 기법은 빠른 시간내에 정책의 영향을 파악해야 하는 교통계획이나 재정적 제약이 따르는 경우에는 적합하지 않다. 가정방문조사가 갖는 이러한 문제점들 때문에 1970년대 초반부터 좀 더 경제적인 방법으로 OD를 추정하는 기법들이 연구되기 시작했다. 현재까지 많이 연구되는 접근법은 링크관측교통량(link traffic counts)으로부터 OD통행표를 추정하는 방법이다.

그러나 관측교통량만을 이용한 OD추정은 유일해를 선정하기가 어려우며, 이를 해결하기 위해 target OD가 추가정보로 이용된다. 일반적으로 target OD로 적용되는 기존 OD도 조사오차와 사회경제적 변화로 인해 신뢰도가 저하되는 경우가 많다. 따라서 해의 정확도를 높이기 위해 조사비용이 비교적 소요되더라도 링크교통량과 target OD외에 추가정보를 확보하여 OD추정을 하는 연구들이 시도되고 있다.

본 연구의 목적은 관측교통량과 target OD외에 비용효율적으로 추가정보를 이용하여 가능해의 범위를 좁힘으로써 추정OD의 정확도를 증진하는 것이다. 본 연구에서는 이를 위해 관측교통량과 target OD외에 표본링크이용비(sample link use proportion)라는 추가정보를 이용하여 OD를 추정하는 기법을 제시하였다. 즉 OD 통행량과 관측교통량과의 관계를 target OD의 통행배분에서 구하지 않고, 도로변 면접조사를 통해 신뢰성높은 정보를 구할 수 있는 표본링크이용비를 이용하였다.

본 연구는 2장에서 기존 연구를 살펴본 후 3장에서는 모형개발과 해도출 알고리즘을 구성한 뒤, 4장에서 시험네트워크에 대해 사례분석을 실시하였다.

II. 기존연구 고찰

1. 추가정보를 이용한 OD 행렬 추정

OD 추정문제는 본질적으로 자료수집비용과 도출 결과의 정확성간에 상충(trade-off) 관계를 갖는다(Kuwahara(1987) 등). 가정방문조사와 링크관측교통량 자료만을 이용하는 기법은 이러한 상충관계의 극단적인 두 예이다.

관측교통량만을 이용한 OD추정은 과소설정문제로 인해 유일한 해를 선정하기가 어렵다. 또한 관측교통량은 상이한 관측시간으로 인한 변동(variation)과 관측오차가 존재한다. 또한 이 기법은 해가 존재하는 경우 관측교통량을 재도출하는 네트워크 균형교통량해를 결정하도록 설계되는데, OD를 통행배분하여 링크교통량을 계산하는 경우 사용자균형 가정을 적합한 배분원칙으로 이용할 수 있지만, 역으로 교통량으로부터 OD를 추정하는 경우에도 이러한 통행배분기법을 그대로 적용하기는 어렵다. 왜냐하면 관측된 교통량을 재발생시키는 여러 개의 OD가 존재할 수 있고, 관측교통량의 변동과 불일치(inconsistency), 관측오차 등으로 인해 균형교통량 자체도 완전한 경우가 거의 없기 때문이다.

과소설정문제를 해결하기 위해 추가정보를 이용하거나 중력모형 또는 엔트로피 극대화 모형과 같이 모형의 구조를 제약하는 기법을 이용한다. 대부분의 관측교통량을 이용한 OD 추정모형은 식(1)의 일반화 최소자승법(General Least Square method)과 같이 링크교통량과 OD에 대해 각각의 자료와 추정치간의 차이를 최소화하는 것이다.

$$\text{Min } \gamma_1 F(v - \bar{v}) + \gamma_2 F(t - \bar{t}) \quad (1)$$

여기서,

v : 추정링크교통량

\bar{v} : 관측링크교통량

t : 모형을 통해 구한 추정 OD

\bar{t} : target OD

γ_1, γ_2 : 교통량자료와 OD자료의 신뢰도

일반적으로 target OD로 이용되는 기존 OD는 대규모 조사대상지역, 낮은 표본율, 조사 및 응답의 부

정확성 등으로 오차나 신뢰도의 범위가 상대적으로 크다. 또한 기존 OD는 대부분 정기적인 대규모 조사를 통해 구하는 것으로, 해당지역의 사회경제적 변수가 안정적이지 못한 곳에서는 신뢰성이 낮다. 따라서 추정해의 정확도를 높이기 위해서는 조사비용이 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.

관측교통량을 이용한 OD 추정연구는 추가정보로 기존OD만 이용하였으나, 최근 여러 가지 다른 추가정보를 이용하는 연구가 진행중이다. OD통행 추정에 이용된 추가정보는 Tamin(1989) 등의 존별 사회경제적 변수, Cascetta(1984) 등의 표본 OD, Carey(1981) 등과 Spiess(1987)의 존별 유입·유출량, Kuwahara(1987) 등의 도로변 면접조사자료, Landau(1982) 등의 코든라인 조사자료 등이 있다. Watling(1994)은 차량번호판 조사를 통해 기중점을 파악하였다.

총통행량과 존 유입·유출량, 표본 OD는 통행발생과 통행분포단계, 즉 존 센트로이드간에 관련된 정보이며, 표본링크이용비, 번호판조사는 링크상에서 도출가능한 정보이다. 특히 표본링크이용비는 OD와 링크간의 관계를 나타내주는 중요한 정보이다. 동적(dynamic) OD 추정시에는 Hellinga(1994)의 시험차량(probe car) 정보와 같이 보다 정밀한 추가정보를 확보할 수 있다. <그림 1>은 이용정보에 따라 OD 추정기법을 구분하여 나타낸 것이다.

