

MCLP

분석데이터

No.	데이터	데이터 건수	전처리 후 건수	출처	비고
1	서울시 100m X 100m 격자 인구	81060건	-	국토정보플랫폼	
2	자동차 등록현황	24300건	-	서울 열린데이터 광장 서울특별시	18~20년
3	서울시 자치구별 수소차 등록현황	288건	-	공공데이터포털 서울특별시	2020.09
4	수소충전소	363건	600건	공공데이터포털 한국가스안전공사	위경도, 20.09
5	전국 LPG충전소	10209건	20410건	공공데이터포털 한국가스안전공사	위경도
6	유치원	21912건	19152건	서울 열린데이터 광장 서울특별시교육청	위경도, 열 삭제
7	초중고 학교	99278건	17028건	공공데이터포털 서울특별시	위경도, 열 삭제, 유일값

입지지수 개발

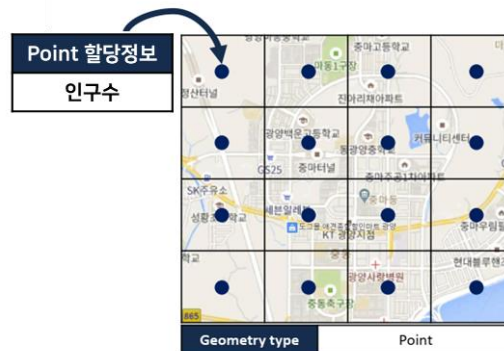


Figure 1 인구선정을 위한 지역특성 요소 추출

- 서울시 인구현황데이터는 현재 100mX100m 기준 64676개의 격자(grid)로 나누어져 있는 상태
- 각 격자의 가운데에 중심점(center point)을 찍음
- 격자의 중심점(Central point)에 정규화된 인구 수 정보 할당

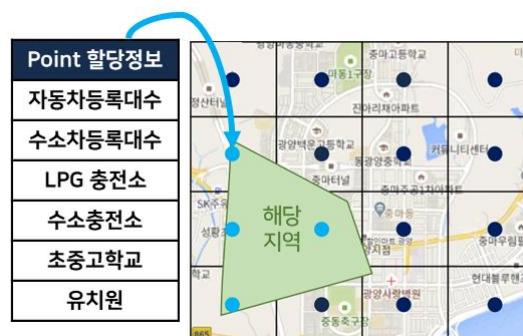


Figure 2 . 인구수 격자 중심점에 정보 할당

- 자동차 등록대수, 수소차 등록대수, LPG 충전소, 수소충전소, 학교(초중고), 유치원 지역에 해당하는 Point를 모두 추출 후, 격자 중심점에 해당 정보 할당

가중치 산정

■ 로지스틱 회귀분석(Logistic regression)

$$y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_i X_i$$

- X = (정규화된) 인구수, 자동차 등록대수, 수소차 등록대수
- Y = LPG 충전소 설치 여부

(정규화) 지역특성		W (가중치)
인구수		0.43195616
자동차등록대수		4.59406385
수소차등록대수		-1.0288214

Figure 3 수소충전소 입지선정 지수

분석모델

■ MCLP(Maximal Coverage Location Problem)

- 제한된 시설물의 개수로 지역 수요를 최대한 커버할 수 있는지 파악하기 위한 입지 선정 모델링
- 전기충전소 입지 연구를 참고하여, 수소충전소가 반경 500M 내의 수요를 커버할 수 있다고 가정
- 이 때 Cover되지 못한 수요는 서비스를 못받는 것이 아닌 서비스를 받긴 하지만 설비로부터의 거리가 멀어 만족할만한 서비스 수준을 제공받지 못하는 수요를 의미

Table 1 MCLP 목적구문, 제약구문

목적구문	$maximize \sum_{i \in I} w_i y_i$	가중치를 고려하여 가장 많은 수요를 포함하는 수소 충전소 위치 선정
제약구문	$\sum_{j \in J} x_j = K$	배치되는 최대 수소 충전소 개수
	$y_i \leq \sum_{j \in N_i} x_j \text{ for all } i \in I$	각 수소 충전소는 반지름(500m)의 영향권을 가짐
변수 설명	<p>i : 수요포인트 index (격자 1개) j : 수소충전소 구축 가능한 후보지역 index I : 수요포인트 집합 (모든 격자) J : 수소 충전소 구축 가능한 후보지역 집합 K : 총 설치해야하는 수소충전소 개수 x : 수소충전소 후보지역 중 위치 j에 수소충전소가 설치되면 1, 그렇지 않으면 0 y : 적어도 하나의 수소충전소로 그 포인트가 커버되면 1, 그렇지 않으면 0 w : 가중치</p>	

