

## 혼합군집분석 기법을 이용한 도로 교통량의 첨두율 산정

김형주\* · 장수은

서울대학교 환경계획학과

### Calculation of the Peak-hour Ratio for Road Traffic Volumes using a Hybrid Clustering Technique

KIM, Hyungjoo\* · CHANG, Justin S.

Department of Environmental Planning, Seoul National University, Seoul 151-050, Korea

#### Abstract

The majority of daily travel demands concentrate at particular time-periods, which causes the difficulties in the travel demand analysis and the corresponding benefit estimation. Thus, it is necessary to consider time-specific traffic characteristics to yield more reliable results. Traditionally, naïve, heuristic, and statistical approaches have been applied to address the peak-hour ratio. In this study, a hybrid clustering model which is one of the statistical methods is applied to calculate the peak-hour ratio and its duration. The 2009 national 24-hour traffic data provided by the Korea institute of Construction Technology are used. The analysis is conducted dividing vehicle types into passenger cars and trucks. For the verification for the usefulness of the methodology, the toll collection system data by the Korea Express Corporation are collected. The result of the research shows lower errors during the off-peak hours and night times and increasing error ratios as the travel distance increases. Since the method proposed can reduce the arbitrariness of analysts and can accommodate the statistical significance test, the model could be considered as a more robust and stable methodology. It is hoped that the result of this paper could contribute to the enhancement of the reliability for the travel demand analysis.

하루 동안 발생하는 교통수요는 대부분 특정 시간대에 집중됨으로써 수요 및 편익 산정에 어려움을 초래한다. 따라서 보다 신뢰성 높은 결과를 산출하기 위해서는 시간대별 특성을 고려할 필요가 있다. 이를 위한 첨두/비첨두의 1시간 통행량으로 환산하는 방법으로는 직관적 방법, 경험적 방법, 통계적 방법 등이 있다. 본 연구에서는 통계적 방법의 일환인 혼합군집분석 기법을 적용하여 첨두/비첨두/심야시간에 대한 지속시간과 집중률을 산정한다. 한국건설기술연구원이 제공하는 2009년 전국 24시간 수시교통량 자료를 이용하였으며, 차종별 특성을 살펴보기 위해 승용차, 트럭, 전차종 등으로 나누어 분석을 실시하였다. 분석결과의 검증에 위해 한국도로공사의 TCS 통행시간 자료를 이용하였다. 검증결과 본 연구결과가 타 연구에 비해 비첨두/심야 시간에는 오차율이 낮으며, 첨두시에는 통행거리가 멀어질수록 오차율이 높아지는 결과를 보였다. 본 연구결과는 임의성을 배제할 수 있으며, 첨두율 추정치에 대한 신뢰성 검증을 수행할 수 있어 보다 안정적인 방법론이라 평가할 수 있을 것이다. 본 연구의 결과가 향후 교통수요 분석의 신뢰성 향상에 일조할 수 있기를 기대한다.

#### Key Words

Peak-Hour Ratio, Road Traffic Volumes, Average Link Method, k-means Clustering, Hybrid Clustering  
첨두율, 도로교통량, 평균연결법, k-평균 군집분석, 혼합군집분석

\* : Corresponding Author  
hkacademy@naver.com, Phone: +82-2-880-5646, Fax: +82-2-871-8847

## I. 서론

교통수요는 대부분 특정 시간대에 집중된다. 이로인해 수요 및 편익 산정에 어려움을 초래한다. 따라서 보다 신뢰성 높은 결과를 산출하기 위해서는 시간대별 특성을 고려할 필요가 있다.

현재 우리나라의 사회기반시설 투자평가제도로는 예비타당성조사와 타당성조사가 있다. 두 투자평가 제도의 표준지침에서는 통행배정을 위한 첨두와 비첨두의 지속시간과 집중률을 지역간 도로와 수도권 지역으로 구분하여 제시하고 있다. 지역간 도로의 경우 첨두 지속시간은 10시간으로, 첨두 1시간 집중률은 7%로 가정하며, 비첨두 지속시간은 9시간으로, 비첨두 1시간 집중률은 2.5%를 가정한다. 그 밖의 심야 시간대는 자유교통류 상태로서 철도 등의 대중교통서비스가 제공되지 않아 사업시행으로 인한 속도 개선 및 수단전환의 효과가 없는 것으로 간주하여 교통분석을 수행하지 않고 있다. 하지만 지역에 따라 교통특성이 서로 다를 수 있음을 언급하며 앞서 제시한 기준과 현저히 차이가 날 경우, 해당 지역의 교통 특성에 적합한 값을 적용하도록 권하고 있다 (한국개발연구원, 2008; 국토해양부, 2009)

본 연구는 위 지침에서 명시한 첨두 지속시간과 집중률에 대한 실증적인 연구이다. 이를 위하여 전국 고속도로, 일반국도, 국지도 및 지방도에서 수집된 24시간 교통량 자료를 이용하여 차량 유형별 첨두 특성을 분석한다. 도로의 위계별로 수집된 교통량을 첨두/비첨두/심야 시간으로 구분하기 위한 방법으로 혼합군집분석 기법 (Hybrid Clustering Technique)을 적용한다. 이를 통하여 해당 시간 동안의 24시간 교통량 대비 1시간 교통량의 비중을 추정한다. 산정된 첨두율을 바탕으로 통행배정 후 추정된 통행시간과 한국도로공사의 TCS 통행시간 자료를 비교한다. 또한 기존방법론의 결과와 비교함으로써 새로운 방법론의 유용성을 검증한다.

## II. 선행연구고찰

통행배정은 첨두 및 비첨두 O/D를 이용하거나, 전일 O/D를 이용하여 수행할 수 있다. 첨두와 비첨두로 나누어 분석할 경우, 현행 KTDB의 O/D 자료가 1일 통행량 기준이므로, 첨두/비첨두의 1시간 교통량으로 환산하여야 한다. 1시간 통행량으로 환산하는 방법으로 첫째, 전

문가의 판단에 기초한 직관적 방법(a naive approach), 둘째, 시행착오(trial and error)를 통한 경험적 방법(a heuristic approach), 셋째, 자료를 토대로 불확실한 사실에 대한 과학적인 판단을 제시하는 통계적 방법(a statistical approach) 등을 고려할 수 있다.

