

(12) 특허협력조약에 의하여 공개된 국제출원

(19) 세계지식재산권기구
국제사무국

(43) 국제공개일

2022 년 10 월 27 일 (27.10.2022) WIPO | PCT



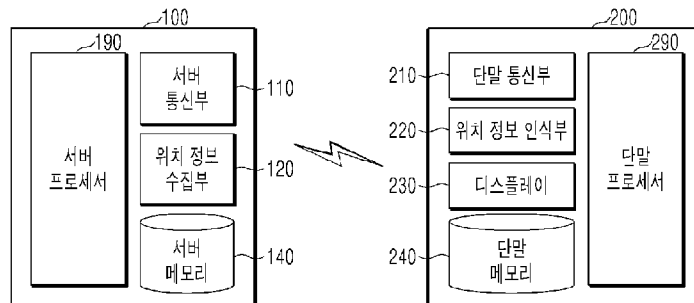
(10) 국제공개번호

WO 2022/225302 A1

- (51) 국제특허분류: *H04L 9/40* (2022.01) *H04W 12/02* (2009.01)
H04L 67/52 (2022.01)
- (21) 국제출원번호: PCT/KR2022/005599
- (22) 국제출원일: 2022 년 4 월 19 일 (19.04.2022)
- (25) 출원언어: 한국어
- (26) 공개언어: 한국어
- (30) 우선권정보:
10-2021-0050628 2021 년 4 월 19 일 (19.04.2021) KR
10-2022-0039007 2022 년 3 월 29 일 (29.03.2022) KR
- (71) 출원인: 서울대학교산학협력단 (SEOUL NATIONAL UNIVERSITY R&DB FOUNDATION) [KR/KR]; 08826 서울특별시 관악구 관악로 1, Seoul (KR).
- (72) 발명자: 심규석 (SHIM, Kyu Seok); 08826 서울특별시 관악구 관악로 1, Seoul (KR). 홍대영 (HONG, Dae Young); 08826 서울특별시 관악구 관악로 1, Seoul (KR).
- (74) 대리인: 특허법인 남앤남 (NAM & NAM); 04516 서울특별시 중구 서소문로 95, 3층, Seoul (KR).
- (81) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.
- (84) 지정국 (별도의 표시가 없는 한, 가능한 모든 종류의 국내 권리의 보호를 위하여): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 유라시아 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 유

(54) Title: METHOD AND SERVER FOR ESTIMATING FREQUENCY DISTRIBUTION FOR LOCATION DATA

(54) 발명의 명칭: 위치 데이터에 대한 빈도 분포 추정 방법 및 서버



110 ... Server communication unit
120 ... Location information collection unit
140 ... Server memory
190 ... Server processor
210 ... Terminal communication unit
220 ... Location information recognition unit
230 ... Display
240 ... Terminal memory
290 ... Terminal processor

(57) Abstract: Disclosed is a server for estimating frequency distribution. The server may comprise: a server communication unit; a location information collection unit for collecting a perturbation location corresponding to the original location of a terminal in a domain area; and a server processor which, when the domain area is mapped to a two-dimensional grid, derives the conditional expectation of a frequency matrix for the perturbation location versus the original location for each cell in the grid by using a prefix-sum matrix, and defines a convex optimization problem involving the derived conditional expectation. Accordingly, estimation of the frequency distribution of the original location can be performed quickly and accurately.

(57) 요약서: 빈도 분포 추정 서버가 개시된다. 본 서버는 서버 통신부, 도메인 영역에서 단말의 원본 위치에 대응하는 교란 위치를 수집하는 위치 정보 수집부 및 도메인 영역이 2 차원의 그리드에 매핑되는 경우, 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도하고, 유도된 조건부 기댓값을 포함하는 볼록 최적화 문제를 정의하는 서버 프로세서를 포함할 수 있다. 이에 따라, 원본 위치의 빈도 분포 추정이 신속하고 정확하게 수행될 수 있다.

[다음 쪽 계속]

WO 2022/225302 A1

럼 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

규칙 4.17에 의한 선언서:

- 신규성을 해치지 아니하는 개시 또는 신규성 상실의 예외에 관한 선언 (규칙 4.17(v))

공개:

- 국제조사보고서와 함께 (조약 제21조(3))

명세서

발명의 명칭: 위치 데이터에 대한 빈도 분포 추정 방법 및 서버 기술분야

- [1] 본 발명은 프라이버시 보호 기준이 적용된 위치 데이터에 대한 빈도 분포를 추정하는 방법 및 이를 수행하는 서버에 관한 것이다.

배경기술

- [2] 이하에서 기술되는 내용은 본 발명의 실시예와 관련되는 배경 정보를 제공할 목적으로 기재된 것일 뿐이고, 기술되는 내용들이 당연하게 종래기술을 구성하는 것은 아니다.
- [3] 현재 스마트폰이나 스마트워치 등으로부터 막대한 양의 사용자 위치 정보가 수집되고 있으며, 사용자의 위치 정보는 광고, 추천 시스템, 웹 검색 등의 개인화 서비스의 품질을 향상시키는데 이용되고 있다. 예를 들어, 수집된 사용자의 위치를 기반으로 사용자에게 가까운 식당 위주로 광고함으로써 해당 식당의 매출이 향상될 수 있다.
- [4] 이러한 국내 위치 정보 산업의 시장 규모는 2020년 기준 2조 331억 원으로 추정되기도 한다. 그러나, 위치 정보로부터 민감한 개인 정보가 유출될 수 있기 때문에, 개인의 위치 정보를 수집하되 프라이버시를 보호하는 방법이 필요하다.
- [5] 이를 위해, 로컬 차분 프라이버시(Local Differential Privacy) 기술이 등장하였다. 로컬 차분 프라이버시란 데이터 수집 서버의 신뢰도에 의존하지 않고 사용자가 스스로 개인 정보를 무작위적(Randomized) 알고리즘으로 교란(Perturbation)시키는 프라이버시 보호의 기준으로서, 구글이나 애플 등의 IT 기업들에서 데이터 수집 시 이용한다.
- [6] 기존의 로컬 차분 프라이버시를 만족하는 위치 정보 수집 기술들 중 그리드 기반 기술, 그리드를 사용하지 않고 2차원 평면 라플라스 노이즈(Planar Laplace Noise)를 삽입하여 교란시키는 기술 및 PM(Piecewise Mechanism)을 위치 데이터를 교란시키는 기술 등은 모두 기술적인 한계점이 있었다.
- [7] 이에, 기존 기술 보다 프라이버시를 보호하면서도 교란 위치의 유용성(utility)을 향상시키는 위치 정보 수집 방법이 필요하며, 원본 위치를 정확하게 모르더라도 교란 위치를 이용하여 원본 위치의 빈도 분포를 추정하는 방법이 필요하다.
- [8] 한편, 전술한 선행기술은 발명자가 본 발명의 도출을 위해 보유하고 있었거나, 본 발명의 도출 과정에서 습득한 기술 정보로서, 반드시 본 발명의 출원 전에 일반 공중에게 공개된 공지기술이라 할 수는 없다.

발명의 상세한 설명

기술적 과제

- [9] 본 발명이 해결하고자 하는 과제는 그리드에 의존하지 않고 프라이버시를

보호하면서도 교란 위치의 유용성(utility)을 향상시키는 위치 정보 수집 방법 및 위치 정보 제공 방법을 제공하는 데에 있다.

[10] 본 발명의 또 다른 과제는 수집된 교란 위치를 이용하여 원본 위치의 빈도 분포를 정확하고 빠르게 추정하는 방법을 제공하는 데에 있다.

[11] 본 발명의 목적은 이상에서 언급한 과제에 한정되지 않으며, 언급되지 않은 본 발명의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있고, 본 발명의 실시 예에 의해 보다 분명하게 이해될 것이다. 또한, 본 발명의 목적 및 장점들은 청구범위에 나타난 수단 및 그 조합에 의해 실현될 수 있음을 알 수 있을 것이다.

과제 해결 수단

[12] 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 추정하는 방법은 도메인 영역 내 단말의 원본 위치에 대응하는 교란 위치(perturbed location)를 수집하는 단계; 상기 도메인 영역이 2차원의 그리드에 매핑되는 경우, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위함 행렬을 이용하여 유도하는 단계; 상기 유도된 조건부 기댓값을 포함하는 볼록 최적화(convex optimization) 문제를 정의하는 단계; 상기 정의된 볼록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에 대한 전위함 기반 원본 위치의 빈도 행렬을 예측하는 단계; 및 상기 예측된 전위함 기반 원본 위치의 빈도 행렬에 기초하여, 원본 위치의 빈도 분포를 추정하는 단계를 포함할 수 있다.

[13] 상기 빈도 분포 추정 방법은 상기 원본 위치가 교란 위치로 샘플링될 때 사용된 기준 영역, 샘플링 파라미터값, 프라이머시 파라미터값 및 상기 셀의 넓이 중 적어도 하나의 정보를 획득하는 단계를 더 포함할 수 있다.

[14] 상기 유도하는 단계는 상기 그리드의 각 셀에 대해, 셀과의 거리에 기초하여 구분된 기준 영역에 대한 서브 그리드 그룹을 설정하는 단계; 및 상기 설정된 서브 그리드 그룹 별로 셀과의 중첩 면적, 해당 서브 그리드 그룹 별 전위함에 기초한 연산 및 상기 획득된 정보에 기초하여, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위함 행렬을 이용하여 유도하는 단계를 포함할 수 있다.

[15] 상기 볼록 최적화 문제는 교란 위치의 빈도 행렬과 상기 조건부 기댓값의 차가 최소화되기 위한 문제일 수 있다.

[16] 상기 빈도 분포 추정 방법은, 상기 원본 위치의 빈도 분포를 추정하는 단계 이후에, 상기 추정된 원본 위치의 빈도 분포를 제공하는 단계를 더 포함할 수 있다.

[17] 아울러, 상기 빈도 분포 추정 방법은, 상기 수집하는 단계 이전에, 상기 교란 위치를 샘플링하기 위한 기준 확률을 정의하는 단계; 상기 기준 확률에 기반하여 상기 원본 위치 및 상기 교란 위치 간의 오차의 기댓값을 유도하는 단계; 및 상기 오차의 기댓값을 최소화하도록 상기 기준 영역의 영역 파라미터를 결정하는

단계를 더 포함할 수 있다.

- [18] 상기 기준 확률은 상기 교란 위치에 대한 확률이고, 상기 교란 위치가 상기 기준 영역 안에 있는 경우를 위한 제 1 기준 확률 및 상기 교란 위치가 상기 기준 영역 밖의 상기 도메인 영역 내에 있는 경우를 위한 제2 기준 확률을 포함할 수 있다.
- [19] 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 추정하는 방법은 도메인 영역에서 단말의 원본 위치에 대응하는 교란 위치를 수집하는 단계; 상기 도메인 영역이 2차원의 그리드에 매핑되는 경우, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 유도하는 단계; 상기 유도된 조건부 기댓값을 포함하는 블록 최적화 문제를 정의하는 단계; 및 상기 정의된 블록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에 대한 원본 위치의 빈도 분포를 추정하는 단계를 포함할 수 있다.
- [20] 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 추정하는 서버는 도메인 영역에서 단말의 원본 위치에 대응하는 교란 위치를 수집하는 위치 정보 수집부; 및 상기 도메인 영역이 2차원의 그리드에 매핑되는 경우, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도하고, 상기 유도된 조건부 기댓값을 포함하는 블록 최적화 문제를 정의하는 서버 프로세서를 포함할 수 있다.
- [21] 상기 서버 프로세서는 상기 정의된 블록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에 대한 전위합 기반 원본 위치의 빈도 행렬을 예측하고, 상기 예측된 전위합 기반 원본 위치의 빈도 행렬에 기초하여, 원본 위치의 빈도 분포를 추정하도록 구성될 수 있다.
- [22] 상기 서버 프로세서는 상기 원본 위치가 교란 위치로 샘플링될 때 사용된 기준 영역, 샘플링 파라미터값, 프라이버시 파라미터값 및 상기 셀의 넓이 중 적어도 하나의 정보를 획득하도록 구성될 수 있다.
- [23] 상기 서버 프로세서는 상기 그리드의 각 셀에 대해, 셀과의 거리에 기초하여 구분된 기준 영역에 대한 서브 그리드 그룹을 설정하고, 상기 설정된 서브 그리드 그룹 별로 셀과의 중첩 면적, 해당 서브 그리드 그룹 별 전위합에 기초한 연산 및 상기 획득된 정보에 기초하여, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도하도록 구성될 수 있다.
- [24] 상기 블록 최적화 문제는 교란 위치의 빈도 행렬과 상기 조건부 기댓값의 차가 최소화되기 위한 문제일 수 있다.
- [25] 상기 위치 데이터에 대한 빈도 분포를 추정하는 서버는 서버 통신부를 더 포함할 수 있으며, 상기 서버 프로세서는 상기 추정된 원본 위치의 빈도 분포를 상기 서버 통신부를 통해 서비스 프로바이더에 제공하도록 구성될 수 있다.
- [26] 상기 서버 프로세서는 상기 교란 위치를 샘플링하기 위한 기준 확률을 정의하고, 상기 기준 확률에 기반하여 상기 원본 위치 및 상기 교란 위치 간의 오차의 기댓값을 유도하며, 상기 오차의 기댓값을 최소화하도록 상기 기준

영역의 영역 파라미터를 결정하도록 구성될 수 있다.

