



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0083489
(43) 공개일자 2022년06월20일

(51) 국제특허분류(Int. Cl.)
G01C 21/34 (2006.01) G01C 21/00 (2006.01)
(52) CPC특허분류
G01C 21/3476 (2013.01)
G01C 21/005 (2013.01)
(21) 출원번호 10-2020-0173784
(22) 출원일자 2020년12월11일
심사청구일자 없음

(71) 출원인
서강대학교산학협력단
서울특별시 마포구 백범로 35 (신수동, 서강대학교)
(72) 발명자
박석
서울특별시 종로구 평창12길 16-6, 2동 202호 (평창동, 평창삼호빌라)
박세화
서울특별시 강서구 화곡로64길 50, 201동 802호 (등촌동, 등촌현대2차아파트)
(74) 대리인
유미특허법인

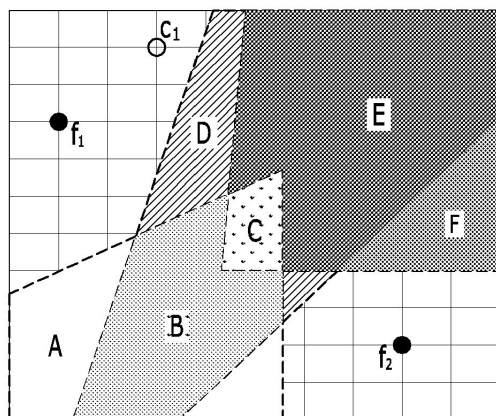
전체 청구항 수 : 총 11 항

(54) 발명의 명칭 차분 프라이버시가 적용된 사용자 위치를 이용한 위치 선택 방법

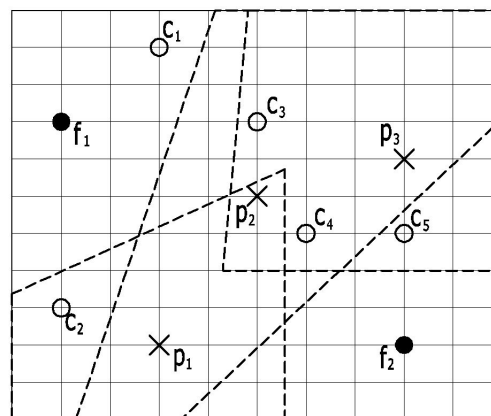
(57) 요약

적어도 하나의 프로세서에 의해 동작하는 위치 선택 시스템이 사용자의 위치를 이용하여 질의에 대한 위치를 선택하는 방법으로서, 질의에 의해 선택될 수 있는 복수의 후보 위치들, 복수의 기존 시설들에 대한 기존 시설 위치들, 그리고 적어도 한 명의 사용자가 대상 영역에 위치하며, 질의를 수신하면 각 후보 위치에 대한 영향 지역을 확인한다. 기존 시설들 각각의 위치를 기초로 대상 영역을 복수의 보로노이 영역들로 구성하고, 임의의 보로노이 영역과 임의의 보로노이 영역 주변의 기존 시설물들을 기초로, 임의의 보로노이 영역에 대한 상한 영역을 구성한 후, 상한 영역에 위치한 사용자 수에 노이즈를 반영하여 수신한 질의에 대한 최적 장소를 선택한다.

대표도



(a)



(b)

(52) CPC특허분류

G01C 21/3438 (2013.01)

G01C 21/3446 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호 108812612

과제번호 2019R1A2C1088126

부처명 과학기술정보통신부

과제관리(전문)기관명 한국연구재단

연구사업명 중견연구자지원사업

연구과제명 개인정보 보호를 위해 차분 프라이버시가 적용된 사용자 정보에 기반한 도로교통망
에서의 최적 장소 선택

기 여 율 1/2

과제수행기관명 서강대학교

연구기간 2020.03.01~2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호 1345315071

과제번호 2017R1D1A1B03036252

부처명 교육부

과제관리(전문)기관명 한국연구재단

연구사업명 이공학개인기초연구지원사업

연구과제명 도로교통망 이동객체에 대한 사용자 의도 추론 기반의 프라이버시를 보호한 예측 질
의처리 기법

기 여 율 1/2

과제수행기관명 서강대학교

연구기간 2017.06.01~2020.11.30

명세서

청구범위

청구항 1

적어도 하나의 프로세서에 의해 동작하는 위치 선택 시스템이 사용자의 위치를 이용하여 질의에 대한 위치를 선택하는 방법으로서,

질의에 의해 선택될 수 있는 복수의 후보 위치들, 복수의 기존 시설들에 대한 기존 시설 위치들, 그리고 적어도 한 명의 사용자가 대상 영역에 위치하며,

질의를 수신하면 각 후보 위치에 대한 영향 지역을 확인하는 단계,

상기 기존 시설들 각각의 위치를 기초로 상기 대상 영역을 복수의 보로노이 영역들로 구성하는 단계,

임의의 보로노이 영역과 상기 임의의 보로노이 영역 주변의 기존 시설물들을 기초로, 상기 임의의 보로노이 영역에 대한 상한 영역을 구성하는 단계,

상기 상한 영역에 위치한 사용자 수에 노이즈를 반영하는 단계, 그리고

적어도 하나의 상한 영역으로 구성된 각 후보 위치별로 노이즈가 반영된 사용자 수를 확인하고, 상기 수신한 질의에 대한 최적 장소를 선택하는 단계

를 포함하는, 위치 선택 방법.

청구항 2

제1항에 있어서,

상기 상한 영역을 구성하는 단계는,

상기 임의의 보로노이 영역에 위치한 제1 기존 시설과 상기 제1 기존 시설 주변의 다른 제2 기존 시설들로 들로네 삼각형을 구성하는 단계,

상기 들로네 삼각형의 제1 외접원을 확인하는 단계,

상기 임의의 보로노이 영역과 중첩되는 제2 외접원을 구하는 단계, 그리고

상기 제1 외접원 및 제2 외접원에 포함된 기존 시설들 각각의 보로노이 셀들을 결합하여 상기 상한 영역으로 구성하는 단계

를 포함하는, 위치 선택 방법.

청구항 3

제2항에 있어서,

상기 보로노이 영역들로 구성하는 단계는,

상기 후보 위치들에 대한 제1 트리, 상기 기존 시설들에 대한 제2 트리, 그리고 프라이버시 예산을 수신하는 단계

를 포함하는, 위치 선택 방법.

청구항 4

제3항에 있어서,

상기 제2 트리는 상기 기존 시설들에 대한 보로노이 영역에 대한 aR-트리로 구성되는, 위치 선택 방법.

청구항 5

제4항에 있어서,

상기 프라이버시 예산을 수신하는 단계는,

상기 입력된 프라이버시 예산을 제1 프라이버시 예산과 제2 프라이버시 예산으로 나누는, 위치 선택 방법.

청구항 6

제5항에 있어서,

상기 제1 프라이버시 예산은 상기 제2 트리를 찾아내기 위해 사용되고, 상기 제2 프라이버시 예산은 보로노이 영역을 분할할 때 사용되는, 위치 선택 방법.

청구항 7

적어도 하나의 프로세서에 의해 동작하는 위치 선택 시스템이 사용자의 위치를 이용하여 질의에 대한 위치를 선택하는 방법으로서,

질의에 의해 선택될 수 있는 복수의 후보 위치들, 복수의 기존 시설들에 대한 기존 시설 위치들, 그리고 적어도 한 명의 사용자가 대상 영역에 위치하며,

질의를 수신하면 각 후보 위치에 대한 영향 지역을 확인하는 단계,

상기 기존 시설들의 위치를 기초로 각 후보 위치마다 보로노이 영역을 구성하고, 각 보로노이 영역간 교차 영역을 확인하는 단계,

임의의 후보 위치에 대해 복수의 보로노이 영역으로 구성된 영향 지역에서, 또 다른 후보 위치에 의한 영향 지역이 중첩되는 중첩 지역과 비 중첩 지역을 분할하는 단계,

상기 중첩 지역에 포함된 적어도 하나의 보로노이 영역에 위치한 사용자 수를 카운트하고, 상기 적어도 하나의 보로노이 영역별로 사용자 수에 노이즈를 반영하여 노이즈 값을 계산하는 단계, 그리고

각 후보 위치별로 노이즈가 반영된 사용자 수를 확인하고, 상기 수신한 질의에 대한 최적 장소를 선택하는 단계를 포함하는, 위치 선택 방법.