### 1. 국내연구

자료에 대한 전문가의 판단에 근거하여 첨두율을 산정하는 직관적 방법의 대표적인 사례는 현행(예비)타당성조사 표준지침을 들 수 있다. 두 표준지침에서는 지역간 통행의 첨두/비첨두/심야시간의 지속시간과 교통량 비중을 <Table 1>과 같이 제시하고 있다. 이 기준은 도로의 등급과 관계없이 고속국도, 일반국도, 지방도 등에 동일하게 적용되고 있다. <Table 2>와 <Table 3>은 수도권 지역의 첨두 집중률을 나타낸 것으로서, 지역 간 통행에서와 마찬가지로 도로의 등급을 고려하지 않고 동일한 값을 적용한다.

경험적 접근법의 사례로는 점진적 방법(an incremental method)과 유전자 알고리즘(a genetic algorithm)을

<Table 1> Duration and concentration ratio for interregional road traffic volumes

Classification	Duration (hour)	Concentration ratio (%)
Peak	10	7
Off-peak	9	2.5
Night	5	-

Source: Korea Development Institute (2008), p. 281.

<Table 2> Duration and concentration ratio for road traffic volumes of the Seoul metropolitan area

Classification	Duration (hour)	Concentration ratio (%)
Peak	4	8.78
Off-peak	16	3.84
Night	4	-

Source: Metropolitan Transport Association (2007), pp. 248-286.

<Table 3> Duration and concentration ratio for road traffic volumes of the Seoul area

Classification	Duration (hour)	Concentration ratio (%)
Peak	4	8.4
Off-peak	15	4.2
Night	5	-

Source: The Seoul government (2003), pp. 181-234.

〈Table 4〉 National average of the duration and concentration ratio for auto traffic volumes

Classification	Duration (hour)	Concentration ratio (%)
Peak	11 (07:00~18:00)	6.58
Off-peak	8 (05:00~07:00) (18:00~24:00)	3.07
Night	5 (00:00~05:00)	0.62

Source: Sung et al. (2009).

〈Table 5〉 National average of the duration and concentration ratio for truck traffic volumes

Classification	Duration (hour)	Concentration ratio (%)
Peak	11 (07:00~18:00)	6.65
Off-peak	8 (05:00~07:00) (18:00~24:00)	2.99
Night	5 (00:00~05:00)	0.59

Source: Sung et al. (2009).

적용한 예가 있다. 철도투자평가 전면개정 연구(한국교통연구원, 2010)에서는 첨두/비첨두율을 조금씩 증가시키는 점진적 방법을 바탕으로, 통행시간 절감편익을 합리적으로 산정하기 위한 첨두율과 1일 평균 통행시간에 가장 가까운 첨두율을 각각 산정하였다. 통행시간 절감편익을 위한 첨두율은 1일 통행량 보다 큰 값을 배정하여 차량운행비 절감편익과 교통사고 저감편익 등이 과다 추정되는 문제가 발생하였다. 결국 합리적인 편익산정을 위해 첨두/비첨두에 대한 구분 없이, 1일 평균통행시간에 근접하는 첨두율을 점진적 방법을 이용하여 최종 제시하였다. 성홍모 외(2009)의 연구는 선택, 교배, 돌연변이의 세 가지 연산자에 바탕을 둔 유전자 알고리즘을 적용하여 첨두/비첨두/심야시간을 구분한 바 있다. (〈Table 4〉, 〈Table 5〉 참조). 시간대의 구분을 위해 시행착오를 통한 유전자 알고리즘 목적값을 최소화 시키도록 총 네 개의 브레이크 포인트를 찾도록 설정하였다. 두 표준지침과 비교했을 때 지속시간은 유사한 것으로 분석되었으나 첨두율은 차이를 보이고 있다.

마지막으로 통계적 방법에 근거한 정성봉 외(2009)의 연구에서는 평균과 표준편차를 이용하여 시간대별 교통량 구분 방법을 제시하였다. 이 연구는 부산, 대전, 울산에 대한 첨두/비첨두/심야의 지속시간과 집중률을 산정하였다.

## 2. 국외연구

영국의 DMRB(Design Manual for Roads and Bridges) 제 12차 Traffic Appraisal of Roads

Schemes(The Highways Agency, The Scottish Office Development Department, The Welsh Office, The Department of the Environment, 1997)에서는 국내의 두 표준지침과 유사하게 교통량의 변화에 따라 분석시간을 달리해야함을 강조하고 있다. 혼잡이 심한 도시 지역과 도시 인접지역 등은 최소 오전첨두(AM peak), 오후첨두(PM peak), 첨두간(inter-peak)로 나눌 것을 제시하고 있으며, 교통량 변화가 명확하게 구분되는 경우에는 분석시간을 더 세분화하여 변화되는 교통패턴의 반응을 권고하고 있다. 미국의 MTF (Metropolitan Travel Forecasting) Current State of the Practice(Transportation Research Board, 2007)에서는 MPO 지역의 크기에 따라서 분석시간을 달리하고 있다. 작은 지역의 MPO에서는 상대적으로 혼잡이 적어 전일배정을 실시하고 있으며, 반대로 복잡한 MPO 지역에서는 하루(24시간)을 2~5개의 시간대 오전첨두(AM peak), 오후첨두(PM peak), 낮 시간(midday), 저녁(evening), 야간(night)으로 분리하고 있다.

경험적 방법의 경우에는 교통신호운영 분야에서 보다 활발하게 진행되었다. 기존의 신호시간 계획은 집계된 교통량 자료에 의거하여 기술자의 판단에 의존하였으며 변화되는 교통패턴을 효과적으로 반영할 수 없었다. 이에 Park et al.(2004)의 연구에서는 유전자 알고리즘을 통하여 최적의 Time of Day 간격을 산출하였다. 산정된 브레이크 포인트의 검증에 위해 브레이크 포인트의 개수를 조정하여 시뮬레이션을 실시하였으며, 총 6개의 브레이크 포인트가 최적으로 산정되었다. 하지만 연산 수행시간이 긴 유전자 알고리즘의 단점이 나타났다.

통계적 방법의 경우에는 계층적 군집분석을 이용한 Smith et al.(2001)의 연구가 있다. 최적의 시간경계를 결정하는 방법으로 계층적 군집분석을 실시하였으며, 이는 신호시간이 변경되는 과정의 비용분석을 통해 산정되었다. 하지만 최적의 시간 결정에 있어 연구자의 판단에 의지하고 있다.