이러한 추가정보 이용기법은 관측교통량과 기존 OD를 이용한 기법에 비해 비용과 시간이 다소 더 소

요되지만, 비교적 신뢰도가 큰 정보를 얻을 수 있다. 본 연구에서는 비교적 자료수집이 용이하고, 상대적으로 정확한 정보수집이 가능한 표본링크이용비를 추가정보로 이용한 OD추정기법을 제시하였다.

2. 표본조사자료를 이용한 OD추정연구

대상지역 전체에 대한 표본OD조사(sample OD survey)는 표본율을 일정수준 이상으로 하는 경우 조사비용과 시간이 전통적인 기법만큼 소요되며, 표본율이 낮을 경우에는 추가정보로서의 신뢰성이 크게 저하된다는 문제를 가지고 있다. 따라서 대상지역 전체에 대한 표본OD를 추가정보로 이용하는 것은 현실적으로 적합하지 않다. 링크상에서의 표본OD조사는 부분적인 링크이용비만 도출하기 때문에, 대상지역 전체에 대한 표본OD조사와는 구별된다.

링크상에서의 표본조사기법에는 도로변 면접조사(roadside interview)와 코든라인조사(cordonline survey)가 있다. 이는 대상지역의 특정지점들(일반적으로 "screenline" 또는 "cordonline"에 위치)에서 운전자 면접조사를 통해 일부 링크의 링크이용비 행렬, 즉 표본링크이용비를 구하는 것이다.

Landau(1982) 등은 코든라인조사에 의해 구한 관측교통량과 코든라인을 통과하는 표본OD를 이용한 최우추정법에 의해 전체 OD를 추정하였다. 그러나 모형이 적합도를 평가할 목적함수를 가지고 있지 않으며, 도출되는 OD도 부분지역(subnetwork) OD라는 문제점을 갖는다.

Kuwahara(1987) 등은 도로변의 특정지점에서 OD 조사와 교통량관측으로 수집된 자료들을 결합하여 분석지역 전체에 대한 OD를 추정하는 기법을 제시하였다. Kuwahara(1987) 등은 Dial 알고리즘을 통해 p_{od}^a 를 먼저 추정하였으며, 식(2)와 같은 목적함수를 설정하였다.

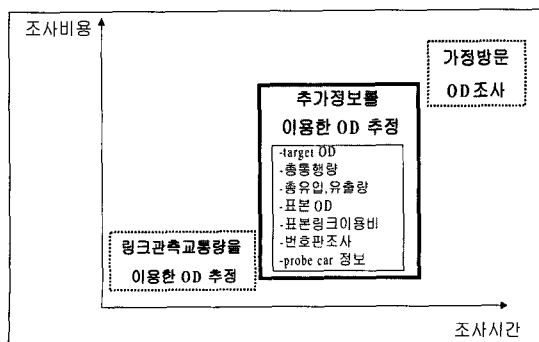
$$Min \sum_a (v_a - \bar{v}_a)^2 = \sum_a \left(\sum_{od} t_{od} p_{od}^a - \bar{v}_a \right)^2 \quad (2)$$

여기서,

v_a : 링크 a에서의 추정교통량

\bar{v}_a : 링크 a에서의 관측교통량

p_{od}^a : 링크 a를 이용하는 OD쌍 od의 통행량



<그림 1> 이용정보에 따른 OD 추정기법 구분

Kuwahara(1987) 등의 기법은 조사된 부분링크 이용량과 모형에서 구한 링크이용비의 비교를 통해 OD를 추정하고, 링크이용비에 의한 추정교통량과 관측교통량의 차이를 최소화하여 표본OD를 추정하는 것이다. 그러나 관측교통량보다 더 자세한 정보인 표본링크이용비를 링크교통량을 추정하는 데에만 이용하여 확보된 정보를 비효율적으로 이용한다는 문제를 가지고 있다.

Lo(1996, 1999) 등은 대부분의 관측교통량으로부터 OD 추정모형이 링크선택비(link choice proportion)를 상수로 가정하고 통행배분모형으로부터 OD를 추정하기 때문에, 잘못된 링크선택비의 영향이 매우 중요하며 상당한 추정오차를 발생시킬 수 있다는 점을 지적하였다. 이러한 문제를 줄이기 위해 Lo 등은 확률적 링크선택비와 OD통행량 정보, 관측교통량을 결합하여 OD추정을 하는 통계적 기법(최우추정법, 베이저안 추론기법)을 제안하였다.

그러나 Lo(1996, 1999) 등이 언급한 표본 OD와 링크선택비는 실제로는 거의 확보가 불가능한(전통적인 OD조사외의 방법으로는) 대상지역 전체에 대한 자료이다. 또한 링크선택비는 non-redundant 링크에서 구해야 하기 때문에, 네트워크관계를 이용하여 링크선택비의 관계를 설정하는데, 이때 링크선택비 보존법칙(conservation of link choice proportion)으로 정의할 수 있는 다음 조건들을 이용한다. (1)한개의 OD쌍(pair)에 대해 출발지와 목적지에 연결된 모든 링크선택비의 합은 1이다. (2)한개의 OD쌍에 대한 특정 노드의 유출, 유입링크선택비의 합은 같다. 하지만 위의 조건은 균형통행배분을 필요로 하며 대규모 네트워크에서는 경로계산(path enumeration)을 이용하지 않고는 설정하기 어렵다.