- [27] 상기 기준 확률은 상기 교란 위치에 대한 확률이고, 상기 교란 위치가 상기 기준 영역 안에 있는 경우를 위한 제 1 기준 확률 및 상기 교란 위치가 상기 기준 영역 밖의 상기 도메인 영역 내에 있는 경우를 위한 제 2 기준 확률을 포함할 수 있다.

- [28] 전술한 것 외의 다른 측면, 특징, 및 이점이 이하의 도면, 청구범위 및 발명의 상세한 설명으로부터 명확해질 것이다.

발명의 효과

- [29] 본 발명의 다양한 실시 예에 의하면, 원본 위치에 대응하는 교란 위치의 사용 신뢰도가 향상될 수 있으며, 프라이버시 보호가 효과적으로 수행될 수 있으며, 수집된 교란 위치를 이용하여 원본 위치의 빈도 분포가 정확하고 신속하게 추정될 수 있다.

- [30] 본 발명의 효과는 이상에서 언급된 것들에 한정되지 않으며, 언급되지 아니한 다른 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

도면의 간단한 설명

- [31] 도 1은 본 발명의 일 실시 예에 따른 복수의 위치 정보 제공 단말로부터 원본 위치에 대응하는 교란 위치를 수집하는 빈도 분포 추정 서버를 개략적으로 설명하기 위한 도면,

- [32] 도 2는 본 발명의 일 실시 예에 따른 빈도 분포 추정 서버 및 위치 정보 제공 단말의 구성을 나타내는 상대 블록도,

- [33] 도 3은 본 발명의 일 실시 예에 따른 도메인 영역을 결정하고 도메인 영역 내의 확률 밀도를 설정하는 빈도 분포 추정 서버의 동작을 설명하기 위한 도면,

- [34] 도 4는 본 발명의 일 실시 예에 따른 원본 위치에 기초하여 교란 위치를 샘플링하는 위치 정보 제공 단말을 설명하기 위한 알고리즘을 나타내고,

- [35] 도 5는 본 발명의 일 실시 예에 따른 빈도 분포 추정 서버의 위치 정보 수집 방법을 나타내는 시퀀스도,

- [36] 도 6은 본 발명의 일 실시 예에 따른 위치 정보 제공 단말의 위치 정보 제공 방법을 나타내는 시퀀스도,

- [37] 도 7은 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 추정하는 방법을 나타내는 시퀀스도,

- [38] 도 8은 본 발명의 일 실시 예에 따른 도메인 영역에 대응하는 그리드를 설명하기 위한 도면,

- [39] 도 9은 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 전위합 기반으로 추정하는 방법을 나타내는 시퀀스도, 그리고,

- [40] 도 10 내지 도 12는 본 발명의 일 실시 예에 따른 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도하는 방법을 설명하기 위한 도면들이다.

발명의 실시를 위한 형태

- [41] 이하에서는 도면을 참조하여 본 발명을 보다 상세하게 설명한다. 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며, 여기에서 설명하는 실시 예들에 한정되지 않는다. 이하 실시 예에서는 본 발명을 명확하게 설명하기 위해서 설명과 직접적인 관계가 없는 부분을 생략하지만, 본 발명의 사상이 적용된 장치 또는 시스템을 구현함에 있어서, 이와 같이 생략된 구성이 불필요함을 의미하는 것은 아니다. 아울러, 명세서 전체를 통하여 동일 또는 유사한 구성요소에 대해서는 동일한 참조번호를 사용한다.
- [42] 이하의 설명에서 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 되며, 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 또한, 이하의 설명에서 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다.
- [43] 이하의 설명에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [44] 본 명세서에서는, 위치 데이터를 수집하는 방법을 먼저 기술한 후, 수집된 위치 데이터의 분포를 추정하는 방법을 후술하기로 한다.
- [45] 도 1은 본 발명의 일 실시 예에 따른 복수의 위치 정보 제공 단말(200A~200N)(200)로부터 원본 위치에 대응하는 교란 위치를 수집하는 빈도 분포 추정 서버(100)를 개략적으로 설명하기 위한 도면이다.
- [46] 여기서, 복수의 위치 정보 제공 단말(200A~200N)(200)의 원본 위치는 해당 단말의 실제 또는 현 위치이며, 해당 단말을 소지한 사용자의 실제 또는 현 위치일 수 있다.
- [47] 또한, 교란(Perturbation) 위치는 원본 위치가 정확하게 특정되지 못하도록, 프라이버시 보호 목적으로 변형/왜곡된 위치이며, 로컬 차분 프라이버시(LDP)를 만족하는 범위 내에서 변형/왜곡될 수 있다.
- [48] 빈도 분포 추정 서버(100)는 복수의 위치 정보 제공 단말(200A~200N)(200)의 교란 위치만을 수집하고 원본 위치를 직접 수집하지 않으므로, 서버의 프라이버시 보호 정책에 상관없이, 개인 프라이버시가 보호될 수 있다.
- [49] 빈도 분포 추정 서버(100)는 위치 교란에 필요한 정보(DI)를 복수의 위치 정보 제공 단말(200A~200N)(200)로 제공할 수 있다. 여기서, 위치 교란에 필요한 정보(DI)는 복수의 위치 정보 제공 단말(200A~200N)(200)의 원본 위치 및 교란 위치가 전부 매핑 가능한 도메인 영역에 대한 정보 및 교란을 위한 각종 파라미터 정보 등을 포함할 수 있으며, 자세한 내용은 후술하기로 한다.
- [50] 복수의 위치 정보 제공 단말(200A~200N)(200)은 빈도 분포 추정 서버(100)로부터 위치 교란에 필요한 정보(DI)를 수신하여, 원본 위치에

대응하는 교란 위치(A~N)를 생성하여, 빈도 분포 추정 서버(100)로 제공할 수 있다.

[51] 빈도 분포 추정 서버(100)는 복수의 위치 정보 제공 단말(200A~200N)(200)의 교란 위치를 수집하고, 교란 위치에 대응하는 원본 위치의 빈도 분포를 추정할 수 있다. 즉, 빈도 분포 추정 서버(100)는 원본 위치를 직접적으로 추정하지는 않지만, 원본 위치의 빈도 분포를 추정할 수 있다.

[52] 빈도 분포 추정 서버(100)는 수집되거나 가공된 위치 데이터를 서비스 프로바이더(SP)에게 제공할 수 있다.

[53] 여기서, 위치 데이터는 단말로부터 수집된 교란 위치 정보, 기준 영역 파라미터, 위치 수집의 대상인 단말의 도메인 영역 정보 및 교란 위치에 대응하는 원본 위치의 추정된 빈도 분포 정보 등을 포함할 수 있으나, 실시 예가 이에 한정되는 것은 아니다.

[54] 서비스 프로바이더(SP)는 원본 위치와 교란 위치 간의 에러율이 낮은 위치 데이터를 빈도 분포 추정 서버(100)로부터 수집하게 됨으로써, 신뢰도가 향상된 개인화 서비스를 효과적으로 수행할 수 있다. 아울러, 서비스 프로바이더(SP)는 교란 위치에 대응하는 원본 위치의 추정된 빈도 분포를 수집하여, 위치에 기반하여 다양한 고객 서비스를 제공할 수 있다.

[55] 선택적 또는 부가적 실시 예로, 복수의 위치 정보 제공 단말(200A~200N)(200)로부터 위치 정보를 수집하는 수집 서버가 별도로 구성되고, 이 경우, 빈도 분포 추정 서버는 상기 수집 서버로부터 교란 위치 정보 및 원본 위치를 교란 위치로 교란하는데 사용된 정보를 수집하여, 교란 위치에 대응하는 원본 위치의 빈도 분포를 추정할 수 있다.

[56] 도 2는 본 발명의 일 실시 예에 따른 빈도 분포 추정 서버(100) 및 위치 정보 제공 단말(200)의 구성을 나타내는 상대 블록도이다.

[57] 빈도 분포 추정 서버(100)는 장치, 디바이스, 클라우드 기반 시스템 및 서버 시스템 등으로 구현될 수 있으나, 실시 예가 이에 한정되는 것은 아니다. 빈도 분포 추정 서버(100)는 위치 정보 제공 단말(200)로부터 위치 정보 제공 단말(200)의 원본 위치에 대응하는 교란 위치를 수집할 수 있다.

[58] 빈도 분포 추정 서버(100)는 서버 통신부(110), 위치 정보 수집부(120), 서버 메모리(140) 및 하나 이상의 서버 프로세서(190)를 포함할 수 있으나, 더 많은 구성 또는 더 적은 구성을 포함하도록 구현될 수 있다.

[59] 서버 통신부(110)는 위치 정보 제공 단말(200)을 포함한 다양한 기기와 통신하기 위한 모듈로, 이동 통신 모듈, 근거리 통신 모듈 등을 포함할 수 있다.

[60] 서버 프로세서(190)는 서버 통신부(110)를 통해 다양한 정보를 위치 정보 제공 단말(200)로 제공하거나, 위치 정보 제공 단말(200)로부터 다양한 정보를 수신할 수 있다.

[61] 위치 정보 수집부(120)는 서버 프로세서(190)의 제어에 따라, 적어도 하나의 위치 정보 제공 단말(200)로부터 각 위치 정보 제공 단말(200)의 원본 위치에

- 대응하는 교란 위치를 서버 통신부(110)를 통해 수집할 수 있다.
- [62] 서버 메모리(140)는 다양한 정보를 저장 가능한 구성이며, 위치 정보 수집부(120)에 의해 수집된 교란 위치를 저장할 수 있다. 서버 메모리(140)는 빈도 분포 추정 서버(100)가 실시 예에 따른 위치 정보 수집 동작을 수행하도록 하는 일련의 코드 또는 명령을 포함한 프로그램과 이에 수반된 입출력 데이터를 서버 프로세서(190)가 읽기/쓰기 가능한 방식으로 저장할 수 있다.
- [63] 서버 프로세서(190)는 하나 이상의 프로세서로 구현될 수 있으며, 단수로 표현하더라도 복수로 간주될 수 있다. 서버 프로세서(190)는 빈도 분포 추정 서버(100)의 각 구성들을 컨트롤하는 모듈이며, 서버 프로세서(190)는 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(microprocessor), 중앙처리장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific integrated circuit), FPGA(field programmable gate array), 그래픽 처리 유닛(Graphics Processing Unit; GPU) 등의 처리 장치를 망라할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다. 서버 프로세서(190)는 인공 지능 연산을 수행하기 위한 러닝 프로세서를 별도로 구비하거나, 자체적으로 러닝 프로세서를 구비할 수 있다.
- [64] 다음으로, 위치 정보 제공 단말(200)은 단말 통신부(210), 위치 정보 인식부(220), 디스플레이(230), 단말 메모리(240) 및 단말 프로세서(290)를 포함할 수 있으나, 더 많은 구성 또는 더 적은 구성을 포함할 수 있다.
- [65] 단말 통신부(210)는 빈도 분포 추정 서버(100)를 포함한 다양한 기기와 통신하기 위한 모듈로, 이동 통신 모듈, 근거리 통신 모듈 등을 포함할 수 있다.
- [66] 위치 정보 인식부(220)는 각종 센서 및 인공위성 기반의 정보(가령, GPS 정보)를 이용하여 위치 정보 제공 단말(200)의 원본 위치를 인식할 수 있다. 가령, 위치 정보 인식부(220)는 위치 정보 제공 단말(200)의 현재 위도 및 경도를 포함하는 위치 정보를 원본 위치로 인식할 수 있다.
- [67] 디스플레이(230)는 각종 정보를 가시적으로 표현하는 모듈이며, 백라이트를 이용한 디스플레이 또는 자발광 디스플레이로 구현될 수 있다.
- [68] 단말 메모리(240)는 각종 정보를 저장할 수 있으며, 원본 위치를 교란 위치로 변경/왜곡하기 위한 알고리즘을 구현한 프로그램을 저장할 수 있다. 또한, 단말 메모리(240)는 위치 정보 제공 단말(200)이 실시 예에 따른 위치 정보 제공 동작을 수행하도록 하는 일련의 코드 또는 명령을 포함한 프로그램 및 이에 수반된 입출력 데이터를 단말 프로세서(290)가 읽기/쓰기 가능한 방식으로 저장할 수 있다.
- [69] 단말 프로세서(290)는 하나 이상의 프로세서로 구현될 수 있으며, 단수로 표현하더라도 복수로 간주될 수 있다. 단말 프로세서(290)는 위치 정보 제공

단말(200)의 각 구성들을 컨트롤하는 모듈이며, 단말 프로세서(290)는 프로그램 내에 포함된 코드 또는 명령으로 표현된 기능을 수행하기 위해 물리적으로 구조화된 회로를 갖는, 하드웨어에 내장된 데이터 처리 장치를 의미할 수 있다. 이와 같이 하드웨어에 내장된 데이터 처리 장치의 일 예로써, 마이크로프로세서(microprocessor), 중앙처리장치(central processing unit: CPU), 프로세서 코어(processor core), 멀티프로세서(multiprocessor), ASIC(application-specific integrated circuit), FPGA(field programmable gate array), 그래픽 처리 유닛(Graphics Processing Unit; GPU) 등의 처리 장치를 망라할 수 있으나, 본 발명의 범위가 이에 한정되는 것은 아니다. 단말 프로세서(290)는 인공 지능 연산을 수행하기 위한 러닝 프로세서를 별도로 구비하거나, 자체적으로 러닝 프로세서를 구비할 수 있다.

