

교통상황 분류를 위한 클러스터링 기법 개발

Development of Traffic State Classification Technique

강우진* · 김영호**

* 주저자 : 서울시립대학교 교통공학과 박사수료

** 교신저자 : 한국교통연구원 모빌리티전환연구본부 본부장

Woojin Kang* · Youngho Kim**

* Ph.D. Candidate, Dept. of Transportation Engineering, Univ. of Seoul

** Chief Director, Dept. of Mobility Transformation, Korea Transport Institute

† Corresponding author : Youngho Kim, ykim@koti.re.kr

Vol. 22 No.1(2023)
February, 2023
pp.81~92

pISSN 1738-0774
eISSN 2384-1729
<https://doi.org/10.12815/kits.2023.22.1.81>

Received 5 December 2022
Revised 19 December 2022
Accepted 30 December 2022

© 2023. The Korea Institute of
Intelligent Transport Systems. All
rights reserved.

요약

교통상황 분류는 신호연동그룹 단위의 정주기식 제어 기법을 효율적으로 적용하기 위하여 TOD 계획을 수립하는데 핵심적인 기술이다. 본 논문에서는 신호연동그룹에 속하는 모든 교차로의 교통 자료 즉, 속도-교통량-밀도를 활용할 수 있는 딥 임베디드 클러스터링(Deep-Embedded Clustering:DEC) 기반 교통상황 분류 방법론을 제시하였다. 기존의 신호계획의 경우 교통량 기반으로 주요 교차로를 선정하고 해당 교차로의 교통자료를 이용하여 단편적인 신호계획을 하였으나, 본 논문에서 제시된 방법론의 경우 신호연동 그룹내 다수 교차로의 종합적인 교통특성에 따라 교통상황을 유연하게 분류하여 신호계획을 할 수 있는 기반을 제공하였다. 본 연구에서 제시된 방법론은 일반적인 군집화 방법론이 입력 자료의 차원이 증가함에 따라 겪는 차원의 저주(Curse of dimensionality) 문제를 완화함으로써 신호연동그룹에 속하는 모든 신호교차로의 교통자료를 고려한 신호시간 계획 수립이 가능하며 기존의 특정교차로 및 교통량만을 이용한 교통상황 분류 방법론의 단점을 극복할 수 있음을 보였다.

핵심어 : 정주기식 교통신호제어, 신호시간계획, 교통상황, 딥임베디드 클러스터링

ABSTRACT

Traffic state classification is crucial for time-of-day (TOD) traffic signal control. This paper proposed a traffic state classification technique applying Deep-Embedded Clustering (DEC) method that uses a high dimensional traffic data observed at all signalized intersections in a traffic signal control sub area (SA). So far, signal timing plan has been determined based on the traffic data observed at the critical intersection in SA. The current method has a limitation that it cannot consider a comprehensive traffic situation in SA. The proposed method alleviates the curse of dimensionality and turns out to overcome the shortcomings of the current signal timing plan.

Key words : Time-of-day traffic signal control, Signal timing plan, Traffic state, Deep-Embedded Clustering

I. 서 론

세계의 도시화율이 2050년에 약 70%로 증가하고 도시 인구가 64억명에 달할 것으로 예측되고 있다 (UN-Habitat, 2020). 인구가 지속적으로 증가한 도시에서 승용차 통행거리도 35조km까지 늘어날 것으로 예상되는데 이렇게 증가된 이동 수요와 이동 거리로 인하여 도시의 교통 혼잡은 더욱 심각해질 것으로 예상된다. 도시부의 교통혼잡은 주변 교차로에 비하여 상대적으로 많은 교통량이 집중되는 주요 교차로(Critical intersection)를 중심으로 퍼져나가는데, 이러한 교통혼잡을 일정 수준 이하로 제어하기 위하여 지역의 교통운영기관은 다양한 교통신호 제어전략 기법을 각 지역의 교통상황을 반영하여 적용하고 있다.

첨단 정보통신기술을 활용하여 취득된 실시간 교통자료를 기반으로 교통운영을 최적화할 수 있는 교통신호 제어전략이 이론적으로는 개발되어 있으나, 현재 세계적으로 가장 많이 활용되고 있는 교통신호 제어전략은 정주기식 (TOD : Time-Of-Day) 제어 기법과 교통 감응식 (Traffic-responsive) 제어 기법이다 (Tang et al., 2019). 우리나라에서는 일부 지자체에서 교통감응식 제어 기법을 적용한 바 있으나 유지관리에 소요되는 많은 비용을 감당하기 어려워 적용 된지 얼마 지나지 않아 대부분 정주기식 제어 기법으로 전환하였다. 현재 우리나라에는 99% 이상의 교통신호제어기가 정주기식으로 운영되고 있다.

도심 교통혼잡은 네트워크의 구조와 신호 운영방식에 따라 다양하게 전파되기 때문에, 교통상황은 하나의 지점에서 관측된 교통자료 뿐만 아니라 폭넓은 지역에서 관측된 자료를 기반으로 판단하는 것이 합리적이다. 교통 운영의 대상 지역을 세분하여 부분별로 독립된(Isolated) 교통제어 전략을 적용하는 것보다 넓은 대상 지역의 교통상황에 따라 유기적으로 조정된 (Coordinated) 교통제어 전략을 적용하는 것이 효과적이지만, 넓은 지역의 교통자료를 기반으로 최적화된 교통제어전략을 수립하는 것은 매우 어려운 작업이다. 우리나라에서는 주요 교통축을 중심으로 신호최적화 및 연동화에 따른 최적신호제어를 위해 신호연동그룹(Signal sub area) 단위로 교차로를 구분하여 정주기식 제어 기법을 적용하고 있다. 신호연동그룹 단위의 정주기식 제어 기법을 효율적으로 적용하기 위하여 신호연동그룹 내에서 관측되는 실제 교통자료를 이용하여 교통 상황을 정확히 분류하고 상황별로 신호시간계획 (Signal timing plan)을 수립하는 방안이 시도되고 있다 (Lee et al., 2011).