청구항 8

제7항에 있어서,

상기 각 보로노이 영역간 교차 영역을 확인하는 단계는,

각 후보 위치별로 상기 기존 시설들과 가장 가까운 점의 집합을 보로노이 영역 분할 기법으로 분할하여 복수의 보로노이 영역들로 생성하는, 위치 선택 방법.

청구항 9

제8항에 있어서,

상기 중첩 지역과 비중첩 지역을 분할하는 단계는,

상기 각 후보 위치의 보로노이 영역 집합, 후보 위치들의 집합, 상기 사용자의 위치, 그리고 프라이버시 예산을 입력으로 수신하는 단계

를 포함하는, 위치 선택 방법.

청구항 10

제9항에 있어서,

상기 노이즈 값을 계산하는 단계는,

상기 보로노이 영역에 위치한 사용자 수 별로 라플라스 노이즈를 추가하는, 위치 선택 방법.

청구항 11

제10항에 있어서,

상기 노이즈 값을 계산하는 단계는,

상기 각 후보 위치별로, 복수의 보로노이 영역 별로 계산된 노이즈 값을 더하여 각 후보 위치별 노이즈가 반영된 사용자 수로 계산하는, 위치 선택 방법.

발명의 설명

기술 분야

[0001] 본 발명은 차분 프라이버시가 적용된 사용자 위치를 이용한 위치 선택 방법에 관한 것이다.

배경 기술

[0002] 최근 모바일 디바이스의 사용과 위치정보 서비스가 보편화된 이후, 관심지점(POI: Point Of Interest) 정보와 사용자들의 위치 정보를 이용한 다양한 서비스가 제공되고 있다. 특히, 지리정보시스템(GIS: Geographic Information System)의 발전으로 인해 지도 검색 서비스가 많이 이용되고 있다.

[0003] 이와 더불어 오픈 플랫폼 서비스(예를 들어, 오픈스트리트 맵 등)는 전 세계 GIS에 대한 오픈 API 를 제공함으로써, 이를 이용한 다양한 위치기반 서비스 개발이 가능해졌다. 이처럼 다양한 플랫폼을 통해 사용자 위치 정보가 쌓여가는 가운데, 사용자 위치 정보를 이용해 사용자의 편의를 향상시키고 새로운 서비스를 효율적으로 처리하기 위한 많은 연구가 진행되고 있다.

[0004] 그 중 하나인 최적장소 선택 질의(optimal location selection query)는 상권 분석이나 공공기관의 설립 계획 등에 사용되는 질의 처리 기법이다. 최적장소 선택 질의란 기존 시설들의 위치와 그 시설들을 사용하는 사용자들의 위치를 고려했을 때, 새로운 시설이 입점하기 위한 최적의 위치를 찾아주는 질의를 말한다.

[0005] 최적장소 선택 질의는 사용자 집합 C와 기존 시설들의 집합 F가 주어졌을 때, 사용자가 선택할 수 있는 후보위치들의 집합 P 중에서 사용자에게 가장 많은 영향을 끼칠 수 있는 위치를 선택해 반환한다. 이 때, 모든 객체 및 사용자는 위치정보를 지니고 있고, 사용자들은 자신의 위치로부터 가장 가까운 시설을 이용한다는 가정을 지닌다.

[0006] 이러한 최적장소 선택 질의는 실용적인 서비스에 적용할 수 있다는 장점을 지닌 반면, 개인 위치 정보의 노출로 인해 많은 문제들이 발생하는 단점이 있다.

발명의 내용

해결하려는 과제

[0007] 따라서, 본 발명은 차분 프라이버시가 적용된 사용자 위치를 이용하여 최대 영향력 위치를 선택하는 위치 선택 방법을 제공한다.

과제의 해결 수단

[0008] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 하나의 특징인 적어도 하나의 프로세서에 의해 동작하는 위치 선택 시스템이 사용자의 위치를 이용하여 질의에 대한 위치를 선택하는 방법으로서,

[0009] 질의에 의해 선택될 수 있는 복수의 후보 위치들, 복수의 기존 시설들에 대한 기존 시설 위치들, 그리고 적어도 한 명의 사용자가 대상 영역에 위치하며, 질의를 수신하면 각 후보 위치에 대한 영향 지역을 확인하는 단계, 상기 기존 시설들 각각의 위치를 기초로 상기 대상 영역을 복수의 보로노이 영역들로 구성하는 단계, 임의의 보로노이 영역과 상기 임의의 보로노이 영역 주변의 기존 시설물들을 기초로, 상기 임의의 보로노이 영역에 대한 상한 영역을 구성하는 단계, 상기 상한 영역에 위치한 사용자 수에 노이즈를 반영하는 단계, 그리고 적어도 하나의 상한 영역으로 구성된 각 후보 위치별로 노이즈가 반영된 사용자 수를 확인하고, 상기 수신한 질의에 대한 최적 장소를 선택하는 단계를 포함한다.

[0010] 상기 상한 영역을 구성하는 단계는, 상기 임의의 보로노이 영역에 위치한 제1 기존 시설과 상기 제1 기존 시설 주변의 다른 제2 기존 시설들로 들로네 삼각형을 구성하는 단계, 상기 들로네 삼각형의 제1 외접원을 확인하는 단계, 상기 임의의 보로노이 영역과 중첩되는 제2 외접원을 구하는 단계, 그리고 상기 제1 외접원 및 제2 외접

원에 포함된 기존 시설들 각각의 보로노이 셀들을 결합하여 상기 상한 영역으로 구성하는 단계를 포함할 수 있다.

- [0011] 상기 보로노이 영역들로 구성하는 단계는, 상기 후보 위치들에 대한 제1 트리, 상기 기존 시설들에 대한 제2 트리, 그리고 프라이버시 예산을 수신하는 단계를 포함할 수 있다.
- [0012] 상기 제2 트리는 상기 기존 시설들에 대한 보로노이 영역에 대한 aR-트리로 구성될 수 있다.
- [0013] 상기 프라이버시 예산을 수신하는 단계는, 상기 입력된 프라이버시 예산을 제1 프라이버시 예산과 제2 프라이버시 예산으로 나눌 수 있다.
- [0014] 상기 제1 프라이버시 예산은 상기 제2 트리를 찾아내기 위해 사용되고, 상기 제2 프라이버시 예산은 보로노이 영역을 분할할 때 사용될 수 있다.
- [0015] 상기 본 발명의 기술적 과제를 달성하기 위한 본 발명의 또 다른 특징인 적어도 하나의 프로세서에 의해 동작하는 위치 선택 시스템이 사용자의 위치를 이용하여 질의에 대한 위치를 선택하는 방법으로서,
- [0016] 질의에 의해 선택될 수 있는 복수의 후보 위치들, 복수의 기존 시설들에 대한 기존 시설 위치들, 그리고 적어도 한 명의 사용자가 대상 영역에 위치하며, 질의를 수신하면 각 후보 위치에 대한 영향 지역을 확인하는 단계, 상기 기존 시설들의 위치를 기초로 각 후보 위치마다 보로노이 영역을 구성하고, 각 보로노이 영역간 교차 영역을 확인하는 단계, 임의의 후보 위치에 대해 복수의 보로노이 영역으로 구성된 영향 지역에서, 또 다른 후보 위치에 의한 영향 지역이 중첩되는 중첩 지역과 비 중첩 지역을 분할하는 단계, 상기 중첩 지역에 포함된 적어도 하나의 보로노이 영역에 위치한 사용자 수를 카운트하고, 상기 적어도 하나의 보로노이 영역별로 사용자 수에 노이즈를 반영하여 노이즈 값을 계산하는 단계, 그리고 각 후보 위치별로 노이즈가 반영된 사용자 수를 확인하고, 상기 수신한 질의에 대한 최적 장소를 선택하는 단계를 포함한다.
- [0017] 상기 각 보로노이 영역간 교차 영역을 확인하는 단계는, 각 후보 위치별로 상기 기존 시설들과 가장 가까운 점의 집합을 보로노이 영역 분할 기법으로 분할하여 복수의 보로노이 영역들로 생성할 수 있다.
- [0018] 상기 중첩 지역과 비중첩 지역을 분할하는 단계는, 상기 각 후보 위치의 보로노이 영역 집합, 후보 위치들의 집합, 상기 사용자의 위치, 그리고 프라이버시 예산을 입력으로 수신하는 단계를 포함할 수 있다.
- [0019] 상기 노이즈 값을 계산하는 단계는, 상기 보로노이 영역에 위치한 사용자 수 별로 라플라스 노이즈를 추가할 수 있다.
- [0020] 상기 노이즈 값을 계산하는 단계는, 상기 각 후보 위치별로, 복수의 보로노이 영역 별로 계산된 노이즈 값을 더하여 각 후보 위치별 노이즈가 반영된 사용자 수로 계산할 수 있다.

