

2022년 추계학술발표대회 : 일반부문

# K-means 클러스터링을 활용한 대학교 기숙사 사생실의 여름철 CO<sub>2</sub> 농도 패턴 분류

## Classification of CO<sub>2</sub> Patterns in a University Dormitory using K-means Clustering for Cooling Season

○조 중 연\*      문 현 준\*\*  
Jo, Joong Yeon      Moon, Hyeun Jun

### Abstract

University dormitories are indoor spaces where students spend most of their time outside of class. Analysis on the IAQ data of the dormitory can contribute to improve students' comfort and health through better understanding of occupants' state based on the IAQ characteristics. The main object of this paper is to cluster IAQ(CO<sub>2</sub>) patterns of dormitory rooms using K-means++ to conduct a basic study to create comfort and healthy indoor environment. In this study, we installed CO<sub>2</sub> sensors at 10 dormitory rooms and collected data on a one-minute basis. As a result, CO<sub>2</sub> patterns were clustered into four groups through the silhouette score(k=4, 0.4507) and each clusters' showed completely different CO<sub>2</sub> profiles.

키워드 : 대학교 기숙사, K-means 클러스터링, 실내공기질

Keywords : University dormitory, K-means clustering, Indoor Air Quality

### 1. 서론

현대인은 대부분의 시간을 실내공간에서 생활한다. 이에 따라 쾌적한 실내환경을 제공하기 위한 연구들이 활발히 진행되고 있다(Jagriti, 2020). 특히 실내공기질(Indoor Air Quality, IAQ) 관리를 통해 재실자의 쾌적도와 건강을 고려한 실내건강환경을 조성하려는 노력이 증대되고 있다.

대학교 기숙사는 대학생들이 수업 외 대부분의 시간을 보내는 공간이므로, 기숙사의 실내환경에 대한 지속적인 연구가 필요하다(록이양, 2018). 쾌적한 실내환경을 조성하기 위해서는 기존 사생실의 IAQ 패턴과 특성을 이해하는 것이 중요한데, IAQ 패턴과 특성에 따라 효과적인 실내환경제어가 가능하기 때문이다. 따라서 본 연구는 쾌적하고 건강한 기숙사 실내환경 조성을 위해 대학교 기숙사 사생실을 대상으로 IAQ 요소 중 CO<sub>2</sub> 데이터를 수집하였고, K-means 알고리즘으로 사생실의 CO<sub>2</sub> 패턴을 분류했다.

\* 단국대 대학원 석사과정

\*\* 단국대 건축공학과 교수, 공학박사

(Corresponding author : Department of Architectural Engineering, Dankook University, hmoon@dankook.ac.kr)

본 연구는 2022년도 산업통상자원부의 재원으로 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구과제입니다(No. 20212020800120)

본 연구는 2022년도 산업통상자원부의 재원으로 한국에너지기술평가원(KETEP)의 지원을 받아 수행한 연구과제입니다(No.20212020800050)

### 2. 연구방법

#### 2.1 데이터 수집

본 연구에서는 경기도 용인시 소재 대학교 내 기숙사 남자동 10개 호실에 CO<sub>2</sub> 센서를 설치하여 1주일(2022.07.21~2022.07.27) 동안 1분 간격으로 수집한 데이터를 활용하였다. 표1은 사생실에 설치된 센서 정보이다.

표1. 기숙사 사생실 설치 IAQ 센서 정보

센서	모델명	측정방식	측정범위
IAQ 측정요소 CO <sub>2</sub>	SenseAir S8	NDIR	400 ~ 2,000 ppm

아래 그림1은 수집된 10개실의 시간별 평균 CO<sub>2</sub> 농도 Data profiling 결과이다. 일일 CO<sub>2</sub> 농도 변화를 살펴보면, 자정(00:00)을 기점으로 새벽 시간대에 1) 농도가 증가하는 패턴을 보이는 호실과 2) 감소하는 패턴을 보이는 호실, 3) 변화량이 없는 호실이 존재한다. 15시부터는 대부분의 사생실이 공통적으로 시간에 따라 CO<sub>2</sub> 농도가 상승하는 경향을 보인다.

#### 2.2 K-means 클러스터링

K-means 클러스터링은 Input Data를 K개 군집으로 분류한 뒤 군집에 포함된 Input의 평균을 클러스터의 중심점(centroid)으로 계산하는 알고리즘이다. 그러나 K-means는

처음에 지정하는 중심점의 위치를 무작위로 결정하여 매번 동일한 결과를 도출하기 어렵고, 각 중심점 사이의 거리가 짧을 경우 적합한 분류가 어려워 최적의 클러스터를 형성하는데 한계가 있다. 본 연구에서는 결과의 개선을 위해, 중심점 무작위 선정 문제를 해결하기 위해 제안된 K-means++ 알고리즘을 사용했다(Arthur, 2006).

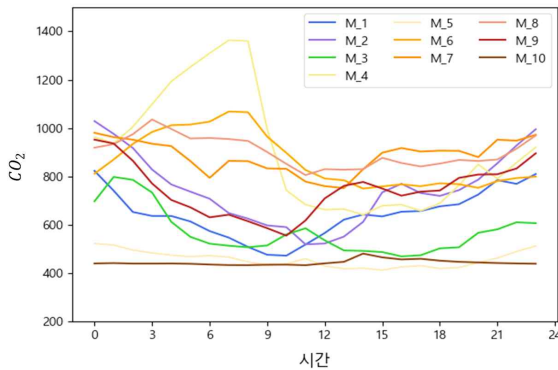


그림1. 10개 호실의 시간별 CO<sub>2</sub> Data Profiling

Silhouette score는 거리 매트릭스를 통해 군집 내에서 데이터가 얼마나 가깝게 형성되어 있는지, 군집 간 구분이 얼마나 뚜렷한지 평가하는 기준이다. 이 값은 -1 ~ 1로 표현되며 1에 가까울수록 최적의 군집 개수를 나타낸다.