신호시간계획은 시간대별, 주간과 주말, 휴일 등에 대하여 다양하게 제시될 수 있으나, 신호시간계획에 대한 지침이 별도로 마련되어 있지 않아 현장에서는 오전첨두, 오후첨두, 비첨두 등과 같이 단순화하여 사용하고 있다. 또한, 교통상황 분류를 위한 교통자료는 신호연동그룹 내의 모든 교차로 자료를 활용하는 것이 바람직하나 모든 교차로에서 관측된 교통자료를 활용하여 상황을 분류하는 것이 어렵기 때문에 주요 교차로에서 관측된 교통 자료만을 이용하여 신호시간계획을 수립한다. 따라서 주요 교차로에서는 동일한 교통상황이 관측되지만 주요 교차로의 상하류에서 상이한 교통상황이 관측될 경우 각 상황에 적합한 신호시간계획이 수립되지 못하는 것이 현실이다.

본 논문에서는 TOD 패턴별 최적 신호시간계획을 수립할 수 있도록 신호연동그룹 내의 모든 교차로별 교통 자료를 활용하여 신호연동그룹의 교통상황을 저비용으로 효율적으로 패턴화할 수 있는 방법론을 제시하고자 한다. 본 논문은 제2장에서 현재 사용되고 있는 군집화 기법과 관련된 선행연구를 검토하고, 제3장에서는 새로운 군집화방법론인 Deep-Embedded Clustering (DEC) 기법을 소개한다. 제4장에서는 본 논문에서 소개된 DEC 기법을 마이크로 시뮬레이션 (AIMSUN)으로 구축한 자료에 적용하여 새로운 군집화 방법론의 현실 적용 가능성을 보였고, 제5장에서는 본 논문을 통하여 도출한 시사점을 정리하고 향후 연구과제를 제시하였다.

II. 선행연구 검토

현재까지 신호 제어에 활용하기 위해 TOD 시간 계획을 구분하는 다양한 군집화 방법론들이 제시되었다. Jeong et al.(2008)은 독립교차로를 기준으로 분할 군집화 방법론 (Point assignment clustering) 중 대표적으로 사용되는 K-means 군집분석을 활용해 TOD의 시간 계획을 분류하는 방법을 제안했다. 최소전이시간을 적용하여 VISSIM 모형을 사용해 군집 분석을 조정했고, 운영적인 측면에서 활용할 수 있는 TOD 시간 계획을 제안해 교차로의 지체를 감소시켰다. Lee et al.(2011)도 K-means 군집 분석에 VPLUSKO라는 새로운 지표를 활용해 교통량으로 군집을 분석하였다. 그리고 현장적용을 통해 교차로의 교통 정체를 완화시키는 성과를 얻었다. K-means는 각 군집의 초기 중심점이 정해지면 개별 데이터는 초기 중심점과 가장 가까운 군집으로 할당한다. 군집의 중심점을 업데이트하고 다시 개별 데이터를 가장 가까운 군집으로 할당하는 과정을 반복적으로 수행하는데 중심점의 변화가 더 이상 발생하지 않으면 군집화가 종료된다. K-means는 직관적이고 구현이 쉽고 수렴성이 보장된다는 장점이 있는 반면에 차원이 높은 데이터를 활용하기 어렵다는 단점이 존재한다. 고차원으로 갈수록 데이터 간 거리가 가까워져 클러스터링의 효과가 없어지기 때문이다. 따라서, 본 논문에서와 같이 많은 수의 입력자료를 이용할 경우 군집화의 성능이 저하되는 단점이 존재한다. Hwang et al.(2021)은 최적의 TOD 시간 분할을 위해 연속된 교차로에서 수집된 교통량과 속도 데이터에 시계열 데이터의 군집 분석을 위한 동적 타임 워핑 모델을 적용하였다. 시간 분할을 위해 활용된 데이터별 군집의 특성을 분석하여 최적의 신호 운영 시나리오를 구성하기 위한 시간 분할 방법론을 제안하였다. 계층적 군집 분석과 계층적 군집 분석에서 활용되는 Euclidean Distance 기반이 아닌 시계열 군집 간의 거리를 데이터의 일부 확장 또는 압축시키는 DTW basic 거리 기반의 DTW 모델을 활용해 TOD 시간을 분할하는 방법을 사용했다.

분할 군집화 이외에도 현장에서 가장 많이 쓰이는 방법론으로는 계층적 군집화 (Hierarchical clustering), 밀도기반 군집화(DBSCAN: Density-Based Spatial Clustering of Application with Noise)와 가우시안 혼합모형 (Gaussian mixture model) 등이 있다. 계층적 군집화는 거리가 가까운 데이터를 계층화된 그룹으로 묶어주는 방법이다 (Nielsen, 2016). 계층적 군집화는 데이터간 거리(Distance metric)를 구한 뒤 결합과정을 거쳐 덴드로그램(Dendrogram)형태로 나타낸다. 거리를 나타내는 방법으로는 유클리드 거리(Euclidean distance), 맨하탄 거리(Manhattan distance), 피어슨 상관 거리(Pearson correlation distance) 등이 사용된다. 계층적 군집화는 군집의 개수를 사전에 정의하지 않고 도출된 덴드로그램에 가상의 선을 적용했을 때 나누어진 그룹의 개수로 군집의 개수를 결정한다.

밀도기반 군집화는 노이즈가 있는 대규모의 데이터에 적용할 수 있는 군집화 방법론이다. 밀도기반 군집화는 군집의 개수를 미리 지정할 필요가 없고 이상치를 효과적으로 제외하여 집단과 충분히 가깝지 않은 데이터는 배제할 수 있다. 또한, 비선형 경계 군집화가 가능하다는 장점 때문에 많이 사용되고 있다. 그러나 밀도기반 군집화는 데이터를 사용하는 순서에 따라 군집화 결과가 상이하게 나타나고 고차원의 데이터를 처리하기 어렵다는 한계가 있다 (Ester et al., 1996).

가우시안 혼합모형은 대상 데이터 집합이 다수의 가우시안 분포를 가진 데이터 집합이 혼합되어 생성되었다는 가정하에 도출된 군집화 방법론이다. 데이터의 분포를 결정하는 모수를 추정하기 위하여 E-step과 M-step을 교대로 반복적으로 사용하는 EM(Expectation Maximization) 알고리즘을 적용한다. 가우시안 혼합모형은 데이터들이 하나의 클러스터에 속할 확률을 나타내기 때문에 군집을 결정론적으로 정하는 K-means에 비하여 유연하며 클러스터의 분산 구조를 반영하여 정확한 클러스터링을 할 수 있다는 장점이 있다. 그러나 초기값에 민감하고 추정할 매개변수가 많아 계산 시간이 길게 소요되고 고차원 데이터를 처리하는데 한계가 있다(Scikit-Learn, 2022).

