

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)(11) 공개번호 10-2021-0085136
(43) 공개일자 2021년07월08일

(51) 국제특허분류(Int. Cl.)

G06Q 50/30 (2012.01) G06N 3/08 (2006.01)
G06Q 10/04 (2012.01) G06Q 50/10 (2012.01)
G09B 29/00 (2006.01)

(52) CPC특허분류

G06Q 50/30 (2013.01)
G06N 3/08 (2013.01)

(21) 출원번호 10-2019-0177870

(22) 출원일자 2019년12월30일

심사청구일자 2019년12월30일

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

허준

서울특별시 서초구 서초대로 65길 13-10

임형준

서울특별시 구로구 신도림로 105, 203동 303호 (신도림동, 신도림우성2차아파트)

마중원

서울특별시 서초구 사임당로23길 58, 301동 408호(서초동, 서초2차현대아파트)

(74) 대리인

김인철

전체 청구항 수 : 총 8 항

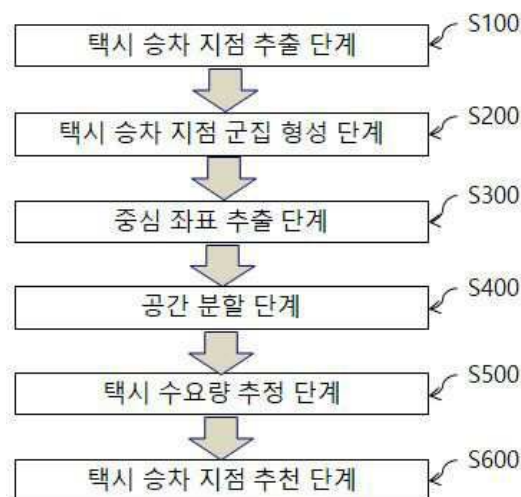
(54) 발명의 명칭 공간 분할 기법을 이용한 택시 수요량 추정방법

(57) 요약

본 발명은 공간 분할 기법을 이용한 택시 수요량 추정방법에 관한 것으로,

본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법은, 컴퓨팅 장치에 의해 수행되며, 상기 컴퓨팅 장치가, 택시 경로 데이터들을 수집하고, 수집된 택시 경로 데이터들로부터 택시 승차 지점들을 추출하는 단계; 군집화 알고리즘을 이용하여 통해 상기 추출된 택시 승차 지점들의 군집들을 형성하는 단계; 상기 형성된 군집들 각각에 속하는 택시 승차 지점들의 중심 좌표들을 추출하는 단계; 상기 추출된 중심 좌표들을 이용하여 공간을 분할하는 단계; 상기 택시 경로 데이터들을 그래프 데이터로 변환하고, 변환된 그래프 데이터를 입력 데이터로 하는 딥러닝을 수행하여 상기 분할된 공간 별로 택시 수요량을 추정하는 단계;를 포함한다.

대표도 - 도3



(52) CPC특허분류

G06Q 10/04 (2013.01)

G06Q 50/10 (2013.01)

G09B 29/007 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1615010455
부처명	국토교통부
과제관리(전문)기관명	국토교통과학기술진흥원
연구사업명	국토공간정보연구사업(R&D)
연구과제명	공간 빅데이터 서비스 활성화 및 실증연구
기 여 율	1/1
과제수행기관명	안양대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

컴퓨팅 장치에 의해 수행되며, 상기 컴퓨팅 장치가,

택시 경로 데이터들을 수집하고, 수집된 택시 경로 데이터들로부터 택시 승차 지점들을 추출하는 단계;

군집화 알고리즘을 이용하여 통해 상기 추출된 택시 승차 지점들의 군집들을 형성하는 단계;

상기 형성된 군집들 각각에 속하는 택시 승차 지점들의 중심 좌표들을 추출하는 단계;

상기 추출된 중심 좌표들을 이용하여 공간을 분할하는 단계;

상기 택시 경로 데이터들을 그래프 데이터로 변환하고, 변환된 그래프 데이터를 입력 데이터로 하는 딥러닝을 수행하여 상기 분할된 공간 별로 택시 수요량을 추정하는 단계;

를 포함하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

청구항 2

청구항 1에 있어서,

상기 택시 경로 데이터는 차량 정보, 운행 일시, 패신저 코드, 좌표 정보를 포함하는 데이터인 것을 특징으로 하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

청구항 3

청구항 1에 있어서,

상기 군집화 알고리즘은 k-means 군집화 알고리즘인 것을 특징으로 하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

청구항 4

청구항 1에 있어서,

상기 중심 좌표들을 추출하는 단계는, K-means 군집화 알고리즘을 통해 형성된 군집의 평균 좌표 (\bar{x}, \bar{y}) 를 하기 식 (1)로 계산하여 추출하는 것을 특징으로 하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

$$\bar{x} = (x_1 + x_2 + \dots + x_N) / N$$

$$\bar{y} = (y_1 + y_2 + \dots + y_N) / N$$

식 (1) :

(N은 같은 군집에 속한 택시 승차 지점 좌표의 개수임)

청구항 5

청구항 1에 있어서,

상기 공간을 분할하는 단계는, 상기 중심 좌표를 기준으로 보로노이 다각형을 이용하여 공간을 분할하는 것을 특징으로 하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

청구항 6

청구항 1에 있어서,

상기 분할된 공간 별로 택시 수요량을 추정하는 단계는,

택시 승차 지점, 하차 지점, 그리고 운행경로를 그래프 데이터 구조로 변환하여 복수의 그래프 데이터를 생성하고, 생성된 그래프 데이터를 학습데이터로 하여 딥러닝 모델을 학습시킨 후, 분할된 공간 기준으로 지역별 평균 승차 횟수 및 주행 거리를 산출하여 지역별 택시 수요량을 추정하는 것을 특징으로 하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

청구항 7

청구항 1에 있어서,

상기 추정된 택시 수요량을 기반으로 택시 승차 지점을 추천하는 단계를 더 포함하는 것을 특징으로 하는 공간 분할 기법을 이용한 택시 수요량 추정방법.