[70] 이상과 같이, 빈도 분포 추정 서버(100) 및 위치 정보 제공 단말(200)의 구성을 설명하였고, 이하에서는 빈도 분포 추정 서버(100) 및 위치 정보 제공 단말(200)의 기능 및 동작을 중점적으로 설명하고자 한다.

[71] 도 3은 본 발명의 일 실시 예에 따른 도메인 영역을 결정하고 도메인 영역 내의 확률 밀도를 설정하는 빈도 분포 추정 서버(100)의 동작을 설명하기 위한 도면이다.

[72] 서버 프로세서(190)는 위치 정보 제공 단말(200)의 원본 위치에 대한 도메인 영역(Dom)을 결정할 수 있는데, 서버 프로세서(190)는 위치 정보 제공 단말(200)의 원본 위치를 정확히 인식하지 못하더라도 위치 정보 제공 단말(200)의 원본 위치를 포함하는 충분한 영역을 도메인 영역(Dom)으로 결정할 수 있다. 여기서, 위치 정보 제공 단말(200)의 원본 위치는 도메인 영역(Dom)에 균등하게 분포된 것으로 가정할 수 있다. 도메인 영역(Dom)은 위치 정보 제공 단말(200)의 원본 위치 및 교란 위치를 모두 포함하는 영역으로, 2 차원 평면 상에 표현될 수 있다. 도메인 영역의 X 축 방향 범위는 -b1에서 b1까지로 표현될 수 있으며, Y축 방향 범위는 -b2에서 b2까지로 표현될 수 있다. 여기서 b1 및 b2는 0 보다 큰 양의 실수이다.

[73] 도메인 영역(Dom)의 내부에는 위치 정보 제공 단말(200)의 원본 위치에 대한 임의의 기준 영역(Sta)이 배치될 수 있는데, 기준 영역(Sta)은 한 변의 길이가 ω 인 정사각형으로 구현될 수 있으나, 실시 예가 이에 국한되는 것은 아니다. 여기서, ω 를 기준 영역(Sta)의 영역 파라미터라고 칭할 수 있다. 여기서 ω 는 0 보다 큰 양의 실수이다.

[74] 여기서, 위치 정보 제공 단말(200)의 위치, 즉 원본 위치 및 교란 위치는 각각 위도 및 경도를 포함할 수 있으며, 도메인 영역(Dom) 상에 매핑될 수 있다. 즉, 위치 정보 제공 단말(200)의 원본 위치를 교란하여 생성한 교란 위치가 도메인 영역(Dom)을 벗어나지 않도록 함으로써, 프라이버시 보호를 달성하면서도 위치 정보 제공 단말(200)로부터 수집된 교란 위치의 유용성, 예측 가능성 및 신뢰성이 향상될 수 있다.

- [75] 서버 프로세서(190)는 도메인 영역(Dom) 내의 기준 영역(Sta)과 위치 정보 제공 단말(200)의 원본 위치에 대한 교란 위치 간의 관계에 따라 도메인 영역(Dom)에서 교란 위치를 샘플링하기 위한 기준 확률을 정의할 수 있으며, 기준 확률은 아래의 [수학식 1]과 같다.

[76] [수식1]

$$\Pr[f_s(p) = \tilde{p}] = \begin{cases} \alpha_w \cdot e^\epsilon, & \text{if } \tilde{p} \in \text{Sta} \\ \alpha_w, & \text{otherwise} \end{cases}$$

- [77] 기준 확률(Pr)은 무작위 변수(Random Variable)를 샘플링하는 샘플링 함수(f_s)의 샘플링 기준을 정의한 확률이다. 기준 확률(Pr)에 의해, 샘플링 함수(f_s)는, 원본 위치(p)에 기반하여 교란 위치($\tilde{p}$)

$\tilde{p}$

)를 샘플링할 때, 교란 위치($\tilde{p}$)

$\tilde{p}$

)가 기준 영역(Sta) 내에 있는 경우,

$$\alpha_w \cdot e^\epsilon$$

의 확률 밀도로 샘플링하며, 교란 위치($\tilde{p}$)

$\tilde{p}$

)가 기준 영역(Sta) 밖에 있는 경우,

$$\alpha_w$$

의 확률 밀도로 샘플링할 수 있다.

- [78] 다만, 서버 프로세서(190)는 원본 위치(p) 및 원본 위치(p)에 대응하는 기준 영역(Sta)의 중심 정보를 인식하지는 못한다. 즉, 서버 프로세서(190)는 원본 위치(p)가 도메인 영역(Dom) 내에 균등하게 분포된 것을 가정하여 기준 확률(Pr)을 정의할 수 있다.

- [79] 기준 확률(Pr)은 교란 위치에 대한 확률 밀도 또는 확률 밀도 함수로 표현될 수 있으며, 교란 위치가 기준 영역(Sta) 안에 있는 경우를 위한 제 1 기준 확률($\alpha_w \cdot e^\epsilon$) 및 교란 위치가 기준 영역(Sta) 밖의 도메인 영역(Dom) 내에 있는 경우를 위한 제2 기준 확률(α_w)

$$\alpha_w \cdot e^\epsilon$$

)를 포함할 수 있다.

$$\alpha_w$$

)를 포함할 수 있다.

[80] 여기서, 제 1 기준 확률(

$$\alpha_w \cdot e^\epsilon$$

)은 상기 제 2 기준 확률(

$$\alpha_w$$

)보다 더 높은 확률로 상기 교란 위치를 샘플링하도록 정의될 수 있다.

[81] 기준 확률(Pr)는 도메인 영역(Dom), 기준 영역(Sta) 및 프라이버시 파라미터(

ϵ

)에 기반하여 정의된 샘플링 파라미터(

$$\alpha_w$$

)에 기반하여 정의될 수 있으며, 샘플링 파라미터(

$$\alpha_w$$

)는 아래 [수학식 2]에 의해 정의될 수 있다.

[82] [수식2]

$$\alpha_w = 1 / (w^2 (e^\epsilon - 1) + 4b_1 b_2)$$

[83] 여기서, 프라이버시 파라미터(

ϵ

)는 프라이버시 보호 수준과 수집된 데이터의 유용성 사이의 트레이드 오프를 조절하는 매개변수이며, 로컬 차분 프라이버시를 충족하기 위한 파라미터일 수 있다.

[84] 서버 프로세서(190)는 결정된 기준 영역(Sta)의 영역 파라미터(w)를 서버 통신부(110)를 통해 위치 수집의 대상인 적어도 하나의 단말(200)에 전송할 수 있다.

[85] 아울러, 서버 프로세서(190)는 위치 정보 수집부(120)를 통해 수집된 교란 위치를 서버 통신부(110)를 통해 서비스 프로바이더(SP)에 제공할 수 있다.

[86] 서버 프로세서(190)는 기준 영역(Sta) 내의 중심(c)을 인식하지 못한 상태에서, 기준 영역(Sta) 내의 중심(c)을 설정하는 관계식을 정의할 수 있는데, 중심(c)을 원본 위치와 최대한 가깝게 배치시키되, 기준 영역(Sta)이 도메인 영역(Dom)을 벗어나지 않게 설정할 수 있는데, 아래 [수학식 3]에 기초하여, 원본 위치와 교란 위치 간의 오차를 최소화하기 위한 중심(c)을 설정하는 관계식을 정의할 수 있다. 여기서, d 는 1 차원(X 축) 및 2 차원(Y축)을 나타낼 수 있다.

[87] [수식3]

$$c[d] = \begin{cases} -b_d + w/2 & \text{if } p[d] \leq -b_d + w/2, \\ p[d] & \text{if } -b_d + w/2 \leq p[d] \leq b_d - w/2, \\ b_d - w/2 & \text{otherwise.} \end{cases}$$

[88] 서버 프로세서(190)는 기준 영역(Sta)의 중심(c)이 결정되면, 기준 확률(Pr)에 기반하여 원본 위치 및 상기 교란 위치 간의 오차의 기댓값을 유도할 수 있다.

[89] 서버 프로세서(190)는 원본 위치(p)가 도메인 영역(Dom)에 균등하게 분포되어 있다고 가정된 상황에서, 상기 중심(c)이 결정된 후, 교란 위치(

$\tilde{p}$

)의 평균 제곱 오차(MSE)의 기댓값을 아래 [수학식 4]와 같이 산출할 수 있다.

[90] [수식4]

$$\begin{aligned} \mathbb{E}[\text{MSE}(\tilde{p}, p)] &= \alpha_w \cdot 4b_1b_2(b_1^2 + b_2^2)/3 \\ &\quad + \alpha_w \cdot w^4(e^\epsilon - 1)/12 \\ &\quad + \alpha_w \cdot w^5(e^\epsilon - 1)(b_1 + b_2)/(48b_1b_2) \end{aligned}$$

[91] 이때, 평균 제곱 오차(MSE)의 기댓값을 최소화시키는 영역 파라미터(w)는 아래 [수학식 5]에 의해 도출될 수 있다.

[92] [수식5]

$$\begin{aligned} &\text{minimize} \quad \mathbb{E}[\text{MSE}(\tilde{p}, p)] \\ &\text{subject to} \quad 0 < w < 2 \cdot \min(b_1, b_2). \end{aligned}$$

[93] 서버 프로세서(190)는 도메인 영역(Dom) 내에 균등하게 분포된 원본 위치(p)에 기반하여, 상기 기댓값을 최소화하도록 상기 기준 영역(Sta)의 영역 파라미터(w)를 결정할 수 있다.

[94] 최적의 영역 파라미터(w)를

w^*

로 표현할 수 있으며, 정해진 기준 영역(Sta)을

$C^*(p)$

으로 표현할 수 있고, 샘플링 파라미터(

α_w

)를

α^*

라고 표현할 수 있다.

[95] 본 발명의 일 실시 예에 의하면, 상술한 바와 같이, 2 차원 결합 분포(2 Dimensional Joint Probability Distribution)를 이용하여, 원본 위치에 대응하는 교란 위치가 높은 정확성 및 신뢰성을 갖도록 생성될 수 있다.

[96] 도 4는 본 발명의 일 실시 예에 따른 원본 위치에 기초하여 교란 위치를 샘플링하는 위치 정보 제공 단말(200)을 설명하기 위한 알고리즘(SC)을 나타낸다.

[97] 먼저, 단말 프로세서(290)는 단말(200)의 원본 위치(p)에 대한 도메인 영역(p) 내의 기준 영역(

 $C^*(p)$

)의 영역 파라미터(

 w^*

)를 단말 통신부(110)를 통해 빈도 분포 추정 서버(100)로부터 획득할 수 있다. 단말 프로세서(290)는 단말 통신부(210)를 통해 도메인 영역(

 D

), 프라이버시 파라미터(

 ϵ

)도 수신할 수 있다.

[98] 알고리즘(SC)의 라인 1에서, 단말 프로세서(290)는 원본 위치(p)에 기반하여 기준 영역(

 $C^*(p)$

)이 상기 도메인 영역(

 D

) 내에 위치하도록 기준 영역(

 $C^*(p)$

)의 중심(c)을 결정할 수 있다. 구체적으로, 단말 프로세서(290)는 각 차원(X 축 및 Y 축)에 대해 원본 위치(p)가 포함된 기준 영역(

 $C^*(p)$

)이 도메인 영역(

 D

)을 넘지 않도록 중심(c)을 결정할 수 있다.