발명의 효과

- [0021] 본 발명에 따르면, 사용자 위치에 차분 프라이버시를 적용함으로써, 사용자 위치 즉, 개인 정보를 보호할 수 있다.
- [0022] 또한, 최적 장소 선택 질의 처리 시 차분 프라이버시를 적용할 때 영향지역 중첩 문제로 인한 정확도 하락 문제를 해결할 수 있다.

도면의 간단한 설명

- [0023] 도 1은 종래의 최적 장소 선택 질의의 예시도이다.
- 도 2는 본 발명의 실시예에 따른 위치 선택 시스템의 구조도이다.
- 도 3은 본 발명의 제1 실시예에 따른 위치 선택 방법에 대한 흐름도이다.
- 도 4는 본 발명의 제1 실시예에 따른 보로노이 영역 분할 방법에 대한 예시도이다.
- 도 5는 본 발명의 제2 실시예에 따른 위치 선택 방법에 대한 흐름도이다.
- 도 6은 본 발명의 실시예에 따라 보로노이 영역의 상한 영역을 구성하는 과정에 대한 예시도이다.

발명을 실시하기 위한 구체적인 내용

- [0024] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.
- [0025] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0026] 이하, 도면을 참조로 하여 본 발명의 실시예에 따른 사용자 위치를 이용한 위치 선택 시스템 및 방법에 대해 상세히 설명한다. 본 발명의 실시예에 대해 설명하기 앞서, 종래에 다양한 방법으로 최적 장소를 선택하는 예에 대해 도 1을 참조로 먼저 설명한다.
- [0027] 도 1은 종래의 최적 장소 선택 질의의 예시도이다.
- [0028] 최적장소 선택 질의란 기존 시설들의 위치와 그 시설들을 사용하는 사용자들의 위치를 고려했을 때, 새로운 시설이 입점하기 위한 최적의 위치를 찾아주는 질의를 의미한다. 최적장소 선택 질의는 사용자 집합 C, 기존 시설들의 집합 F가 주어지면, 사용자가 선택할 수 있는 후보 위치들의 집합 P 중에서 사용자에게 가장 많은 영향을 끼칠 수 있는 후보 위치를 선택해 반환한다.
- [0029] 이 때, 모든 객체 및 사용자는 위치정보를 지니고 있고, 사용자들은 자신의 위치로부터 가장 가까운 시설을 이용한다는 가정을 지닌다. 최적장소 선택 질의는 응용 서비스에 따라 각기 다른 목적함수를 가질 수 있는데, 주로 사용자들의 선택 목적에 따라 다음과 같이 세 가지 방법으로 나뉜다.
- [0030] 첫 번째 방법은 상권 분석과 같이 고객 유치를 위해 영향을 받는 사용자를 최대화하는 최적장소 선택 질의 방법(Max-inf)이다. 또 다른 방법으로, 새로운 프랜차이즈 레스토랑의 입점과 같이 모든 고객과의 거리를 줄이기 위해 사용되는 방법(Min-sum)이 있으며, 마지막으로 소방서나 경찰서와 같이 공공기관 설치를 위해 가장 멀리 있는 사용자의 거리를 최소화하기 위해 사용되는 방법(Min-max)이 있다.
- [0031] 도 1의 (a)에 도시된 바와 같이 첫 번째 방법(Max-inf)은, 질의 요청자가 경쟁적인 상권에 위치한 경우, 후보위치들 중에서 최대한 많은 사용자들에게 영향을 끼칠 수 있는 위치를 찾아내길 원한다. 이 때, 모든 장소 객체 및 사용자들이 유클리드 평면상에 위치한 점으로 가정하고, 목적 함수로 두 객체 사이의 거리는 유클리드 거리를 사용하여 수학적 1을 이용하여 주어진 후보 집합 중에서 사용자에게 영향력이 가장 높은 위치를 찾아낸다.

수학적 1

$$MaxInf(P) = \arg \max_{p \in P} \sum_{c \in IS(p)} w(c)$$

[0032]

[0033] 여기서, w(c)는 사용자들의 영향력을 의미하는데, 응용 서비스마다 다르게 적용될 수 있다. 하지만, 근본적으로 이 값이 질의처리 방법에 영향을 끼치지 않기 때문에 1로 가정하여 설명한다.

[0034] 그리고, IS(p)는 사용자들이 최근 최근점 시설을 이용한다는 가정하에 후보위치에 영향을 받는 사용자 집합을 의미한다. 후보 위치에 영향을 받는 사용자 집합은 다음 수학적 2와 같이 정의할 수 있다.

수학적 2

$$IS(p) = \{c \in C | \forall f \in F, d(c, p) \leq d(c, f)\}$$

[0035]

[0036] 도 1의 (a)에서 P₁의 영향을 받는 사용자는 {c₂}이고, P₂의 영향을 받는 사용자들은 {c₃, c₄}이며, P₃의 영향을 받는 사용자들은 {c₃, c₄, c₅}이다. 수학적 1에 따라 Max-inf 방법에서는 P₃가 가장 많은 사용자에게 영향을 끼칠 수 있기 때문에 최적 장소로 반환된다.

[0037] 한편, 도 1의 (b)에 도시된 바와 같이, 두 번째 방법인 Min-sum 방법은 브랜드 경영자의 입장에서 사용자에게

서비스의 편의를 제공하기 위한 새로운 지점을 선택해 반환한다. 기존 시설들이 모두 상호 보완적인 역할을 한다고 가정하기 때문에, 사용자까지의 거리의 합이 가장 작은 후보위치를 찾는 것을 목적으로 한다.

[0038] Min-sum은 Max-inf와 마찬가지로 사용자들은 모두 최근접 시설을 이용한다고 가정하고, Min-sum의 목적 함수는 다음 수학적식 3과 같이 정의된다.

수학적식 3

[0039]
$$MinSum(P) = \arg \min_{p \in P} \sum_{c \in C} w(c) \cdot \min\{d(c, f) | f \in \{p\}\}$$

[0040] Max-inf와 같은 예제에 대해 목적 함수가 Min-sum으로 변경된 경우의 결과는 도 1의 (b)와 같다.

[0041] 그리고, 수학적식 3을 이용해 각각의 후보 위치 중 p_1 에 대한 비용 함수 값을 구하면 $d(f_1, c_1) + d(f_1, c_3) + d(f_2, c_4) + d(f_2, c_5) + d(p_1, c_2) = 8 + 16 + 13 + 9 + 5 = 51$ 이다. 반면, p_2 는 48, p_3 는 51이므로 Min-sum의 경우 p_2 가 최적 장소가 된다.

[0042] 또 다른 방법인 Min-max에서는 사용자들에게 공공 서비스를 제공하는 것을 목적으로 한다. 따라서 각 시설로부터 사용자까지의 거리 중 가장 큰 거리 값이 작아지는 방향으로 후보 시설을 선택한다. Min-sum과 마찬가지로 각각의 시설들은 상호 보완적인 역할을 하며, 사용자들은 최근접 시설을 이용한다고 가정한다. Min-max 방법의 목적 함수는 다음 수학적식 4와 같이 정의된다.

수학적식 4

[0043]
$$MinMax(P) = \arg \min_{p \in P} \{\max_{c \in C} \{w(c) \cdot \min\{d(c, f) | f \in \{p\}\}\}\}$$

[0044] 도 1의 (c)에 도시된 바와 같이 Max-max로 최적 장소를 선택할 경우, p_1 과 p_3 은 모두 거리가 가장 큰 c_2 에 영향을 주지 못하기 때문에 비용 함수가 $d(c_2, f_1) = 25$ 로 유지된다. 반면, p_1 은 c_2 의 최근접 시설이 되기 때문에, 비용 함수 값을 $d(c_2, f_1) = 25$ 에서 $d(c_3, f_1) = 16$ 로 낮추게 된다. 따라서, Min-max의 경우 p_1 이 최적 장소가 된다.