청구항 8

컴퓨팅 장치에 의해 수행되며,

택시 경로 데이터들을 수집하고, 수집된 택시 경로 데이터들로부터 택시 승차 지점들을 추출하는 단계;

군집화 알고리즘을 이용하여 통해 상기 추출된 택시 승차 지점들의 군집들을 형성하는 단계;

상기 형성된 군집들 각각에 속하는 택시 승차 지점들의 중심 좌표들을 추출하는 단계;

상기 추출된 중심 좌표들을 이용하여 공간을 분할하는 단계;

상기 택시 경로 데이터들을 그래프 데이터로 변환하고, 변환된 그래프 데이터를 입력 데이터로 하는 딥러닝을 수행하여 상기 분할된 공간 별로 택시 수요량을 추정하는 단계;

를 실행시키기 위하여, 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 공간 분할 기법을 이용한 택시 수요량 추정방법에 관한 것이다.

배경 기술

[0003] 도 1은 종래 기술에 따른 택시 수요량 추정방법에 사용되는 공간 분할 기법을 설명하기 위한 도면이다.

[0004] 도 1에 도시된 바와 같은, 종래 기술에서는 택시 위치 정보를 그래프 형태로 변환할 때, 정사각형 형태의 격자 또는 정육각형 격자로 대상 지역을 나누는 경우가 대부분이었다. 이 경우 지역을 쉽게 분할할 수 있으나, 택시 수요가 거의 없는 지역(산 또는 하천 등)만을 포함할 수 있어 추정 정확도가 떨어지게 된다.

[0005] 또한, 택시 수요가 많은 지역은 공간 분할을 더 작게, 수요가 적은 지역은 분할을 더 크게 할 필요가 있는데, 균일한 격자를 적용하면 이를 해결할 수 없다. 따라서 택시 수요량에 따른 효율적인 공간 분할을 통해 더 정확하고 실용적인 택시 수요량을 추정할 수 있는 택시 수요량 추정방법의 개발이 필요하다.

선행기술문헌

특허문헌

[0007] (특허문헌 0001) 한국등록특허 제1993551호 (2019.06.20)

발명의 내용

해결하려는 과제

[0008] 본 발명은 택시 수요량에 따른 효율적인 공간 분할을 통해 더 정확하고 실용적인 택시 수요량을 추정할 수 있는 공간 분할 기법을 이용한 택시 수요량 추정방법을 제공하는 것을 목적으로 한다.

과제의 해결 수단

[0010] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법은, 컴퓨팅 장치에 의해 수행되며, 상기 컴퓨팅 장치가, 택시 경로 데이터들을 수집하고, 수집된 택시 경로 데이터들로부터 택시 승차 지점들을 추출하는 단계; 군집화 알고리즘을 이용하여 통해 상기 추출된 택시 승차 지점들의 군집들을 형성하는 단계; 상기 형성된 군집들 각각에 속하는 택시 승차 지점들의 중심 좌표들을 추출하는 단계; 상기 추출된 중심 좌표들을 이용하여 공간을 분할하는 단계; 상기 택시 경로 데이터들을 그래프 데이터로 변환하고, 변환된 그래프 데이터를 입력 데이터로 하는 딥러닝을 수행하여 상기 분할된 공간 별로 택시 수요량을 추정하는 단계;를 포함한다.

[0011] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 있어서, 상기 택시 경로 데이터는 차량 정보, 운행 일시, 패신저 코드, 좌표 정보를 포함하는 데이터일 수 있다.

[0012] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 있어서, 상기 군집화 알고리즘은 k-means 군집화 알고리즘일 수 있다.

[0013] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 있어서, 상기 중심 좌표들을 추출하는 단계는, K-means 군집화 알고리즘을 통해 형성된 군집의 평균 좌표 (\bar{x}, \bar{y}) 를 하기 식 (1)로 계산하여 추출할 수 있다.

[0014] 식 (1) :
$$\bar{x} = (x_1 + x_2 + \dots + x_N) / N$$

$$\bar{y} = (y_1 + y_2 + \dots + y_N) / N$$

[0015] (N은 같은 군집에 속한 택시 승차 지점 좌표의 개수임)

[0016] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 있어서, 상기 공간을 분할하는 단계는, 상기 중심 좌표를 기준으로 보로노이 다각형을 이용하여 공간을 분할할 수 있다.

[0017] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 있어서, 상기 분할된 공간 별로 택시 수요량을 추정하는 단계는, 택시 승차 지점, 하차 지점, 그리고 운행경로를 그래프 데이터 구조로 변환하여 복수의 그래프 데이터를 생성하고, 생성된 그래프 데이터를 학습데이터로 하여 딥러닝 모델을 학습시킨 후, 분할된 공간 기준으로 지역별 평균 승차 횟수 및 주행 거리를 산출하여 지역별 택시 수요량을 추정할 수 있다.

[0018] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 있어서, 상기 추정된 택시 수요량을 기반으로 택시 승차 지점을 추천하는 단계를 더 포함할 수 있다.