[99] 알고리즘(SC)의 라인 2에서, 단말 프로세서(290)는 샘플링 기준 파라미터(α^*)

)를 산출할 수 있다.

[100] 알고리즘(SC)의 라인 3 내지 7에서, 단말 프로세서(290)는 소정의 확률 구간에서 결정된 샘플링 값(x)에 기반하여, 도메인 영역(D)

) 또는 기준 영역($C^*(p)$)

)으로부터 교란 위치($\tilde{p}$)

)를 샘플링할 수 있다. 여기서, 확률 구간은 샘플링 값(x)의 표본 공간으로, 샘플링 값(x)은 균등하게 분포(uniform distribution)된 [0, 1] 구간으로부터 샘플링된 값이다.

[101] 구체적으로, 단말 프로세서(290)는 영역 파라미터(w^*)

) 및 도메인 영역(D)

)에 기반하여, 샘플링 기준 파라미터(α^*)

)를 결정하고, 상기 도메인 영역(D)

) 및 상기 샘플링 기준 파라미터(α^*)

)에 기반하여 결정된 기준 값($4\alpha^*b_1b_2$)

)과 결정된 샘플링 값(x)의 비교 결과에 따라 교란 위치($\tilde{p}$)

)를 샘플링할 영역을 결정할 수 있다. 해당 샘플링할 영역은 도메인 영역(D)

)를 샘플링할 영역을 결정할 수 있다. 해당 샘플링할 영역은 도메인 영역(D)

D

) 또는 기준 영역(

$C^*(p)$

)이다.

[102] 알고리즘(SC)의 라인 6 및 7에서, 샘플링 값(x)이 기준 값(

$4\alpha^*b_1b_2$

) 보다 크면, 단말 프로세서(290)는 기준 영역(

$C^*(p)$

)에서 균등 분포에 따라 교란 위치(

$\sim$
 $\tilde{p}$

)를 샘플링한다.

[103] 도 5는 본 발명의 일 실시 예에 따른 빈도 분포 추정 서버(100)의 위치 정보 수집 방법을 나타내는 시퀀스도이다. 상기 각 단계들은 빈도 분포 추정 서버(100)의 서버 프로세서(190)에 의해 수행될 수 있다.

[104] 상기 빈도 분포 추정 서버(100)는 단말의 원본 위치에 대한 도메인 영역을 결정한다(S510).

[105] 빈도 분포 추정 서버(100)는 도메인 영역 내의 기준 영역과 원본 위치에 대한 교란 위치 간의 관계에 따라 상기 도메인 영역에서 교란 위치를 샘플링하기 위한 기준 확률을 정의한다(S520).

[106] 원본 위치 및 교란 위치는 각각 위도 및 경도를 포함하고, 도메인 영역 상에 매핑될 수 있다.

[107] 그 후에, 빈도 분포 추정 서버(100)는 기준 확률에 기반하여 원본 위치 및 상기 교란 위치 간의 오차의 기댓값을 유도한다(S530).

[108] 그 후에, 빈도 분포 추정 서버(100)는 원본 위치에 기반하여, 상기 기댓값을 최소화하도록 상기 기준 영역의 영역 파라미터를 결정한다(S540).

[109] 여기서, 기준 확률은 도메인 영역, 기준 영역 및 프라이버시 파라미터에 기반하여 정의된 샘플링 파라미터(

α_w

)에 기반하여 정의될 수 있다. 기준 영역은 정사각형의 영역이며, 기준 영역의 영역 파라미터는 정사각형의 변 길이를 포함할 수 있으나, 기준 영역의 모양은 정사각형에 한정되는 것은 아니다.

[110] 또한, 기준 확률은 교란 위치에 대한 확률 밀도 함수이고, 교란 위치가 기준

- 영역 안에 있는 경우를 위한 제 1 기준 확률 및 교란 위치가 기준 영역 밖의 도메인 영역 내에 있는 경우를 위한 제2 기준 확률을 포함할 수 있다.
- [111] 또한, 제 1 기준 확률은 제 2 기준 확률보다 더 높은 확률로 교란 위치를 샘플링하도록 정의될 수 있다.
- [112] 빈도 분포 추정 서버(100)는 결정된 기준 영역의 영역 파라미터를 위치 수집의 대상인 적어도 하나의 단말에 전송할 수 있다.
- [113] 빈도 분포 추정 서버(100)는 적어도 하나의 단말로부터 원본 위치에 대응하는 교란 위치를 수집하고, 수집된 교란 위치 및/또는 교란 위치에 대응하는 원본 위치의 빈도 분포 추정 정보를 서비스 프로바이더로 제공할 수 있다.
- [114] 도 6은 본 발명의 일 실시 예에 따른 위치 정보 제공 단말(200)의 위치 정보 제공 방법을 나타내는 시퀀스도이다.
- [115] 위치 정보 제공 단말(200)은 단말의 원본 위치에 대한 도메인 영역 내의 기준 영역의 영역 파라미터를 획득한다(S610).
- [116] 그 다음, 위치 정보 제공 단말(200)은 원본 위치에 기반하여 기준 영역이 도메인 영역 내에 위치하도록 기준 영역의 중심을 결정한다(S620).
- [117] 그 후에, 위치 정보 제공 단말(200)은 소정의 확률 구간에서 결정된 샘플링 값에 기반하여, 도메인 영역 또는 기준 영역으로부터 교란 위치를 샘플링한다(S630).
- [118] 상기 S630 단계에서, 위치 정보 제공 단말(200)은 영역 파라미터 및 도메인 영역에 기반하여, 샘플링 기준 파라미터를 결정하고, 도메인 영역 및 샘플링 기준 파라미터에 기반하여 결정된 기준 값과 결정된 샘플링 값의 비교 결과에 따라 교란 위치를 샘플링할 영역을 결정할 수 있다.
- [119] 도 7은 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 추정하는 방법을 나타내는 시퀀스도이고, 도 8은 본 발명의 일 실시 예에 따른 도메인 영역에 대응하는 그리드를 설명하기 위한 도면이다.
- [120] 빈도 분포 추정 서버(100)의 서버 프로세서(190)는 도메인 영역에서 단말의 원본 위치에 대응하는 교란 위치(perturbed location)를 수집할 수 있다(S710).
- [121] 여기서, 서버 프로세서(190)는 도메인 영역을 2차원의 그리드에 매핑할 수 있다. 이에, 위치 정보 제공 단말(200)의 원본 위치 및 교란 위치가 모두 그리드 내에 포함될 수 있다.
- [122] 도 8을 참고하면, 서버 프로세서(190)는 도메인 영역(Dom)을 $n1 \times n2$ 사이즈의 균등하게 나뉜 그리드로 매핑할 수 있는데, 그리드의 각 셀($R_{1,1}$ 내지 $R_{n2,n1}$)은 동일한 면적이며, 해당 면적은 a_{cell} 로 표현될 수 있다. 각 셀($R_{1,1}$ 내지 $R_{n2,n1}$) 각각은 중심점($c_{1,1}$ 내지 $c_{n2,n1}$)을 가진다.
- [123] 서버 프로세서(190)는 원본 위치의 빈도 분포만 추정하므로, 임의의 셀($R_{i,j}$)에 배치된 위치 정보 제공 단말(200)의 원본 위치 또는 교란 위치를 계산상 편의를 위해 임의의 셀($R_{i,j}$)의 중심점($c_{i,j}$)에 배치된 것으로 간주할 수 있다.
- [124] 상술한 기준 영역(

$$C^*(p)$$

) 관련하여, 임의의 셀($R_{i,j}$)에 대한 기준 영역($C^*(i,j)$)을 $C^*(c_{i,j})$ 이라 정의하고, 임의의 셀($R_{k,l}$)에 대한 기준 영역을 $C^*(k,l)$ 이라 정의할 수 있다. 또한, 임의의 영역(R)의 넓이를 $area(R)$ 이라 할 때,

$$\sigma_{k,l}^{(i,j)}$$

를 임의의 셀($R_{i,j}$) 및 임의의 기준 영역($C^*(k,l)$)이 겹치는 영역의 넓이로 아래 [수학식 6]과 같이 정의될 수 있다.

[125] [수식6]

$$\sigma_{k,l}^{(i,j)} = area(R_{i,j} \cap C^*(k,l))$$

[126] 또한, $O(i,j)$ 는 기준 영역($C^*(k,l)$)이 임의의 셀($R_{i,j}$)과 일부라도 겹치는 셀들의 집합으로, 아래 [수학식 7]과 같이 정의될 수 있다.

[127] [수식7]

$$O(i,j) = \{(k,l) | \sigma_{k,l}^{(i,j)} > 0\}$$

[128] 또한, 임의의 셀($R_{i,j}$)에 속한 위치들의 개수를 전체 개수로 나눈 빈도수 비율을 $\theta_{i,j}$ 라고 정의할 수 있으며, 원본 위치의 i 번째 행의 j 번째 열의 성분이 $\theta_{i,j}$ 인 행렬 $\Theta=(\theta_{i,j})$ 을 원본 위치의 빈도 행렬이라 할 수 있으며, 교란 위치의 i 번째 행의 j 번째 열의 성분이 $\theta_{i,j}$ 인 행렬 $\Theta=(\theta_{i,j})$ 을 교란 위치의 빈도 행렬이라 할 수 있다. 원본 위치의 빈도 행렬의 합은 1일 수 있으나, 실시 예가 이에 한정되는 것은 아니다.

[129] 서버 프로세서(190)는 도메인 영역(Dom)의 교란 위치의 빈도 행렬($\Theta=(\theta_{i,j})$)에 기초하여, 원본 위치의 빈도 행렬($\Theta=(\theta_{i,j})$)을 추정함으로써, 원본 위치의 빈도 분포를 추정할 수 있다.

[130] 도 7의 S710 단계 이후에, 서버 프로세서(190)는 도메인 영역이 2차원의 그리드에 매핑되는 경우, 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 유도할 수 있다(S720).

[131] 구체적으로, 서버 프로세서(190)는 원본 위치의 빈도 행렬(Θ)이 주어진 경우, 교란 위치의 임의의 셀($R_{i,j}$)의 빈도수 비율($\theta_{i,j}$)을 아래 [식 8]과 같이 유도할 수 있다.

[132] [수식8]

$$E[\tilde{\theta}_{i,j} | \Theta] = \alpha^* a_{cell} \left(1 + (e^{\epsilon} - 1) \sum_{(k,l) \in O(i,j)} \sigma_{k,l}^{(i,j)} \theta_{k,l} \right),$$

[133] 여기서, 로컬 차분 프라이버시(LDP)를 만족하는, 상술한 프라이버시 파라미터(ϵ)

) 및 샘플링 기준 파라미터(α^*)

$$\alpha^*$$

)가 유도식에 사용될 수 있다.

[134] S720 단계 이후에, 서버 프로세서(190)는 유도된 조건부 기댓값을 포함하는 블록 최적화 문제를 정의할 수 있다(S730).

[135] 구체적으로, 서버 프로세서(190)는 교란 위치의 빈도 행렬(Θ)과 가장 비슷한 조건부 기댓값을 가지는 원본 위치의 빈도 행렬(Θ)을 찾음으로써, 원본 위치의 빈도 행렬을 추정할 수 있는데, 블록 최적화 문제는 아래 [수학식 9]와 같이 표현될 수 있다.

[136] [수식9]

$$\begin{aligned} & \underset{\Theta}{\text{minimize}} \quad (\tilde{\theta}_{i,j} - \mathbb{E}[\tilde{\theta}_{i,j} \mid \Theta])^2 + \lambda \|\Theta\|_F^2 \\ & \text{subject to} \quad \theta_{i,j} \geq 0, \quad i = 1, \dots, n_1, \quad j = 1, \dots, n_2 \\ & \quad \sum_{i,j} \theta_{i,j} = 1, \end{aligned}$$

[137] 여기서, λ 는 0보다 크며, 작은 프로베니우스 노름(frobenius norm)을 가지는 빈도 행렬을 찾기 위한 파라미터이며, 소수의 셀에 빈도가 지나치게 몰리는 현상을 막기 위해 사용될 수 있다.

[138] S730 단계 이후, 서버 프로세서(190)는 정의된 블록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에 대한 원본 위치의 빈도 분포를 추정할 수 있다(S740).

[139] 도 7에 도시된 방법을 이용하는 경우, 위치 데이터의 총 개수가 유지되는 제약 조건 및/또는 빈도수가 음수가 될 수 없다는 제약 조건을 모두 준수하면서, 교란 위치에 기반하여 원본 위치의 빈도 분포가 정확하게 연산될 수 있다. 다만, $n_1 \times n_2$ 의 도메인 영역(Dom)의 각 셀에 대한 조건부 기댓값이 최악의 경우 $O(n_1 n_2)$ 의 시간 복잡도를 필요로 하므로, 보다 개선된 방법이 필요하다.