## 3. 시사점

선행연구검토에서 확인할 수 있듯이 첨두율 산정과 관련한 실증연구는 부족한 실정이다. 첨두율 산정 시 도로위계 및 차종별 통행패턴에 대한 고려도 부족하다. 적용된 방법론 또한 여러 한계를 내포하고 있다. 우선 직관적 방법은 첨두율 산정의 근거가 부족한 임의적인 접근

법이다. 경험적 방법은 직관적 방법에 비해 현실 적합성이 높은 침투율 추정치를 제시할 수 있으나, 추정된 값의 신뢰성 검증을 위한 결정론적 규칙을 제시할 수 없다. 마지막으로 통계적 방법은 임의성을 배제할 수 있고, 침투율 추정치에 대한 신뢰성 검증도 가능하나, 정성봉 외(2009)가 가정한 “평균  $\pm 1$  표준편차”에 의한 침투율 판단은 보편적 동의를 얻기 어렵다. 이에 본 연구는 통계적 방법의 일환인 혼합군집분석 기법을 적용하여 기존 연구의 한계를 완화·개선하고자 한다.

### III. 연구방법론

본 연구에 적용된 방법론은 혼합군집분석 기법이다. 우선 계층적 군집분석의 일환인 평균연결법을 이용하여 적정 군집의 개수를 선정하며, 이어서 비계층적 군집분석인 k-평균 군집분석으로 침투/비침투/심야의 지속시간과 집중률을 산정한다.

#### 1. 군집분석의 개요

군집분석이란 관찰대상인 개체들을 유사한 동류집단으로 분류하는 다변량분석을 말한다. 이는 개체들의 소속집단이 알려진 다음에 소속집단을 결정하는 변수를 탐색하는 판별분석과는 달리, 개체들의 다양한 특성(변수)들 사이의 관계를 분석하여 동류의 집단들로 분류한다. 즉, 각 군집을 대표하는 군집 중심이나 데이터 구조를 이용해 자료의 성격을 파악하는 개체공간의 차원축소를 통한 탐색적 자료분석 방법이라 할 수 있다. 이러한 유사성의 척도에 기초해 집단내의 변량에 대한 집단간의 변량을 최대화시키는 방법에 따라 계층적(hierarchical)군집화 방법과 비계층적(non-hierarchical, disjoint, 또는 k-평균)군집화 방법이 있다(Hair et al., 1987).

#### 2. 계층적 군집분석

계층적 방법은 가까운 개체끼리 차례로 묶거나 멀리 떨어진 개체를 차례로 분리해가는 군집방법으로, 한 번 병합된 개체는 다시 분리되지 않는 것이 특징이다. 이러한 계층적 군집분석에는 단일연결법(single linkage method), 완전 연결법(complete linkage method), 평균연결법(average linkage method), 중심연결법(centroid linkage method), 메디안연결법(median linkage

method), Ward의 방법 등이 있으나, 본 연구는 이 중 가장 보편적으로 적용되는 평균연결법을 이용한다.

평균연결법은 모든 개체들 사이의 평균거리를 기초로, 평균거리가 작은 군집부터 차례로 병합시키는 방법이다. 이 방법을 적용할 경우, 최적의 군집수에 대한 정보를 얻을 수 있는 장점이 있다. 이때, 군집의 변화에 따른 거리의 변화를 덴드로그램에서 확인하여 상대적으로 큰 변화를 나타내거나, 군집의 개수와 통계량을 비교하여 급격한 변화를 보일 때를 최적 군집의 개수로 결정(Kaufman and Rousseeuw, 1990)할 수 있다. 여기서 덴드로그램이란 계층적 군집의 단계를 시각적으로 표현한 것으로 각 단계에서 조합되는 군집과 거리값을 보여준다. 평균연결법의 세부 알고리즘은 다음과 같다.

- [단계1] 맨 처음 단일 개체를 포함하는  $n$ 개 군집과 대칭인  $n \times n$  거리행렬  $d_{ij}$ 를 구성
- [단계2] 최단거리가 되는 두 개 군집 쌍이  $k, l$ 일 때 이 둘 사이의 거리  $d_{kl}$ 을 계산
- [단계3] 두 군집  $k, l$ 을 병합하고 이들로 구성된 군집 ( $kl$ )의 거리행렬에서 두 군집에 대응하는 행과 열을 모두 삭제하고, ( $kl$ )에 대응하는 행과 열에 대해 나머지 군집들과의 거리를 새로 계산
- [단계4] 앞의 단계2와 단계3을 총  $N-1$ 번 반복, 매 회 반복 후 서로 병합된 군집 및 각 병합 단계에서의 수준(거리)을 표시

#### 3. 비계층적 군집분석

비계층적 방법은 최적분리 군집분석이라고도 하는데, 이는 군집의 개수 또는 반경을 이미 정하고, 각 군집의 일부 또는 전부를 판정기준에 의해 재할당하기 때문이다. 이 방법은 계층적 방법과 달리 한 번 분리된 개체도 반복적 시행을 거쳐 재분류할 수 있는 것이 특징이다. 본 연구는 비계층적 군집분석 중 k-평균법을 적용한다. k-평균 군집분석은 먼저  $k$ 개의 군집을 나누고 각 자료들과 선택된  $k$ 개 군집과의 거리를 계산하여 초기 군집을 형성한다. 이렇게 생성된 군집의 중심점과 각 자료들의 거리를 다시 계산한 후 각 자료를 거리가 가장 가까운 군집에 포함시킨다. 이 과정은 기존의 중심점과 새로운 중심점의 차이가 없을 때까지 반복·시행되며 구체적인 알고리즘은 다음과 같다.

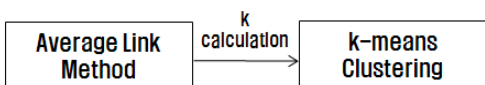
- [단계1] 주어진 표본에서 각 개체들을 초기 k개 군집의 중심으로 선택
- [단계2] 임의로 k개의 점들을 선정하여 초기값(initial value)으로 설정
- [단계3] k개의 초기 중심점(centroid) 선택
- [단계4] 각 개체를 가장 가까운 중심점을 갖는 군집으로 할당 후 새로운 군집의 중심점 계산
- [단계5] 각 개체의 할당에 변화가 없을 때까지 위의 단계를 반복하여 최종적으로 k개 군집 형성

#### 4. 혼합군집분석 모형 설계

본 연구에서는 <Figure 1>과 같이 혼합군집분석 모형을 설계하였다. 여기서 혼합군집분석 모형이란 평균연결법과 k-평균 군집분석의 장점을 활용한 방법론으로 먼저 수집된 도로교통량을 계층적 군집분석인 평균연결법을 이용하여 적정 k군집의 개수를 선정한다. 평균연결법에 사용된 변수는 교통량과 시간으로 군집을 Euclidean 거리에 기반하여 동류집단으로 분류한다. 시간과 교통량만을 변수로 군집분석을 실시하였기 때문에 다른 변수의 영향을 통제할 수 있다. 적정 k군집은 보편적으로 사용되는 덴드로그램과 통계량을 비교하여 산출한다. 산출된 3개의 군집이 첨두, 비첨두, 심야시간을 판단하는 기준으로 각 군집에 포함된 시간과 교통량의 비중을 확인하여 분류를 실시한다.