이상에서 살펴본 분할 군집화, 계층적 군집화, 밀도기반 군집화, 가우시안혼합 모형은 데이터 셋의 특성과 군집화의 목적에 따라 선별적으로 적용된다. 이러한 군집화 방법들이 많은 문제에 효과적으로 적용되고 있으나 차원이 높은 데이터를 처리하는데 한계가 있다는 점이 공통적으로 지적되고 있다.

교통상황은 시시각각으로 지역에 따라 많은 변화가 관측되는 복잡한 현상이다. 이것을 효율적으로 구분하여 상황별로 적합한 제어전략을 구사하는 것이 중요하다. 도심내에서 발생하는 교통혼잡을 제어하기 위하여 신호연동그룹 단위로 신호 제어전략을 도출한다. 하나의 신호연동그룹에는 다수의 신호교차로가 포함되고, 하나의 신호교차로에는 4개의 접근로가 존재하며 하나의 접근로에는 직진, 좌회전, 우회전과 같이 3개의 방향별 속도, 교통량, 밀도(점유율) 자료가 존재한다. 즉 하나의 신호교차로에는 $4 \times 3 \times 3 = 36$ 개의 데이터가 도출될 수 있다. 신호연동그룹에 포함되는 신호교차로의 수에 비례하여 데이터 수가 증가하고, 흔히 말하는 차원의 저주 (Curse of dimensionality) 문제 때문에 일반적으로 사용되는 군집화 방법론을 적용하기에는 많은 한계가 존재한다. 본 연구에서는 활용 자료의 개수가 많은 경우에도 교통상황을 효과적으로 군집화할 수 있는 방법론을 개발하고자 한다.

Ⅲ. 연구방법론

1. 새로운 군집화 방법론 적용

본 연구에서는 하나의 신호연동그룹에 속하는 다수의 교차로로부터 관측된 교통자료를 기반으로 딥러닝을 이용한 군집화 방법론인 Deep-Embedded Clustering (DEC)을 이용하여 교통상황을 분류하였다. 본 연구에서 적용한 DEC는 앞에서 소개한 분할 군집화, 계층적 군집화, 밀도기반 군집화, 가우시안혼합 모형을 직접 적용하기 어려운 고차원 입력값을 활용하는 군집화 방법론이다. 각 교차로에서 수집된 다차원 자료를 정해진 시간 간격으로 집계하여 교차로 (y축)와 교차로의 교통변수 측정값(x축) 축으로 정렬하여 2차원 이미지로 형상화한 후 딥러닝 기법 중의 하나인 Autoencoder를 적용하여 이미지를 축소하고 복원하는 과정에서 도출된 저차원의 잠재변수 (Latent variables)를 대상으로 K-means를 적용하는 군집화 방법이다 (Xie et al., 2016).

일반적인 군집화 방법이 입력데이터를 직접 활용하는 반면에 DEC 모형은 입력데이터를 심층신경망으로 표현된 비선형 매핑을 통하여 입력변수보다 훨씬 저차원의 잠재속성공간 (Latent variable space)으로 변환한다. 변환된 잠재속성공간에서 생성된 속성값들을 기반으로 군집화를 반복적으로 수행하면서 군집들의 중심값 (Cluster center)과 비선형매핑의 매개변수 값들이 학습된다.

DEC는 오토인코더 모형의 매개변수 초기화와 최적화로 구성된다. 매개변수 초기화는 심층오토인코더의 학습을 통하여 이루어지고, 매개변수 최적화는 보조목표분포 (Auxiliary target distribution)와 KL 다이버전스 (Kullback-Leibler divergence)의 반복적인 계산을 수행하면서 이루어진다. 이 과정에 대한 구체적인 설명은 다음과 같다.

2. 매개변수 초기화

DEC의 매개변수 초기화에 사용되는 심층오토인코더는 다층구조로 이루어져 있다. 심층오토인코더의 각 층은 다음 수식으로 표현되는 디노이즈 오토인코더로 구성되어 있다.

$$\begin{aligned} \tilde{x} &\sim \text{Dropout}(x) \dots\dots\dots (1) \\ h &= g_1(W_1\tilde{x} + b_1) \dots\dots\dots (2) \\ \tilde{h} &\sim \text{Dropout}(h) \dots\dots\dots (3) \\ y &= g_2(W_2\tilde{h} + b_2) \dots\dots\dots (4) \end{aligned}$$

여기서 Dropout은 입력 값을 일정 비율에 따라 무작위적으로 0으로 변환하는 함수이고, g_1 과 g_2 는 활성화 함수(Activation function)이다. 학습은 입력변수 x 와 출력값 y 의 차이인 $\|x - y\|$ 를 최소화하는 과정이다. 탐욕 (Greedy) 알고리즘을 순방향으로 적용하여 학습된 매개변수들은 역방향으로 재현 오차 (Reconstruction)를 최소화하도록 다시 한번 학습하여 최종적으로 결정된다. 이 과정을 통하여 도출된 오토인코더의 Encoder Layer와 Decoder layer 중에서 Encoder layer가 DEC 모형에 사용된다. <Fig. 1>과 같이 Encoder layer를 통과한 속성변수를 K-means를 적용하여 k개의 군집으로 분류한다.

3. 매개변수 최적화

비선형매핑의 매개변수와 군집의 중심에 대한 초기값이 정해지면 비지도 학습을 통해 2가지 과정을 반복적으로 수행함으로써 군집을 보다 정교하게 분류할 수 있다. 첫 번째 과정은 군집의 중심값을 기준으로 잠재속성공간의 값을 가장 근접한 군집으로 일차적으로 할당하는 것이다. 두 번째 과정은 비선형 매핑의 매개변수를 업데이트하고 보조목표분포를 활용하여 군집화의 신뢰도를 향상 시킬 수 있도록 군집의 중심값을 조정하는 것이다. 이러한 과정은 수렴 기준을 만족할 때까지 반복적으로 수행된다. 본 논문에서는 매개변수를 최적화하기 위한 이터레이션의 error 임계값은 0.001설정하였다.

