

# 링크교통정보의 신뢰도를 고려한 링크별 적정 프로브대수 산정

## The Optimum Number of Vehicle Probes With Accounting the Reliability of Link Travel Times

장 정 아

한국전자통신연구원  
텔레매틱스연구단 연구원

최 기 주

아주대학교 환경건설교통공학부  
교수

심 상 우

아주대학교 건설교통공학과  
박사과정

### 목 차

- I. 서론
- II. 이론적 배경
  - 1. 기존연구의 고찰
  - 2. 연구의 방향성
- III. 링크교통정보의 신뢰도 문제
  - 1. 조사개요
  - 2. 조사결과
- 3. 중요한시사점
- IV. 링크별 적정프로브대수 산출
  - 1. 개념
  - 2. 알고리즘
  - 2. 산정결과
- V. 결론 및 향후과제

## I. 서론

차량에 의한 혼잡 문제는 오늘날 가장 어려운 도시 문제로 신호 시스템, 차로의 증설, 지능형 교통시스템(ITS; Intelligent Transportation System) 등으로 그 해결책을 모색하고 있다. 이중 첨단 여행자 정보 시스템(ATIS; Advanced Traveler Information System)은 운전자의 경로와 수단 선택에 있어서 현재의 교통정보를 제공함으로써 그들의 의사결정을 돕는 시스템이다. ATIS에 있어 실시간 정보를 바탕으로 링크 통행시간은 주로 자동차의 경로 계획과 선택 기법에 중요한 역할을 하기 때문에 필수적으로 제공이 된다.

이러한 실시간 링크 통행시간의 정보는 다양한 방법으로 얻어질 수 있다. 고정적인 검지기를 이용하여 정보를 제공받을 수 있고, GPS(Global Positioning System) 수신기를 장착하여 운행하는 차량을 프로브 차량<sup>1)</sup>으로 하여 교통관리센터와의 통신을 함으로서 교통정보를 제공하는 방법이 있다. 이중에서 프로브 차량을 이용한 방법은 앞으로 더욱 발전될 텔레매틱스 기술을 통해 정보의 실시간 송수신이 가능한 시스템으로 쉽게 확장이 되어 교통 정보 수

집 및 제공 체계의 큰 역할을 차지하리라 여겨진다. 연구에서는 먼저 기존의 다양한 프로브대수 산출 알고리즘을 검토하였고, 이를 현실적으로 적용하기 위해 서울시 지역의 차량번호판법 조사를 기반으로 확인된 링크교통정보의 신뢰도 문제를 생각해 보고 있다. 이를 통해 많은 연구자들이 고안하였던 다양한 프로브 대수 산정 알고리즘의 현실적 문제점들을 파악하였다. 본 연구는 이러한 취지에서 링크별 적정 신뢰수준을 만족해야 한다는 제약조건하에 네트워크의 프로브 대수를 산출하는 알고리즘을 개발할 목적으로 한다.

## II. 이론적 배경

### 1. 기존연구의 고찰

통행시간조사는 ITE(Institute of Transportation Engineers)의 매뉴얼에서 제시한 방법론을 따라 적정 신뢰도를 얻기 위한 최소 표본수를 산정하였다. 최소 표본 수는 평균속도, 속도의 평균 범위, 신뢰수준에 따른 허용오차의 함수로 구성이 된다. 이때의 기본 가정은 모집단의 평균  $\mu$ 와 표준편차  $\sigma$ 의 추정치 신뢰성이 있고, 표본의 평균 범위  $\bar{R}$ 와  $\sigma$ 사이의 관계성이 보장된다는 것이다. 이것은 통행시간 조사에서 모든 통행시간이 동일한 교통류와 동일한 환경조건에서 수행이 되어야 한다는 것을 의미한다. 그러나 현실적으로 그러하지 못하다. 이를 기반으로 프로브 차량 대수 산출에 관한 연구가 진행되었다.

A. May(1990)는 통행시간 조사에서 표본 수를50에서

1) 프로브(Probe)차량: Probe라 함은 탐침으로서 광택을 발견하거나 하는데 필요한 장비의 일종이나 ITS에서 언급하는 의미는 통행에 관련한 계반데이터-통행속도 및 시간 등-를 산출하기 위해서 네트워크를 통행하는 차량을 의미함. 본 연구에서 제안하는 Probe차량은 GPS를 장착한 것으로서 주로 차량의 순간속도, 위치정보를 직접적으로 알 수 있고, 통행시간은 부수적으로 알 수 있는 차량을 일컫음.