선정된 k군집을 바탕으로 비계층적 군집분석인 k-평균 군집분석을 이용하여 첨두율을 산정한다. 선정된 k값(k=3,4,5,6)에 따라서 첨두, 비첨두, 심야의 지속시간이 연속적인 결과도 있고, 그렇지 않은 결과도 존재한다. 본 연구에서는 k=3일 때의 결과로 첨두시간과 심야시간은 인접시간들이 연속적으로 분류되었고, 비첨두시간은 비연속적으로 분류되었다.

계층적 군집분석의 단점은 한번 병합된 개체는 다시 분리되지 않지만 비계층적 군집분석은 한 번 분리된 개체도 반복적인 시행을 거쳐 재분류할 수 있다. 이러한 비계층적 군집분석의 장점을 활용하여 최종적인 군집분석의 결과를 도출하고자 혼합군집분석 모형을 설계하였다.



<Figure 1> A design of hybrid clustering

#### IV. 첨두율 분석

본 연구에서는 혼합군집분석을 적용하여 첨두율을 산정한다. 이 결과의 검증에 위해 현행 (예비)타당성조사 표준지침과 성홍모 외(2009)에서 제시한 결과와 비교·분석한다.

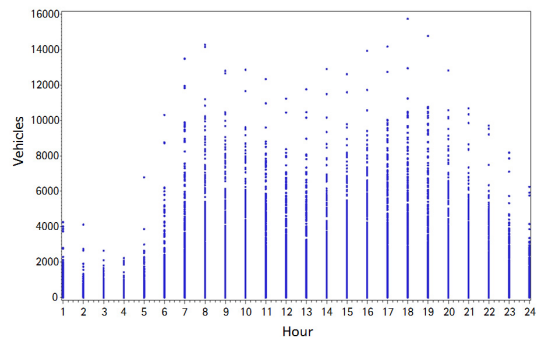
##### 1. 분석자료

본 연구는 한국건설기술연구원이 제공하는 전국의 고속국도, 일반국도, 국지도, 지방도에 대한 24시간 수시조사 교통량 자료를 이용한다. 수시조사란 기본 교통량 자료가 필요하다고 판단되는 모든 구간에 대해서 광범위하게 실시하는 조사로 전체적인 도로 이용 상황을 파악하기 위한 것이다. 수시조사의 목적은 전국의 고속국도, 일반국도, 국지도, 지방도에 대하여 도로의 계획, 건설 및 관리를 위한 기초 자료를 수집함과 동시에 조사 지점의 통과 교통량을 측정하여 평균 일 교통량(AADT)의 산출을 위한 기본 교통량을 구하는데 있다. 본 연구에서는 버스를 제외한 승용차, 트럭, 전차종 등으로 분류하여

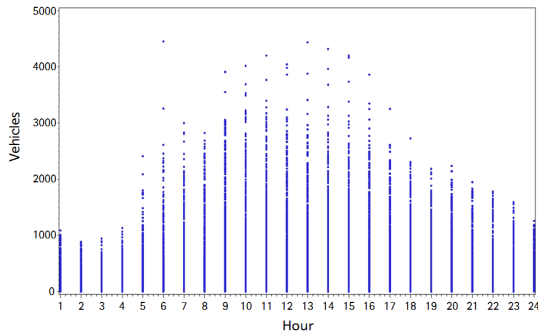
<Table 6> Descriptive statistics of the data

(unit : vehicles/hour)

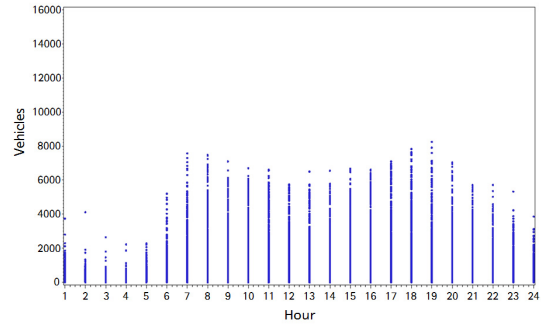
Classification		Average	Standard deviation	Max.	Min.
Expressways	Autos	1511.6	1851.7	11834.0	3.0
	Trucks	597.5	620.1	4454.0	0.0
	Total	2187.3	2468.3	15355.0	3.0
National highways	Autos	275.7	615.0	15747.0	0.0
	Trucks	127.6	184.2	2372.0	0.0
	Total	413.4	785.4	17928.0	0.0
Government-aided provincial roads	Autos	240.2	401.1	4569.0	0.0
	Trucks	98.7	139.5	1559.0	0.0
	Total	350.3	530.2	5574.0	0.0
Provincial roads	Autos	165.4	350.5	5375.0	0.0
	Trucks	70.7	119.1	1695.0	0.0
	Total	243.9	464.3	6732.0	0.0



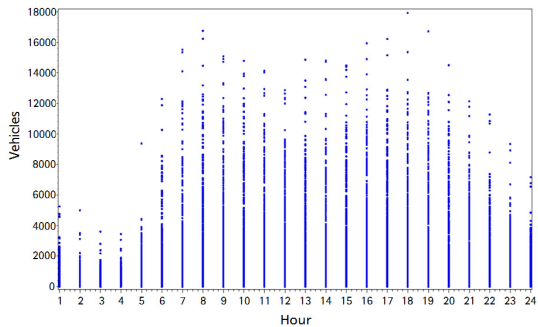
<Figure 2> Temporal distribution of auto traffic volumes