[0020] 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법은, 택시 경로 데이터들을 수집하고, 수집된 택시 경로 데이터들로부터 택시 승차 지점들을 추출하는 단계; 군집화 알고리즘을 이용하여 통해 상기 추출된 택시 승차 지점들의 군집들을 형성하는 단계; 상기 형성된 군집들 각각에 속하는 택시 승차 지점들의 중심 좌표들을 추출하는 단계; 상기 추출된 중심 좌표들을 이용하여 공간을 분할하는 단계; 상기 택시 경로 데이터들을 그래프 데이터로 변환하고, 변환된 그래프 데이터를 입력 데이터로 하는 딥러닝을 수행하여 상기 분할된 공간 별로 택시 수요량을 추정하는 단계;를 실행시키기 위하여, 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨

터 프로그램에 의해 수행될 수 있다.

[0022] 기타 본 발명의 다양한 측면에 따른 구현예들의 구체적인 사항은 이하의 상세한 설명에 포함되어 있다.

발명의 효과

[0024] 본 발명의 실시 형태에 따르면, 택시 수요량에 따른 효율적인 공간 분할을 통해 더 정확하고 실용적인 택시 수요량 추정이 가능하다. 또한, 정확한 택시 수요량 추정을 통해 택시 투입 대수를 탄력적으로 운영하여 운행 효율을 최적화할 수 있다.

도면의 간단한 설명

[0026] 도 1은 종래 기술에 따른 택시 수요량 추정방법에 사용되는 공간 분할 기법을 설명하기 위한 도면이다.
 도 2는 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법을 설명하기 위한 도면이다.
 도 3은 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법이 도시된 순서도이다.
 도 4는 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 사용되는 택시 경로 데이터를 예시한 도면이다.
 도 5는 택시 경로 데이터에서 택시 승차 지점들만을 추출하여 도시한 택시 승차 지점 예시도이다.
 도 6은 추출된 택시 승차 지점들에 대해 군집화 알고리즘을 적용하여 군집화(clustering)한 것을 예시하는 예시도이다.
 도 7은 추출된 택시 승차 지점들에 대해 군집화 알고리즘을 적용한 군집화 결과를 지도 상에 반영한 예시도이다.
 도 8은 도 7의 일부를 확대 도시한 도면이다.
 도 9는 도 7의 군집화로부터 각 군집 별로 추출된 중심 좌표들을 예시하는 예시도이다.
 도 10은 도 9의 중심 좌표들을 이용하여 공간을 분할하는 것을 예시하는 예시도이다.
 도 11은 도 4의 택시 경로 데이터를 그래프 데이터로 변환하여 택시 수요량 추정을 위한 딥러닝 모델의 입력 데이터로 변환하는 과정을 설명하기 위한 도면이다.
 도 12는 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 의해 추정된 택시 수요량을 기반으로 택시 운전자에게 택시 승차 지점을 추천하는 예시적인 유저 인터페이스가 도시된 도면이다.
 도 13은 종래 기술의 공간 분할 기법과 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 의해 추정된 택시 수요량의 정확도를 비교 예시한 도면이다.
 도 14는 본 발명의 일 실시예에 따른 컴퓨팅 장치가 도시된 도면이다.

발명을 실시하기 위한 구체적인 내용

[0027] 본 발명은 다양한 변환을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예를 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0029] 본 발명에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 발명에서, '포함하다' 또는 '가지다' 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소,

부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다. 이하, 도면을 참조하여 본 발명의 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법을 설명한다.

- [0031] 도 2는 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법을 설명하기 위한 도면이다.
- [0032] 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법(이하, “택시 수요량 추정방법” 이라고도 함)은 컴퓨팅 장치에 의해 수행된다. 컴퓨팅 장치는 외부의 DB로부터 택시 경로 데이터들을 수집하고, 수집된 택시 경로 데이터들을 이용하여 공간을 분할하고, 택시 경로 데이터를 그래프 데이터로 변환하여 구축하고, 그래프 데이터를 입력 데이터로 하는 딥러닝을 수행하여 분할된 공간 별로 택시 수요량을 예측한다. 컴퓨팅 장치에 대해서는, 도 14를 참조하여 후술한다.
- [0033] 본 명세서에서 정의하는 택시는, 정해진 노선이 없이 승객을 태우고 이동하는 이동체를 의미하는 것으로, 노선 없이 승객을 태우고 이동체면 자동차, 오토바이, 수송용 드론 등 종류를 불문한다. 다만, 설명의 편의를 위해 “택시”로 칭한다. 이는 청구범위 해석에 있어서도 동일하다.
- [0035] 도 3을 참조하여 본 발명의 일 실시예에 따른 택시 수요량 추정방법을 상세히 설명한다. 도 3은 본 발명의 일 실시예에 따른 택시 수요량 추정방법이 도시된 순서도이다.
- [0036] 도 3에 도시된 바와 같이, 본 발명의 일 실시예에 따른 택시 수요량 추정방법은, 택시 승차 지점 추출 단계(S100)와, 택시 승차 지점 군집 형성 단계(S200)와, 중심 좌표 추출 단계(S300)와, 공간 분할 단계(S400)와, 택시 수요량 추정 단계(S500)와, 택시 승차 지점 추천 단계(S600)를 포함한다.
- [0038] S100 : 택시 승차 지점 추출 단계
- [0039] 먼저, 컴퓨팅 장치는 외부의 DB로부터 택시 경로 데이터들을 수집한다. 외부의 DB는 택시 사업과 관련된 DB로, 콜택시 서비스를 제공하는 업체에서 운영하는 데이터베이스일 수 있다.
- [0040] 택시 경로 데이터는 차량 정보, 운행 일시, 패신저 코드(Passenger code), 승하차 좌표 정보 등을 포함하는 데이터로서, 택시의 2차원 좌표 정보 뿐만 아니라 고객의 승차/하차 여부를 나타낼 수 있다.
- [0041] 패신저 코드는 공차(손님이 없는 상태)일 때 ‘0’, 승차 시 ‘1’, 하차 시 ‘2’, 손님을 태우고 운행 중일 때 ‘3’으로 표시될 수 있다. 예시적으로 택시 경로 데이터는 다음과 같이 쉼표로 구분된 텍스트 파일 형식으로 될 수 있다.
- [0042] 서울31바2061,20170423003208,1,126.991799,37.570501
- [0043] 택시 경로 데이터는 쉼표를 기준으로 순서대로 차량 번호, 운행 일시(2017년-4월-23일-00:32:08), 패신저 코드(Passenger code), 해당 코드 시점의 X좌표, Y좌표를 포함할 수 있다.
- [0044] 도 4에 택시 경로 데이터의 다른 예시가 개시되어 있다. 여기서, 차량 정보는 중요한 요소가 아니므로, 생략될 수 있다. 도 4에서 D1은 택시 경로 데이터의 일련 번호이고, D2는 패신저 코드이며, D3는 승하차 좌표 정보이다.
- [0045] 상기와 같은 택시 경로 데이터들을 수집한 컴퓨팅 장치는 수집된 택시 경로 데이터들로부터 택시 승차 지점들만을 추출한다. 구체적으로, 택시 경로 데이터들 중에서 패신저 코드가 “1”인 데이터의 좌표 정보만을 추출한다.
- [0046] 도 5는 택시 경로 데이터에서 택시 승차 지점들만을 추출하여 도시한 택시 승차 지점 예시도이다. 패신저 코드가 “1”인 데이터의 좌표 정보들은 도 5와 같이 표시될 수 있다.
- [0048] S200 : 택시 승차 지점 군집 형성 단계
- [0049] S100 단계 이후에, 컴퓨팅 장치는 다양한 군집화 알고리즘을 이용하여 S100 단계에서 추출된 택시 승차 지점들의 군집(clustering)들을 형성한다.