100개가 적절하다고 하였다. 그러나 각 링크에서의 프로브 차량의 기준은 좀 더 보수적인 척도가 필요하다. 통행시간의 더 큰 값에서 허용할 수 있는 분산은 평균 통행시간의 더 작은 값에서 좀 더 강하다. 더욱이 이러한 지표는 평균 통행시간에서의 표준 편차와 종속성을 무시할 수 있다.

Shrinivasan 과 Jovanis(1996)는 신뢰성 있는 통행시간 정보를 수집하는데 필요한 프로브차량대수와 특정 네트워크상의 적정한 총 프로브차량 대수를 고려한 방법을 제시하였다. 그들은 중심극한 정리를 이용한 통계적 산출 식에 의하여 미리 결정된 허용오차와 신뢰수준을 만족하는 구간당 최소 프로브 차량 대수를 구하였다.

M.Chen 과 S. Chien(2000) 은 미시적 시뮬레이션인 CORSIM 자료를 이용한 고속도로의 통행시간 추정을 위한 프로브 차량대수를 결정하였다. 기존의 연구가 적은 샘플 사이즈 일 때는 적절하지 못한 점을 보완하여 고속도로 링크 통행시간을 추정하기 위한 최소프로브 대수의 산출하였다.

Douglas J. Holdener 와 Shawn M. Turner(1994)의 경우 네트워크에서 프로브 차량을 운영하기 위한(즉 어느 정도 신뢰성을 바탕으로) 최소한의 프로브 대수를 산정하는데 목적이 있다. 물론 이 연구는 일정 지역의 특성을 반영한 내용이기 때문에 우리나라에 직접적으로 적용하기는 어렵다고 보겠다.

적정 프로브의 대수를 산정하는 데에는 두 가지 중요한 입력 변수가 있다. 하나는 travel time variation과 다른 하나는 accuracy requirement이다. 휴스턴의 자체 적인 데이터로 산정된 식은 다음과 같다.

$$n = \frac{t_{n-1}^2 s^2}{E^2}$$

- n = 필요한 sample size
- t<sub>n-1</sub> = n-1의 자유도를 가진 t 분포 값
- s = 표본 표준 편차
- E = 허용오차

위 식에서 S값은 travel speed의 분산계수를 이용하여 구하며, t 분포 값 역시 정규 분포를 이용하여 값을 추정할 수 있다. 이 값 역시 통계학적인 접근 방법을 고려하여 신뢰성을 부여한 것이다

이정희(2001)의 연구에서는 단속류 도시간선도로를 대상으로 실제 현장실험에 의한 통행시간 데이터의 수집하여 한산과 혼잡에 대한 구간 통과 차량들의 속도 편차와 교통량, 구간길이 등의 요인에 대한 관계를 먼저 규명하였다. 이후 계산된 최소 표본수 만큼을 랜덤 추출한 통행속도의 평균과 모평균을 비교하여 최소 표본수 결정방법에 대한 신뢰성을 분석하였다.

## 2. 연구의 방향성

대체로 위의 기존 연구는 모집단의 분포를 가정한 통계

적 접근 방법을 기반으로 적정 프로브 대수를 산정하는 방식을 채택하고 있다. 이때 때론 시뮬레이션적방법이나 휴리스틱한 방법으로 결과를 도출하고 있으며 결과에 대한 검증은 대체로 이루어지지 않았다. 본 연구에서는 이전의 통계적 기반의 연구 성과가 현실적으로 수용가능한지를 판단하여 보고 가능한 알고리즘을 제시하려 한다.

## III. 링크교통정보의 신뢰도 문제

### 1. 조사개요

#### 1) 시공간적 범위

▶ 조사 기간: 2002.12.17~2002.12.20

▶ 조사 시간:

-첨두시간: 7:30~10:30

-비첨두시간: 14:00~14:00

▶ 공간적 범위: 서울시 지역의 편도 4차로 이상의 도로를 대상으로 하였으며 구체적 지역은 <표 1>과 같다.