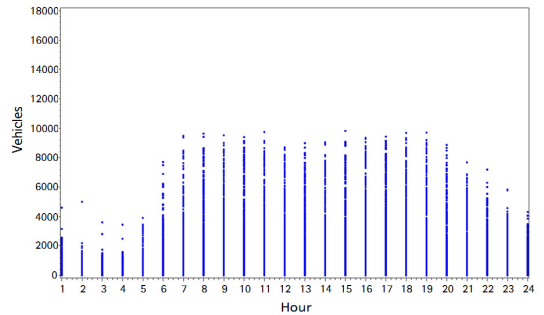
〈Figure 3〉 Temporal distribution of truck traffic volumes



〈Figure 5〉 Temporal distribution of auto traffic volumes after excluding outliers



〈Figure 4〉 Temporal distribution of total traffic volumes



〈Figure 6〉 Temporal distribution of total traffic volumes after excluding outliers

분석을 실시하였으며, 도로유형별 지점 수는 고속국도 451지점, 일반국도 1,171지점, 국지도 338지점, 지방도 1,132지점이다. 버스는 샘플의 수가 적어 유의한 결과를 도출하기 어려울 것으로 판단하여 분석에서 제외하였다.

## 2. 이상치 제거

일반적으로 교통량 자료는 조사과정에서 포함된 오류와 자료의 집계 및 처리과정에서 발생하는 오류를 포함하고 있다. 이러한 데이터의 특성으로 인하여 첨두율을 산정하기 전에 자료의 정확도를 향상시키기 위해서 이상치를 제거하였다. 이상치를 제거하기 위해 중위수 절대편차(MAD, median absolute deviation) 방법을 사용하였으며, 계산된 MAD값이 3 이상인 경우 이상치로 판단(이현석·남궁성, 2009)하고 분석에서 제외시켰다.

$$MAD = \text{median} |x_i - x_{med}|$$

$MAD$  : 중위수 절대편차

$x_i$  : 집계간격 내에서 관측된  $i$ 번째 교통량

$x_{med}$  : 집계간격 내에서 관측된 교통량 중앙값

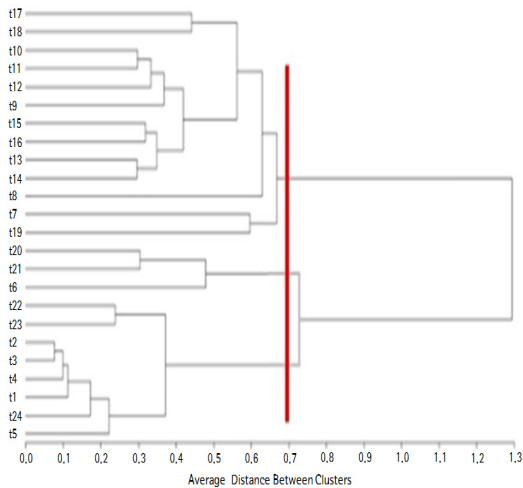
MAD의 방법으로 〈Figure 5〉와 〈Figure 6〉과 같이 승용차와 전차종의 이상치를 제거하였다. 트럭의 경우에는 이상치가 발견되지 않았다.

## 3. 첨두율 산정

교통량 자료의 이상치 제거 후 시간대별 교통량 자료를 바탕으로 첨두/비첨두/심야시간의 지속시간 및 첨두율을 혼합군집분석 기법을 이용하여 산정하였다. 앞서 언급한 바와 같이 비계층적 군집화 방법인  $k$ -평균 군집 분석을 적용하기 위해서는 적절한  $k$ 값의 선정이 필수적이다. 이를 위해 계층적 군집화 방법인 평균연결법을 사용하였다. 적정 군집의 개수 결정에는 덴드로그램과 통계치를 사용하였다. 덴드로그램에서 산정된 군집을 평가하기 위해  $R^2$ 를 비교하여 최종 군집  $k$ 를 산출하였다. 통

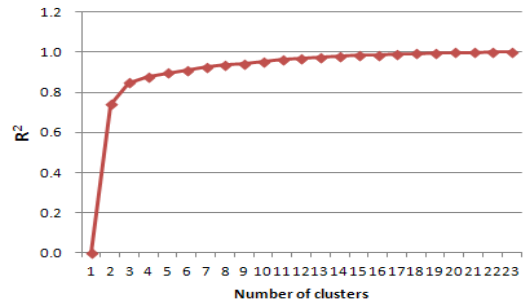
〈Table 7〉  $R^2$  by vehicle type from the average link method

Number of clusters	Autos	Trucks	Total
1	0	0	0
2	0.740342	0.748956	0.740342
3	0.848398	0.829182	0.848398
4	0.877299	0.856287	0.877299
5	0.895757	0.874897	0.895757
6	0.90789	0.888997	0.90789
7	0.926489	0.900546	0.926489
8	0.935417	0.916541	0.935417
9	0.941715	0.926945	0.941715
10	0.95348	0.937147	0.95348
11	0.964921	0.947672	0.964921
12	0.97016	0.955453	0.97016
13	0.975331	0.962161	0.975331
14	0.979612	0.968539	0.979612
15	0.983174	0.974646	0.983174
16	0.986381	0.980707	0.986381
17	0.989479	0.984116	0.989479
18	0.992557	0.991188	0.992557
19	0.995674	0.995725	0.995674
20	0.997614	0.997379	0.997614
21	0.998939	0.99842	0.998939
22	0.999436	0.99928	0.999436
23	0.999817	0.99973	0.999817

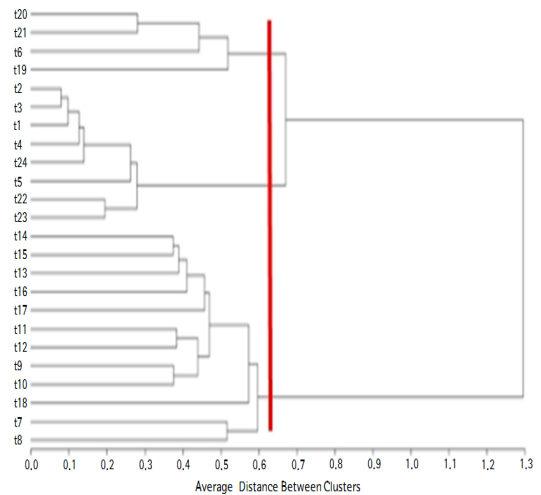


〈Figure 7〉 Dendrogram of auto traffic volumes by the average link method

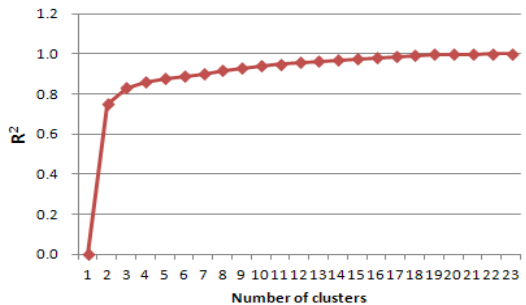
계치에서는 군집의 개수가 증가함에 따라  $R^2$ 의 증가분이 완만하게 되는 지점이 적절한 군집의 후보가 될 수 있다. 〈Table 7〉의 차종별 분석결과를 보면 군집수가 3인 구간에서 이전 군집의 값과 비교해 변화량이 급격히 감소하고 있다.