Hyodo(2001) 등은 식(3)과 같이 각 링크에서의 설문조사자료와 관측교통량의 차이를 최소화하는 최소자승법을 제시하였는데, 이는 조사된 각 경로별 교통량이라는 상세한 정보를 제대로 이용하지 않고 링크교통량으로 집단화시켜서 정보이용의 비효율문제를 초래한다. 또한 OD 통행량 표본오차와 링크교통량 관측오차를 고려하지 않아 일반적인 상황에서의 모형 추정력을 보인다고 하기 어렵다.

$$\min_{t_k} F = \sum_a \left(\sum_k \delta_{ak} t_k - v_a \right)^2 \quad (3)$$

여기서,

δ_{ak} : 링크 a가 경로 k의 부분이면 1, 그렇지 않으면 0

t_k : 경로 k의 통행량 $(od\ t = \sum_{k=1}^K t_k)$

v_a : 링크 a의 관측교통량

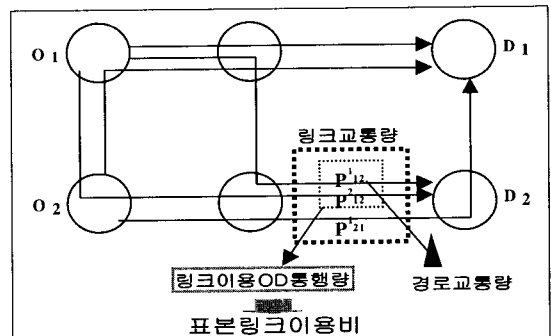
앞에서 살펴본 바와 같이 표본조사자료를 이용한 기존 OD추정기법은 대부분 추가로 확보한 정보의 특성 또는 장점을 유지하는 목적함수 구성 및 분석을 수행하지 못하여, 확보된 정보를 비효율적으로 이용하는 문제를 가지고 있다. 이는 추가정보의 정밀도(precision)가 높아질수록 보다 중요한 문제가 될 수 있다.

III. 모형 및 알고리즘의 개발

1. 표본링크이용비를 이용한 OD추정모형 개발

<그림 2>는 링크상에서 도출가능한 정보들을 나타내는 사례네트워크로, O_1, O_2 는 기점, D_1, D_2 는 종점을 나타낸다.

P_{12}^1, P_{12}^2 는 기점 O_1 에서 출발하여 종점 D_2 로 가는 통행의 경로 1,2의 통행량이며, P_{21}^1 은 기점 O_2 에서 출발하여 종점 D_1 로 가는 통행의 경로 1의 통행량을 나타낸다. 여기서 P_{12}^1, P_{12}^2 은 OD쌍 $O_1 D_2$ 간의 링크이용 OD 통행량이며, $P_{12}^1, P_{12}^2, P_{21}^1$ 은 모두 합하여 관측교통량을 나타낸다. 따라서 링크상에서 도출 가능한 정보들은 관측교통량, 링크이용OD 통행량(표본링크이용교통량), 경로교통량 순으로 세부적인 정보가 되며, 이러한 점에서 기존 OD통행량은 상당히 개략적인 정



<그림 2> 링크상에서 도출가능한 정보(예)

보이다. 그러므로 관측교통량으로부터 OD 추정기법의 추정력을 향상시키기 위한 추가정보는 관측교통량보다 더 세부적인 정보를 이용하는 것이 효과적이라고 할 수 있다.

〈그림 3〉은 관측교통량으로부터의 OD 추정기법과 표본링크이용비를 이용한 OD추정기법의 이용정보유형을 예로써 비교한 것이다. 여기서 행은 OD쌍, 열은 링크를 나타내며, 각 셀은 각 OD통행량이 관측링크를 통행하는 비율을 나타낸다. 행의 합은 각 OD쌍의 통행량, 열의 합은 각 링크의 교통량을 나타낸다. 따라서 이론적으로 모든 링크에서 모든 OD통행량을 조사한다면 OD통행량을 직접 구할 수 있다. 그러나 현실적으로는 〈그림 3(b)〉처럼 일부 링크에 대해서만 관측교통량과 링크이용비를 조사할 수 있으며, 조사 링크의 OD통행량도 전체 OD 통행량중 일부 OD통행량이다. 따라서 이러한 표본자료를 이용하여 전체 OD를 추정하는 기법이 필요하게 된다.

〈그림 3〉에서 음영부분은 교통량 관측 또는 표본링크이용비 조사가 이루어진 곳을 나타내는데, 본 예제에서는 총 6개 링크중 4개 링크에서 교통량이 관측되었으며, 2개 링크에서 표본링크이용비가 조사되는 것을 가정하였다. 〈그림 3(a)〉는 관측교통량만을 이용하여 OD추정을 한다. 따라서 관측교통량과 OD간의 관계는 이용자균형 등 단일하고 고정적인 통행배분 map에 의해서만 설정할 수 있다. 그러므로 matrix end 값은 관측된 교통량이지만, 셀 내부는 통행배분 모형 등에 의해 채우게 된다.

모형기반(model-based) 기법과 달리 〈그림 3(b)〉는 matrix end값과 셀 내부의 값 모두를 관측자료 또는 조사자료를 이용하여 구축하는 자료기반(data-based) 기법이다. 즉 관측교통량 뿐만 아니라 표본이지만 특정링크를 이용하는 OD통행의 규모를 구할 수 있다. 관측교통량만을 이용하는 것보다 표본링크이용비를 추가정보를 이용하는 것이 더 정확한 OD를 추정할 가능성이 크다. 왜냐하면 링크상에 관측교통량은 하나만 존재하지만 표본링크이용비는 여러 개가 존재하며, 관측교통량과 달리 링크와 OD와의 관계를 한꺼번에 나타내주기 때문에 과소설정문제를 감소시키며, 보다 정확한 추정을 가능하게 할 수 있다. 또한 표본링크이용비 정보를 이용하면 통행거리분포도 파악할 수 있으며, 이를 기존OD의 통행거리분포와 비교하여 OD추정력을 증진시킬 수도 있다.