1) 일차 할당 (Soft assignment)

잠재속성변수값 z_i 를 군집 j 에 할당할 확률 (Soft assignment)인 q_{ij} 를 계산하기 위하여 잠재속성변수값과 군집의 중심값간의 유사성을 측정할 수 있도록 다음과 같은 Student's t-Distribution을 계산한다.

$$q_{ij} = \frac{(1 + \|z_i - \mu_j\|^2)^{-1}}{\sum_j (1 + \|z_i - \mu_j\|^2)^{-1}} \dots\dots\dots (5)$$

2) KL 다이버전스 최소화 (KL divergence minimization)

일차 할당을 통해 도출된 군집을 보조목표분포에 적합되도록 지속적으로 조정한다. 이를 위하여 일차로 할당된 군집의 분포 q_i 와 보조목표분포 p_i 간 KL 다이버전스 손실을 다음과 같이 정의한다.

$$L = KL(P \parallel Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \dots\dots\dots (6)$$

여기서 목표분포인 P의 정의가 DEC 알고리즘의 성능을 결정한다. 본 연구에서는 아래와 같이 목표분포를 정의하였다. 여기서 $f_j = \sum_i q_{ij}$ 이고 군집 j 의 빈도수를 의미한다.

$$p_{ij} = \frac{\frac{q_{ij}^2}{f_j}}{\sum_{j'} \frac{q_{ij'}^2}{f_{j'}}} \dots\dots\dots (7)$$

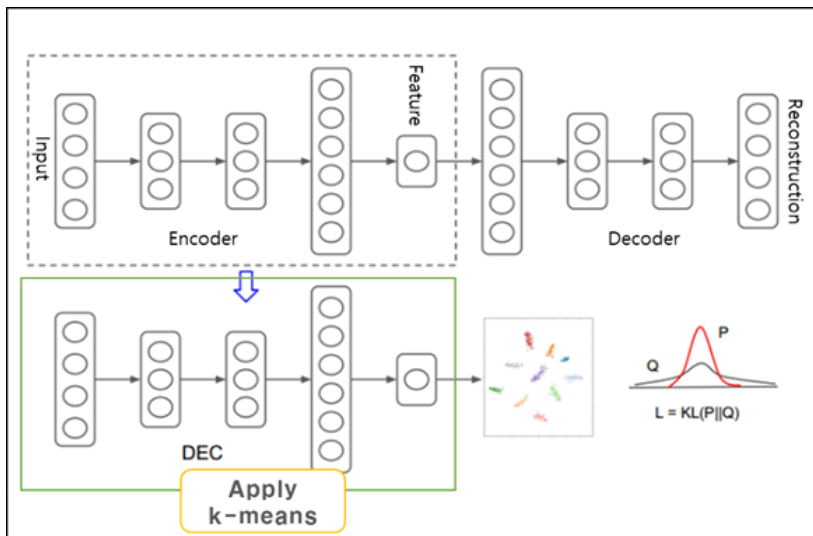
3) 최적화 (Optimization)

Stochastic Gradient Descent (SDG)를 활용하여 비선형매핑의 매개변수와 군집의 중심을 최적화한다. KL 다이버전스 손실 L 의 잠재속성공간의 값(z_i)과 군집이 중심(μ_j)에 대한 그래디언트는 다음과 같이 계산된다.

$$\frac{\partial L}{\partial z_i} = \frac{\alpha + 1}{\alpha} \sum_j \left(1 + \frac{\|z_i - \mu_j\|^2}{\alpha}\right)^{-1} \times (p_{ij} - q_{ij})(z_i - \mu_j) \dots\dots\dots (8)$$

$$\frac{\partial L}{\partial \mu_j} = -\frac{\alpha + 1}{\alpha} \sum_i \left(1 + \frac{\|z_i - \mu_j\|^2}{\alpha}\right)^{-1} \times (p_{ij} - q_{ij})(z_i - \mu_j) \dots\dots\dots (9)$$

위의 그래디언트 값은 비선형매핑의 매개변수 그래디언트 값을 계산하는 Backpropagation에 적용되고, 군집 중심값의 변화가 기준치 이하로 감소할 경우 최적화가 종료된다.