- [0050] 군집화 알고리즘은 가까운 것끼리 작은 군집을 형성하고, 가까이에 있는 작은 군집이 결합해서 더 큰 군집을 형성하는 형태의 계층적 군집화(hierarchical clustering) 알고리즘일 수 있다. 또한, 군집화 알고리즘은 계층적 구조를 만들지 않고 전체 데이터를 유사한 것들끼리 나누어서 묶는 분할 군집화(partitioning clustering) 알고리즘일 수 있다. 계층적 군집화 알고리즘은 단일연결법, 완전 연결법, 평균 연결법 중 적어도 어느 하나일 수 있고, 분할 군집화 알고리즘은 k-means 군집화 알고리즘일 수 있다. 본 명세서에서는 k-means 군집화 알고리즘을 예시로 하여 설명하나, 본 발명의 범위가 k-means 군집화 알고리즘에 한정되는 것은 아니다.
- [0051] k-means 군집화 알고리즘은 주어진 데이터를 k개의 클러스터로 묶는 알고리즘으로, 각 클러스터와 거리 차이의 분산을 최소화하는 방식으로 동작한다. k-means 군집화 알고리즘의 동작 과정은 다음과 같다.
- [0052] ① 전체 데이터 중에서 무작위로 k개의 데이터를 선택해서 각각을 군집의 초기 중심으로 정한다.
- [0053] ② 각 데이터와 각 군집 중심 간의 거리를 계산한다.
- [0054] ③ 각각의 점들을 방금 구한 중심 가운데 제일 가까운 것에 연결지어서 새로운 집합을 나눌 수 있다.
- [0055] ④ 각 군집에 대해 배정된 모든 데이터의 평균을 구해서, 이것을 새로운 군집 중심으로 정한다.
- [0056] ** 이전 상태와 비교하여 군집의 변화가 있다면, 과정 ②로 돌아가고, 그렇지 않으면 종료한다.
- [0058] 상기의 과정을 통해 택시 승차 지점들을 K개의 군집으로 나눌 수 있다. 군집의 개수 K는 일반적으로 3 이상으로 설정하고 사용자가 조절 가능하다.
- [0059] 도 6은 추출된 택시 승차 지점들에 대해 k-means 군집화 알고리즘을 적용하여 군집화(clustering)한 것을 예시하는 예시도이다. 도 6에서 V10과 V11은 각각 택시 승차 지점의 X좌표, Y좌표를 의미하며, K-means 군집화 알고리즘을 통해 나온 결과물은 기존의 좌표에 군집 번호(cluster)가 부여된다. 예를 들어, K를 300으로 설정한 경우, 1부터 300까지의 군집 번호가 부여된 형태의 파일이 결과물로 도출된다.
- [0060] 도 7은 추출된 택시 승차 지점들에 대해 k-means 군집화 알고리즘을 적용한 군집화 결과를 지도 상에 반영한 예시도이다. 도 7은 도 6에서 도출된 군집 번호가 부여된 형태의 결과물을 시각화한 것이다.
- [0062] S300 : 중심 좌표 추출 단계
- [0063] 도 8은 도 7의 일부를 확대 도시한 도면이다. 도 7에서 군집의 경계만으로 공간을 분할할 수 있을 것으로 보이지만, 이를 확대한 도 8을 살펴보면 도로 위를 이동하는 이동 객체의 특성상 좌표가 도로 위에서만 형성되기 때문에 도 7과 같은 결과만으로는 경계를 특정하기 어렵다. 따라서, 본 S300 단계에서는 추출된 택시 승차 지점들에 대해 군집화 알고리즘을 적용한 군집들의 중심 좌표를 추출하여 경계를 특정할 수 있도록 한다.
- [0064] 구체적으로 S200 단계 이후에, 컴퓨팅 장치는 형성된 군집들 각각에 속하는 택시 승차 지점들의 중심 좌표들을 추출한다. 예를 들어, K-means 군집화 알고리즘을 통해 형성된 군집의 평균 좌표 (\bar{x}, \bar{y}) 를 아래 식 (1)로 계산하여 추출한다.
- $$\bar{x} = (x_1 + x_2 + \dots + x_N) / N$$
- $$\bar{y} = (y_1 + y_2 + \dots + y_N) / N$$
- [0065] 식 (1) :
- [0066] 여기서, N은 같은 군집에 속한 택시 승차 지점 좌표의 개수이다. 최종적으로 군집 개수 만큼의 중심 좌표가 도출된다. 예를 들어, K를 300으로 설정하여 300개의 군집이 형성된 경우, 300개의 중심 좌표가 추출된다.
- [0067] 도 9는 도 7의 군집화로부터 각 군집 별로 추출된 중심 좌표들을 예시하는 예시도이다.
- [0069] S400 : 공간 분할 단계
- [0070] 컴퓨팅 장치는 S300 단계에서 추출된 중심 좌표들을 이용하여 공간을 분할한다. 구체적으로 중심 좌표를 기준으로 보로노이 다각형(Voronoi polygon)을 이용하여 공간을 분할한다. 보로노이 다각형은 평면 위에 점들이 있을