P matrix		Trips by OD pairs					
		1	2	3	4	5	sum
Link Traffic Counts	1						
	2						
	3						
	4						
	5						
	6						

(a) 관측교통량으로부터의 OD 추정

P matrix		Trips by OD pairs					
		1	2	3	4	5	sum
Link Traffic Counts	1						
	2						
	3						
	4						
	5						
	6						

(b) 표본링크이용비를 이용한 OD추정

〈그림 3〉 OD 추정기법별 이용정보유형 비교(예)

주) 음영은 관측 또는 조사가 이루어지는 정보임.

링크이용비는 OD와 링크, 즉 존으로 대표되는 특정지역과 네트워크의 특정지점간의 관계를 나타내는 정보이다. 따라서 표본링크이용비는 일부 링크에 대해서만 조사하여도 다른 링크의 링크이용비를 알 수 있으며, OD 추정력을 비용효율적으로 증진시킬 수 있다. 그러므로 기존 OD, 링크관측교통량 외에 링크이용비 정보를 이용함으로써 OD추정에 대한 보다 완전한 자료를 구축할 수 있게 된다.

본 연구에서는 기존 OD를 가용정보로 하고 추가정보로 표본링크이용비를 이용하는 OD추정문제를 검토한다. 제시하는 모형식은 식(4)와 같다.

$$Min F \sum_a \sum_{od} (t_{od}^a - \bar{t}_{od}^a)^2 = \sum_a \sum_{od} (t_{od}^a - \bar{v}_a \cdot \bar{p}_{od}^a)^2 \tag{4}$$

여기서,

\bar{t}_{od}^a : 링크 a에서 조사된 od쌍의 통행량

t_{od}^a : 링크 a에서 추정된 od쌍의 통행량

\bar{v}_a : 링크 a의 관측교통량

\bar{p}_{od}^a : 링크 a에서 조사된 od쌍의 통행량비 (=표본링크이용비)

본 연구에서 이용하는 표본링크이용비는 Lo 등(1996, 1999)이 제시한 링크선택비와는 다른 것이다. 링크선택비는 OD 통행량이 특정링크를 이용하는 비율이기 때문에, OD 통행량이 이용하는 모든 링크의 비율을 알아야 하지만, 표본링크이용비는 단순히 링크를 이용하는 교통량중 특정OD 통행량들의 비율을 구하는 것이다. 따라서 링크선택비를 통해 OD를 추정하기 위해서는 링크선택비의 역함수를 취해 링크교통량에 곱해주어야 하나, 표본링크이용비를 이용한 기법은 식(4)와 같이 관측교통량에 표본링크이용비만을 곱해주어도 OD 추정이 가능하게 된다. 또한 조사자료로 확보할 수 있는 표본링크이용비는 추정교통량에 대한 OD 통행량이 아니라 관측오차를 가지고 있는 관측교통량에 대한 OD 통행량의 비율이다.

2. 해도출 알고리즘

관측교통량으로부터 OD행렬을 추정하는 모형은 대부분 이용자균형기반의 통행배분모형을 이용하며, 해가 존재하는 경우 링크교통량을 재도출하는 통행량해를 결정하도록 설계된다. 그러나 본 연구에서는 OD와 링크와의 관계를 기존 OD에 대한 통행배분에서 구하지 않고 표본링크이용비를 이용하여 구하기 때문에, 경로기반 비균형 통행배분(path-based non-equilibrium assignment) 개념 하에서 링크교통량 보존법칙 등을 고려할 필요가 없는 OD추정 해도출 알고리즘을 제시하였다. 비균형 통행배분을 이용하면 링크선택비보존법칙과 non-redundant의 구분이 필요치 않게되며, 보다 일반화된 모형을 구성할 수 있다.

그 이유는 첫째, 관측된 교통량을 재발생시키는 여러 개의 OD가 존재할 수 있고, 관측교통량의 변동과 관측오차 등으로 관측교통량이 균형교통량과 완전히 일치되기는 거의 불가능하다. 이와같이 관측교통량은 변동과 오차를 포함하기 때문에, 관측교통량을 기준으로 OD를 추정할 경우에 링크교통량 보존법칙이 정확한 추정을 오히려 제약하게 될 수도 있다. Sherali 등(1994)도 개별 이용자들은 최소화경로를 선택하지만, 불완전한 정보로 인해 실제 교통량은 이용자균형해와 반드시 일치하지 않을 수도 있으며, 관측교통량 자료의 내재된 불일치(inconsistency)로 인해 링크교통량을 도출하는 통행표가 존재하지 않을 수도 있다는 점을 지적하였다. 이에따라 Sherali 등(1994)

은 목적함수 내에 적합하게 내재시킨 페널티(penalty) 하에서 균형조건의 침해(violation)뿐만 아니라 관측 링크교통량의 편차(deviation)도 허용하는 모형을 제시하였지만, 관측교통량의 변동 또는 오차가 큰 경우 이용자균형 통행배분에 의해 단일 OD를 생성하기는 어렵다. 확률적 통행배분 또는 통계적 기반의 OD 추정기법을 통해 이러한 문제를 완화하려는 연구(Lo 등, 1996)도 있었다. 그러나 확률적 통행배분의 경우 근본적으로 이용자균형개념을 이용하기 때문에 동일한 문제가 발생하며, 통계적 기반의 OD 추정은 현실적으로 OD의 공분산을 구할 수 없기 때문에 적용 가능성이 낮다.

둘째, 링크이용비보다 경로선택비의 민감도가 낮으므로, 이용자 균형배분에서 발생할 수 있는 링크교통량보존제약이 완화되며, 따라서 표본링크이용비를 효율적으로 이용할 수 있기 때문이다. 셋째, OD와 링크와의 관계를 기존 OD의 통행배분에서 구하지 않고, 도로변 면접조사를 통해 직접 구하기 때문이다. 넷째, path enumeration을 하는 대신 통행배분을 수행하지 않으므로 대규모 네트워크에서 보다 효율적으로 알고리즘을 수행할 수 있다.