[140] 이하에서는, 도 7의 방법보다 더욱 개선된 방법을 설명하기로 한다.

[141] 도 9은 본 발명의 일 실시 예에 따른 위치 데이터에 대한 빈도 분포를 전위합 기반으로 추정하는 방법을 나타내는 시퀀스도이고, 도 10 내지 도 12는 본 발명의 일 실시 예에 따른 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도하는 방법을 설명하기 위한 도면들이다.

[142] 서버 프로세서(190)는 도메인 영역 내 단말의 원본 위치에 대응하는 교란 위치를 수집한다(S910). 상기 단계는 도 7의 경우와 동일하므로, 자세한 설명은 생략하기로 한다.

[143] 그 후에, 서버 프로세서(190)는 도메인 영역이 2차원의 그리드에 매핑되는 경우, 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도할 수 있다(S920).

[144] S920 단계는 도 10 내지 도 12를 참고하여 보다 구체적으로 설명하기로 한다. 도 10은 도 9의 S920 단계를 보다 구체적으로 설명하기 위한 서브 시퀀스도이고, 도 11a 내지 도 11c는 본 발명의 일 실시 예에 따른 임의의 셀($R_{i,j}$)에 대한 서브 그리드 그룹을 설정하는 방법을 설명하기 위한 도면들이며, 도 12는 본 발명의

일 실시 예에 따른 임의의 셀($R_{i,j}$)에 대한 완성된 서브 그리드 그룹을 나타낸다.

[145] 먼저, 서버 프로세서(190)는 그리드의 각 셀에 대해, 셀과의 거리에 기초하여 구분된 기준 영역에 대한 서브 그리드 그룹을 설정할 수 있다(도 10의 S921).

[146] 도 11a 내지 도 11c에 도시된 기준 영역들(가령, $C^*(i-2, j+2)$, $C^*(i-1, j+2)$, $C^*(i+1, j+1)$ 등)의 영역 파라미터(w^*)

w^*

)는 동일 사이즈이며, 소정의 셀($R_{i,j}$)과 모서리만 겹치는 서브 그리드 그룹(도 11a), 변이 겹치는 서브 그리드 그룹(도 11b) 및 소정의 셀($R_{i,j}$) 전부가 겹치는 서브 그리드 그룹(11c)으로 구분될 수 있다. 이러한 서브 그리드 그룹들은 각 셀마다 설정될 수 있다.

[147] 도 12를 참고하면, 서버 프로세서(190)는 소정의 셀($R_{i,j}$) 마다 모서리만 겹치는 서브 그리드의 그룹(제1 서브 그리드 그룹, 1N1~1N4), 변이 겹치는 서브 그리드의 그룹(제2 서브 그리드 그룹, 2N1~2N4) 및 전부 겹치는 서브 그리드 그룹(제3 서브 그리드 그룹, 4N)으로 설정함으로써, 각 셀마다 서브 그리드의 그룹을 세팅할 수 있다. 여기서, 제2 서브 그리드 그룹 및 제3 서브 그리드 그룹 내의 각 셀의

$\sigma_{k,l}^{(i,j)}$

값은 동일하다.

[148] 또한, 서버 프로세서(190)는 원본 위치의 빈도 행렬(Θ)에 대한 전위합 행렬

$\Phi = (\phi_{i,j})$

을 아래 [수학식 10]에 기초하여 정의할 수 있다.

[149] [수식10]

$$\phi_{i,j} = \begin{cases} 0 & \text{if } i = 0 \text{ or } j = 0, \\ \sum_{i'=1}^i \sum_{j'=1}^j \theta_{i',j'} & \text{if } 1 \leq i \text{ and } 1 \leq j. \end{cases}$$

[150] 전위합 연산시에, $s_1 \leq i \leq t_1$, $s_2 \leq j \leq t_2$ 조건 하에서, 셀(i,j)을 포함하는 직사각형 모양의 서브 그리드를 $G(s_1, s_2, t_1, t_2)$ 라고 정의될 수 있다. 해당 서브그리드 $G(s_1, s_2, t_1, t_2)$ 에 해당하는 빈도 행렬의 성분들의 합을 $\text{sum}_{\Phi}(s_1, s_2, t_1, t_2)$ 라 할 때, 서버 프로세서(190)는 전위합 행렬을 이용하면 $\text{sum}_{\Phi}(s_1, s_2, t_1, t_2)$ 를 아래 [수학식 11]과 같이 $O(1)$ 의 시간 복잡도로 계산할 수 있다.

[151] [수식11]

$$\text{sum}_{\Phi}(s_1, s_2, t_1, t_2) = \sum_{i=s_1}^{t_1} \sum_{j=s_2}^{t_2} \theta_{i,j} = \phi_{t_1, t_2} - \phi_{s_1-1, t_2} - \phi_{t_1, s_2-1} + \phi_{s_1-1, s_2-1}$$

[152] $O(i,j)$ 에 속하는 셀들에 대하여, 9개의 서브 그리드(제1 서브 그리드 그룹의 4개(1N1~1N4), 제2 서브 그리드 그룹의 4개(2N1~2N4) 및 제3 서브 그리드 그룹

전체 1개(4N))로 이루어진 집합을 $\Gamma(i,j)$ 이라 정의할 수 있으며, 각각의 서브 그리드 $T \in \Gamma(i,j)$ 내의 셀들 (k,l) 에 대해

$$\sigma_{k,l}^{(i,j)}$$

의 값이 동일하므로, 서브그리드 그룹 $T \in \Gamma(i,j)$ 내의 모든 셀들의

$$\sigma_{k,l}^{(i,j)}$$

를 $\sigma(T)$ 로 정의할 수 있다.

[153] 도 12의 임의의 셀($R_{i,j}$)의 경우, $\Gamma(i,j)$ 는 아래 [수학식 12]와 같이 유도될 수 있다.

[154] [수식12]

$$\Gamma(i,j) = \left\{ \begin{array}{l} G(i-2, j-2, i-2, j-2), G(i-2, j+2, i-2, j+2), \\ G(i+2, j-2, i+2, j-2), G(i+2, j+2, i+2, j+2), \\ G(i-1, j-2, i+1, j-2), G(i-1, j+2, i+1, j+2), \\ G(i-2, j-1, i-2, j+1), G(i+2, j-1, i+2, j+1), \\ G(i-1, j-1, i+1, j+1) \end{array} \right\}$$

[155] 살핀 바와 같이, 서버 프로세서(190)는 설정된 서브 그리드 그룹 별로 셀과의 중첩 면적, 해당 서브 그리드 그룹 별 전위함에 기초한 연산 및 획득된 정보(원본 위치가 교란 위치로 샘플링될 때 사용된 기준 영역, 샘플링 파라미터값, 프라이버시 파라미터값 및 상기 셀의 넓이)에 기초하여, 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위함 행렬을 이용하여 유도할 수 있다(S923).

[156] 이에, 상술한 [수학식 8]은 아래 [수학식 13]과 동치관계를 만족할 수 있다.

[157] [수식13]

$$\mathbb{E}[\tilde{\theta}_{i,j} | \boldsymbol{\theta}] = \alpha^* a_{\text{cell}} \left(1 + (e^\epsilon - 1) \sum_{T \in \Gamma(i,j)} \sigma(T) \cdot \text{sum}_{\mathbf{F}}(T) \right),$$

[158] [수학식 13]의 시그마 연산의 시간 복잡도는 $O(1)$ 이다.

[159] 서버 프로세서(190)는 유도된 조건부 기댓값을 포함하는 블록 최적화(convex optimization) 문제를 정의할 수 있다(S930).

[160] 이에, [수학식 9]는 아래 [수학식 14]와 동치관계를 만족할 수 있다.

[161] [수식14]

$$\begin{aligned} & \underset{\mathbf{F}}{\text{minimize}} \quad (\tilde{\theta}_{i,j} - \mathbb{E}[\tilde{\theta}_{i,j} | \boldsymbol{\theta}])^2 + \lambda \sum_{i,j} \text{sum}_{\mathbf{F}}(i,i,j,j) \\ & \text{subject to} \quad \text{sum}_{\mathbf{F}}(i,i,j,j) \geq 0, \quad i=1, \dots, n_1, \quad j=1, \dots, n_2 \\ & \quad \quad \quad \phi_{r_1, r_2} = 1. \end{aligned}$$

[162] [수학식 14]의 블록 최적화 문제의 경우, [수학식 9]보다 연산 효율이 크게 개선될 수 있다. 그리드 셀의 개수가 늘어나더라도 원본 위치의 빈도 분포 추정이 빠르고 용이하게 수행될 수 있으며, 원본 위치의 빈도 분포 추정이 신속하고 정확하게 수행될 수 있다.

[163] 상기 블록화 문제는, 교란 위치의 빈도 행렬과 상기 조건부 기댓값의 차가

최소화되기 위한 문제일 수 있다.

- [164] S930 단계 이후, 서버 프로세서(190)는 정의된 블록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에 대한 전위합 기반의 원본 빈도 행렬을 예측할 수 있다(S940).
- [165] 그 후에, 서버 프로세서(190)는 예측된 전위합 기반 원본 위치의 빈도 행렬에 기초하여, 원본 위치의 빈도 분포를 추정할 수 있다(S950).
- [166] 상술한 빈도 분포 추정 방법은, 원본 위치가 교란 위치로 샘플링될 때 사용된 기준 영역, 샘플링 파라미터값, 프라이버시 파라미터값 및 상기 셀의 넓이 중 적어도 하나의 정보를 획득하는 단계를 더 포함할 수 있다.
- [167] 또한, 빈도 분포 추정 방법은, S950 단계 이후, 상기 추정된 원본 위치의 빈도 행렬 또는 빈도 분포를 서비스 프로바이더 등에 제공할 수 있다(S960).
- [168] 또한, 빈도 분포 추정 방법은, 수집하는 단계(S910) 이전에, 상술한 바와 같이, 교란 위치를 샘플링하기 위한 기준 확률을 정의하고, 기준 확률에 기반하여 원본 위치 및 교란 위치 간의 오차의 기댓값을 유도하며, 오차의 기댓값을 최소화하도록 기준 영역의 영역 파라미터를 결정할 수 있다.
- [169] 여기서, 상기 기준 확률은 교란 위치에 대한 확률이고, 교란 위치가 기준 영역 안에 있는 경우를 위한 제 1 기준 확률 및 교란 위치가 기준 영역 밖의 도메인 영역 내에 있는 경우를 위한 제2 기준 확률을 포함할 수 있다.
- [170] 전술한 본 발명의 일 실시예에 따른 위치 정보 수집 방법 및 위치 정보 제공 방법은 프로그램이 기록된 매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 매체는, 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 매체의 예로는, HDD(Hard Disk Drive), SSD(Solid State Disk), SDD(Silicon Disk Drive), ROM, RAM, CD-ROM, 자기 테이프, 플로피 디스크, 광 데이터 저장 장치 등이 있다.
- [171] 이상 설명된 본 발명의 실시 예에 대한 설명은 예시를 위한 것이며, 본 발명이 속하는 기술분야의 통상의 지식을 가진 자는 본 발명의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시 예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.
- [172] 본 발명의 범위는 상기 상세한 설명보다는 후술하는 청구범위에 의하여 나타내어지며, 청구범위의 의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 발명의 범위에 포함되는 것으로 해석되어야 한다.
- [173] 본 발명은 과학기술정보통신부의 지원을 받은

개인기초연구(과기정통부)(R&D)(과제고유번호: 1711129668, 과제명: 인공지능 기반 고성능 데이터베이스 관리 시스템 관련 연구) 및 클라우드로봇복합인공지능기술개발(R&D)(과제고유번호: 1711159644, 과제명: 다수 로봇의 지능을 통합 고도화하는 클라우드 로봇 지능 증강·공유 및 프레임워크 기술 개발)의 일환으로 수행한 연구로부터 도출된 것이다.