<표 1> 공간적 범위

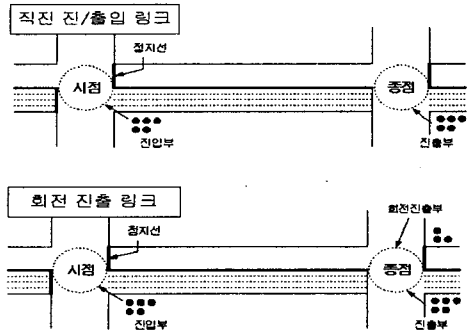
축	From Node	To Node	축	From Node	To Node
강남대로	뽕뽕사거리	서초우성APT	도산대로	청담동사거리	학동사거리
	서초우성APT	뽕뽕사거리		학동사거리	청담동사거리
	서초우성APT	강남역		학동사거리	도산공원
	강남역	서초우성APT		도산공원	학동사거리
	강남역	제일생명사거리		도산공원	안세병원
	제일생명사거리	강남역		안세병원	도산공원
	제일생명사거리	논현역		안세병원	신사역
	논현역	제일생명사거리		신사역	안세병원
종로	종로5가	종로4가	신촌~양화대로	합정역	서교동사거리
	종로4가	종로5가		서교동사거리	합정역
	종로4가	종로3가		서교동사거리	홍대입구사거리
	종로3가	종로4가		홍대입구사거리	서교동사거리
	종로3가	종로2가		홍대입구사거리	동교동삼거리
	종로2가	종로3가		동교동삼거리	홍대입구사거리
	종로2가	종각		동교동삼거리	신촌교회삼거리
종각	종로2가	신촌교회삼거리	동교동삼거리		
도봉로	강북구청사거리	우이교사거리	도봉로	창동시정입구	정의여중입구
	우이교사거리	강북구청사거리		정의여중입구	창동시정입구
	우이교사거리	창동시정입구		정의여중입구	도봉보건소
	창동시정입구	우이교사거리		도봉보건소	정의여중입구

#### 2) 조사방법

실측조사방법은 검지기를 이용하는 방법, 실험차량을 이용하는 방법, 번호판매칭기법 등 여러 가지가 있으나 본 연구에서는 번호판매칭기법(License Plate Matching)을 통하여 통행시간 조사를 실시하였다.

통행시간 실측값은 교통류를 대표할 수 있는 값이어야 하므로 많은 표본을 수집하여야 한다. 따라서 본 연구에서는 다수의 표본을 수집하여 교통류상황을 반영할 수 있는 번호판매칭기법을 이용하였으며, 조사방식은 Manual 방식을 이용하였다. Manual 방식은 조사원이 직접 기록해야 하므로 <그림 1>과 같이 조사원을 (차로수+1)명을 배치하여 각 차로를 통과하는 승용차의 번호판 4자리를 5초마다

조사 Sheet에 기록하였다.



<그림 1> 조사원 배치지점

## 2. 조사결과

### 1) 링크별 프로브 대수 산출식

링크별 프로브 대수 산출식은 Shrinivasan 과 Jovanis(1996) 등 다양한 연구문헌에서 검토한 바 있는 다음의 식을 적용하였다.

$$n_p = \left\lceil \frac{\Phi^{-1}\left(\frac{1+r}{2}\right) \left(\frac{\sigma_p}{\mu_p}\right)}{\epsilon_{max}} \right\rceil$$

- $\mu_p$  = 시간 t동안 링크 l를 이등한 모집단의 통행시간 평균
- $\sigma_p$  = 모집단의 표준편차
- $n_p$  = 필요한 최소 프로브 대수
- $\epsilon_{max}$  = 최대 허용오차
- $r$  = 최대 허용오차보다 작은 오차 절대값을 갖는 시간의 비율
- $\Phi$  = 누적 확률밀도함수

본 연구에서 이를 적용하기 위해서는 다음의 내용이 필요하다.

첫째, 최대 허용오차의 정의 ::교통운명을 위해 요구되는 오차범위(속도: 2-4mph)를 기준으로 합- Manual of Transportation Engineering Studies 속도 5kph 로 가정하였다(통행시간의 오차 = 링크길이/(5kph) 라 함)

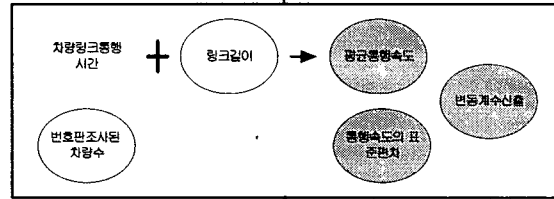
둘째, 최대 허용오차보다 작은 오차 절대 값을 갖는 시간의 비율의 정의(신뢰수준): 신뢰수준90%로 한다.