때 이 점을 하나씩 포함하는 다각형들로 면을 분할하되 다음 조건을 만족시키는 경우를 말한다.

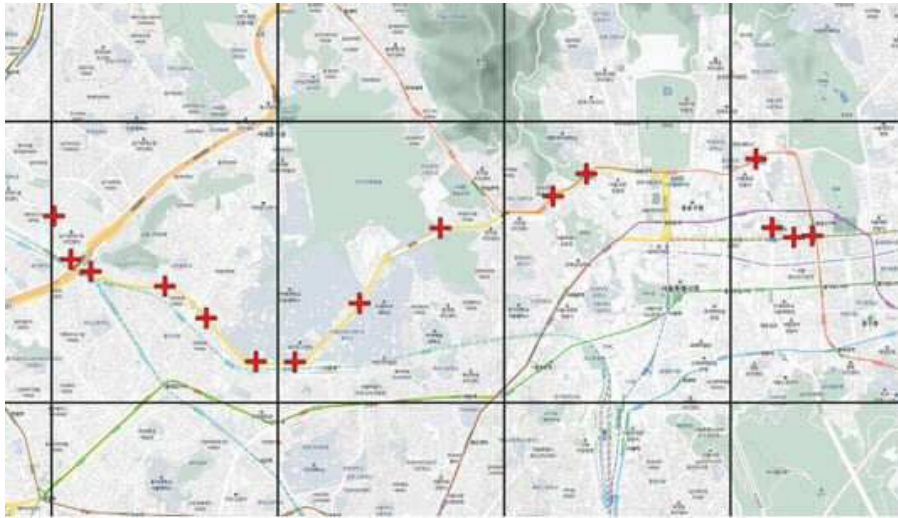
- [0072] 조건 : 처음에 주어진 점들을 보로노이 다각형을 만들어내는 ‘생성점(generating point)’ 이라고 할 때, 분할된 다각형 내부의 임의의 점과 그 다각형이 포함하고 있는 생성점 사이의 거리는 그 다각형 외부의 생성점과의 거리보다 가까워야 한다.
- [0074] 이러한 조건이 만족되도록 분할하기 위해서는 인접한 두 생성점을 선분으로 연결하고 이 선분의 수직이등분선을 긋는다. 이러한 방식으로 수직이등분선을 그려 가면 수직이등분선을 변으로 하는 다각형이 만들어지면서 면이 다각형들로 분할된다. 이때 생기는 다각형을 ‘보로노이 다각형’ 이라고 한다.
- [0075] S300 단계에서 추출된 중심 좌표들이 생성점이 된다. 도 10은 도 9의 중심 좌표들을 이용하여 공간을 분할하는 것을 예시하고 있다.
- [0077] S500 : 택시 수요량 추정 단계
- [0078] 컴퓨팅 장치는 S400 단계에서 분할된 공간을 통해 택시 수요량을 추정한다. 컴퓨팅 장치는 택시 경로 데이터들을 도 11과 같은 그래프 형태의 데이터로 변환하여 구축하고, 그래프 데이터를 입력 데이터로 사용하는 딥러닝 모델을 적용한다.
- [0079] 컴퓨팅 장치는 분할된 공간으로부터 그래프 데이터 구조를 생성한 후 택시 승차 지점과 하차 지점, 그리고 그 운행경로를 이 그래프 데이터 구조로 변환하여 복수의 그래프 데이터를 생성하고, 생성된 그래프 데이터를 학습 데이터로 하여 딥러닝 모델을 학습시킨다. 또한 딥러닝 모델은 학습 데이터를 이용한 학습의 결과, 분할된 공간을 기준으로 지역별 예상 승차 횟수 및 예상 주행 거리 등을 산출할 수 있다. 따라서, 지역별 택시 수요량을 추정할 수 있게 된다.
- [0081] S600 : 택시 승차 지점 추천 단계
- [0082] 컴퓨팅 장치는 S500 단계에서 추정된 지역별 택시 수요량을 기반으로 택시 운전자에게 택시 승차 지점을 추천할 수 있다.
- [0083] 도 12는 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 의해 추정된 택시 수요량을 기반으로 택시 운전자에게 택시 승차 지점을 추천하는 예시적인 유저 인터페이스가 도시된 도면이다.
- [0084] 택시 운전자가 현재 위치를 입력하거나, GPS 등을 통해 현재 위치가 자동 수신된 경우, 택시의 현재 위치 주변에 본 발명의 일 실시예에 따른 공간 분할 기법이 적용되어 분할된 공간이 표시된다. 분할된 공간은 지역별 택시 수요량이 반영되어 각각 다른 색상 처리될 수 있다. 유저 인터페이스는 분할된 공간 별로 예상 승차 횟수, 해당 승차 지점까지의 거리 등을 표시할 수 있다. 유저 인터페이스는 예상 승차 횟수 및 거리를 이용하여 택시 운전자에게 유리한 승차 지점을 순위화하여 표시할 수 있다.
- [0085] 도 12의 예시도를 참조하면, 현재 연세대 정문에 위치한 택시가 맑은 날씨일 때 반경 1km 내에서 다음 택시 승차 지점을 추천받고자할 때, 분할된 공간을 기반으로 다음 택시 승차 지점 세 곳(연세대 세브란스, 연세대 북문, 산울림 소극장)을 추천받을 수 있다. 또한, 같은 조건 하의 택시가 무조건 손님을 태우기보다는 장거리 손님을 태우고 싶을 때, 분할된 공간을 기반으로 산출된 예상 주행 거리를 통해 또 다른 승차 지점을 추천받을 수 있다.
- [0087] 도 13은 종래 기술의 공간 분할 기법과 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 의해 추정된 택시 수요량의 정확도를 비교 예시한 도면이다.
- [0088] 도 13에 도시된 바와 같이, 종래 기술의 격자형 공간 분할 기법에 의하면, 택시 수요량의 추정 오차가 19.36%로 매우 높아서 실제 이를 사용하기가 어려우나, 본 발명의 일 실시예에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법에 의하면, 택시 수요량의 추정 오차가 10.88%로 획기적으로 개선되어 실사용이 가능하다는 효과가