본 연구에서는 경로기반의 알고리즘을 통해 경로교통량과 OD통행량을 동시에 추정하는 기법을 제시하였다. 경로기반의 알고리즘을 이용하면 부분지역 OD에 대한 정보를 보다 손쉽게 구할 수 있다. 다양한 path 선택비의 조합최적화를 수행하기 위해 본 연구에서는 유전 알고리즘을 이용하였다. 본 연구에서 이용한 표본링크이용비로부터의 OD 추정에 대한 상세한 알고리즘은 다음과 같다.

■ OD estimation algorithm using GA and Link Use Proportions(GALUP)

(Step 0) All Feasible Path Enumeration(AFPE)

Define

O : 기점집합

D : 종점집합

$a(i, j)$: 노드 i 에서 노드 j 를 연결하는 링크,
 $i, j \in N$ (노드집합)

① 초기화

각 노드별로 연결된 링크설정 : $a(i, j)$

노드 i 와 연결된 링크수 : $n_i = 1, 2, \dots, m$

② 각 기점별로 종점을 향하여 tree-building

```

For each origin node O
  For each no,
    k = no
    Path(k) = { a(O, j) }
    여기서, 노드 j는 기점과 연결된 다음 노드
      (sequential node)
  End
End

```

③ 경로 탐색(반복탐색)

```

next node j ∈ D 이면 정지 : Path(k)도출
Otherwise
  For each nj,
    k = k + nj - 1
    링크 a(j, l)을 Path(k)에 추가 ; 즉,
    Path(k) = { a(O, j), ..., a(j, l) }
    여기서, 노드 l은 노드 j와 연결된 다음
      노드(sequential node)
  End

```

[Step 1] 초기화(Random Generation)

- ① 파라메터값 등 설정
 - mutratio (돌연변이 확률)
 - max iter (최대반복수)
 - n = 1 (세대수 또는 반복수)
- ② (0, 1) 사이의 값을 가지는 T_n[m][od], P_n[m][od][k] random values 발생
- ③ 식(5)를 통해 n 세대중 m 번째 유전자의 추정 OD 계산

$$t_{n,od}^a[m] = T_n[m][od] * P_n[m][od][k] * \bar{t}_{od} \quad (5)$$

여기서,

t_{od}^a[m] : n세대중 m번째 유전자의 기종점 od 통
 행발생량중 링크 a를 이용하는 교통량
 T_n[m][od] : n세대중 m번째 유전자의 기종점
 od 통행발생비율
 P_n[m][od][k] : n세대 중 m번째 유전자의 기종
 점간 od통행의 경로 k 이용비율

[Step 2] Fitness Calculation

- ① 식(5)에서 도출된 t_{od}^a[m]를 이용하여 링크 a에
 대한 t_{od}^a[m] 계산
- ② 식(4)를 이용하여 적합도 계산

③ Sort : 적합도값이 낮은 순서로 sorting

[Step 3] Cross-over

- ① 전체 유전자중 최상위 유전자 일정 비율 선택
- ② 일정비율 이하의 유전자값을 하나씩 삭제하고, 새
 로운 염색체값을 상위값에서부터 일정비율 까지의
 개체 중 임의로 2개의 값을 선택해 재설정
- ③ 마지막 염색체값까지 값이 재설정되면 정지

[Step 4] Mutation

- ① 각 T_{n+1}^{*}[m][od]과 P_{n+1}^{*}[m][od][k]에 대해 (0,1)
 사이의 확률값 r 발생
- ② r ≤ mutratio 이면
 (0,1) 사이의 값을 가지는 T_{n+1}^{*}[m][od]과
 X_{n+1}^{*}[m][od][k]의 random values 발생

[Step 5] Fitness Calculation

- ① 식 (6)을 계산

$$t_{n+1,od}^a[m] = T_{n+1}[m][od] * P_{n+1}[m][od][k] * \bar{t}_{od} \quad (6)$$

- ② 이하는 [Step 2]와 동일

[Step 6] Stopping Criteria

n ≤ max iter이면 n = n + 1 하고 [Step 2]로 가고,
 그렇지 않으면 정지.

위 [Step 0]의 AFPE에서, 경로의 순환(loop)을
 방지하기 위하여 노드나 링크의 중복 금지제약을 줄 수도
 있다. 이런 경우, 과정 ③에서 Path(k)에 동일한 링크
 나 노드가 추가되면, 해당 링크를 추가하지 않으면 된다.

제시된 알고리즘은 표본링크이용비 자료가 확보가능해
 짐에 따라 백승걸(2000) 등이 제시한 기법에서 다중통행
 배분(multiple assignment) map을 수행하지 않아도 되
 기 때문에, 알고리즘을 보다 효율적으로 수행할 수 있다.

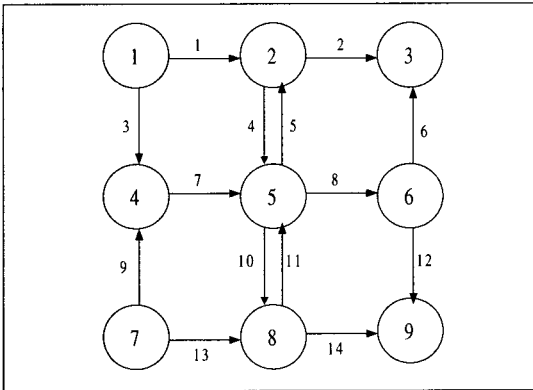
IV. 사례분석

1. 사례 네트워크 및 입력자료

본 연구에서 이용한 실제 OD 및 사례 네트워크는
 <표 1>과 <그림 4>와 같다.