<Fig. 1> Concept of the DEC model

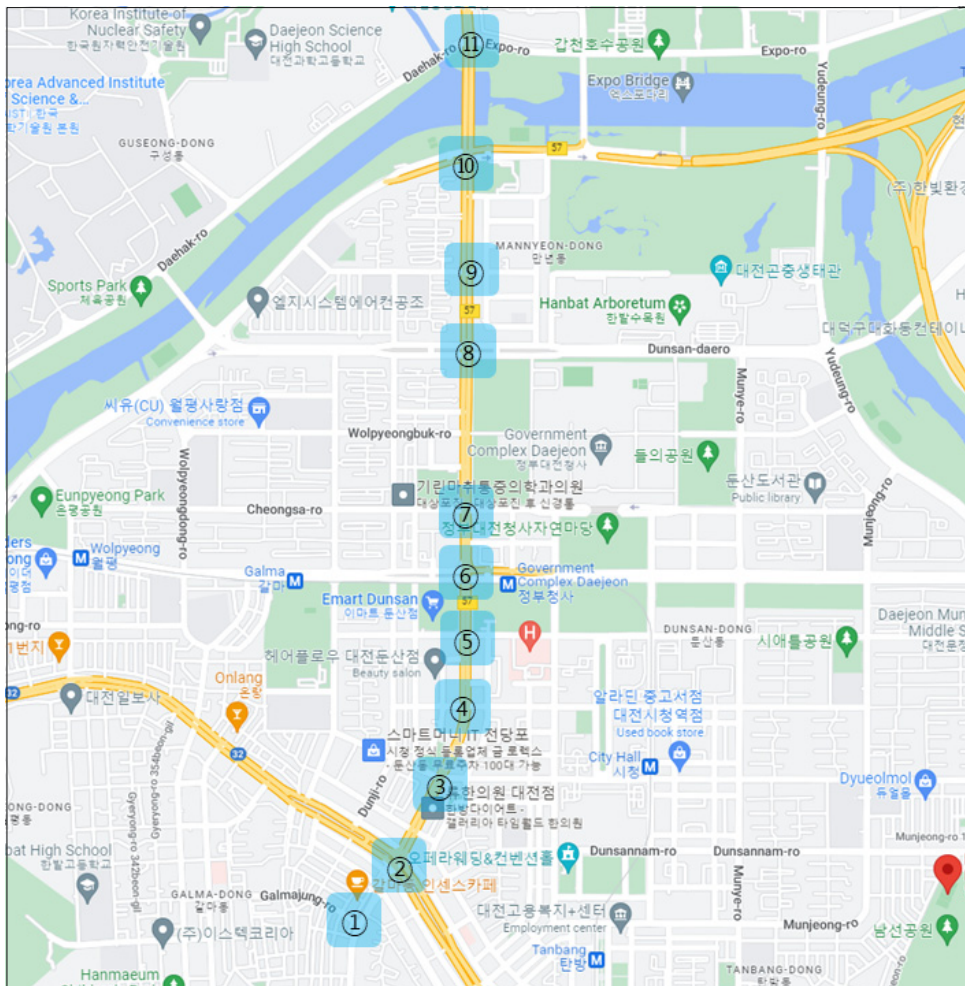
IV. 적용 사례

본 연구에서는 대전시의 실제 도시부 신호연동그룹(SA)을 마이크로 시뮬레이션툴 (AIMSUN)로 구축하였고, 시뮬레이션 틀에 의하여 생성된 교통자료에 DEC 모형을 적용하여 군집화를 수행하였다. DEC 모형 적용 구간은 대전시 대덕대교 네거리 및 정부청사역 네거리, 큰마을 네거리를 지나는 대덕대로의 일부 구간이다.

총 연장은 약 3.2km이며 총 11개의 교차로로 구성되어 있으며, 대상지는 대전의 중심을 남북으로 가로지르는 간선도로로 다른 주간선도로 두 개와 만나는 등 대전의 중심이 되는 도로중 하나이다. 2019년 8월 ~ 9월 동안 12개의 교차로 중 11개의 교차로에 카메라를 설치하여 교통량을 관측하였고 AIMSUN의 OD 교통량자료로 사용하였다.

1. DEC의 입력자료

AIMSUN을 통하여 생성된 DEC 입력자료는 속도, 교통량, 밀도 자료이다. 대덕대로 44개 검지기 (11개 교차로 4개 방향)에서 2019년 8월 18일 (일)부터 9월 21일(토)까지 06:00부터 21:00까지 10분 단위로 속도, 교통량, 밀도를 생성하였다. 10분단위 데이터를 총 5주간 생성하여 총 3,360개의 행 데이터와 각 교차로 진입부에 설치된 총 44개의 검지기를 통해 수집한 속도, 교통량, 밀도 데이터를 6Set씩 중첩시켜 생성한 총 792개의 열 데이터로 구성된 DEC 모형의 최종입력 데이터로 다음과 같이 생성하였다.