있다.

- [0090] 도 14는 본 발명의 실시예에 따른, 컴퓨팅 장치를 나타내는 도면이다. 도 14의 컴퓨팅 장치(TN100)는 본 명세서에서 기술된 공간 분할 기법을 이용한 택시 수요량 추정방법을 수행하기 위한 장치일 수 있다.
- [0091] 도 14의 실시예에서, 컴퓨팅 장치(TN100)는 적어도 하나의 프로세서(TN110), 송수신 장치(TN120), 및 메모리(TN130)를 포함할 수 있다. 또한, 컴퓨팅 장치(TN100)는 저장 장치(TN140), 입력 인터페이스 장치(TN150), 출력 인터페이스 장치(TN160) 등을 더 포함할 수 있다. 컴퓨팅 장치(TN100)에 포함된 구성 요소들은 버스(bus)(TN170)에 의해 연결되어 서로 통신을 수행할 수 있다.
- [0092] 프로세서(TN110)는 메모리(TN130) 및 저장 장치(TN140) 중에서 적어도 하나에 저장된 프로그램 명령(program command)을 실행할 수 있다. 프로세서(TN110)는 중앙 처리 장치(CPU: central processing unit), 그래픽 처리 장치(GPU: graphics processing unit), 또는 본 발명의 실시예에 따른 방법들이 수행되는 전용의 프로세서를 의미할 수 있다. 프로세서(TN110)는 본 발명의 실시예와 관련하여 기술된 절차, 기능, 및 방법 등을 구현하도록 구성될 수 있다. 프로세서(TN110)는 컴퓨팅 장치(TN100)의 각 구성 요소를 제어할 수 있다.
- [0093] 메모리(TN130) 및 저장 장치(TN140) 각각은 프로세서(TN110)의 동작과 관련된 다양한 정보를 저장할 수 있다. 메모리(TN130) 및 저장 장치(TN140) 각각은 휘발성 저장 매체 및 비휘발성 저장 매체 중에서 적어도 하나로 구성될 수 있다. 예를 들어, 메모리(TN130)는 읽기 전용 메모리(ROM: read only memory) 및 랜덤 액세스 메모리(RAM: random access memory) 중에서 적어도 하나로 구성될 수 있다.
- [0094] 송수신 장치(TN120)는 유선 신호 또는 무선 신호를 송신 또는 수신할 수 있다. 송수신 장치(TN120)는 네트워크에 연결되어 통신을 수행할 수 있다.
- [0096] 한편, 본 발명은 컴퓨터프로그램으로 구현될 수도 있다. 본 발명은 하드웨어와 결합되어, 본 발명에 따른 공간 분할 기법을 이용한 택시 수요량 추정방법을 실행시키기 위하여, 컴퓨터가 판독 가능한 기록매체에 저장된 컴퓨터 프로그램으로 구현될 수 있다.
- [0097] 본 발명의 실시예에 따른 방법들은 다양한 컴퓨터수단을 통하여 판독 가능한 프로그램 형태로 구현되어 컴퓨터로 판독 가능한 기록매체에 기록될 수 있다. 여기서, 기록매체는 프로그램 명령, 데이터 파일, 데이터구조 등을 단독으로 또는 조합하여 포함할 수 있다.
- [0098] 기록매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다.
- [0099] 예컨대 기록매체는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CDROM, DVD와 같은 광 기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치를 포함한다.
- [0100] 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어를 포함할 수 있다.
- [0101] 이러한 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.
- [0103] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 통상의 기술자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