〈Figure 8〉  $R^2$  by the number of clusters for auto traffic volumes

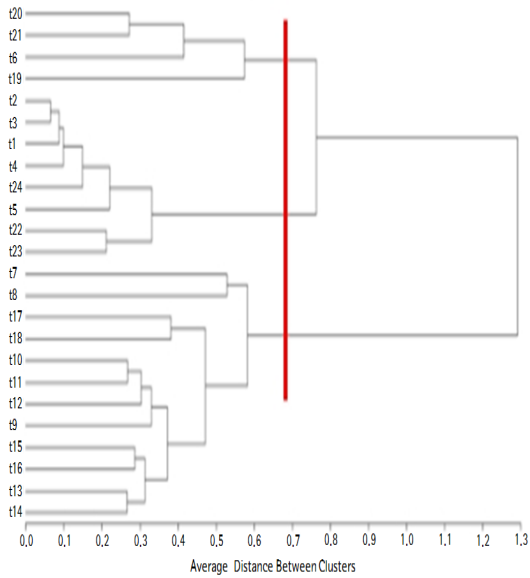


〈Figure 9〉 Dendrogram of truck traffic volumes by the average link method

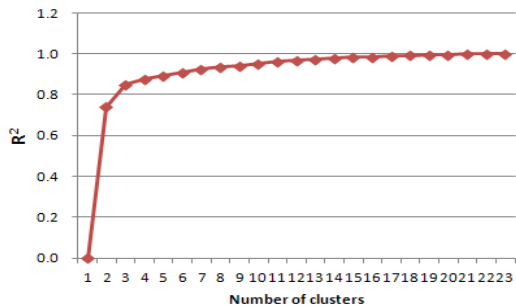


〈Figure 10〉  $R^2$  by the number of clusters for truck traffic volumes

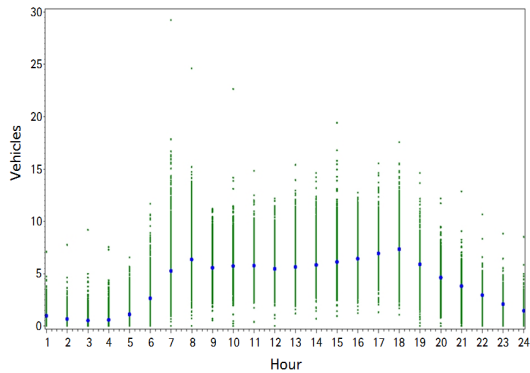
평균연결법을 사용하여 덴드로그램의 유사성(거리)과 산출통계량을 기준으로 군집의 수(k)를 3개로 나누었다. 이어서 산출된 군집수를 바탕으로 비계층적 군집화



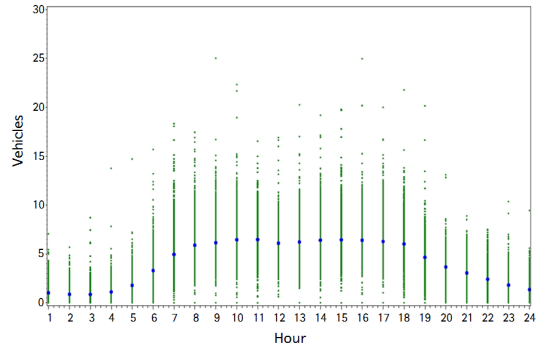
〈Figure 11〉 Dendrogram of total traffic volumes by the average link method



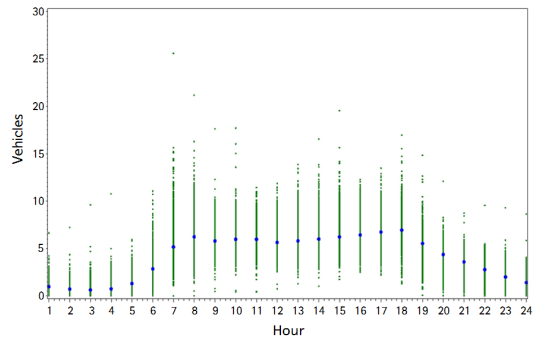
〈Figure 12〉  $R^2$  by the number of clusters for total traffic volumes



〈Figure 13〉 Ratio of one-hour to daily auto traffic volumes



〈Figure 14〉 Ratio of one-hour to daily truck traffic volumes



〈Figure 15〉 Ratio of one-hour to daily total traffic volumes

〈Table 8〉 Duration and concentration ratio for auto traffic volumes by the k-means Clustering

Classification	Duration (hour)	Concentration ratio (%)
Peak	12 (07:00~19:00)	6.05
Off-peak	4 (06:00~07:00) (19:00~22:00)	4.25
Night	8 (22:00~06:00)	1.30

〈Table 9〉 Duration and concentration ratio for truck traffic volumes by the k-means Clustering

Classification	Duration (hour)	Concentration ratio (%)
Peak	11 (08:00~19:00)	6.27
Off-peak	4 (06:00~08:00) (19:00~21:00)	4.16
Night	9 (21:00~06:00)	1.60

〈Table 10〉 Duration and concentration ratio for total traffic volumes by the k-means Clustering

Classification	Duration (hour)	Concentration ratio (%)
Peak	12 (07:00~19:00)	6.08
Off-peak	4 (06:00~07:00) (19:00~22:00)	4.09
Night	8 (22:00~06:00)	1.33



방법인 k-평균 군집분석을 적용하였다. k-평균 군집분석으로 침두/비침두/심야시간의 지속시간 및 집중률을 산정한 결과는 <Figure 13>~<Figure 15> 및 <Table 8>~<Table 10>과 같다.