[0045] 이상에서 설명한 바와 같이 종래의 방법으로 최적장소 선택을 정의할 경우, 후보 위치들이 근접해 있는 경우 사용자 정보가 중복으로 사용되기 때문에, 사용자 위치가 노출되는 문제가 발생한다.

[0046] 따라서, 본 발명의 실시예에서는 최적 장소 선택 정의 처리 시, 차분 프라이버시를 적용하여 개인 위치 정보를 보호하면서도 정확도가 높은 장소를 선택할 수 있는 시스템 및 방법을 제시한다. 본 발명의 실시예에서는 설명의 편의를 위하여 정의 기법으로 Max-inf를 예로 하여 설명하나, 반드시 이와 같이 한정되는 것은 아니다.

[0047] 본 발명의 실시예에서는 위치 선택 시스템(100)이 최적 장소 선택 정의에 대해 위치를 선택하기 위해 두 가지 실시예를 예로 하여 설명한다. 본 발명의 실시예에 대해 설명하기 앞서, 본 발명의 실시예에서 언급되는 용어들, 즉, 차분 프라이버시, 민감도, 순차구성, 영향지역에 대해 먼저 다음과 같이 정의한다.

[0048] 차분 프라이버시는 개인 정보보호 기법으로, 임의의 메커니즘 M 과 이벤트 집합 $S \subseteq \text{Range}(M)$ 가 주어졌을 때, $\|D_1 - D_2\|_1 \leq 1$ 를 만족하는 두 이웃 데이터베이스 D_1, D_2 에 대해 아래 수학적식 5를 만족하면 메커니즘 M 은 차분 프라이버시를 만족한다.

수학적식 5

[0049]
$$\Pr[M(D_1) \in S] \leq \exp(\epsilon) \Pr[M(D_2) \in S]$$

[0050] 수학적식 5에서 ϵ 은 프라이버시 보호 예산을 의미하며, 시스템 파라미터로 보호하고자 하는 데이터의 보호 정도

를 의미한다. 또한 $\| \cdot \|_1$ 은 L_1 노름(norm)을 의미하며, 이웃 데이터베이스란 임의의 한 튜플의 존재 유무만이 차이를 지닌 데이터베이스 쌍을 말한다.

[0051] 차분 프라이버시의 개념에서 또 하나의 중요한 파라미터로 질의의 정확도에 영향을 끼치는 민감도(sensitivity)는 메커니즘 M 에 따라 결정되는 값이다. 즉, 이웃 데이터베이스 D_1, D_2 가 주어졌을 때, 메커니즘 M 의 민감도란, 한 튜플의 존재유무에 따라 최악의 경우 M 의 결과에 얼마나 많은 영향을 끼치는가를 나타내는 척도를 의미하며, 다음 수학적 식 6과 같이 표현된다.

수학적 식 6

$$\Delta M = \max_{D_1, D_2} \|M(D_1) - M(D_2)\|$$

[0052]

[0053] 수학적 식 6에서 민감도는 적용하는 메커니즘에 따라 L_1 노름과 L_2 노름 등으로 다르게 정의될 수 있다. 그러나 본 발명의 실시예와 관련된 계산 질의(counting query) 질의의 경우, L_1 노름으로 정의되므로 추후 기술되는 모든 민감도는 특별한 언급이 없는 한 L_1 노름으로 정의된 민감도를 의미한다.

[0054] 차분 프라이버시는 그 정의의 특성상 같은 데이터베이스에 대한 질의가 반복되는 경우 개인 정보가 유출될 가능성이 커지게 된다. 이런 문제를 순차구성이라 하며, 어떤 메커니즘을 사용한다고 하더라도 아래와 같은 성질을 지니게 된다.

[0055] 임의의 두 메커니즘 M_1, M_2 와 각각의 메커니즘 별로 프라이버시 예산 ϵ_1, ϵ_2 가 주어졌을 때, 두 메커니즘의 조합 $M_{1,2}$ 는 $(\epsilon_1 + \epsilon_2)$ -차분 프라이버시를 만족하게 된다. 따라서, 연속되는 메커니즘의 조합은 프라이버시 예산이 선형적으로 증가하게 되고, 전체 메커니즘의 결과가 총 만큼의 프라이버시 예산을 사용하고자 하는 경우에는 각각의 메커니즘에 더 많은 노이즈를 추가해야 한다.

[0056] 차분 프라이버시 기반 최적 장소 선택 질의를 처리하기 위해서는 최적 장소 선택할 때 각각의 목적함수에 맞는 차분 프라이버시를 만족하는 메커니즘을 적용해 질의를 처리해야 한다. 그러나 후보위치들이 근접해 있는 경우에는 사용자 정보가 중복되어 노출되는 문제가 공통적으로 발생하는데 편의를 위해 Max-inf 방법을 통해 이를 설명한다.

[0057] Max-inf 방법에 차분 프라이버시를 적용하기 위한 직관적인 방법은 질의 처리시 라플라스 메커니즘을 적용하는 것이다. 라플라스 메커니즘은 질의 결과에 라플라스 분포로부터 노이즈를 생성해 추가하는 기법으로, 차분 프라이버시를 만족한다.

[0058] Max-inf 방법에 라플라스 메커니즘을 적용하기 위해서는 우선 후보위치가 영향을 끼치는 영역을 구성해야 한다. 이를 후보 위치의 영향 지역(influence region)이라 한다.

[0059] 영향 지역 내에 위치한 사용자들은 모두 후보 지역 p 를 최근접 시설로 하게 되고, 이는 곧 이런 사용자들의 집합이 상술한 수학적 식 2의 $IS(p)$ 임을 의미한다. 따라서 모든 후보 지역마다 영향 지역을 구성한다면, 최적 장소 선택 질의를 수행할 수 있다.

[0060] 도 2는 본 발명의 실시예에 따른 위치 선택 시스템의 구조도이다.

[0061] 도 2에 도시된 바와 같이, 적어도 하나의 프로세서에 의해 동작하는 위치 선택 시스템(100)에서, 본 발명의 동작을 실행하도록 기술된 명령들(instructions)이 포함된 프로그램을 실행한다. 프로그램은 컴퓨터 판독 가능한 저장매체에 저장될 수 있고, 유통될 수 있다.

[0062] 위치 선택 시스템(100)의 하드웨어는 적어도 하나의 프로세서(110), 메모리(120), 스토리지(130), 통신 인터페이스(140)를 포함할 수 있고, 버스를 통해 연결될 수 있다. 이외에도 입력 장치 및 출력 장치 등의 하드웨어가 포함될 수 있다. 위치 선택 시스템(100)은 프로그램을 구동할 수 있는 운영 체제를 비롯한 각종 소프트웨어가 탑재될 수 있다.

[0063] 프로세서(110)는 위치 선택 시스템(100)의 동작을 제어하는 장치로서, 프로그램에 포함된 명령들을 처리하는 다양한 형태의 프로세서일 수 있고, 예를 들면, CPU(Central Processing Unit), MPU(Micro Processor Unit),

MCU(Micro Controller Unit), GPU(Graphic Processing Unit) 등 일 수 있다.