<Fig. 2> Daejeok-daero, Test site and 11 intersections

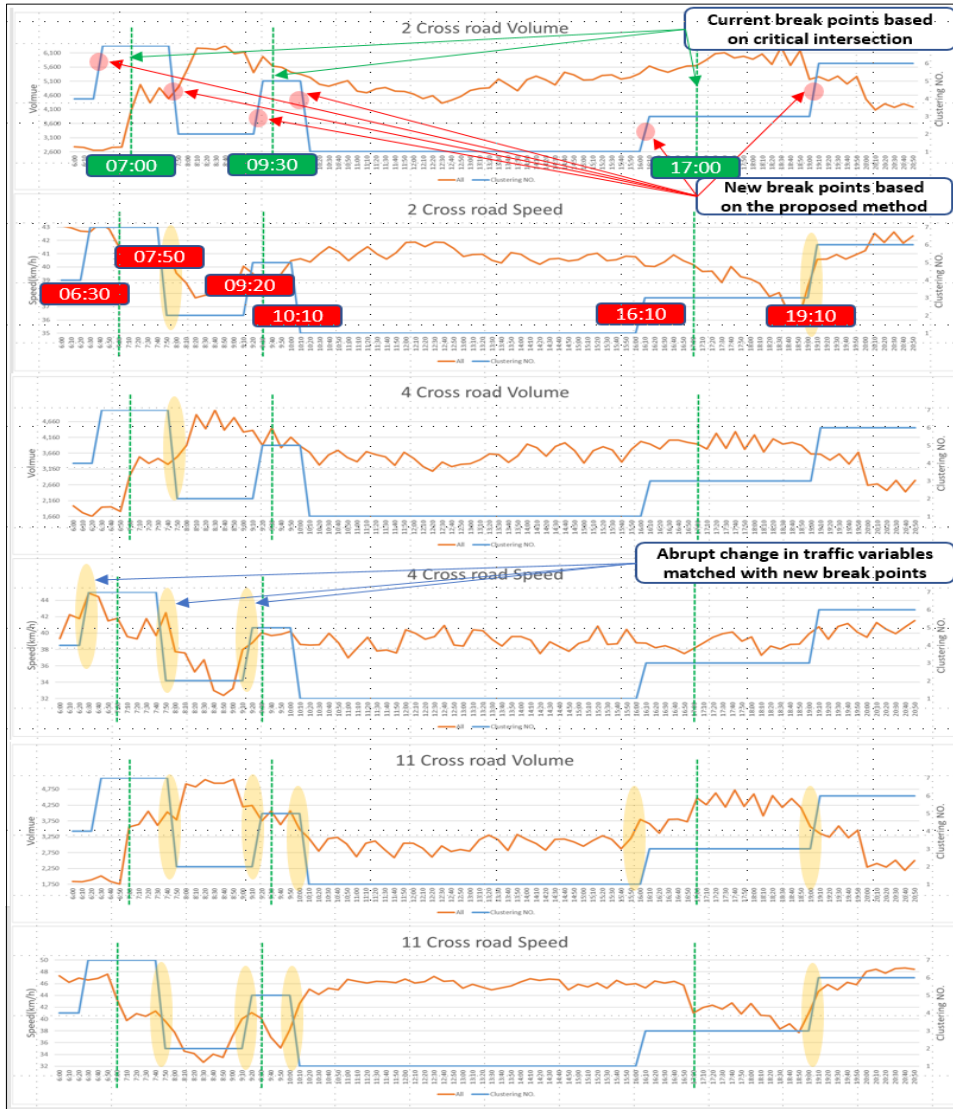
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	Date	Weekday	Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	20190818	1	6:00	0.5497	0.5921	0.6909	0.8152	0.8464	0.7088	0.6415	0.8361	0.7391	0.7255	0.7941	0.6485	0.7092	0.6418	0.6562	0.538	0.8157	0.6199	0.6952
3	20190818	1	6:10	0.5353	0.5562	0.7263	0.8128	0.8459	0.6381	0.5895	0.8362	0.7042	0.7212	0.8079	0.5745	0.7652	0.6114	0.6514	0.5505	0.8052	0.5316	0.7002
4	20190818	1	6:20	0.5534	0.5329	0.6948	0.8076	0.8461	0.6577	0.6437	0.8371	0.7047	0.7959	0.8344	0.714	0.6945	0.6815	0.8821	0.5817	0.7717	0.6231	0.6954
5	20190818	1	6:30	0.5094	0.5967	0.7204	0.8226	0.8568	0.6631	0.662	0.817	0.7124	0.7504	0.7645	0.6117	0.6524	0.6641	0.8269	0.5687	0.8136	0.6725	0.6231
6	20190818	1	6:40	0.5582	0.637	0.7535	0.8269	0.8672	0.6239	0.6058	0.8515	0.7026	0.7516	0.8179	0.6653	0.7458	0.5354	0.8461	0.7045	0.7818	0.6479	0.6487
7	20190818	1	6:50	0.5555	0.5615	0.76	0.81	0.8493	0.6956	0.6391	0.8267	0.6325	0.7759	0.7943	0.53	0.6492	0.6087	0.8607	0.5383	0.7567	0.5685	0.5716
8	20190818	1	7:00	0.5078	0.5414	0.6543	0.8045	0.8194	0.6666	0.6282	0.8219	0.7293	0.7314	0.816	0.6885	0.7031	0.6892	0.8709	0.6364	0.7509	0.6586	0.7016
9	20190818	1	7:10	0.5058	0.5828	0.7119	0.8223	0.8619	0.6861	0.6628	0.8528	0.7055	0.7244	0.7596	0.5905	0.6034	0.5957	0.8258	0.6335	0.7058	0.7376	0.6682
10	20190818	1	7:20	0.5574	0.5316	0.7061	0.8352	0.8583	0.6863	0.6429	0.8355	0.7106	0.7504	0.7949	0.5383	0.6311	0.5975	0.8055	0.6255	0.7532	0.6647	0.6007
11	20190818	1	7:30	0.5343	0.5754	0.762	0.8314	0.8666	0.6613	0.6044	0.81	0.6682	0.7365	0.8045	0.6177	0.6325	0.5993	0.8858	0.6594	0.768	0.6226	0.6409
12	20190818	1	7:40	0.5174	0.5418	0.7213	0.8299	0.8429	0.6146	0.6319	0.8245	0.6869	0.74	0.7992	0.5828	0.6893	0.6988	0.8699	0.586	0.7479	0.6605	0.701
13	20190818	1	7:50	0.4959	0.5861	0.7037	0.8229	0.8392	0.6756	0.6468	0.81	0.69	0.7044	0.7564	0.6181	0.5082	0.7031	0.8627	0.653	0.7274	0.6703	0.6834
14	20190818	1	8:00	0.537	0.5825	0.7124	0.8176	0.8391	0.6642	0.5786	0.8231	0.7093	0.7285	0.7762	0.5823	0.4857	0.698	0.8068	0.6514	0.6361	0.662	0.6807
15	20190818	1	8:10	0.5837	0.5804	0.7024	0.7935	0.8269	0.6482	0.5935	0.7997	0.6442	0.6908	0.7759	0.5122	0.4492	0.6681	0.8192	0.6399	0.5949	0.7216	0.6244
16	20190818	1	8:20	0.5169	0.5569	0.7365	0.7927	0.8515	0.6631	0.6009	0.785	0.706	0.7244	0.7644	0.6305	0.4796	0.6855	0.8071	0.6218	0.5847	0.7125	0.7008
17	20190818	1	8:30	0.507	0.5415	0.7312	0.8277	0.868	0.6589	0.6039	0.8186	0.6925	0.7831	0.5383	0.6312	0.5777	0.6841	0.8286	0.638	0.618	0.6588	0.6841
18	20190818	1	8:40	0.5014	0.5327	0.6839	0.7884	0.8288	0.653	0.6025	0.7778	0.694	0.7074	0.7796	0.5394	0.4566	0.6679	0.8127	0.5881	0.6093	0.7204	0.7308
19	20190818	1	8:50	0.5262	0.5673	0.6954	0.7901	0.8357	0.6124	0.5837	0.808	0.6277	0.6541	0.7616	0.6421	0.4874	0.6322	0.8136	0.6049	0.602	0.6943	0.7053
20	20190818	1	9:00	0.5671	0.5641	0.7024	0.769	0.8131	0.6669	0.6023	0.8131	0.6949	0.6711	0.7424	0.5887	0.5156	0.6841	0.8296	0.6725	0.6409	0.674	0.6444
21	20190818	1	9:10	0.5685	0.5706	0.7052	0.7604								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
22	20190818	1	9:20	0.5901	0.5446	0.6888	0.7762								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
23	20190818	1	9:30	0.6057	0.5641	0.6986	0.7946								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
24	20190818	1	9:40	0.5897	0.5515	0.7093	0.7925								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
25	20190818	1	9:50	0.5938	0.5781	0.7362	0.769								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
26	20190818	1	1:00	0.6183	0.5817	0.7189	0.7519								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
27	20190818	1	1:10	0.6269	0.5826	0.7152	0.7893								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
28	20190818	1	1:20	0.6127	0.5532	0.6999	0.7405								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
29	20190818	1	1:30	0.6082	0.573	0.6845	0.75								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
30	20190818	1	1:40	0.6317	0.5727	0.7111	0.7948								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
31	20190818	1	1:50	0.6241	0.5754	0.6993	0.7528								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
32	20190818	1	1:100	0.6362	0.5399	0.7045	0.7637								0.871	0.482	0.657	0.8565	0.6508	0.6377	0.6951	0.6445
33	20190818	1	1:110	0.6186	0.5858	0.6967	0.7581	0.7517	0.665	0.6261	0.8007	0.6111	0.5663	0.7876	0.6159	0.6833	0.6802	0.8037	0.6991	0.6962	0.6369	0.6714
34	20190818	1	1:120	0.6209	0.5682	0.7039	0.7645	0.7405	0.6607	0.6501	0.8016	0.5908	0.6066	0.777	0.5997	0.6845	0.6762	0.8034	0.6896	0.6805	0.6604	0.7149
35	20190818	1	1:130	0.6093	0.5802	0.6826	0.7645	0.7252	0.6679	0.6373	0.7831	0.5908	0.5575	0.7839	0.5625	0.6588	0.6573	0.7888	0.6517	0.6882	0.6386	0.6935