#### 4. 침두율 검증

본 연구결과의 검증을 위해 현행 (예비)타당성조사 표준지침, 성홍모 외(2009) 등에서 제시한 침두/비침두/심야 지속시간과 집중률을 비교·분석하였다. 표준지침의 경우 지속시간과 차종별 침두율을 제시하고 있지 않다. 그러므로 표준지침의 침두/비침두/심야의 지속시간은 <Table 11>, <Table 12>, <Table 13>과 같이 가정한다. 그리고 성홍모 외(2009) 연구결과는 전차종의 결

과가 제시되지 않아 차종의 대부분을 차지하는 승용차 침두율을 전차종의 결과로 가정하고 비교하였다.

국가교통DB가 제공하는 도로부문 기종점 통행량 자료를 바탕으로 각 연구들의 침두율을 적용하여 통행배정을 실시하였다. 각 연구에서 산정된 통행시간 결과의 오차를 검증을 위해 한국도로공사의 TCS 통행시간 자료를 활용하였다. 한국도로공사 자료는 2009년 10월 한달간의 TCS자료를 구간으로 분석하였다. 거리대별 영향을 확인하기 위하여 4개 기종점(서울발 천안/대전/대구/부

<Table 11> Comparison of peak-hour travel time

Seoul-Cheonan	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	10 (07:00~17:00)	7	49.5	55.4	11.92
Sung et al. (2009)	11 (07:00~18:00)	6.58	49.0	54.3	10.82
This study	12 (07:00~19:00)	6.08	49.3	53.2	7.91
Seoul-Daejeon	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	10 (07:00~17:00)	7	96.3	89.9	-6.65
Sung et al. (2009)	11 (07:00~18:00)	6.58	96.8	88.5	-8.57
This study	12 (07:00~19:00)	6.08	97.1	87.1	-10.30
Seoul-Daegu	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	10 (07:00~17:00)	7	178.4	176.5	-1.07
Sung et al. (2009)	11 (07:00~18:00)	6.58	179.2	174.2	-2.79
This study	12 (07:00~19:00)	6.08	179.2	171.5	-4.30
Seoul-Busan	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	10 (07:00~17:00)	7	284.4	241.5	-15.08
Sung et al. (2009)	11 (07:00~18:00)	6.58	284.1	238.5	-16.05
This study	12 (07:00~19:00)	6.08	283.8	235.2	-17.12

<Table 12> Comparison of off-peak-hour travel time

Seoul-Cheonan	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	9 (05:00~07:00) (17:00~24:00)	2.5	45.1	46.0	2.00
Sung et al. (2009)	8 (05:00~07:00) (18:00~24:00)	3.07	44.2	47.1	6.56
This study	4 (06:00~07:00) (19:00~22:00)	4.09	45.7	49.1	7.44
Seoul-Daejeon	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	9 (05:00~07:00) (17:00~24:00)	2.5	88.3	77.6	-12.12
Sung et al. (2009)	8 (05:00~07:00) (18:00~24:00)	3.07	86.5	78.9	-8.79
This study	4 (06:00~07:00) (19:00~22:00)	4.09	88.5	81.6	-7.80
Seoul-Daegu	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	9 (05:00~07:00) (17:00~24:00)	2.5	168.3	154.0	-8.50
Sung et al. (2009)	8 (05:00~07:00) (18:00~24:00)	3.07	167.4	156.4	-6.57
This study	4 (06:00~07:00) (19:00~22:00)	4.09	169.6	161.2	-4.95
Seoul-Busan	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	10 (07:00~17:00)	7	284.4	241.5	-15.08
Sung et al. (2009)	11 (07:00~18:00)	6.58	284.1	238.5	-16.05
This study	12 (07:00~19:00)	6.08	283.8	235.2	-17.12

<Table 13> Comparison of night-hour travel time

Seoul-Cheonan	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	5 (00:00-05:00)	-	37.6	-	-
Sung et al. (2009)	5 (00:00-05:00)	0.62	37.6	44.5	18.35
This study	8 (22:00-06:00)	1.33	38.3	44.8	16.97
Seoul-Daejeon	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	5 (00:00-05:00)	-	78.0	-	-
Sung et al. (2009)	5 (00:00-05:00)	0.62	78.0	75.5	-3.21
This study	8 (22:00-06:00)	1.33	79.1	76.0	-3.92
Seoul-Daegu	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	5 (00:00-05:00)	-	145.3	-	-
Sung et al. (2009)	5 (00:00-05:00)	0.62	145.3	149.8	3.10
This study	8 (22:00-06:00)	1.33	152.4	150.7	-1.12
Seoul-Busan	Peak hour ratio		Travel time		
	Duration (hour)	Concentration (%)	Observed (min.)	Estimated (min.)	Error (%)
KDI (2008)	5 (00:00-05:00)	-	265.8	-	-
Sung et al. (2009)	5 (00:00-05:00)	0.62	265.8	210.0	-20.99
This study	8 (22:00-06:00)	1.33	260.2	211.1	-18.87

산)간 통행시간을 비교하였다. 또한 TCS 통행시간 자료의 이상치를 제거하는 방법으로 정규분포를 사용하여 “평균 ± 2 표준편차” 구간을 벗어나는 값을 이상치로 판단하였다. 그리고 0 및 음수의 통행시간 값은 검지 오류로 판단하여 제거하였다.

침두율 검증결과를 살펴보면, 침두/비침두/심야의 오차율이 대부분 비슷하게 도출되었다. 침두시에 서울~천안의 경우 본 연구의 오차율이 7.91%로 가장 낮았으며, 비침두의 경우에는 서울~대전, 서울~대구, 서울~부산에서 오차율이 타연구들에 비해서 낮게 산정되었다. 심야의 경우에도 서울~천안, 서울~대구, 서울~부산에서 오차율이 낮았다. 표준지침은 집중률이 제공되지 않으므로 오차율 산정에서 제외시켰다.

본 연구결과는 타 연구와 비교해서 비침두/심야 시간

에는 오차율이 낮으며, 침두시에는 통행거리가 길어질수록 오차율이 커지는 결과를 보였다. 하지만 전반적으로 보았을 때 검증을 실시한 연구결과가 모두 비슷한 경향의 오차율을 나타내고 있다. 다만 본 연구결과는 타연구들에 비해 임의성을 배제할 수 있고 침두율 추정치에 대한 신뢰성 검증이 가능하므로 보다 안정적인 연구결과라 평가할 수 있을 것이다.