- [0064] 프로세서(110)는 차분 프라이버시 기반 최적 장소 선택 질의를 처리하기 위해, 유클리드 공간 상의 임의의 점에 대해 후보 위치와 영향 지역을 구성한다. 임의의 점에 대해 후보 위치와 영향 지역을 구성하는 방법에 대해서는 이후 상세히 설명한다.
- [0065] 그리고 프로세서(110)는 영향 지역을 보로노이 영역 분할(voronoi region partition) 기법을 이용하여 복수의 보로노이 영역으로 분할한다. 또한, 프로세서(110)는 개인 위치 정보를 보호하기 위해 적어도 하나의 중복 후보 위치를 확인하고, 확인한 중복 후보 위치에 차분 프라이버시를 적용한다.
- [0066] 또한, 프로세서(110)는 R-트리 구조 기반의 확장 보로노이 영역 가지치기 기법을 이용하여, 후보 위치의 영향 지역에 대한 상한 영역을 계산한다. 프로세서(110)가 상한 영역을 계산할 때, 들로네 삼각화(Delaunay triangulation)를 이용하며, 이에 대해서는 이후 상세히 설명한다.
- [0067] 메모리(120)는 본 발명의 동작을 실행하도록 기술된 명령들이 프로세서(110)에 의해 처리되도록 해당 프로그램을 로드한다. 메모리(120)는 예를 들면, ROM(read only memory), RAM(random access memory) 등 일 수 있다. 스토리지(130)는 본 발명의 동작을 실행하는데 요구되는 각종 데이터, 프로그램 등을 저장한다. 통신 인터페이스(140)는 유/무선 통신 모듈일 수 있다.
- [0068] 이상에서 설명한 위치 선택 시스템(100)이 최적 장소 선택 질의를 처리하여 최적 장소를 선택하는 방법에 대해 도 3 내지 도 6을 참조로 설명한다. 도 3에서는 보로노이 영역 분할 기법을 이용하여 최적 장소 선택 질의를 처리하는 제1 실시예에 따른 방법을 설명하며, 도 5에서는 R-트리 구조 기반 확장 보로노이 영역 가지치기 기법을 이용하여 최적 장소 선택 질의를 처리하는 제2 실시예에 따른 방법을 설명한다.
- [0069] 도 3은 본 발명의 제1 실시예에 따른 위치 선택 방법에 대한 흐름도이다.
- [0070] 종래의 기술을 통해 순차 구성을 적용해 최적 장소 선택 질의를 수행할 경우, 정확도가 지수적(exponentially)으로 감소하는 문제가 발생한다. 이를 해결하기 위해 본 발명의 실시예에서는 보로노이 영역을 분할해 병렬 구성(parallel composition)을 적용하는 보로노이 영역 분할 방법(VPM: Voronoi region Partition Method)을 제안한다. 여기서, 보로노이 영역은 복수의 기존 시설들이 위치한 2차원 평면 상의 임의의 점으로부터 기존 시설까지의 거리가 가장 가까운 점의 집합으로 2차원 평면을 분할했을 때의 영역을 의미한다.
- [0071] 도 3에 도시된 바와 같이 질의에 대한 최적 위치를 반환하는 과정에서 보로노이 영역 분할 방법을 적용하기 위해, 위치 선택 시스템(100)은 먼저 질의에 대한 최적 장소를 선택해야 할 유클리드 공간 즉, 2차원의 좌표 평면으로 구성된 전체 영역(이하, 설명의 편의를 위하여 '대상 영역'이라 지칭함)에서 후보 위치를 확인한다(S100). 여기서 후보 위치는 대상 영역에 임의의 위치에 기 설정되어 있다.
- [0072] 위치 선택 시스템(100)은 각 후보 위치를 기초로 영향 지역들을 생성한다(S110). 여기서, 영향 지역은 각 후보 위치 별로, 어느 하나의 후보 위치와 기존 시설들을 결합하여 만든 보로노이 영역을 해당 후보 위치의 영향 지역이라 한다. 즉, 2차원 평면 상의 임의의 점들 중에서 후보 위치까지의 거리가 가장 가까운 점들의 집합을 영향 지역이라 할 수 있다. 위치 선택 시스템(100)이 후보 위치를 기초로 영향 지역을 생성하는 방법은 다양한 방법으로 생성할 수 있으므로, 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다.
- [0073] 위치 선택 시스템(100)은 각 후보 위치마다 기존 시설들과 보로노이 영역을 구성하고, 보로노이 영역 간 교차 영역을 확인한다(S120). 그리고, 위치 선택 시스템(100)은 각 후보 위치들의 보로노이 영역 집합, 후보 위치들의 집합, 사용자 위치, 프라이버시 예산을 입력으로 수신한다(S130).
- [0074] 위치 선택 시스템(100)은 S130 단계에서 수신한 각 후보 위치의 영향 지역들 중, 다른 후보 위치에 의해 확인된 영향 지역과 중첩되는 중첩 지역과 비 중첩 지역을 분할한다(S140). 그리고 위치 선택 시스템(100)은 확인한 중첩 지역의 보로노이 영역에 위치한 사용자 수를 카운트한다(S150).
- [0075] 위치 선택 시스템(100)은 중첩 지역의 보로노이 영역에 사용자 수를 카운트한 후, 각 보로노이 영역별로 사용자 수에 노이즈를 반영하여 보로노이 영역별 노이즈 값을 계산한다(S160). 본 발명의 실시예에서는 라플라스 메커니즘을 적용하여 중첩 지역의 보로노이 영역별로 위치한 사용자 수에 노이즈를 반영하는 것을 예로 하여 설명하며, 라플라스 메커니즘은 이미 알려진 것으로 본 발명의 실시예에서는 상세한 설명을 생략한다.
- [0076] 위치 선택 시스템(100)은 중첩 지역의 보로노이 영역별 노이즈 값을 더하여 해당 후보 위치에 대한 노이즈 값을 계산하고, 각 후보 위치에 대해 계산된 노이즈 값을 기초로 복수의 후보 영역 중 최적 장소를 선택하여 제공한

다(S170).

[0077] 도 4는 본 발명의 제1 실시예에 따른 보로노이 영역 분할 방법에 대한 예시도이다.

[0078] 도 4에 도시된 바와 같이, 위치 선택 시스템(100)은 모든 후보 위치에 대해 영향 지역을 계산한다. 위치 선택 시스템(100)은 후보 위치들의 보로노이 영역 집합과 후보 위치들의 집합, 사용자 위치와 프라이버시 예산을 입력으로 받아, 다음 표 1의 알고리즘을 수행한다.

[0079] 표 1의 알고리즘은 보로노이 영역마다 노이즈 카운트를 생성하는 과정을 나타낸 알고리즘이고, 표 2의 알고리즘은 분할 영역을 계산하는 과정에 대한 알고리즘이다.

표 1

Input: V - Voronoi regions of potential locations, P - potential locations, C - user locations, ϵ - privacy budget
Output: $p_r \in P$ - near-optimal location
1: $PQ \leftarrow$ priority queue sorted by user count in decreasing order
2: **for** $p_i \in P$ **do**
3: $OP \leftarrow []$
4: $v_i \leftarrow V[p_i]$
5: **for** $v_j \in V$ **do**
6: **if** v_i intersects with v_j **then**
7: add v_j to OP
8: $L \leftarrow GetPV(v_i, 0, OP, V)$
9: $cnt_i \leftarrow 0$
10: **for** $pv_j \in L$ **do**
11: $cnt \leftarrow$ count the number of users that pv_j contains
12: $cnt_i \leftarrow cnt_i + cnt + Laplace(\frac{1}{\epsilon})$
13: $PQ.enqueue(cnt_i, p_i)$
14: **Return** $PQ.top$

[0080]

표 2

Input: pv - Partitioned influence region, j - index of overlapped potential location set, OP - overlapped potential location set, V - Voronoi regions of potential locations
Output: $PV(p)$ - A set of partitioned influence regions
1: $L \leftarrow []$
2: **if** j is greater than $OP.size$ **then**
3: insert pv into L
4: **else**
5: $p_j \leftarrow OP[j]$
6: $v_j \leftarrow V[p_j]$
7: **if** pv intersects with v_j **then**
8: $pv1 \leftarrow pv.intersection(v_j)$
9: $pv2 \leftarrow pv.difference(v_j)$
10: $L1 \leftarrow GetPV(pv1, j + 1, OP, V)$
11: $L2 \leftarrow GetPV(pv2, j + 1, OP, V)$
12: $L \leftarrow L1 + L2$
13: **else**
14: $L \leftarrow GetPV(pv, j + 1, OP, V)$
15: **Return** L

[0081]