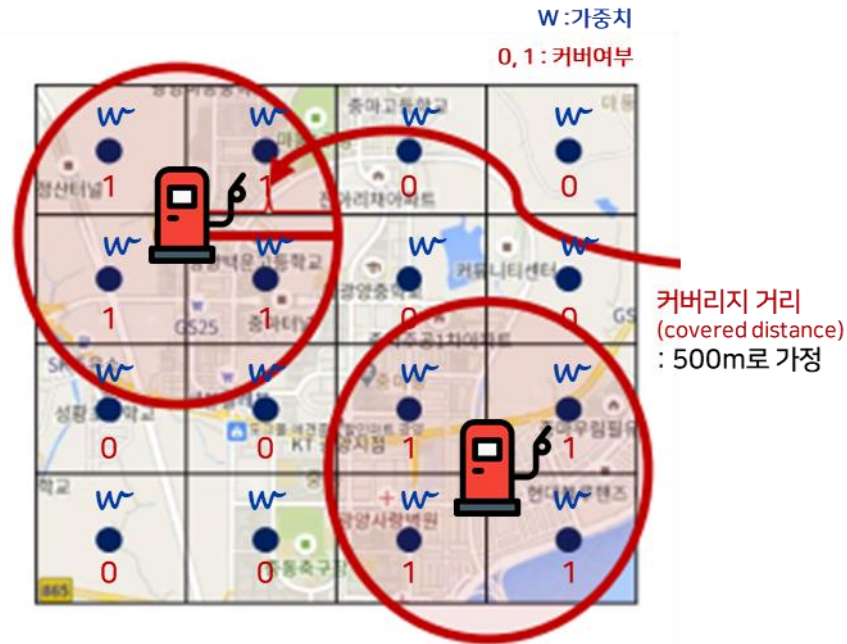


Figure 4 MCLP 방법론

1. 정규화된 인구수, 자동차 등록대수, 수소차 등록대수에 가중치를 곱해서 통합된 가중치 정보(w)를 산출하여 격자에 할당
2. 기존 LPG 충전소 위치를 기반으로 1종 보호시설(학교, 유치원) 및 기존 수소충전소 위치를 제외한 후보지역(J) 선정
3. 총 설치할 수소충전소 개수 K 에 맞춰 랜덤하게 수소충전소 할당 후 후보지역에 수소충전소가 설치되면 $x = 1$, 그렇지 않으면 $x = 0$ 을 할당
4. 세번째 식 에 따라 수소충전소로부터 covered distance(500m)내에 수요포인트를 커버할 수 있으면 $y = 1$, 없으면 $y = 0$ 을 할당
5. 가중치를 고려하여 가장 많은 수요를 포함하는 수소 충전소 위치 선정

K-medoids clustering

분석데이터

No.	데이터	데이터 건수	전처리 후 건수	출처	비고
1	전국 LPG충전소	10209건	20410건	공공데이터포털 한국가스안전공사	위경도

분석모델

■ K-medoids clustering

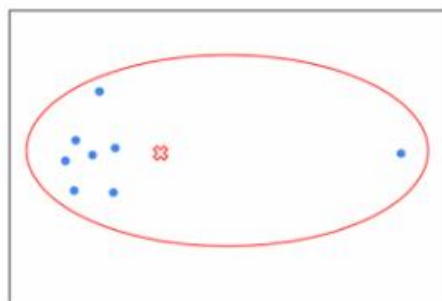
- K-medoids clustering은 K-means를 변형한 것으로, 군집의 무게 중심을 구하기 위해서 데이터의 평균 대신 중간점(medoids)를 사용한 분석
- 초기에 제일 처음 랜덤하게 포인트를 하나 잡아서 그 포인트에 가까운 데이터들을 같은 군집으로 할당. 그 다음 아래와 같은 방법으로 반복함.
- 다음과 같은 목적함수 값이 최소화될 때까지 군집의 중심위치와 각 데이터가 소속될 군집을 반복해서 찾음. 이 값을 관성(inertia)이라 함.

$$J = \sum_{k=1}^K \sum_{i \in C_k} d(x_i, \mu_k)$$

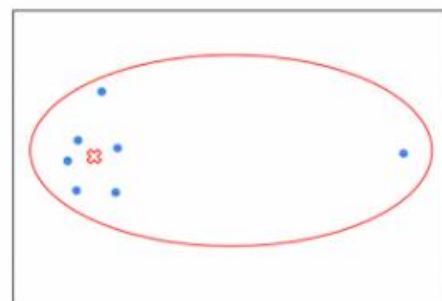
- 이 식에서 K는 군집의 갯수이고 C_k 는 k번째 군집에 속하는 데이터의 집합, μ_k 는 k번째 군집의 중심위치(centroid), d 는 x_i , μ_k 두 데이터 사이의 거리 혹은 비유사도(dissimilarity)로 정의. 만약 유클리드 거리를 사용한다면 다음과 같음.

$$d(x_i, \mu_k) = \|x_i - \mu_k\|^2$$

- 세부 알고리즘
 1. 임의의 중심위치 $\mu_k(k=1,...,K)$ 를 고른다. 보통 데이터 표본 중에서 K개를 선택
 2. 모든 데이터 $x_i(i=1,...,N)$ 에서 각각의 중심위치 μ_k 까지의 거리를 계산
 3. 각 데이터에서 가장 가까운 중심위치를 선택하여 각 데이터가 속하는 군집을 정함
 4. 각 군집에 대해 중심위치 μ_k 를 다시 계산
 5. 2~4를 반복



(a) Mean



(b) Medoid

Figure 5 k-means clustering과 k-medoids clustering 비교