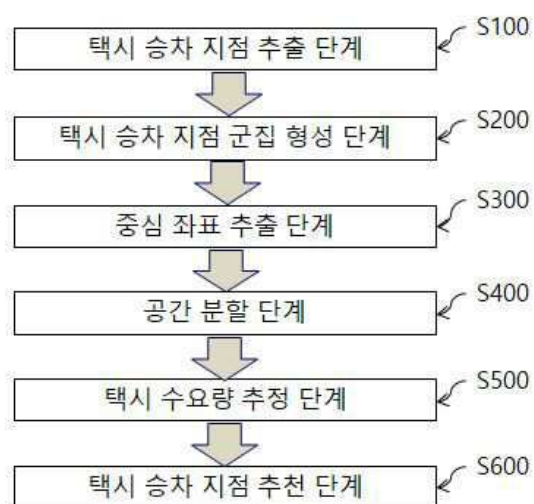
도면1



도면2



도면3

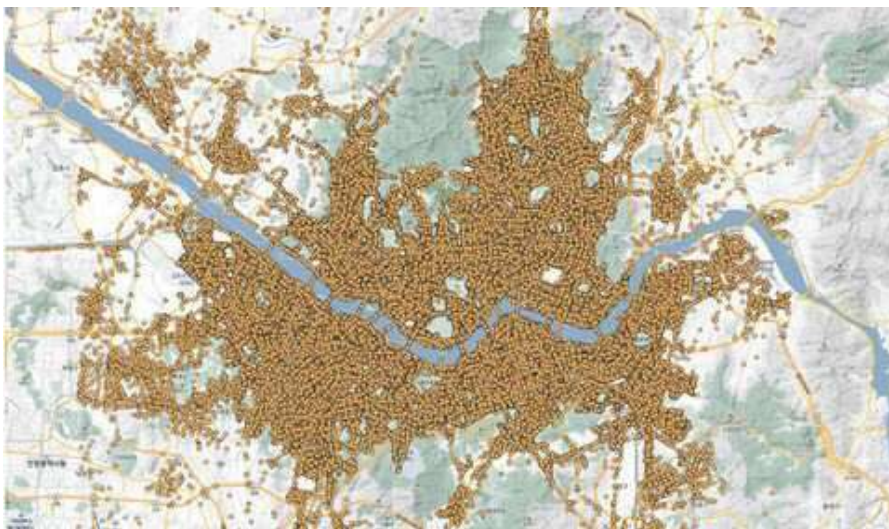


도면4

D1	D2	D3
"1"	27, 1,	27.018462, 37.602442↓
"2"	36, 2,	27.043099, 37.613885↓
"3"	44, 1,	27.044237, 37.619079↓
"4"	59, 2,	27.027212, 37.581139↓

1: 손님 탑승
2: 손님 하차

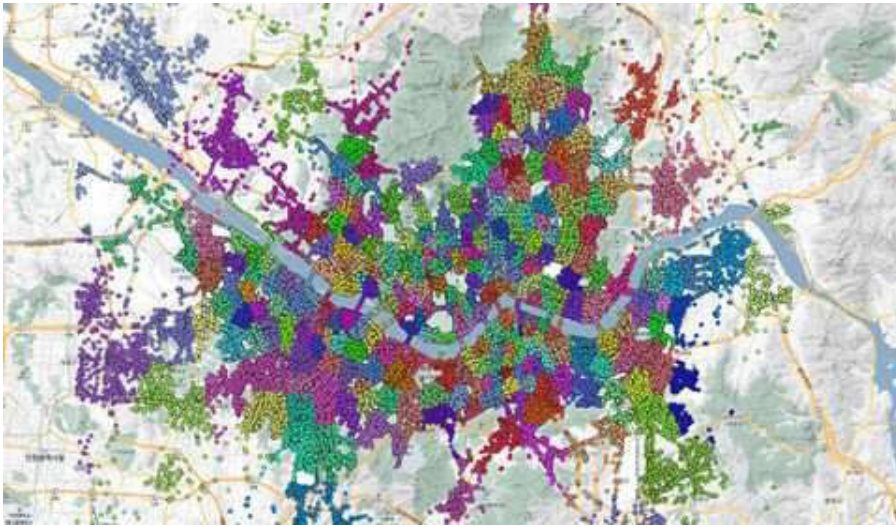
도면5



도면6

V10	V11		V10	V11	cluster
126.9741	37.57075		126.9741	37.57075	104
126.8529	37.52101		126.8529	37.52101	52
127.0548	37.70013		127.0548	37.70013	82
127.0856	37.56056		127.0856	37.56056	34
127.0506	37.57189		127.0506	37.57189	142
127.0113	37.56814		127.0113	37.56814	254
126.9823	37.4251		126.9823	37.4251	135
126.9528	37.48054		126.9528	37.48054	216
126.9011	37.57853		126.9011	37.57853	1
126.8921	37.57742		126.8921	37.57742	1
126.9634	37.52428		126.9634	37.52428	113
126.9575	37.5337		126.9575	37.5337	89