청구범위

- [청구항 1] 위치 데이터에 대한 빈도 분포를 추정하는 방법으로서,
 도메인 영역 내 단말의 원본 위치에 대응하는 교란 위치(perturbed location)를 수집하는 단계;
 상기 도메인 영역이 2차원의 그리드에 매핑되는 경우, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위함 행렬을 이용하여 유도하는 단계;
 상기 유도된 조건부 기댓값을 포함하는 볼록 최적화(convex optimization) 문제를 정의하는 단계;
 상기 정의된 볼록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에 대한 전위함 기반 원본 위치의 빈도 행렬을 예측하는 단계; 및
 상기 예측된 전위함 기반 원본 위치의 빈도 행렬에 기초하여, 원본 위치의 빈도 분포를 추정하는 단계를 포함하는, 빈도 분포 추정 방법.
- [청구항 2] 제1항에 있어서,
 상기 빈도 분포 추정 방법은,
 상기 원본 위치가 교란 위치로 샘플링될 때 사용된 기준 영역, 샘플링 파라미터값, 프라이머시 파라미터값 및 상기 셀의 넓이 중 적어도 하나의 정보를 획득하는 단계를 더 포함하는, 빈도 분포 추정 방법.
- [청구항 3] 제2항에 있어서,
 상기 유도하는 단계는,
 상기 그리드의 각 셀에 대해, 셀과의 거리에 기초하여 구분된 기준 영역에 대한 서브 그리드 그룹을 설정하는 단계; 및
 상기 설정된 서브 그리드 그룹 별로 셀과의 중첩 면적, 해당 서브 그리드 그룹 별 전위함에 기초한 연산 및 상기 획득된 정보에 기초하여, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위함 행렬을 이용하여 유도하는 단계를 포함하는, 빈도 분포 추정 방법.
- [청구항 4] 제1항에 있어서,
 상기 볼록 최적화 문제는,
 교란 위치의 빈도 행렬과 상기 조건부 기댓값의 차가 최소화되기 위한 문제인, 빈도 분포 추정 방법.
- [청구항 5] 제1항에 있어서,
 상기 원본 위치의 빈도 분포를 추정하는 단계 이후에,
 상기 추정된 원본 위치의 빈도 분포를 제공하는 단계를 더 포함하는, 빈도 분포 추정 방법.
- [청구항 6] 제2항에 있어서,
 상기 수집하는 단계 이전에,

상기 교란 위치를 샘플링하기 위한 기준 확률을 정의하는 단계;
 상기 기준 확률에 기반하여 상기 원본 위치 및 상기 교란 위치 간의
 오차의 기댓값을 유도하는 단계; 및
 상기 오차의 기댓값을 최소화하도록 상기 기준 영역의 영역 파라미터를
 결정하는 단계를 더 포함하는, 빈도 분포 추정 방법.

[청구항 7] 제6항에 있어서,
 상기 기준 확률은,
 상기 교란 위치에 대한 확률이고, 상기 교란 위치가 상기 기준 영역 안에
 있는 경우를 위한 제 1 기준 확률 및 상기 교란 위치가 상기 기준 영역
 밖의 상기 도메인 영역 내에 있는 경우를 위한 제2 기준 확률을 포함하는,
 빈도 분포 추정 방법.

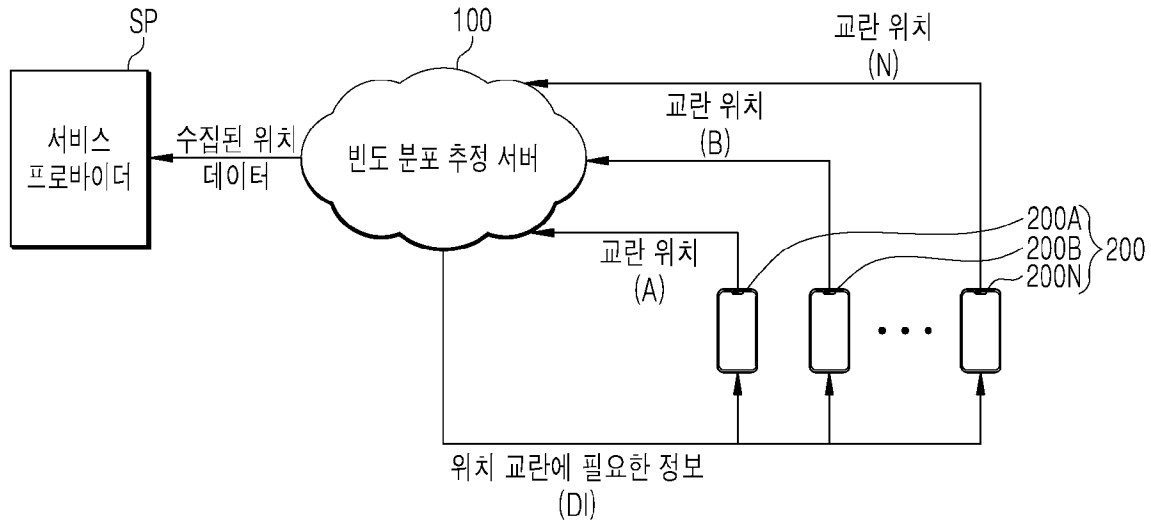
[청구항 8] 위치 데이터에 대한 빈도 분포를 추정하는 방법으로서,
 도메인 영역에서 단말의 원본 위치에 대응하는 교란 위치(perturbed
 location)를 수집하는 단계;
 상기 도메인 영역이 2차원의 그리드에 매핑되는 경우, 상기 그리드의 각
 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을
 유도하는 단계;
 상기 유도된 조건부 기댓값을 포함하는 블록 최적화 문제를 정의하는
 단계; 및
 상기 정의된 블록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에
 대한 원본 위치의 빈도 분포를 추정하는 단계를 포함하는, 빈도 분포 추정
 방법.

[청구항 9] 위치 데이터에 대한 빈도 분포를 추정하는 서버로서,
 도메인 영역에서 단말의 원본 위치에 대응하는 교란 위치를 수집하는
 위치 정보 수집부; 및
 상기 도메인 영역이 2차원의 그리드에 매핑되는 경우, 상기 그리드의 각
 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을
 전위합 행렬을 이용하여 유도하고, 상기 유도된 조건부 기댓값을
 포함하는 블록 최적화 문제를 정의하는 서버 프로세서를 포함하며,
 상기 서버 프로세서는,
 상기 정의된 블록 최적화 문제의 해에 기초하여, 상기 그리드의 각 셀에
 대한 전위합 기반 원본 위치의 빈도 행렬을 예측하고, 상기 예측된 전위합
 기반 원본 위치의 빈도 행렬에 기초하여, 원본 위치의 빈도 분포를
 추정하도록 구성되는, 빈도 분포 추정 장치.

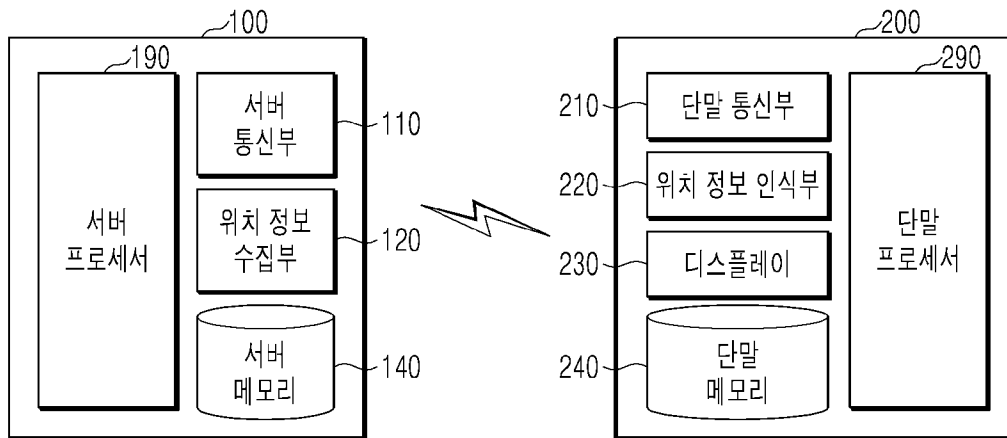
[청구항 10] 제9항에 있어서,
 상기 서버 프로세서는,
 상기 원본 위치가 교란 위치로 샘플링될 때 사용된 기준 영역, 샘플링
 파라미터값, 프라이버시 파라미터값 및 상기 셀의 넓이 중 적어도 하나의

- 정보를 획득하도록 구성되는, 빈도 분포 추정 서버.
- [청구항 11] 제10항에 있어서,
상기 서버 프로세서는,
상기 그리드의 각 셀에 대해, 셀과의 거리에 기초하여 구분된 기준 영역에 대한 서브 그리드 그룹을 설정하고, 상기 설정된 서브 그리드 그룹 별로 셀과의 중첩 면적, 해당 서브 그리드 그룹 별 전위합에 기초한 연산 및 상기 획득된 정보에 기초하여, 상기 그리드의 각 셀에 대한 원본 위치 대비 교란 위치에 대한 빈도 행렬의 조건부 기댓값을 전위합 행렬을 이용하여 유도하도록 구성되는, 빈도 분포 추정 서버.
- [청구항 12] 제9항에 있어서,
상기 블록 최적화 문제는,
교란 위치의 빈도 행렬과 상기 조건부 기댓값의 차가 최소화되기 위한 문제인, 빈도 분포 추정 서버.
- [청구항 13] 제9항에 있어서,
서버 통신부를 더 포함하며,
상기 서버 프로세서는,
상기 추정된 원본 위치의 빈도 분포를 상기 서버 통신부를 통해 서비스 프로바이더에 제공하도록 구성되는, 빈도 분포 추정 서버.
- [청구항 14] 제10항에 있어서,
상기 서버 프로세서는,
상기 교란 위치를 샘플링하기 위한 기준 확률을 정의하고, 상기 기준 확률에 기반하여 상기 원본 위치 및 상기 교란 위치 간의 오차의 기댓값을 유도하며, 상기 오차의 기댓값을 최소화하도록 상기 기준 영역의 영역 파라미터를 결정하도록 구성되는, 빈도 분포 추정 서버.
- [청구항 15] 제14항에 있어서,
상기 기준 확률은,
상기 교란 위치에 대한 확률이고, 상기 교란 위치가 상기 기준 영역 안에 있는 경우를 위한 제 1 기준 확률 및 상기 교란 위치가 상기 기준 영역 밖의 상기 도메인 영역 내에 있는 경우를 위한 제2 기준 확률을 포함하는, 빈도 분포 추정 서버.

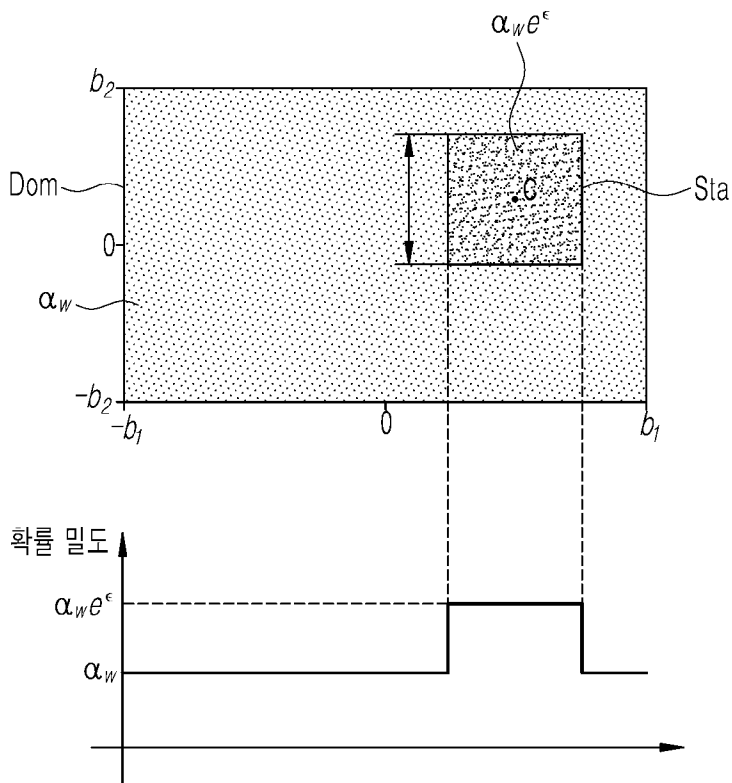
[도1]



[도2]



[도3]



[도4]

SC

Input: An original point $p \in \mathcal{D}$, domain $\mathcal{D}$, privacy budget ϵ , side length w^*

Output: The perturbed point $\tilde{p}$

1 Compute the center c of $C^*(p)$ by the following equation:

$$c[d] = \begin{cases} -b_d + \frac{w}{2} & \text{if } p[d] \leq -b_d + \frac{w}{2}, \\ p[d] & \text{if } -b_d + \frac{w}{2} < p[d] \leq b_d - \frac{w}{2}, \\ b_d - \frac{w}{2} & \text{otherwise.} \end{cases}$$

2 $\alpha^* \leftarrow 1/(w^{*2}(e^\epsilon - 1) + 4b_1b_2)$.