〈표 1〉 사례 네트워크의 실제 OD

구분		종점		
		3	5	9
기점	1	1800	1000	1500
	5	1200	-	1000
	7	1600	1400	2000



〈그림 4〉 사례네트워크

GAMLUP의 파라미터는 다음과 같다.

- population size는 40.
- T는 30% 변동 ($0.7 \leq T_n[m][od] \leq 1.3$)
- P는 30% 변동 ($0.7 \leq P_n[m][od][p] \leq 1.3$)
- 돌연변이율은 반복수에 따라 다르게 설정
- 최대반복수는 10,000번

기존 OD 통행행렬 \bar{t}_w 과 관측링크교통량 \bar{v}_a 는 백승걸(2000) 등과 동일하게 실제값에 변동치를 부여하여 발생시켰다. 다만 제시기법의 추정력을 명확히 보기 위해 각 링크별 표본링크이용비는 모두 균일한 표본오차를 부여하여 발생시켰다. 계산편의를 위해 표본링크이용비의 표본율은 모든 링크에서 동일한 것으로 설정하였다.

비교모형은 Yang(1995)이 제시한 IEA(Iterative Equilibrium Assignment)와 백승걸(2000) 등의 GAM(Genetic Algorithm Method)이다. 여기서 IEA와 GAM은 관측교통량과 target OD를 이용한다. 각 모형의 추정OD와 실제 OD간의 차이를 비교하기 위해 식(7)의 Relative Mean Absolute Error(MAE)를 추정자로 사용하였다.

$$MAE(\%) = \left(\frac{\sum_{w \in W} |t_w - t_w^+|}{\sum_{w \in W} t_w^+} \right) \times 100 \quad (7)$$

여기서,

t_w 와 t_w^+ : 추정 OD와 실제 OD 행렬

W : 기종점의 수

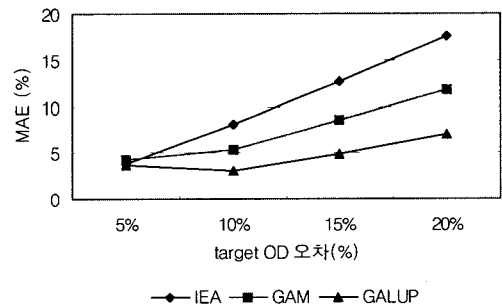
2. 분석결과

1) target OD오차에 대한 모형추정력

〈그림 5〉는 관측교통량오차 6%, 표본이용비 오차 10%, 교통량 관측지점수 및 표본링크이용비 조사지점수가 8개소일 때 target OD오차에 대한 모형추정력을 나타낸다.

분석결과 target OD오차가 증가할수록 각 모형의 추정오차는 증가하나, GALUP 추정오차는 상대적으로 적게 증가하였다. 이는 표본링크이용비를 추가정보로 이용하는 경우 target OD의 오차변화에 상당히 안정적인 OD 추정력을 확보할 수 있다는 것을 나타낸다.

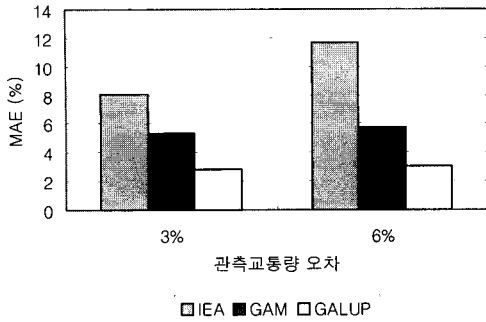
〈그림 5〉에서 특이한 점은 target OD의 오차가 10%일 때보다 5%일 때 GALUP의 추정오차가 약간 더 높다는 것이다. 표본링크이용비의 오차는 모든 target OD오차에서 10%였다. 본 연구에서는 이러한 결과를 이용정보간의 정밀도 문제로 해석하였다. 즉 상대적으로 상세한 정보를 가지는 표본링크이용비의 오차는 상대적으로 덜 상세한 정보를 가지는 target OD의 오차보다는 작아야 한다는 것이다.



〈그림 5〉 target OD오차에 대한 모형추정력

2) 관측교통량 오차에 대한 모형추정력

〈그림 6〉은 target OD 오차 10%, 교통량 관측지점수 및 표본링크이용비 조사지점수 8개소, 표본이용비 오차 10%일 때 관측교통량오차에 대한 모형추정



〈그림 6〉 관측교통량오차에 대한 모형추정력

정력을 나타낸다.

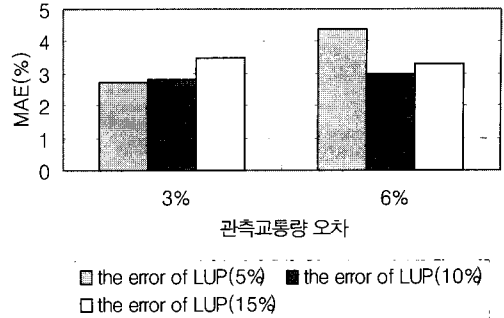
분석결과 관측교통량 오차 3%와 6% 모두에서 GALUP의 추정오차가 타모형의 추정오차보다 낮았다. GALUP의 추정오차는 관측교통량 오차 3%일 경우보다 6%일 경우에 더 높았지만, 그 차이는 상당히 작았다. 따라서 GALUP는 target OD의 오차뿐만 아니라 관측교통량 오차에 대해서도 상당히 안정적인 OD 추정력을 나타낸다고 할 수 있다. 이러한 사례분석결과로만 볼 때 표본링크이용비는 관측교통량의 오차나 변동이 큰 곳일수록 기존 모형보다 적용성이 크다고 할 수 있다.