<Fig. 3> DEC Model input data sets

2. DEC모형을 통한 교통상황 분류 결과

현장에서 정주기식 신호제어전략을 적용할 경우, 하나의 신호연동그룹으로 묶인 군집교차로 내에서 교통량이 가장 많은 주요교차로 (Critical intersection)의 교통량을 토대로 TOD 계획을 수립하고 있다. 문헌 고찰에서와 같이 주요교차로의 교통량이 급격하게 변화하는 시점을 기준(Break point)으로 교통상황을 분류한다. 본 논문의 대상구간인 대전시 대덕대로의 경우 평일 TOD는 새벽시간(0:00~07:00), 출근시간(07:00~09:30), 낮시간(09:30~17:00), 퇴근시간(17:00~21:00), 심야시간(21:00~24:00)과 같이 5개의 시간으로 구분하여 운영되고 있다. 토요일 TOD는 새벽(0:00 ~07:00), 출근시간(07:00~09:30), 오전시간(09:30~11:30), 오후시간 (11:30~21:00), 심야시간 (21:00~24:00)이다. 일요일 TOD는 새벽(0:00~09:30), 전일시간(09:30~21:00), 심야시간(21:00~24:00)이며, 공휴일은 일요일과 동일하다.

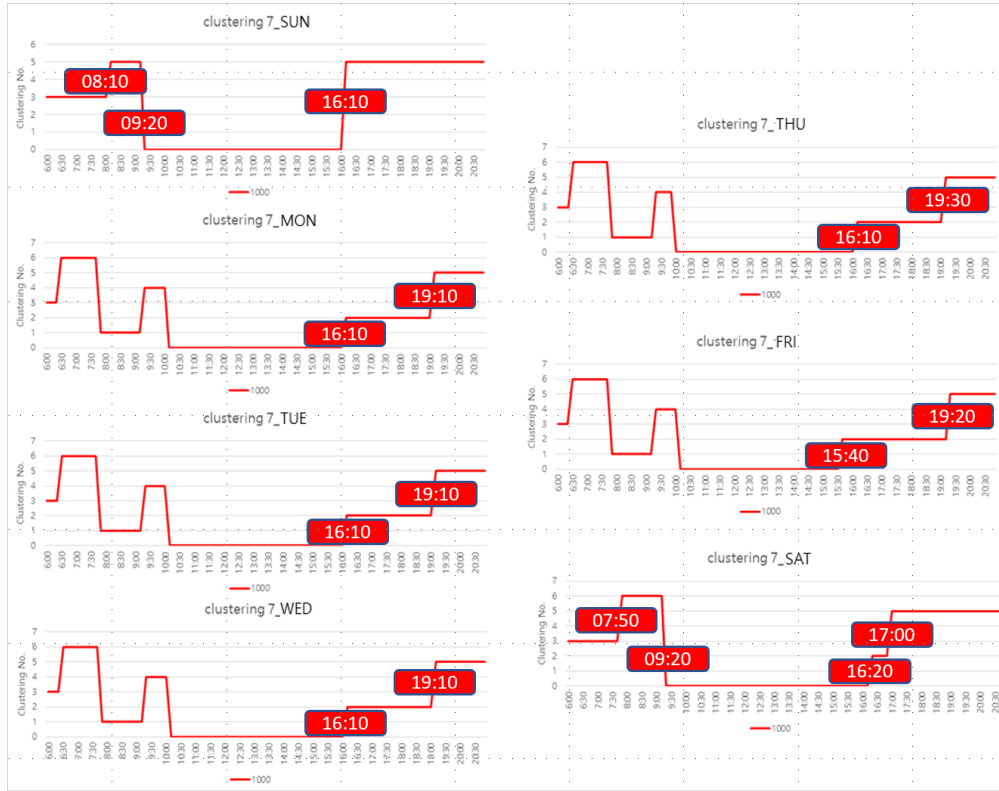
본 논문에서는 DEC 모형을 이용하여 신호연동그룹내의 모든 교차로에서 관측된 교통량과 속도, 밀도 자료를 활용하여 교통상황을 평일과 공휴일에 관계없이 7개의 군집으로 구분하였다. 본 연구의 DEC 모형을 활용하여 06:00~20:00까지의 교통상황을 분류한 결과는 기존의 Break point와는 다른 양상을 보인다. 2019년 8월26일(월요일) 기준으로 06:00~21:00 동안 기존 방식에 의하면 4개의 교통상황으로 구분되지만, 본 논문에서 제안한 방법에 따르면 7개 교통상황으로 구분되었다. 동일한 교통상황이라고 생각되는 10분단위 시간대별로 1~7번까지 Clustering 번호로 결과값이 생성되었다.



<Fig. 4> Comparison of break points from the current and the proposed method

<Fig. 4>에서 주요교차로 (2번 교차로)를 기준으로 지정된 Break point와 본 논문의 DEC 모형으로 제시된 Break point는 상이하였다. 기존 방법론에 의한 교통상황 구분은 주요교차로에서 출근시간의 시작 시각 (07:00)에 교통량이 급격히 증가하는 것을 잘 표현하고 있으나, 출근시간의 종료시각 (09:30)과 퇴근 시간의 시작시각 (17:00)은 교통량의 급격한 변화를 나타내기 보다는 일정한 기준 교통량을 상회하거나 하회하는 시점을 표현하는 것으로 해석하는 것이 합리적이다. 반면,본 논문에서 제안한 DEC 모형은 06:20이전과 06:30이 후를 구분하였다. 주요교차로에서는 교통량과 속도 관점에서 큰 변화가 감지되지 않았지만 4번 교차로의 속도가 크게 변화한 것이 두 시간대를 구분하는 주요 요인으로 작용한 것으로 판단된다. 07:00에 여러 교차로에서 교통량의 큰 변화가 관측되었으나 DEC모형에서는 07:00 전후의 교통상황은 서로 다르지 않은 것으로 판단되었다. 07:40 이전과 07:50 이후의 교통상황이 다른 것으로 구분되었는데, 이 시각에 주요 교차로에서

관측된 속도의 변화, 4번과 11번 교차로에서 관측된 속도와 교통량의 급격한 변화가 두 교통상황을 구분한 주요 요인으로 판단된다. 09:10시 이전과 09:20이후의 교통상황이 구분된 것은 4번 교차로의 속도, 11번 교차로의 속도와 교통량의 변화가 주 원인인 것으로 판단된다. 이 밖에도 10:00이전과 10:10이후의 구분은 11번 교차로의 속도와 교통량 변화, 16:00이전과 16:10이후의 구분은 11번 교차로 속도의 변화, 19:00 이전과 19:10 이후는 11번 교차로의 속도와 교통량의 변화가 주요 원인으로 판단된다.