이러한 가정에 의하면 위의 식은 다음과 같은 식으로 변형된다.

$$n_p = \left\lceil (1.65 * \frac{\sigma_p}{\mu_p} / 1.39)^2 \right\rceil$$

실측현황자료 - 12월 34개 링크

링크ID	링크길이	차로 수	신호횟단	제공 보도수	주기	통행 시간	통행 속도	수집차 량대수	표준편차	최소 필요대수
17412	781	4	1	91	126	6.198413	6	160	4.89125	1
17412	781	4	1	92	125	6.248	11	124.7142857	6.26231386	7
17412	781	4	1	93	112	6.973214	10	83.19899997	9.380781953	6
17412	781	4	1	94	93	7.968819	6	133.6153648	5.876267515	13



$$n_p = \frac{\Phi^{-1}\left(\frac{1+r}{2}\right) \left(\frac{\sigma_p}{\mu_p}\right)}{\epsilon_{max}} \text{ 산출}$$

결과 분석

<그림 2> 프로브대수 분석과정

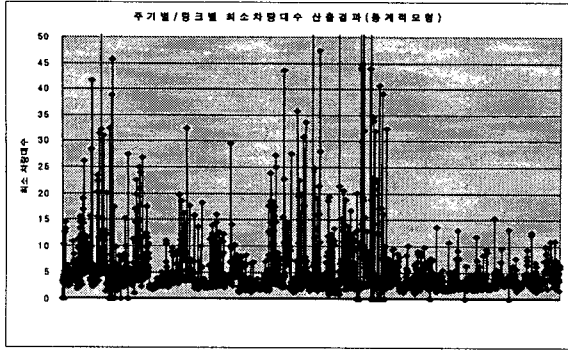
### 2) 산출결과

<그림 2>의 과정을 5분주기별로 모든 링크를 대상으로 차량번호판법에 의한 산출된 통행시간(속도로 환산)을 이용하여 실측표준편차를 이용하여 최소필요대수를 산출하면 <표 2>와 같은 형태로 각 주기별, 링크별로 산출하여 <그림 3>과 <표 3>과 같은 주기별/링크별 최소차량대수가 산출이 되었다.

2대에서 7대 사이의 최소 프로브대수를 요구하는 시간이 전체의 82%정도가 나타남을 알 수 있다. 그 외에 8대 이상 요구하는 시간대도 18%이상이 된다.

<표 2> 실측지역 기초데이터와 통계적모형결과

링크ID	실측지역 기초 데이터							통계적 모형 결과		
	링크 길이	차로 수	신호횟단	제공 보도수	주기	통행 시간	통행 속도	수집차 량대수	표준편차	최소 필요대수
17412	781	4	1	94	124.7	6.3	17	17.7	3.4	
17412	781	4	1	95	83.2	9.4	16	23.2	2.9	
17412	781	4	1	96	130.6	6.0	13	22.7	4.5	
17412	781	4	1	97	160.2	4.9	12	42.2	10.3	
17412	781	4	1	98	141.0	5.5	8	20.8	4.5	
...										
17412	781	4	1	100	170.9	4.6	20	48.4	12.6	



<그림 3> 주기별/링크별 최소차량대수 산출결과

<표 3> 산출된 최소차량대수별 빈도분포

적정(최소)프로브 대수	개수(주기/링크)	적정(최소)프로브 대수	개수(주기/링크)
1	2	17	9
2	242	18	9
3	674	19	9
4	481	20	9
5	293	21	6
6	171	22	2
7	109	23	6
8	81	24	4
9	55	25	2
10	45	26	1
11	40	27	2
12	34	28	3
13	39	29	2
14	18	30	1
15	21	31이상	29
16	14	총합계	2413

### 3. 중요한 시사점

첫째, 기 연구자들이 제시하였던 적정 프로브 대수에 대한 요구조건을 만족하기 위해서는 참 통행시간과 통행시간의 편차를 확인 가능해야 한다.

즉, 시간대별(혹은 주기별로 통행시간의 편차)에 대한 고민이 필요하며 이는 단위 시간(하루, 한 시간, 혹은 한 주기)별로 적정 프로브 대수 산출 값이 달라진다.

둘째, 링크별 혹은 축별로 더 많은 최소 프로브 대수를 요구하는 특성이 서울시 전체에 나타날 수 있을 것으로 사료되며 본 연구에선 5개 축 위주로 분석하여 결론을 도출하였고, 이러한 결론을 일반화시키기는 현재의 정보의 제한으로 어렵다고 판단된다.

셋째, 서울시 전체적으로 특정 시간대에 링크별 적정 프로브 대수를 더 많이 요구하는 시간대가 나타날 수 있으나 본 연구 대상의 링크에서는 오전시간이 보다 오후 시간보다 프로브 대수를 더 많이 필요로 하고 있다.