3 Sample x uniformly from $[0, 1]$.

4 if $x \leq 4\alpha^*b_1b_2$ then

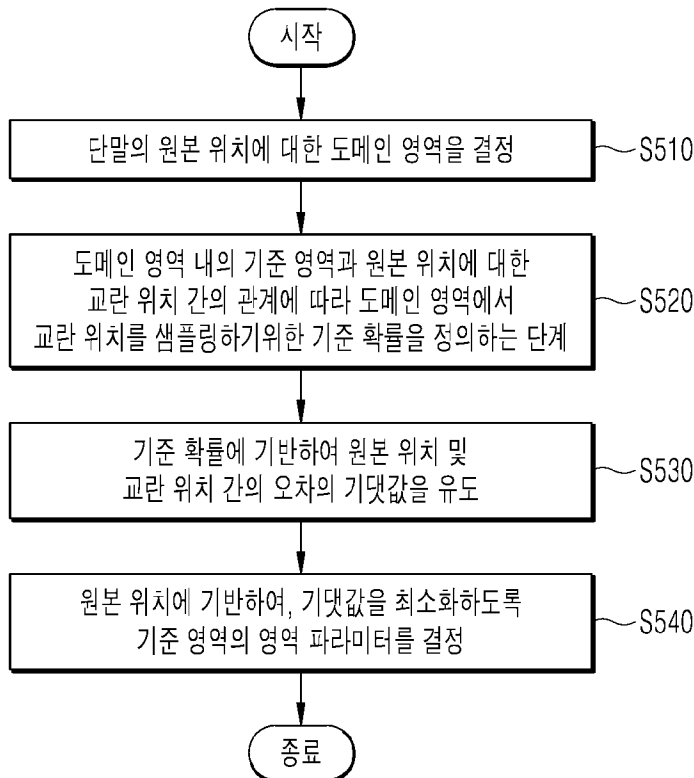
5 └ Sample $\tilde{p}$ uniformly from $\mathcal{D}$

6 else

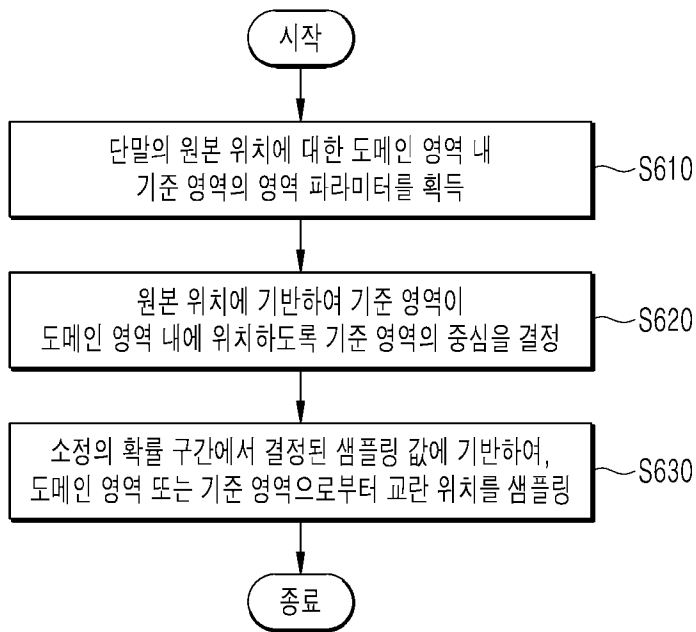
7 └ Sample $\tilde{p}$ uniformly from $C^*(p)$.

8 return $\tilde{p}$

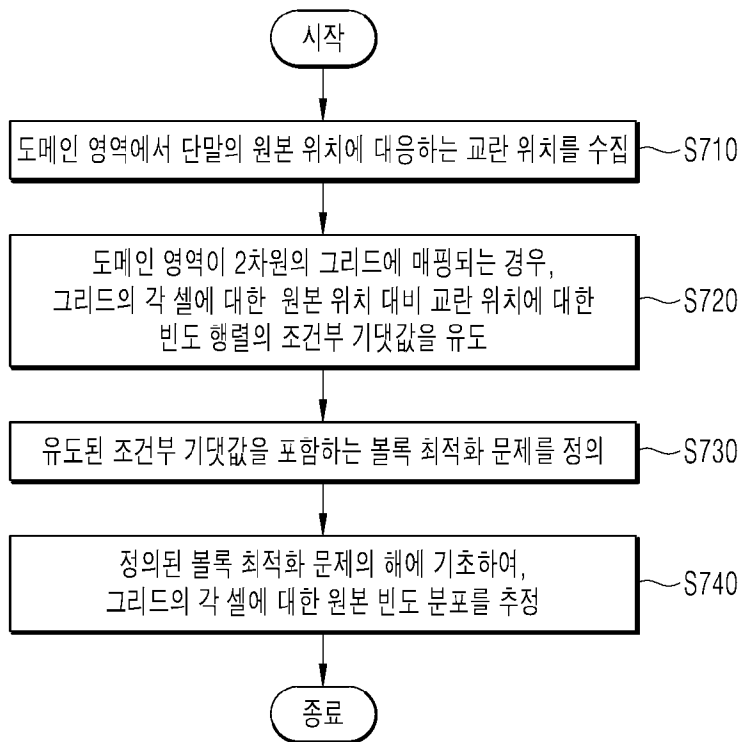
[도5]



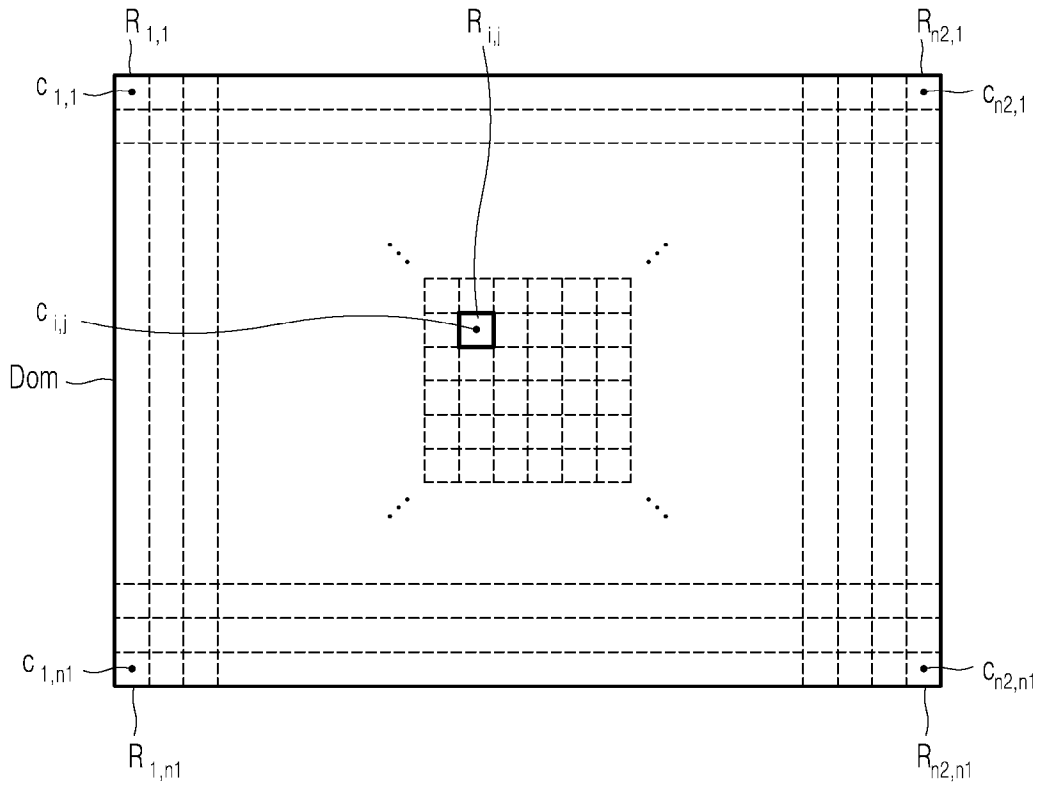
[도6]



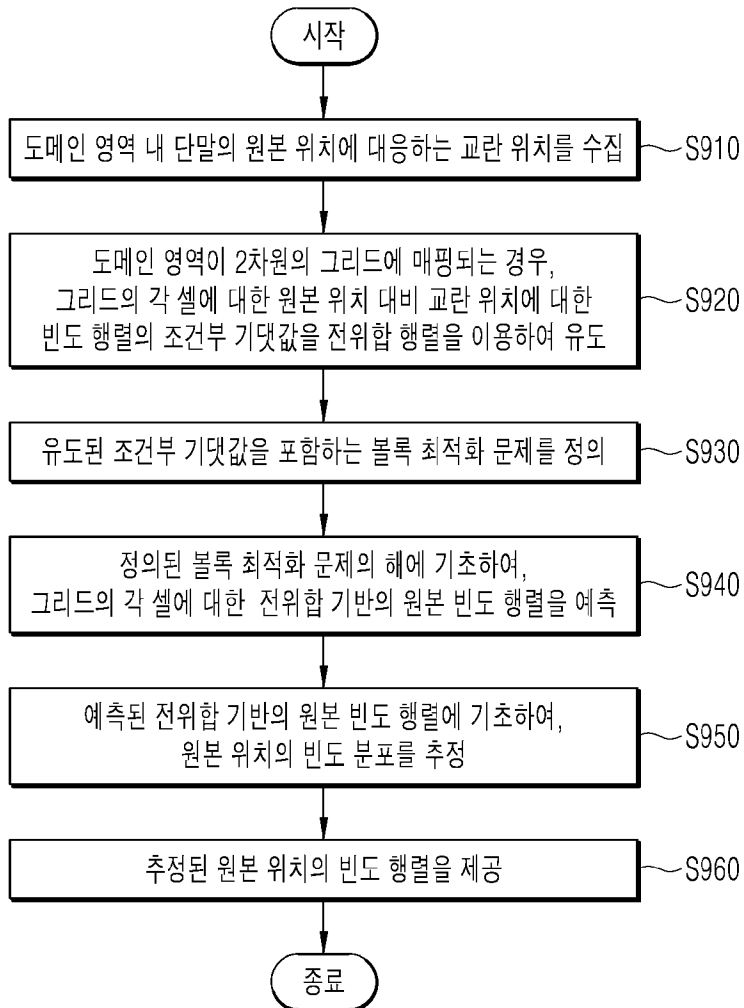
[도7]



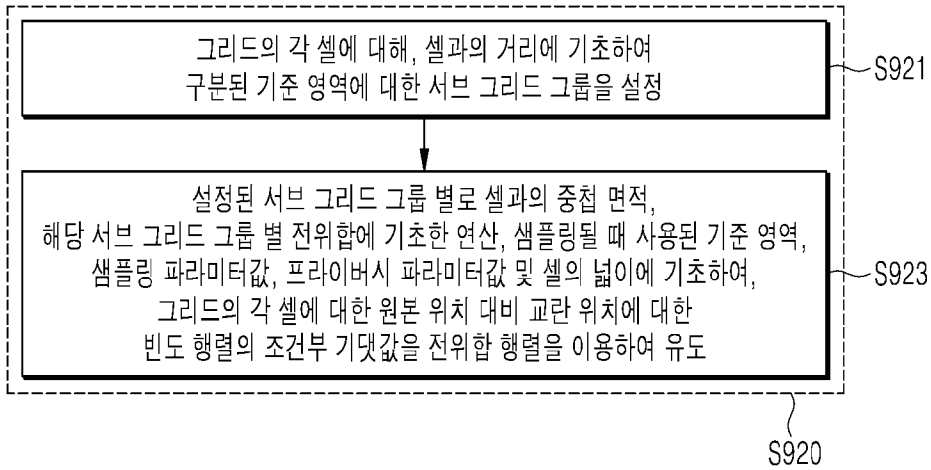
[도8]



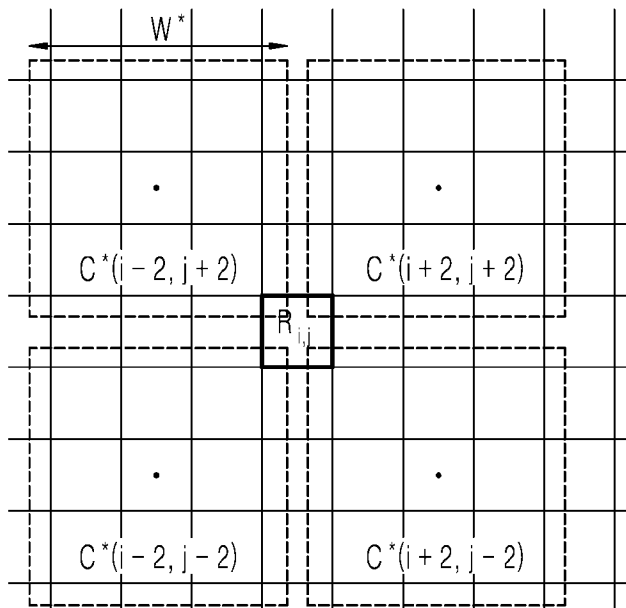
[도9]