3) 표본링크이용비 오차에 대한 GALUP의 추정력

〈그림 7〉는 target OD 오차 10%, 관측지점수 및 조사지점수가 8개소인 경우, 관측교통량 오차별로 표본링크이용비 오차에 대한 모형추정력을 나타낸다.

분석결과 전반적으로 관측교통량오차와 표본링크이용비 오차가 클수록 추정력은 점차 낮아지는 것으로 나타났다. 그러나 관측교통량 오차 6%에서 표본링크이용비 오차가 5%인 경우는 예외인데, 이는 동일한 가중치로 본 연구의 모형에 적용되는 링크관측교통량과 표본링크이용비의 이용정보간의 정밀도문제와 관련된다. 즉 관측교통량의 오차가 기본 정보인 표본링크이용비의 오차보다 큰 경우에는 모형의 추정력 향상을 기대하기 어렵다는 것이다. 따라서 관측교통량을 이용한 OD추정에서 추가정보를 확보하기 위한 조사를 하는 경우에도 기본적인 정보인 관측교통량의 오차를 줄이는 것이 상당히 중요하다고 할 수 있다.

백승걸(2000)은 기존 OD오차와 관측교통량 오차 중 낮은 값의 오차를 추정OD 오차의 하한값으로 설



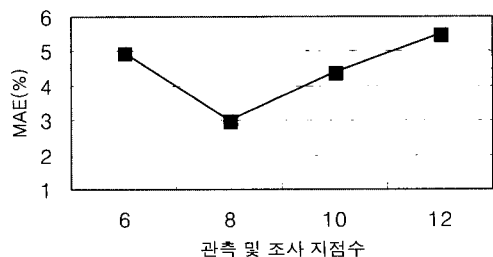
〈그림 7〉 관측교통량 및 표본링크이용비 오차에 대한 모형추정력

정할 수 있으며, 기존 OD를 갱신하기 위해서는 관측교통량 오차가 최소한 기존 OD의 오차보다는 적어야 합리적인 추정이 가능함을 보였다. 즉 추정 OD오차에 대한 이용정보 오차의 “하한값(low limit value)” 기준을 제시한 바 있다. 본 연구결과는 표본링크이용비를 추가정보를 이용하는 경우 다른 정보들의 가시적인 오차 하한값보다 더 낮게 추정오차를 줄일 수 있다는 것을 나타낸다.

4) 표본링크이용비 조사지점수에 대한 모형 추정력

〈그림 8〉은 관측교통량 오차 6%, target OD 오차는 10%, 표본이용비 오차 10%일 때 교통량 관측 및 표본링크이용비 조사지점수에 대한 모형의 추정력을 나타내는 것이다.

분석결과 모든 조사 및 관측지점수에서 추정 OD의 오차가 관측교통량 오차인 6%보다 낮아, 표본링크이용비가 추가정보로서 추정력을 향상시키는 것으로 나타났다. 그러나 표본링크이용비 조사지점수 증가에 따라 추정OD의 오차는 오히려 증가하거나 유사하였다. 따라서 링크상에서 추가정보는 링크교통량



〈그림 8〉 관측 및 조사지점수에 대한 모형추정력

관측지점이 많아도 반드시 최적해를 보장하지는 못하는 관측교통량의 최적관측위치문제(Yang, 1998; 백승걸 등, 1998)처럼 "최적조사위치문제"가 발생하는 것으로 파악된다. 특히 정보의 정밀도문제 및 백승걸(2000) 등의 분석결과와 비교할 때 링크교통량의 최적관측위치문제보다 표본링크이용비의 조사위치문제가 더 중요한 영향을 미칠 수 있는 것으로 판단된다.

V. 결론

본 연구에서는 관측교통량자료와 기존 OD자료 외에 표본링크이용비라는 추가정보를 이용하여 OD를 추정하는 기법을 제시하였다. 추가정보를 이용한 OD추정 연구들을 통해 기존 OD와 관측교통량의 변동과 오차가 클 것으로 예상되는 경우 저비용, 단시간에 조사 분석이 가능한 코든라인조사, 도로변조사 등의 표본 조사를 적극적으로 도모하여 OD 추정의 정확성을 증진할 수 있다.

본 연구에서는 OD와 링크와의 관계를 기존 OD에 대한 통행배분에서 구하지 않고 표본링크이용비를 이용하여 구하기 때문에, 경로기반 비균형 통행배분 개념 하에서 교통량 보존법칙 등을 고려할 필요가 없는 OD추정 해도출 알고리즘을 제시하였다.

분석결과 표본링크이용비는 조사의 정확도가 낮아도 OD 통행량과 링크교통량간의 관계를 나타내는 정보이므로 OD 추정력을 효율적으로 향상시킬 수 있었다. 그리고 기존 OD 오차나 관측교통량 오차에 크게 영향을 받지 않고 안정적인 추정력을 나타내어 교통량이나 OD 행렬이 변화된 곳에서도 비교적 안정적인 OD 추정이 가능하였다.

또한 표본링크이용비를 추가정보로 이용할 경우에는 이용정보간의 정밀도문제가 발생하기 때문에 다른 이용자료와의 정밀도 수준에 따라 추가정보의 조사수준을 설정하여야 한다는 점을 제시하였다. 이러한 결과를 통해 정적 OD 및 동적 OD 추정에서의 정보이용에 대한 시사점을 찾을 수 있다. 동적 OD 추정 등 첨단교통체계관리를 위해서는 다양한 정보가 필요하며, 교통량 검지기, 위치추적 비콘, 영상카메라 등 많은 장치를 통해 이러한 정보를 획득할 수 있다. 그러나 추가정보를 이용하는 경우, 특히 보다 정밀한 정보들을 이용하는 경우에는 정보의 종류와 양이 많아진다고 보다 양호한 의사결정이나 제공정보를 향상

확보할 수 있는 것이 아니며, 자료간의 오차의 민감도 등을 고려한 적절한 자료병합방법론 등에 대한 검토가 선행되어야 할 것이다. 이것은 정보의 수가 많아질수록 추정의 안정성은 높아질 수 있으나, 특정 오차가 낮은 경우에는 추정력이 향상되지는 않을 수도 있다는 점을 나타낸다고도 할 수 있다. 이러한 관점에서 관측교통량을 기본정보로 하는 추정기법은 관측교통량에 대한 관측수준을 정밀하게 하여야 추가정보의 조사수준 및 자료활용성을 높일 수 있다.