즉 통계적 분석법에 의한 적정 프로브 대수 산출 알고리즘은 동적인 형태로 그 결과가 시간대별, 링크별 다른 결과를 제시하며 통계적 가정에 기반을 두고 있다. 통계적 모형에서는 통행시간의 분포를 단일 정규분포로 가정(중심극한정리에 의해 모집단이 정규분포로 가정될 수 있음)

근거로 함)하였으나, 실제 통행시간의 분포에서는 쌍봉 혹은 여러 집단으로 통행시간 분포가 나타나고 있어 이러한 기본 가정이 현실적이지 못하다. 또한  $E_{MAX}$  값이라는 허용 오차(구간통행시간 혹은 통행속도의 허용범위)가 혼잡도별, 링크별로 다르기 때문에 이를 고려하기에도 쉽지 않다.

더욱이 실제 시스템에 적용하기 위해서는 이러한 모형을 구현하기 위해서는 실제 통행시간과 통행시간의 표준편차를 알아야 한다는 매우 어려운 과제가 있어 즉 본 연구에서 실측하였던 실제 참 통행시간을 알지 못하고 통계적 모형을 적용하기는 어렵다. 프로브 기반으로 모형을 산출하기에는 변동 계수값(표준편차/평균)을 산출하는 데 그 변이가 매우 커질 수 있다. 따라서 통계적 모형을 적용하기에는 현실적으로 쉽지 않으리라 사료된다.

## IV. 링크별 적정프로브대수 산출

### 1. 개념

본 연구에서 제시하는 알고리즘은 앞에서의 통계적 알고리즘과는 약간의 차이가 있다. 기본적인 개념으로 “향후 프로브차량을 활용하는 시스템이 발전되어 적절한 데이터 베이스가 확보된다면” 각 도로 및 교통데이터들을 근거로 데이터마이닝 기법을 적용하여 프로브 대수를 산출할 수 있다는 가설에서 출발하였다. 즉 <그림 4>와 같이 링크별 변동계수, 링크길이, 차로수, 신호조건등 기타 자료 등이 확보된다면 최소수집프로브 대수가 결정될 수 있다.

수집주기	수집 프로브 대수 산출도(신뢰도)	수집 속도(평균 속도)	변동계수	차로수	기타
1	9	87	41	71	100
2	12	72	46	72	100
3	15	68	53	73	100
...	...	...	...	...	...
5	5	61	49	80	100

수집프로브대수 = (신뢰도, C.V., 링크길이, 차로수, 기타) 링크별 자료

<그림 4> 개념도

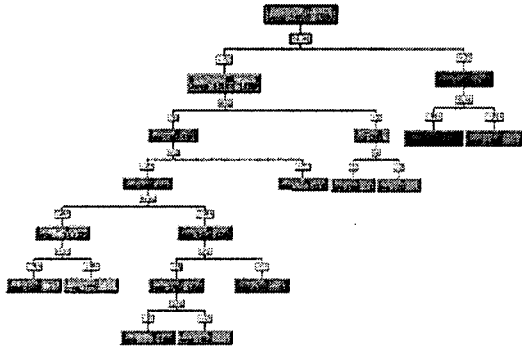
### 2. 알고리즘

전술한 가정을 뒷받침하기 위해 사용한 알고리즘은 데이터마이닝 기법의 하나인 의사결정나무모형에 의한 알고리즘의 사용이다. 의사결정나무 모형은 데이터가 많아질수록 예측력이 우수하다고 알려져 왔기 때문에 향후 링크별 신뢰도에 대한 정의와 수집되는 데이터들이 확장될 경우 모형의 구축 및 적용이 매우 용이하다. 이때 사용한 주요 조건 및 변수는 다음과 같다.

- 조건: 신뢰도가 75-100% 사이

- 목표 변수 : 적정 프로브 대수
- 종속 변수 : 링크 길이, 차로수, 횡단보도 신호의 수, 링크 통행시간

결과로 산출되는 의사 결정 나무 구조는 <그림 5>와 같다.