<Fig. 5> Clustering of traffic states during weekends and weekdays based on the DEC model

기존의 방법론은 주말/공휴일과 주중을 완전히 다른 군집체계에서 분류하지만 본 논문의 DEC 모형은 주말/공휴일도 주중과 동일한 군집체계에서 분류한다. <Fig. 5>에서 주중과는 다르게 일요일은 3개의 군집으로 토요일은 5개의 군집으로 구분하였다. 일요일의 경우 기존 방법이 새벽(0:00~09:30)과 전일시간(09:30~21:00)과 같이 2개의 군집으로 구분한 반면, DEC 모형은 새벽시간(06:00~08:00), 오전첨두(08:10~09:10), 일과시간(09:20~16:00), 저녁시간(16:10~21:00)과 같이 4개의 군집으로 구분하였다. 기존 방법론에서 구분된 전일시간을 DEC 모형은 오전첨두, 일과시간, 저녁시간으로 세분한 것으로 판단된다. 토요일의 경우 기존 방법이 새벽시간(0:00~07:00), 출근시간(07:00~09:30), 오전시간(09:30~11:30), 오후시간(11:30~21:00)과 같이 4개의 군집으로 분류한 반면, DEC 모형은 새벽시간(06:00~07:50), 출근시간(07:50~09:20), 일과시간(09:20~16:10), 오후첨두(16:20~16:50), 저녁시간(17:00~21:00)과 같이 5개의 군집으로 분류하였다. 기존 방법론이 11:30을 기준으로 오전시간과 오후시간으로 구분한 반면, DEC 모형은 16:10을 기준으로 일과시간과 교통량이 급격히 증가하는 오후첨두시간 (16:20~16:50)을 구분하고 17:00 이후를 저녁시간으로 구분하였다. 본 논문에서 활용한 시간적

범위인 06:00~21:00중 Break point를 정리하여 <Table 1>와 <Table 2>로 나타내었다.

<Table 1> Break point time of Weekday

Break Point Time	Monday~Thursday		Friday	
	Current method	Proposed method	Current method	Proposed method
1 st	07:00	06:30	07:00	06:30
2 nd	09:30	07:50	09:30	07:50
3 rd	17:00	09:20	17:00	09:20
4 th	21:00	10:10	21:00	10:10
5 th	-	16:10	-	15:40
6 th	-	19:10	-	19:20

<Table 2> Break point time of Weekend

Break Point Time	Sunday		Saturday	
	Current method	Proposed method	Current method	Proposed method
1 st	09:30	08:10	07:00	07:50
2 nd	21:00	09:20	09:30	09:20
3 rd	-	16:10	11:30	16:20
4 th	-	-	21:00	17:00

DEC 모형으로 수행한 군집화의 결과를 보면 주말과 주중의 교통상황의 차이가 명확히 확인되었다. 또한 주중의 요일별로 오전 교통상황은 정체의 시작과 종료뿐만 아니라 군집도 동일하게 분류되었으나, 목요일과 금요일의 오후 교통상황은 다른 요일과 미세하게 상이함도 표현할 수 있음을 확인하였다.

V. 결 론

본 연구에서는 현재 우리나라에서 가장 널리 이용되고 있는 정주기식 신호제어전략의 TOD계획을 효율적으로 수립할 수 있도록 신호연동 그룹내에 속하는 모든 교차로의 자료를 활용한 교통상황 분류 기법을 제시하였다. 본 논문에서 제시된 DEC모형을 활용할 경우 기존의 주요 교차로의 교통자료를 이용한 TOD계획 대비, 교통상황을 보다 세밀하게 분류할 수 있었다. 기존에는 주요 교차로(Critical intersection)의 교통자료(교통량)만을 활용하여 교통량이 급격하게 변하는 Break point 시점을 기준으로 교통상황 분류 및 TOD 계획했다면, 신호연동그룹 내 전체 교차로 및 교통자료(교통량, 속도, 밀도)를 활용하여 교통상황을 분류할 수 있는 방법론을 제시하였다는 것이 본 연구의 주요 성과이다. 본 연구에 제시된 교통상황 분류 방법론은 주요 교차로가 아닌 교차로의 교통상황의 변화를 고려한 신호 제어전략이 필요에 따라 수립이 가능하다는 것을 보여주었다. 도시부의 교통혼잡에 선제적으로 대응할 수 있도록 본 연구에서 도출한 교통상황 분류기법과 교통상황예측 기법의 연계를 향후 연구과제로 제시한다.

REFERENCES

- Easter, M., Kriegel, H. P., Sander, J. and Xu, X.(1996), “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, AAAI(American Association for the Artificial Intelligence), KDD'96*, pp.226-231.
- Hwang, J., Parj, M., Kim, Y. and Kang, W.(2021), “Optimize TOD Time-Division with Dynamic Time Warping Distance-based Non-Hierarchical Cluster Analysis,” *The Journal of The Korea Institute of Intelligent Transport Systems*, Vol. 20, no. 5, pp.113-129.
- Jeong, Y., Kim, D., Kim, Y., Hwang, K. and Kim, Y.(2008), “Determination of TOD Plan Breaking Points Using a Clustering Analysis”, *Korean Society of Road Engineers*, pp.341-347.
- Lee, I. G., Lee, H. S. and Kim, Y. C.(2011), “Optimization of TIME-OF-DAY and Estimation on the Field Application for Arterial Road”, *Journal of Korean Society of Transportation*, vol. 29, no. 4, pp.113-123.
- Nielsen, F.(2016), *Introduction to HPC with MPI for Data Science*, Springer, pp.195-211.
- Scikit-Learn, <https://scikit-learn.org/stable/modules/mixture.html>, 2022.11.28.
- Tang, K., Boltze, M., Nakamura, H. and Tian, Z.(2019), “Global Practices on Road Traffic Signal Control”, *World Conference on Transport Research Society*, pp.7-17.
- UN-Habitat(2020), *World Cities Report 2020*, p.97.
- Xie, J., Cirshick, R. and Farhadi, A.(2016), “Unsupervised Deep Embedding for Clustering Analysis”, *Proceedings of the 33rd International Conference on Machine Learning, PMLR(Proceedings of Machine Learning Research)*, vol. 48, pp.478-487.