마지막으로 링크상의 추가정보는 최적조사위치문제를 고려하여야 하며, 특히 정보의 정밀도 측면에서 볼 때 링크교통량의 최적관측위치문제보다 표본링크이용비의 최적조사위치문제가 추정에 더 중요한 영향을 미칠 수 있는 것으로 파악되었다.

표본링크이용비를 추가정보로 이용한 OD추정과 관련하여 다음과 같은 향후연구가 필요하다. 우선 링크교통량 관측위치문제와 더불어 표본링크이용비 조사지점 선정기법이 필요하다. 둘째, 추가정보와 다른 이용정보와의 상호관계, 특히 이용정보의 정밀도에 의한 모형추정력 변화를 보다 명확히 분석할 필요가 있다. 셋째, 각 링크에서 조사되는 표본링크이용비의 표본을 다양하게 설정하여 모형의 현실적용성을 확보할 필요가 있다. 마지막으로 도시부에서의 표본링크이용비는 예산 제약 및 교통지체 등으로 대상네트워크에서 동시에 조사하기 어렵기 때문에, 상이한 시간대에 조사한 자료들을 결합하는 기법이 필요하다.

참고문헌

1. 백승걸(2001), "유전알고리즘을 이용한 링크관측 교통량으로부터의 기종점 통행행렬 추정", 서울대학교 대학원 박사학위논문.
2. 백승걸·임강원·이승재(1998), "기종점 추정을 위한 최적관측위치로서의 존 코든라인", 대한국토 및도시계획학회지, 제33권 제6호.
3. 백승걸·김현명·임용택·임강원(2000), "링크관측교통량을 이용한 도시부 OD 통행행렬 추정(GA와 SAB 알고리즘의 비교를 중심으로)", 대한교통학회지, 제18권 제6호, 대한교통학회, pp.89~99.
4. 백승걸·김현명·임용택·임강원(2001), "관측교통량을 이용한 다차종 OD 통행량 추정", 대한교통학회지, 제19권 제2호, 대한교통학회, pp.61~72.

5. Carey, M. C. Hendrickson K. Siddharthan (1981), "A Method for Direct Estimation of Origin/Destination Trip Matrices", *Transp. Sci.* pp.32~49.
6. Cascetta E. and Nguyen S.(1988), "A unified framework for estimating or updating OD matrices from traffic counts", *Transpn. Res.* 22B, pp.437~455.
7. Hellinga B. R.(1994), "Estimating dynamic OD estimation demands from link and probe counts", thesis of Ph.D, Queen's University.
8. Hendrickson C., McNeil S.(1984b), "Matrix entry estimation errors, International Symposium on Transportation and Traffic Theory", The Netherlands, Delft, pp.413~430.
9. Hyodo T., Takahashi Y. and Tsuboi T. (2001), "A study on estimation method of table by traffic count data and on-street questionnaire survey", *journal of EASTS* 2001, Vol. 4, No. 3.
10. Kim H. M., Baek S. K., and Lim Y. T (2001), "OD Matrices Estimated with a Genetic Algorithm from link traffic counts", *Transpn. Res. Record* 1771, pp.156~163.
11. Kuwahara M, Sullivan E. C.(1987), "Estimating OD matrices from roadside survey data", *Transpn. Res.* 21B, pp.233~248.
12. Landau U., Hauer E., and Geva I.(1982), "Estimation of cross-cordon OD flows from cordon studies", *Transportation Research Record*, 891, pp.5~10.
13. Lo H. P. and N. Zhang and Lam W. H. K.(1996), "Estimation of and OD matrix with random link choice proportions : a statistical approach", *Transpn. Res.* 30B, pp.309~324.
14. Lo H. P. N. Zhang, W. H. K. Lam(1999), "Decomposition Algorithm for statistical estimation of OD matrix with random link choice proportions from traffic counts", *TR33B*, pp.369~385.
15. Sherali H. D. R. Sivanadan and A. G. Hobeika(1994), "A linear Programming approach for synthesizing OD trip tables from link traffic volumes". *Transpn. Res.* 28B, pp.213~233.
16. Spiess H.(1987), "A Maximum likelihood model for estimating O-D matrices", *Transpn. Res.* 21B, pp.395~412.
17. Tamin O. Z. and Willumsen G.(1989), "Transport demand model estimation from traffic counts", *Transpn.* Vol.16, pp.3~26.
18. Watling D. P.(1994), "Maximum likelihood Estimation of OD matrices from a partial registration plate number", *Transpn. Res.* 28B, pp.289~314.
19. Yang H(1995), "Heuristic algorithms for the bilevel O-D matrix estimation problem", *Transpn. Res.* 29B, pp.231~242.
20. Yang. H(1998), "Optimal traffic counting locations for OD estimation", *Transpn. Res.* 32B, pp.109~126.

✉ 주 작 성 자 : 백승걸
 ✉ 논문투고일 : 2002. 2. 6
 논문심사일 : 2002. 4. 2 (1차)
 2002. 6. 3 (2차)
 2002. 7. 20 (3차)
 심사판정일 : 2002. 7. 20
 ✉ 반론접수기간 : 2002. 12. 31