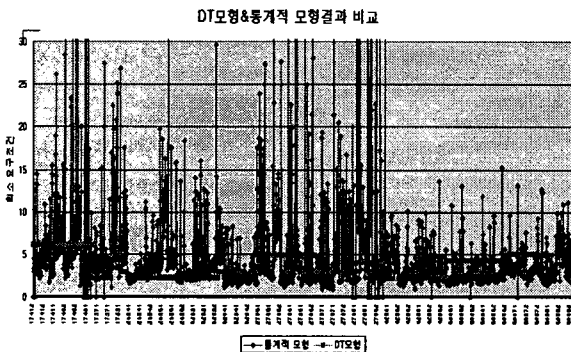


<그림 5> 의사결정나무구조

주요 링크별 최소차량대수와 만족한 주요 링크별/주기 수는 <표 4>와 같다. 이때 통계적 모형결과와 비교한 것은 <그림 6>과 같다.

<표 4> 의사결정나무모형으로 산출된 최소차량대수별 빈도분포

적정(최소)프로브대수	개수(주기/링크)
2	799
3	561
4	137
4.2	196
4.9	295
5	147
6	279
총합계	2413



<그림 6> 모형의 결과 비교

## V. 결론 및 향후과제

링크통행시간 교통정보는 운전자의 경로계획과 선택에 있어 매우 중요한 역할을 한다. 따라서 이는 교통정보체계의 구축에 있어 교통정보의 수집 및 가공이란 단계에 있어 핵심적 역할을 담당한다. 현재 링크통행시간 정보의 추정을 위해서는 GPS 및 AVI(Automotive Vehicle Identification)과 같은 방식의 구간검지체계를 이용한 수집방안이 최근 많이 구현되고 있으며 이러한 프로브를 이용한 방법에서 아직도 해결해야 하는 문제는 교통정보의 신뢰도와 이에 따른 적정 프로브 대수를 유지 계획하는 문제이다.

이러한 프로브 대수 산출에 관한 연구는 크게 두 가지로 나뉜다. 하나는 도시 도로망의 교통정보제공을 위해 요구되는 최소 프로브차량대수 결정에 관한 연구이며, 다른 하나는 교통류 특성에 따라 해당 링크 구간의 신뢰성 있는 정보 산출을 위해 필요한 표본수를 결정하는 연구이다. 본 연구는 이 중 후자에 관한 실증적 연구 결과를 제시하고 있다.

첫째로 그동안 많이 사용하였던 통계적 모형에 대하여 차량번호관법에 의한 실제 통행시간을 근거로 링크별로 적정(최소)프로브대수를 산출하였다.

둘째로 도로/교통조건을 근거로 의사결정나무에 의한 최소차량대수를 산정하였다.

첫 번째 결과는 현실적인 적정(최소)프로브차량 대수를 제시하기는 하였으나, 주기별로(시간대별로) 링크별로 다양한 형태의 결과물을 가져와 현실적으로 시스템에 사용하기 어렵다.

두 번째 결과는 ATIS 및 텔레매틱스 시스템의 확장 등에 용이하게 사용할 수 있을 것이나 검증이 좀더 필요하다.

본 연구는 국내외 실증적 프로브대수산출 연구에서 실증적 연구결과라는 부문에 의의가 있으나, 향후 네트워크 확장성 문제나 시스템 적용을 위해서는 보정 및 추가 검증이 요구된다.

## 참고문헌

아주대학교 교통연구센터(2003), "SK 구간교통정보 보정 알고리즘 개발"

이정희(2001), 교통정보제공을 위한 구간통행시간 산출 방법론 연구 - 적정표본수 결정방법을 중심으로-, 서울시립대학교 석사학위 논문

최기주, 신치현(1998), GPS와 GIS를 이용한 링크통행시간 예측기법, 대한교통학회지 제16권 제2호, pp 197-207

Adolf D. May(1990), Traffic Flow Fundamentals, Prentice Hall

Alan Toppen and Karl Wunderlich(2004), Travel Time Data Collection for Measurement of Advanced Traveller Information Systems Accuracy, ITS America 2004 14th Annual Meeting

FHWA(1998), "Travel Time Data Collection Handbook"

H. Douglas Robertson and Shawn M. Turner(1994) Manual of Transportation Engineering Studies, Prentice Hall

Karthik K. Srinivasan And Paul P. Jovanis, (1996) "Determination of Number of Probe Vehicles Required for Reliable Travel Time Measurement in Urban Network", Transportation Research Record 1517, pp15-22

Mei Chen and Steven I.J. Chein(2000), "Determining the Number of Probe Vehicles for Freeway Travel Estimation Using Microscopic Simulation", Transportation Research

Nagui M.Rouphail(1991) "Travel Time Distribution on Signalized Links : Applications for ADVANCE", Advance Working Paper Series

Piet H.L. Bovy and Remmelt Thijs(2000), "Estimators of Travel Time for Road Networks", Delft University Press