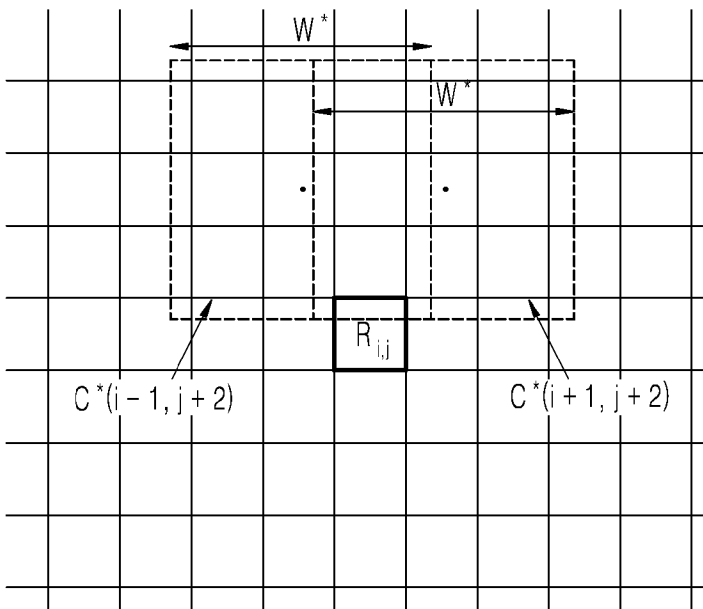
[도 10]



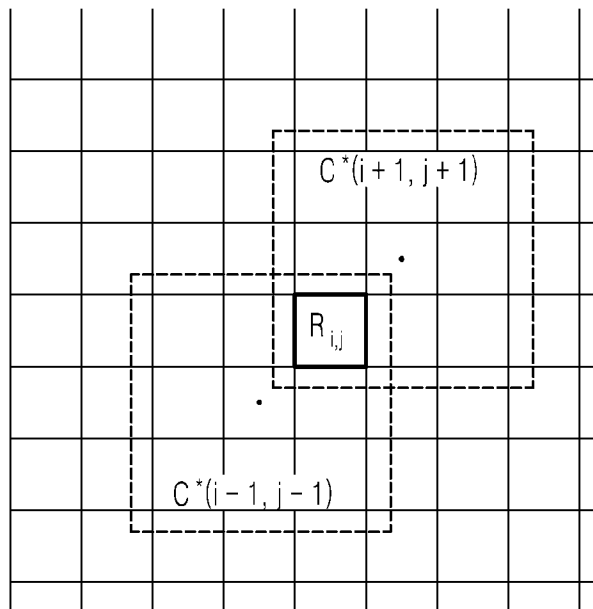
[도 11a]



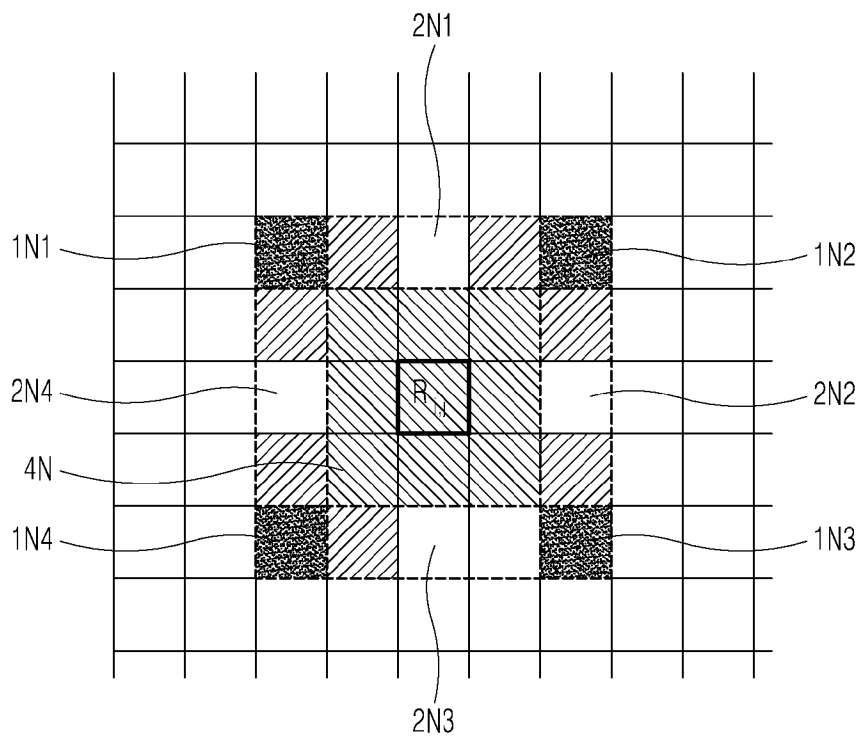
[도 11b]



[도11c]



[도12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2022/005599

A. CLASSIFICATION OF SUBJECT MATTER H04L 9/40(2022.01)i; H04L 67/52(2022.01)i; H04W 12/02(2009.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) H04L 9/40(2022.01); G06F 21/62(2013.01); G06F 7/58(2006.01); H04W 84/02(2009.01) Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Korean utility models and applications for utility models: IPC as above Japanese utility models and applications for utility models: IPC as above Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) local differential privacy(LDP), 교란 정보(perturbed information), 평균 제곱 오차(MSE), 2차원(2-dimensional), 블록 최적화(convex optimization), 빈도 분포(frequency distribution)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WANG, Ning et al. Collecting and Analyzing Multidimensional Data with Local Differential Privacy. arXiv:1907.00782v1 [cs.CR]. 28 June 2019. See sections II-III.	1-15
A	GU, Xiaolan et al. Providing Input-Discriminative Protection for Local Differential Privacy. arXiv:1911.01402v2 [cs.CR]. 24 February 2020. See sections II-V.	1-15
A	JIANG, Bo et al. Local Information Privacy with Bounded Prior. 2019 IEEE International Conference on Communications (ICC). 15 July 2019. See section III.	1-15
A	US 2019-0272388 A1 (ETRON TECHNOLOGY, INC.) 05 September 2019 (2019-09-05) See paragraphs [0033]-[0051].	1-15
A	US 2011-0149929 A1 (KLEIDER, John E. et al.) 23 June 2011 (2011-06-23) See paragraph [0026].	1-15
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 18 July 2022		Date of mailing of the international search report 19 July 2022
Name and mailing address of the ISA/KR Korean Intellectual Property Office Government Complex-Daejeon Building 4, 189 Cheongsaro, Seo-gu, Daejeon 35208 Facsimile No. +82-42-481-8578		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/KR2022/005599

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	HONG, Daeyoung et al. Collecting Geospatial Data with Local Differential Privacy for Personalized Services. 2021 IEEE 37th International Conference on Data engineering (ICDE). 19-22 April 2021. See sections I-V. * This document is a document declaring exceptions to lack of novelty in an earlier application that serves as a basis for claiming priority of the present international application.	1-15

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/KR2022/005599

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2019-0272388	A1	05 September 2019	CN	110221809	A	10 September 2019
				TW	201937389	A	16 September 2019
				TW	202046138	A	16 December 2020
				TW	1702505	B	21 August 2020
US	2011-0149929	A1	23 June 2011	US	2009-0003308	A1	01 January 2009
				US	2009-0011722	A1	08 January 2009
				US	2009-0052561	A1	26 February 2009
				US	2010-0172427	A1	08 July 2010
				US	2012-0121003	A1	17 May 2012
				US	8098713	B2	17 January 2012
				US	8131218	B2	06 March 2012
				US	8331420	B2	11 December 2012
				US	8331863	B2	11 December 2012
				US	8358644	B2	22 January 2013
				US	8379752	B2	19 February 2013

A. 발명이 속하는 기술분류(국제특허분류(IPC)) H04L 9/40(2022.01)i; H04L 67/52(2022.01)i; H04W 12/02(2009.01)i		
B. 조사된 분야 조사된 최소문헌(국제특허분류를 기재) H04L 9/40(2022.01); G06F 21/62(2013.01); G06F 7/58(2006.01); H04W 84/02(2009.01) 조사된 기술분야에 속하는 최소문헌 이외의 문헌 한국등록실용신안공보 및 한국공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 일본등록실용신안공보 및 일본공개실용신안공보: 조사된 최소문헌란에 기재된 IPC 국제조사에 이용된 전산 데이터베이스(데이터베이스의 명칭 및 검색어(해당하는 경우)) local differential privacy (LDP), 교란 정보(perturbed information), 평균 제곱 오차(MSE), 2차원(2-dimensional), 블록 최적화(convex optimization), 빈도 분포(frequency distribution)		
C. 관련 문헌		
카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
A	NING WANG 등, 'Collecting and Analyzing Multidimensional Data with Local Differential Privacy', arXiv:1907.00782v1 [cs.CR], 2019.06.28 섹션 II-III	1-15
A	XIAOLAN GU 등, 'Providing Input-Discriminative Protection for Local Differential Privacy', arXiv:1911.01402v2 [cs.CR], 2020.02.24 섹션 II-V	1-15
A	BO JIANG 등, 'Local Information Privacy with Bounded Prior', 2019 IEEE International Conference on Communications (ICC), 2019.07.15 섹션 III	1-15
A	US 2019-0272388 A1 (ETRON TECHNOLOGY, INC.) 2019.09.05 단락 [0033]-[0051]	1-15
A	US 2011-0149929 A1 (JOHN E. KLEIDER 등) 2011.06.23 단락 [0026]	1-15
☒ 추가 문헌이 C(계속)에 기재되어 있습니다. ☒ 대응특허에 관한 별지를 참조하십시오.		
* 인용된 문헌의 특별 카테고리: “A” 특별히 관련이 없는 것으로 보이는 일반적인 기술수준을 정의한 문헌 “D” 본 국제출원에서 출원인이 인용한 문헌 “E” 국제출원일보다 빠른 출원일 또는 우선일을 가지나 국제출원일 이후에 공개된 선출원 또는 특허 문헌 “L” 우선권 주장에 의문을 제기하는 문헌 또는 다른 인용문헌의 공개일 또는 다른 특별한 이유(이유를 명시)를 밝히기 위하여 인용된 문헌 “O” 구두 개시, 사용, 전시 또는 기타 수단을 언급하고 있는 문헌 “P” 우선일 이후에 공개되었으나 국제출원일 이전에 공개된 문헌 “T” 국제출원일 또는 우선일 후에 공개된 문헌으로, 출원과 상충하지 않으며 발명의 기초가 되는 원리나 이론을 이해하기 위해 인용된 문헌 “X” 특별한 관련이 있는 문헌. 해당 문헌 하나만으로 청구된 발명의 신규성 또는 진보성이 없는 것으로 본다. “Y” 특별한 관련이 있는 문헌. 해당 문헌이 하나 이상의 다른 문헌과 조합하는 경우로 그 조합이 당업자에게 자명한 경우 청구된 발명은 진보성이 없는 것으로 본다. “&” 동일한 대응특허문헌에 속하는 문헌		
국제조사의 실제 완료일 2022년07월18일(18.07.2022)		국제조사보고서 발송일 2022년07월19일(19.07.2022)
ISA/KR의 명칭 및 우편주소 대한민국 특허청 (35208) 대전광역시 서구 청사로 189, 4동 (둔산동, 정부대전청사) 팩스 번호 +82-42-481-8578		심사관 김성희 전화번호 +82-42-481-3516

C. 관련 문헌

카테고리*	인용문헌명 및 관련 구절(해당하는 경우)의 기재	관련 청구항
PX	DAEYOUNG HONG 등, 'Collecting Geospatial Data with Local Differential Privacy for Personalized Services', 2021 IEEE 37th International Conference on Data engineering (ICDE), 2021.04.19-22 섹션 I-V * 위 문헌은 본 국제출원의 우선권주장의 기초가 되는 선출원에서 '공지예외주장'된 문헌임.	1-15

국제조사보고서에서 인용된 특허문헌	공개일	대응특허문헌	공개일
US 2019-0272388 A1	2019/09/05	CN 110221809 A	2019/09/10
		TW 201937389 A	2019/09/16
		TW 202046138 A	2020/12/16
		TW I702505 B	2020/08/21
US 2011-0149929 A1	2011/06/23	US 2009-0003308 A1	2009/01/01
		US 2009-0011722 A1	2009/01/08
		US 2009-0052561 A1	2009/02/26
		US 2010-0172427 A1	2010/07/08
		US 2012-0121003 A1	2012/05/17
		US 8098713 B2	2012/01/17
		US 8131218 B2	2012/03/06
		US 8331420 B2	2012/12/11
		US 8331863 B2	2012/12/11
		US 8358644 B2	2013/01/22
		US 8379752 B2	2013/02/19