- [0082] 위치 선택 시스템(100)은 표 1의 알고리즘을 이용하여 먼저 후보 위치 별로 교차 영역이 발생하는 중복 후보 위치 집합을 구한 후, 표 2의 알고리즘을 호출하여 분할 영역을 계산한다.
- [0083] 분할 영역을 구하면 위치 선택 시스템(100)은 분할된 영역마다 노이즈 카운트를 생성한 후, 가장 큰 노이즈 카운트를 지닌 후보 위치를 반환한다. 표 1의 알고리즘의 5~7줄은 중복 후보 위치 집합을 구하는 부분이고, 8번째 줄은 분할 영역을 계산하기 위해 표 2의 알고리즘을 호출하는 부분이다.
- [0084] 마지막으로 분할 영역을 반환 받아 10~12번째 줄까지 노이즈 카운트를 생성한다. 표 2의 알고리즘은 분할-정복 방식으로 분할 영역을 계산하는 코드로 재귀적으로 호출된다. 여기서, p_v 는 현재까지 분할된 영역, j 는 중복 후보 위치 집합에서 몇 번째 후보 위치까지 중복 영역을 계산했는지에 대한 인덱스를 의미한다.
- [0085] 위치 선택 시스템(100)은 중복 후보 위치 집합 내의 모든 후보 위치들의 중복 영역을 계산할 때까지 표 2의 알고리즘을 반복한다. 그리고, 표 2의 알고리즘에서 8번째 줄에서 교차 영역을 계산하고, 9번째 줄에서 교차 영역과 그렇지 않은 영역으로 나눈다. 그리고 나서, 위치 선택 시스템(100)은 10~11번째 줄에서 각각의 경우에 대해 다른 후보 위치들과의 중복 영역을 계산하기 위해 알고리즘을 재귀적으로 호출한다. 마지막으로 분할된 영역들을 결합해서 반환한다.
- [0086] 상술한 도 1의 (a)에서 p_1 의 경우를 예로 하여 설명하면, 위치 선택 시스템(100)은 도 4의 (a)에 도시된 바와 같이 p_1 에 대한 보로노이 영역(A~C)을 구성한다. 그리고, A, B, C 영역에 포함된 사용자를 구하면 도 4의 (b)와 같이 $A:\{c_2\}$, $B:\{\}$, $C:\{\}$ 이 된다. 그리고 p_1 의 노이즈 값은 $(2+n_A)+(0+n_B)+(0+n_C)=(2+n_A+n_B+n_C)$ 가 된다. 여기서 n_X 는 각각 영역 X에 추가된 노이즈를 의미한다.
- [0087] 이렇게 보로노이 영역을 분할해 각각의 보로노이 영역마다 노이즈 카운트를 생성할 경우, 후보 위치의 영향 지역끼리 교차되지 않기 때문에 영향 지역 중복 문제를 해결할 수 있다.
- [0088] 또한, 분할된 보로노이 영역에 노이즈를 추가할 경우, 차분 프라이버시의 병렬 구성을 적용할 수 있다. 보로노이 영역 분할 방법에서 분할된 영역의 수를 N 이라 하면, 병렬 구성의 특성에 따라 추가된 전체 노이즈는 $\sqrt{N}$ 에 비례하여 증가하기 때문에 정확도 면에서 순차구성보다 더 효과적이다.
- [0089] 다음은 사용자에게 영향력이 적은 후보 위치를 사전에 걸러내어 불필요한 계산 비용을 줄이기 위하여, 프라이버시 예산을 나눠 사용함으로써 정확도 높고 효율적인 질의 처리 방법을 제2 실시예로 하여 설명한다. 즉, 보로노이 영역 분할 기법이 아닌 R-트리 구조 기반 확장 보로노이 영역 가지치기 기법을 이용하여 최적 위치를 선택하는 방법에 대해 도 5를 참조로 설명한다.
- [0090] 도 5는 본 발명의 제2 실시예에 따른 위치 선택 방법에 대한 흐름도이다.
- [0091] 도 5에 도시된 바와 같이, 위치 선택 시스템(100)은 먼저 대상 영역에서 기존 시설들에 대한 위치를 각각 확인한 후(S200), 대상 영역에 위치한 후보 위치를 기초로 영향 지역들을 확인한다(S210).
- [0092] 그리고 각 기존 시설들을 기초로 보로노이 영역들로 분할한다(S220). 이때, 위치 선택 시스템(100)이 기존 시설들을 기초로 보로노이 영역들을 분할하는 방법은 다양하게 수행될 수 있으므로 본 발명의 실시예에서는 어느 하나의 방법으로 한정하지 않는다. 그리고 후보 위치는 복수의 보로노이 영역 중 어느 하나의 영역에 포함된다.
- [0093] 위치 선택 시스템(100)은 후보 위치들에 대한 제1 트리(R_p), 각 기존 시설들에 대한 제2 트리(R_f), 그리고 프라이버시 예산을 입력으로 수신한다(S230). 여기서 제1 트리와 제2 트리의 형태를 어느 하나의 형태로 한정하지 않는다.
- [0094] 그리고 위치 선택 시스템(100)은 수신한 프라이버시 예산을 제1 프라이버시 예산(ϵ_1)과 제2 프라이버시 예산(ϵ_2)으로 나눈다(S240). ϵ_1 은 영향력이 적은 후보 위치 즉, 제2 트리인 R_f 를 찾아내기 위해 사용하고, ϵ_2 는 질의 처리 과정에서 보로노이 영역 분할 기법을 적용할 때 사용한다. 본 발명의 실시예에서는, 위치 선택 시스템(100)이 프라이버시 예산을 제1 프라이버시 예산과 제2 프라이버시 예산으로 나누는 방법, 그리고 나누는 비율을 어느 하나로 한정하여 설명하지 않는다.
- [0095] 위치 선택 시스템(100)은 복수의 보로노이 영역들 중 상한 영역을 구성하고자 하는 임의의 보로노이 영역을 선택한다(S250). 위치 선택 시스템(100)이 임의의 보로노이 영역을 선택하는 방법은 다양하며, 모든 보로노이 영

역들에 대해 상한 영역이 구성되기 때문에 선택되는 보로노이 영역을 어느 하나로 특정하여 설명하지 않는다.

- [0096] 위치 선택 시스템(100)은 선택한 보로노이 영역 주변의 보로노이 영역에 위치한 다른 기존 시설들과 들로네 삼각형을 구성한다(S260). 즉, 위치 선택 시스템(100)은 선택한 보로노이 영역과 중첩되는 들로네 삼각형의 외접원을 구한다.
- [0097] 여기서 들로네 삼각형은 보로노이 다이어그램이 구성된 이산 포인트 집합이 있을 때, 보로노이 영역의 정점을 기준으로 이웃한 세 포인트를 연결해 삼각형을 만들고, 그 삼각형 외접원 내부에는 다른 포인트가 들어오지 않도록 평면을 분할하는 방법이다. 이 때, 이 외접원의 중심이 보로노이 영역의 정점이 된다. 만약 새로운 점이 입력된다면, 입력된 점이 외접원 내에 위치하게 되는 모든 들로네 삼각형들이 영향을 받게 되고, 이 삼각형과 관련된 포인트들이 새로 입력된 점의 보로노이 이웃이 된다.
- [0098] 위치 선택 시스템(100)은 선택한 보로노이 영역의 상한 영역을 구성한다(S270). 즉, S260 단계에서 구한 외접원들과 관계된 기존 시설들의 보로노이 셀들을 결합하여 선택한 보로노이 영역의 상한 영역을 구성한다. 이때, 구성한 상한 영역은 선택한 보로노이 영역에 위치한 후보의 영향 지역보다 크게 구성된다. 그리고 후보 지역이 기존 시설들의 보로노이 셀 내에 다른 곳에 위치하더라도, 영향 지역이 상한 영역보다 작다.
- [0099] 위치 선택 시스템(100)은 S270 단계에서 구성한 상한 영역에 위치한 사용자 수에 노이즈를 반영한다(S280). 이를 위해, 위치 선택 시스템(100)은 R-트리를 두 개 이용하는데, 후보 위치에 대한 R-트리(R_p)와 기존 시설 위치에 대한 R-트리(R_f)로 구분할 수 있다.
- [0100] 여기서, 기존 시설 위치에 대한 R-트리인 R_f 는 기존 시설들의 보로노이 영역에 대한 aR-트리로 구성된다. aR-트리는 R-트리의 변형된 형태로, 노드들마다 어그리게이션(aggregation) 값들을 지니고 있다. 본 발명의 실시예에서는 노이즈 카운트의 상한 값을 지니고 있는 것으로 예를 들어 설명한다..
- [0101] 위치 선택 시스템(100)은 S280 단계에서 각 후보 위치의 상한 영역별로 노이즈가 반영된 사용자 수를 기초로 최적 장소를 선택하여 반환한다(S290).
- [0102] 다음은 보로노이 영역의 상한 영역을 구성하는 과정에 대해 도 6을 참조로 설명한다.
- [0103] 도 6은 본 발명의 실시예에 따라 보로노이 영역의 상한 영역을 구성하는 과정에 대한 예시도이다.
- [0104] 도 6의 (a)에 도시된 바와 같이 위치 선택 시스템(100)이 기존 시설들로 보로노이 다이어그램을 구성한다. 이는 임의의 후보 위치 p_1 로 상한 영역을 구성하게 되면 해당 후보 위치의 영향 지역이 기존 시설 f_2 의 영향 지역보다 넓다는 것을 보여주기 위해 도시한 것이다. 본 발명의 실시예에서는 복수의 후보 위치 중 p_1 을 예로 하여 설명하며, 복수의 기존 시설 중 f_2 의 상한 영역을 구성하는 것을 예로 하여 설명한다.
- [0105] 도 6의 (b)에 도시된 바와 같이, 음영으로 표시된 보로노이 영역은 상한 영역을 구성하려는 대상이 되는 기존 시설의 보로노이 영역을 표시한 것이다.
- [0106] 위치 선택 시스템(100)은 선택한 기존 시설 f_2 의 상한 영역을 구성하기 위해 다음의 단계를 거친다. 먼저, 주변의 다른 기존 시설들과 들로네 삼각형을 구성한다.
- [0107] 이를 위해 위치 선택 시스템(100)은 들로네 삼각형의 외접원을 구성하는데, 도 6의 (c)에 도시된 바와 같이 f_2 의 보로노이 영역과 중첩되는 외접원을 구한다. 그리고, 위치 선택 시스템(100)은 각 들로네 삼각형들의 외접원을 구하고, 도 6의 (c)에 도시된 음영으로 표시된 영역과 중첩되는 외접원들을 찾는다.
- [0108] 그리고, 위치 선택 시스템(100)은 도 6의 (d)에 도시된 바와 같이, f_2 의 상한 영역을 구성한다. 이때, 도 6의 (c)에서 구한 외접원들과 관계된 기존 시설들의 보로노이 셀들을 결합하여 상한 영역으로 구성한다. 위치 선택 시스템(100)이 구성한 상한 영역은 도 6의 (e)에 도시된 바와 같이, 후보 위치의 영향 지역(빛금 영역, ①)보다 큰 것을 알 수 있다.
- [0109] 이때, 새로운 후보 위치가 점으로 입력되면, 입력된 점이 외접원 내에 위치하게 되는 모든 들로네 삼각형들이 영향을 받게 된다. 이 삼각형과 관련된 포인트들이 새로 입력된 점의 보로노이 이웃이 된다. 위치 선택 시스템(100)은 후보 보로노이 이웃을 다음 수학적 식 7과 같이 정의한다.