## V. 결론

본 연구에서는 혼합군집분석 기법을 활용하여 침두/비침두/심야시간에 대한 지속시간과 집중률을 산정하였다. 2009년 전국 24시간 수시교통량 자료를 바탕으로 분석을 수행하였으며, 차종별 특성을 살펴보기 위해서 승용차, 트럭, 전차종 등으로 나누어 분석을 실시하였다. 산정된 결과를 승용차 및 트럭의 경우 성홍모 외(2009)의 연구결과와 비교하였고, 전차종은 표준지침에서 제시하고 있는 결과와 비교하였다. 분석결과의 검증을 위해 표준지침, 성홍모 외(2009), 그리고 본 연구에서 제시한 침두율을 적용하여 통행배정을 실시하였다. 산정된 통행시간 결과의 오차율 검증을 위해 한국도로공사의 TCS 통행시간 자료를 이용하였다. 검증의 결과 분석 구간별로 모든 연구가 비슷한 오차율을 나타내는 가운데 본 연구결과가 타 연구에 비해 비침두/심야 시간에는 오차율이 낮으며, 침두시에는 거리가 멀어질수록 오차율이 커지는 결과를 보였다.

본 연구는 통계적 방법에 근거하여 침두/비침두/심야의 지속시간과 집중률 산정방법을 제시하였다. 이는 선행연구들의 직관적이고 경험적인 접근법에 비해 임의성을 배제할 수 있으며, 침두율 추정치에 대한 신뢰성 검증을 수행할 수 있어 보다 안정적인 방법론이라 평가할 수 있을 것이다.

향후 연구과제로 버스(2종)의 침두/비침두/심야의 지속시간과 집중률을 산정해볼 필요가 있다. 또한 5대 지방광역권 중 24시간 교통량 자료의 수집이 가능한 지역을 대상으로 본 연구의 방법론을 적용하여 침두율을 분석할 필요가 있다. 이는 지방광역권 교통수업을 추진함에 있어 교통수요 분석결과의 신뢰성 향상에 기여할 수 있을 것으로 판단된다.

알림 : 본 논문은 서울대학교 제14회 교통연구세미나(2011.4.23)에서 발표된 내용을 수정·보완하여 작성된 것입니다.

## REFERENCES

1. Chung, S., S. Kim, and J. Kim(2009), "The Korea Transport Institute", Transportation Research, Vol.16, No.3, The Korea Transport Institute, pp.41~50.
2. DO, Myungsik·LEE, Hyang Mee·NAMKOONG, Seong(2008), "Outlier Filtering and Missing Data Imputation Algorithm using TCS Data", Journal of Korean Society of Transportation, Vol.26, No.4, Korean Society of Transportation, pp.241~250.
3. Huh, M., K. Song, and D. Kang(2002), "Implementation and Efficiency Comparison of Clustering Algorithms", The Korean Journal of Applied Statistics, Vol.10, SungKyunKwan University Research Institute of Applied Statistics, pp.92~106.
4. Hair J. F., R. E. Anderson, and R. L. Tatham(1987), "Multivariate Data Analysis with Readings", Macmillan, pp.149~161.
5. Korea Development Institute(2008), "The study on revision and supplement of standard guideline for preliminary feasibility study on road and railroad projects(the 5th edition)".
6. Kaufman, L., P. J. Rousseeuw(1990), "Finding Groups in Data: An Introduction to Cluster Analysis", Wiley Series in Probability and Mathematical Statistics.
7. Lee, H., S. Namkoong(2009), "A Development of Preprocessing Models of Toll Collection System Data for Travel Time Estimation", The Journal of The Korea Institute of Intelligent Transport Systems, Vol.8, No.5, Intelligent Transport Systems, pp.1~11.
8. Ministry of Land, Transport and Maritime Affairs(2009), "Guideline for investment assessment of transport facilities(the 3rd edition)".
9. Metropolitan Transportation Authority(2007), "2006 Seoul Metropolitan Household Travel Survey Data", Metropolitan Transportation Authority, Seoul, Incheon, Gyeonggi-do.
10. MacQueen, J.(1967), "Some Methods for Classification and Analysis of Multivariate Observations, Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability", Vol.1, pp.281~297.
11. Park, B., P. Santra, I. Yun, D. Lee(2004), "Optimization Time-of-Day Breakpoints for Better Traffic Signal Control", Transportation Research Record, No.1867, pp.217~223.
12. Seoul(2003), "2002 Seoul Household Travel Survey Data", Seoul.
13. SAS Institute Inc(1988), "SAS/STAT User's Guide: Release 6.03 edition".
14. Sung, H., I. Yun, M. Park, C. Kim(2009), "Analysis of Hourly Variation for the Highway Traffics in Korea", Transportation Research, Vol.16, No.2, The Korea Transport Institute, pp.1~11.
15. Smith, B. L., W. T. Scherer, and T. A. Hauser(2001), "Data-mining tools for the support of signal-timing plan development", Transportation Research Record: Journal of the Transportation Research Board, No.1768, pp.141~147.
16. Smith, B. L., W. T. Scherer, T. A. Hauser, B. Park(2002), "Data-Driven Methodology for Signal Timing Plan Development: A Computational Approach", Computer-Aided Civil and Infrastructure Engineering, Vol.17, Issue.6, pp.387~395.
17. The Korea Transport Institute(2010), "The study on revision of entire guideline for railroad investment assessment", Korea Rail Network Authority.
18. The Highways Agency, The Scottish Office Development Department, The Welsh Office, The Department of the Environment

- (1997), "Design Manual for Roads and Bridges", Vol.12, Traffic Appraisal of Roads Schemes.
19. Transportation Research Board(2007), "Metropolitan Travel Forecasting", Current State of the Practice.
20. Wang, X., W. Cottrell, and S. Mu(2005), "Using K-means clustering to identify time-of-day break points for traffic signal timing plans", Proc. 8th International IEEE Conference on Intelligent Transportation Systems, Vienna, Austria, pp.519~524.

✉ 주 작성 자 : 김형주  
✉ 교 신 저 자 : 김형주  
✉ 논문투고일 : 2011. 7. 25  
✉ 논문심사일 : 2011. 8. 27 (1차)  
2011. 11. 15 (2차)  
2011. 12. 2 (3차)  
✉ 심사판정일 : 2011. 12. 2  
✉ 반론접수기한 : 2012. 6. 30  
✉ 3인 익명 심사필  
✉ 1인 abstract 교정필