### 3. 연구 결과

그림2의 Silhouette score는 클러스터가 4개일 때 0.4507로 가장 높았으며, 이어서 5개(0.4340), 3개(0.3785), 6개(0.3593) 순으로 나타났다. 결과값 중 가장 1에 근접하는 k=4를 클러스터 수로 선정하여 군집화를 수행하였다.

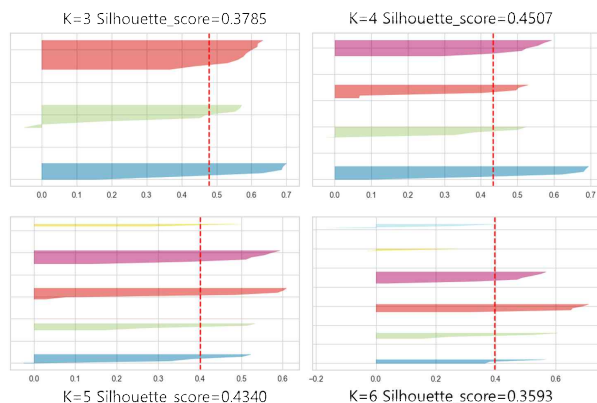


그림2. K 값에 따른 Silhouette score 시각화

그림3은 기숙사 10개 호실의 IAQ(CO<sub>2</sub>)를 4개의 클러스터로 분류한 결과이다. 클러스터에 포함된 호실들의 시간별 CO<sub>2</sub> 농도 평균을 빨간 그래프로 나타내어 각 클러스터의 중심값을 알 수 있도록 하였다. 클러스터링 결과, Cluster 0은 평균 CO<sub>2</sub> 농도가 451ppm으로 데이터 수집 기

간 동안 재실자가 없어 공실로 유지가 된 것으로 판단된 군집이다. 반면, Cluster 1~3은 평균 CO<sub>2</sub> 농도가 각각 883, 676, 907ppm으로, 재실자가 거주하는 사생실 군집으로 판단된다. Cluster 2는 새벽부터 낮(00:00~10:00)까지 CO<sub>2</sub> 농도가 감소하다가 이후 자정(10:00~24:00)까지 농도가 증가하는 유형의 군집이다. Cluster 3의 경우 낮부터 자정(10:00~24:00)까지 CO<sub>2</sub> 평균 농도는 Cluster 1에 못 미치나, 새벽부터 아침까지(00:00~08:00) 농도가 동 시간대 다른 군집에 비해 최대 2.65배 높은 유형의 군집으로 나타나 새벽 시간대에 재실자가 주로 거주하는 군집으로 판단된다.

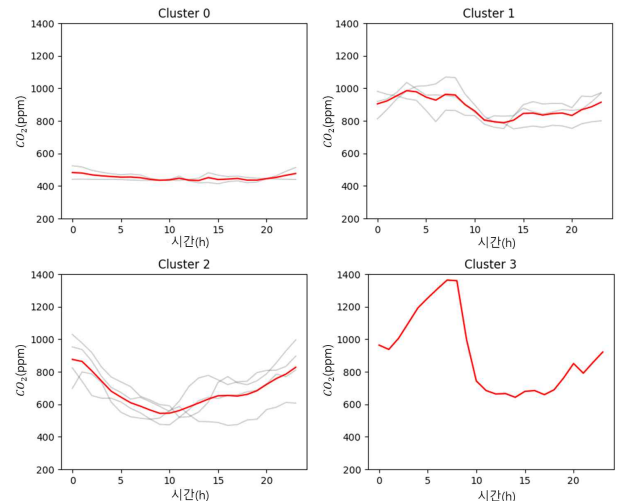


그림3. 남자동 10개 호실의 IAQ(CO<sub>2</sub>) 패턴 클러스터링 결과

### 4. 결론

본 연구에서는 대학교 기숙사 남자동 10개 사생실의 공기질 패턴과 특성을 이해하기 위해 CO<sub>2</sub> 센서를 설치하여 1분 간격으로 농도 데이터를 수집하였다. 수집한 CO<sub>2</sub> 농도 데이터의 분석 및 시각화를 위해 Data Profiling을 실시하였으며 각 호실별 유사한 CO<sub>2</sub> 농도 패턴을 군집화하기 위해 K-means 클러스터링을 통해 총 4개의 군집을 형성하였다. 그 결과, 군집마다 서로 다른 CO<sub>2</sub> 농도 패턴과 특성을 확인할 수 있었다. 향후 초미세먼지(PM2.5), 포름알데히드 등의 IAQ 데이터를 이용한 추가적인 클러스터 분석을 통해, 사생실의 재실상태를 유추하고 이를 반영한 군집별 HVAC 시스템 제어 전략으로 보다 더 쾌적한 실내건강환경을 제공할 수 있을 것으로 기대된다.

### 참고문헌

1. Saini J., Dutta, M. & Marques, G. A. (2020) comprehensive review on indoor air quality monitoring systems for enhanced public health. Sustain Environ Res
2. 록이양, 대학교 기숙사의 실내환경 실태와 주관적 반응 석사학위논문, 전북대학교원 주거환경학과, 2018
3. Arthur David, Sergei Vassilvitskii. (2006) K-means++: The advantages of careful seeding. Stanford