수학식 7

$$[0110] \quad CVN(p) = \{f' \in t | t \in DT \wedge Disk(t) \cap V(f) \neq \emptyset\}$$

[0111] 여기서, DT는 기존 시설들의 집합으로 구성된 들로네 삼각형 집합, Disk()를 들로네 삼각형의 외접원이라고 하면, 후보 보로노이 이웃은 기존 시설들의 부분집합으로 정의된다.

[0112] 후보 보로노이 이웃은 임의의 후보 위치 p가 V(f)에 위치한 경우 영향을 받게 되는 기존 시설들의 상한 집합이다. 따라서, 후보 위치 p의 영향 지역은 후보 보로노이 이웃에 포함된 기존 시설들의 보로노이 영역보다 커질 수 없다. 따라서, CVN(p)에 포함된 기존 시설들의 보로노이 영역 집합을 확장 보로노이 영역이라 하고, ϵ_1 을 사용해 노이즈 카운트를 생성한다. 위치 선택 시스템(100)이 노이즈 카운트를 생성할 때, 다음 수학식 8 및 수학식 9를 이용한다.

수학식 8

$$[0113] \quad ncount(f) = |IS(f)| + Lap(1/\epsilon_1)$$

수학식 9

$$[0114] \quad \sum_{f' \in CVN(f)} ncount(f')$$

[0115] 수학식 9는 라플라스 노이즈 부분에 중복 후보 위치 집합의 크기 대신 1이 들어갔다는 것을 제외하고는 상술한 수학식 6과 같은 형태를 지닌다.

[0116] 이렇게 모든 기존 시설의 보로노이 영역에 대한 후보 보로노이 이웃을 찾고 노이즈 카운트를 생성한 후에는, 불필요한 후보 위치들을 질의 처리 과정 중에 가지치기를 수행할 수 있다. 확장 보로노이 영역 방법에서는 aR-tree를 이용해 기존 시설들의 인덱스 R_t 를 구성하고, 최적 우선 탐색 방식(best first search)으로 R_t 를 탐색하며 질의 처리를 수행한다.

[0117] R_t 의 말단 노드에는 기존 시설들의 보로노이 영역과 노이즈 카운트의 상한 값들 중 가장 큰 값을 저장한다. 그리고 루트 노드 및 중간 노드에는 자신의 자식 노드들에 대한 상한 값을 저장한다. 각 노드 별 노이즈 카운트 상한 값은 노드가 루트 노드를 포함한 중간 노드인 경우에는 수학식 10과 같이 계산되고, 그 외의 경우에는 수학식 11과 같이 계산된다.