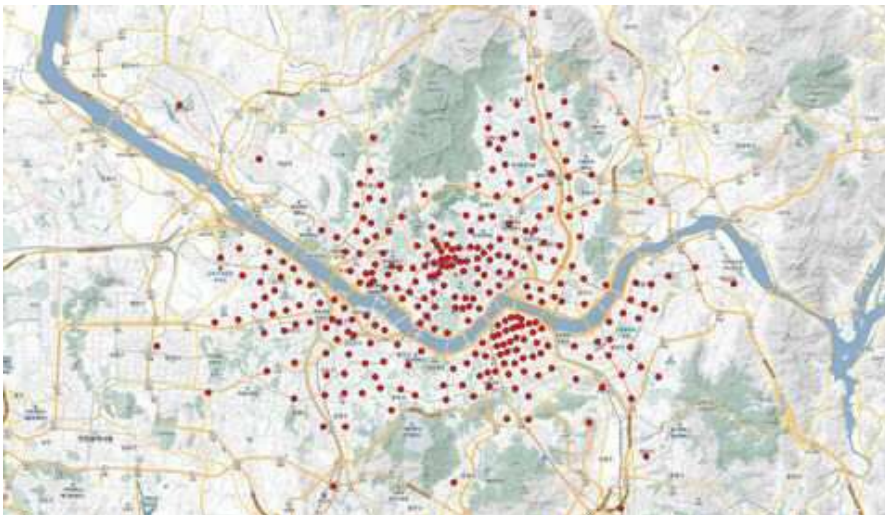
도면7



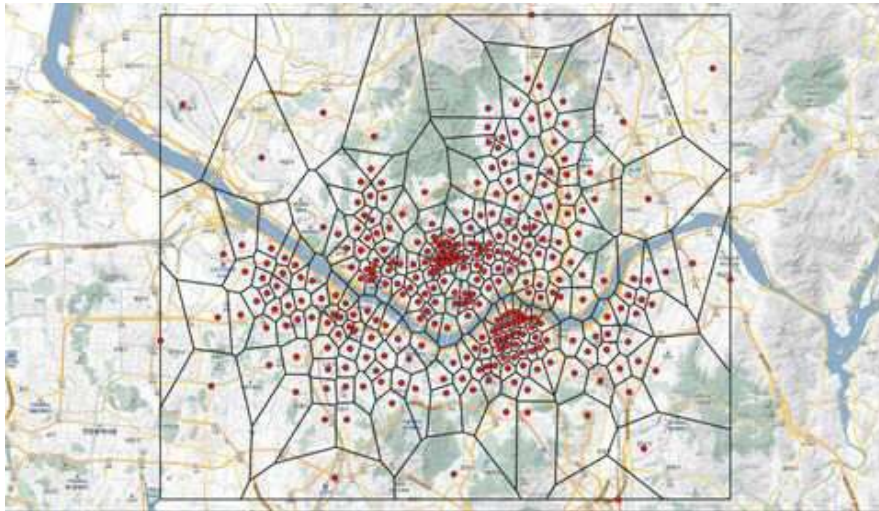
도면8



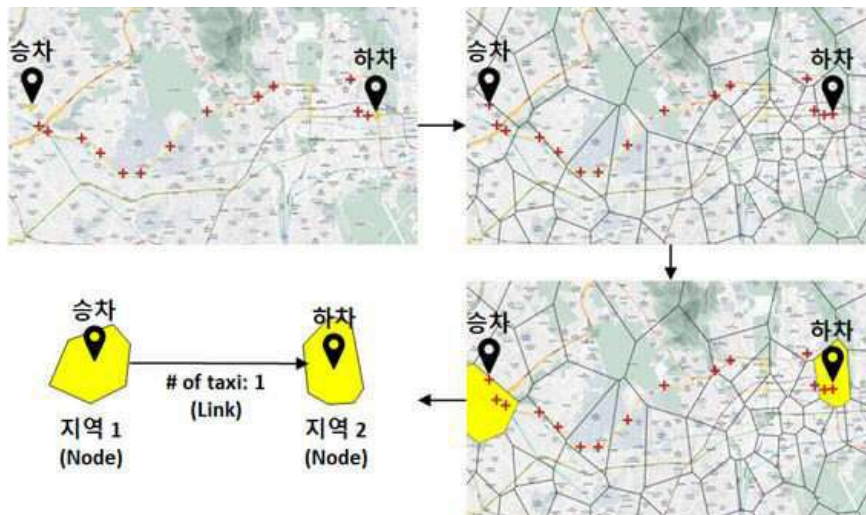
도면9



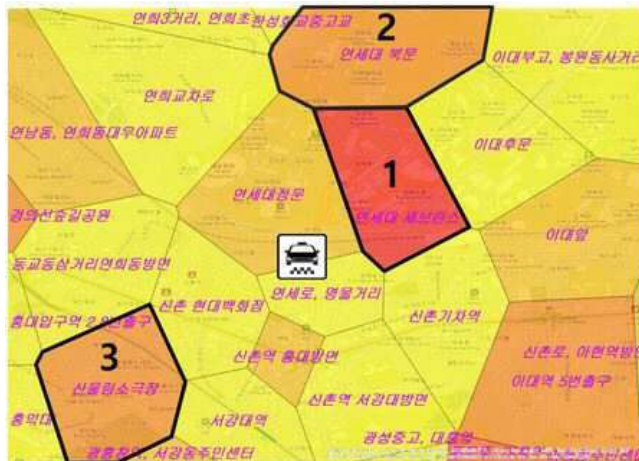
도면10



도면11

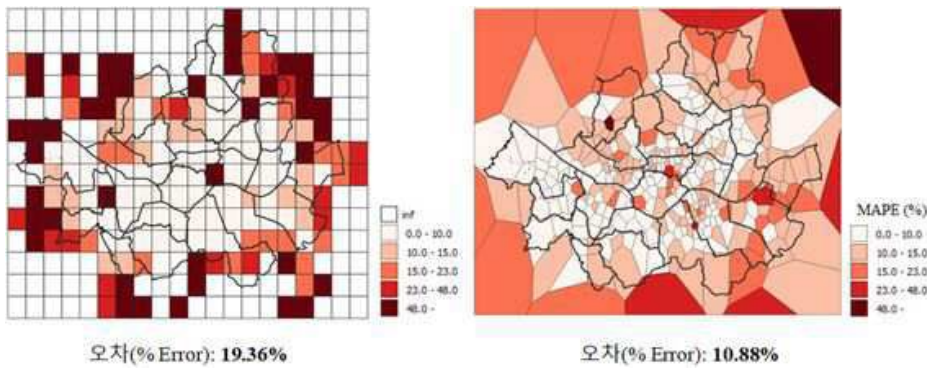


도면12



순위	지역명	평균 탑승 횟수	거리(km)
1	연세대 세브란스	20.24	0.362
2	연세대 북문	10.78	0.955
3	산울림소극장	9.28	0.864

도면13



도면14