수학식 10

$$[0118] \quad UBncount(N_f) = \max_{f \in N_f.children} ncount(f)$$

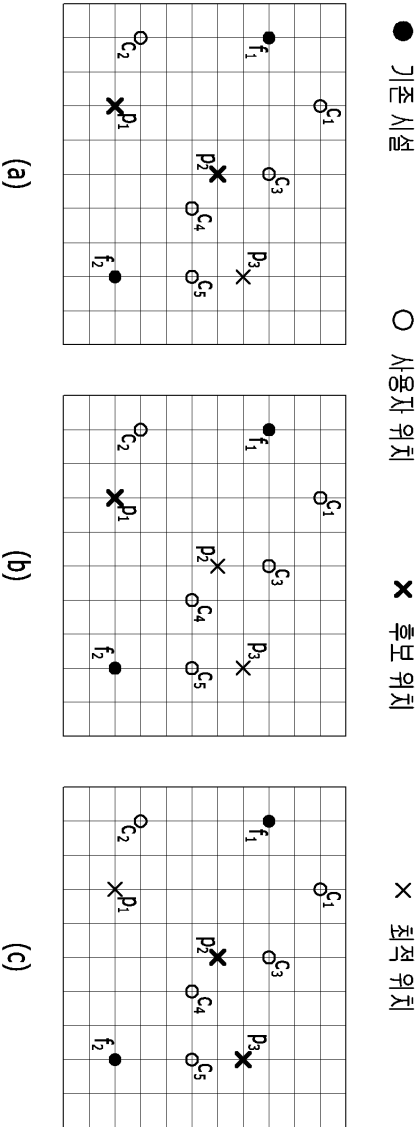
수학식 11

$$[0119] \quad UBncount(N_f) = \max_{N'_f \in N_f.children} UBncount(f)$$

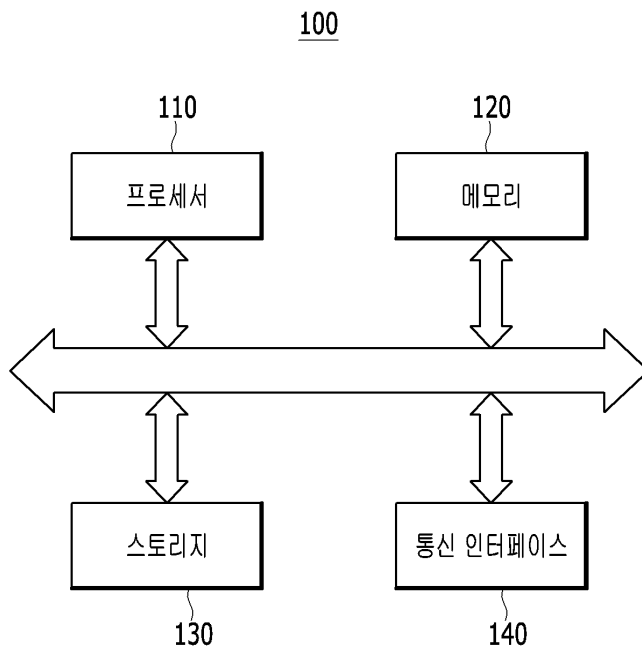
- [0120] 보로노이 영역 기반의 질의 처리 방법은 후보 위치들의 R-트리 R_p 와 사용자 위치에 대한 R-트리 R_c , 마지막으로 기존 시설들에 대한 aR-트리 R_f 의 세 개의 R-tree 구조를 사용한다. VEM(Voronoi envelope based filtering method)은 기본적으로 R_f 를 탐색하면서 관련된 R_p 의 중간 노드들을 찾아나가고, 그와 동시에 노이즈 카운트 상한 값이 현재까지 계산한 최적 카운트 보다 낮은 경우에는 가지치기 한다.
- [0121] 만약 말단 노드에 도달한 경우에는 관련된 후보 위치들에 대해 보로노이 기반으로 영역을 분할한다. $E_p(E_f)$ 를 $R_p(R_f)$ 의 노드 또는 객체들이라고 할 때, 전체적인 질의 처리 과정을 노드들의 타입에 따라 아래의 4가지 경우로 나눌 수 있다.
- [0122] i) E_p , E_f 가 모두 R-트리의 노드일 때,
- [0123] ii) E_p 는 후보 위치 객체이고, E_f 는 R-트리의 노드일 때,
- [0124] iii) E_p 는 R-트리 노드이고, E_f 는 기존 시설의 보로노이 영역일 때,
- [0125] iv) E_p 는 후보 위치 객체이고, E_f 는 기존 시설의 보로노이 영역일 때,
- [0126] 위치 선택 시스템(100)은 우선순위 큐를 이용해 위의 네 가지 경우에 후보위치들과 관련된 후보 보로노이 이웃을 찾아나가며, 동시에 노이즈 카운트 상한 값을 최적 카운트 값과 비교해 큐에 삽입할지 말지를 결정한다.
- [0127] i)의 경우에는 E_p , E_f 의 자식 노드들을 탐색하고, 그들의 MBR들이 교차되는 자식 노드 쌍들을 찾아 노이즈 카운트 상한이 최적 카운트 값 보다 큰 경우에만 큐에 삽입한다.
- [0128] ii)의 경우에는 E_f 의 자식 노드들을 탐색하고 E_p 를 포함하고 있는 자식 노드를 찾아 노이즈 카운트 상한이 최적 카운트 값 보다 큰 경우에만 큐에 삽입한다.
- [0129] iii)의 경우에는 E_p 의 자식 노드들을 탐색하고 E_f 의 노이즈 카운트 상한이 최적 카운트 값 보다 큰 경우에만 큐에 삽입한다.
- [0130] 마지막으로 iv)의 경우에는 보로노이 기반 영역 분할을 수행하고, 필요한 경우 최적 카운트 값을 갱신한다.
- [0131] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

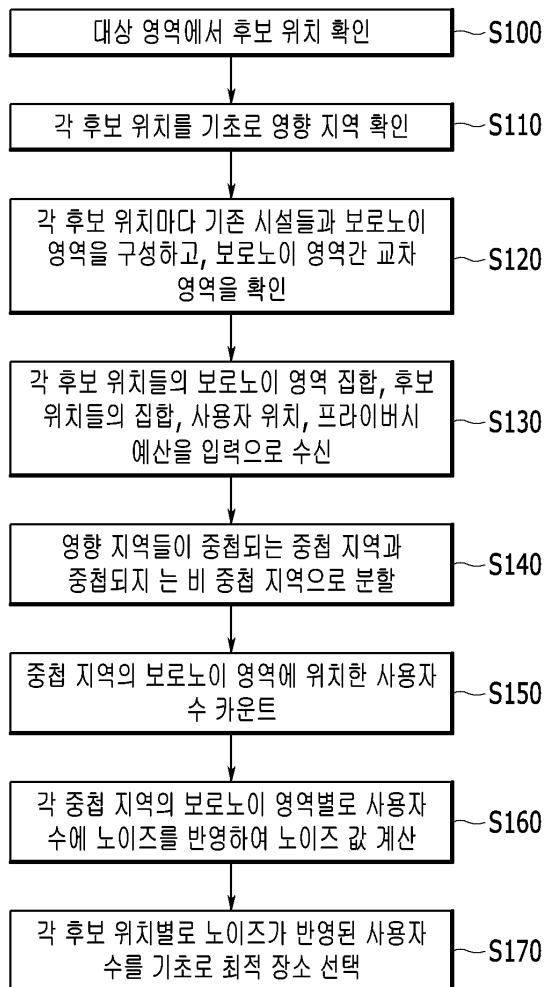
도면1



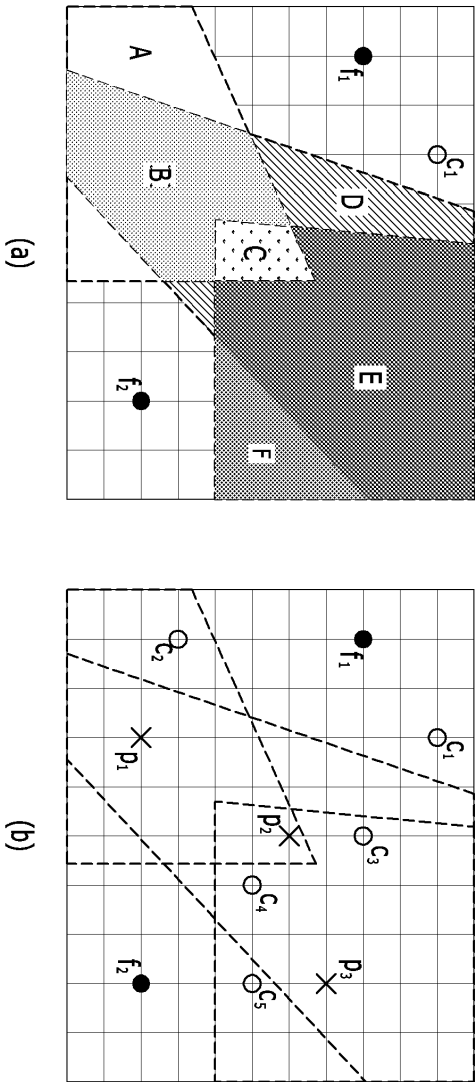
도면2



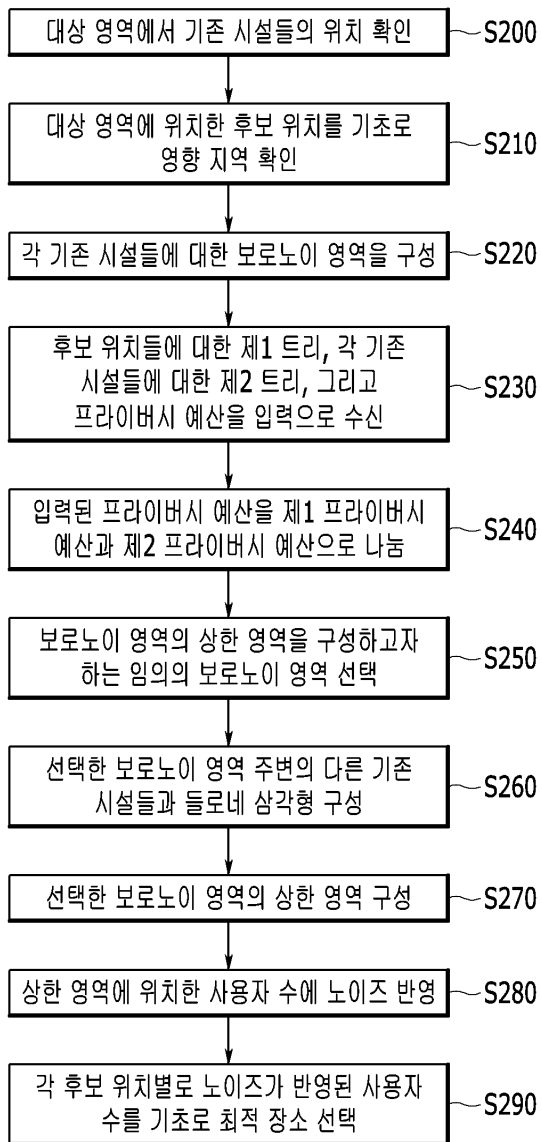
도면3



도면4



도면5



도면6

