

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7520792号  
(P7520792)

(45)発行日 令和6年7月23日(2024.7.23)

(24)登録日 令和6年7月12日(2024.7.12)

(51)国際特許分類		F I		
G 0 8 G	1/00 (2006.01)	G 0 8 G	1/00	A
G 0 8 G	1/005(2006.01)	G 0 8 G	1/005	
G 0 8 G	1/13 (2006.01)	G 0 8 G	1/13	
G 0 1 C	21/26 (2006.01)	G 0 1 C	21/26	P

請求項の数 11 (全14頁)

(21)出願番号	特願2021-159425(P2021-159425)	(73)特許権者	000208891 K D D I 株式会社 東京都新宿区西新宿二丁目 3 番 2 号
(22)出願日	令和3年9月29日(2021.9.29)	(74)代理人	100165179 弁理士 田 崎 聡
(65)公開番号	特開2023-49602(P2023-49602A)	(74)代理人	100175824 弁理士 小林 淳一
(43)公開日	令和5年4月10日(2023.4.10)	(74)代理人	100114937 弁理士 松本 裕幸
審査請求日	令和5年7月18日(2023.7.18)	(72)発明者	米川 慧 埼玉県ふじみ野市大原二丁目 1 番 1 5 号 株式会社 K D D I 総合研究所内
		(72)発明者	黒川 茂莉 埼玉県ふじみ野市大原二丁目 1 番 1 5 号 株式会社 K D D I 総合研究所内 最終頁に続く

(54)【発明の名称】 P O I 系列データ拡張装置、P O I 系列データ拡張方法、およびプログラム

(57)【特許請求の範囲】

【請求項 1】

ユーザごと P O I の履歴であるユーザ P O I 系列を取得するユーザ P O I 系列取得部と、  
乱数を用いて前記ユーザ P O I 系列から一の P O I を決定し、当該一の P O I の近傍に存在する近傍 P O I を示す近傍 P O I 情報を取得する近傍 P O I 情報取得部と、

近傍 P O I 情報に含まれる前記近傍 P O I から 1 つの近傍 P O I を選択する選択部と、

前記選択部が選択した近傍 P O I と前記一の P O I とを置換するか、前記選択部が選択した近傍 P O I を前記ユーザ P O I 系列に挿入する置換挿入部と、

前記置換挿入部が置換または挿入した前記近傍 P O I を含む前記ユーザ P O I 系列を出力するユーザ P O I 系列出力部と、

を備える P O I 系列データ拡張装置。

【請求項 2】

前記近傍 P O I 情報には、複数の P O I どうしの物理的距離に基づいて、前記近傍 P O I が定められている客観的近傍 P O I 情報が含まれる請求項 1 に記載の P O I 系列データ拡張装置。

【請求項 3】

前記近傍 P O I 情報には、ユーザの行動に基づく複数の P O I どうしの関連の程度に基づいて、前記近傍 P O I が定められている主観的近傍 P O I 情報が含まれる請求項 2 に記載の P O I 系列データ拡張装置。

【請求項 4】

前記選択部は、前記客観的の近傍 P O I 情報と、前記主観的の近傍 P O I 情報とを所定の方法で選択することにより、前記客観的の近傍 P O I 情報または前記主観的の近傍 P O I 情報のいずれかから前記近傍 P O I を選択する

請求項 3 に記載の P O I 系列データ拡張装置。

【請求項 5】

前記所定の方法とは、ユーザの行動に基づく複数の P O I どうしの関連の程度を示す遷移確率に基づいて選択する方法である請求項 4 に記載の P O I 系列データ拡張装置。

【請求項 6】

前記遷移確率に基づいて選択する方法は、前記遷移確率の有効度合いが大きいほど、前記主観的の近傍 P O I 情報が選択される確率が大きくなる方法である請求項 5 に記載の P O I 系列データ拡張装置。

10

【請求項 7】

前記所定の方法とは、無作為に選択する方法である請求項 4 に記載の P O I 系列データ拡張装置。

【請求項 8】

前記主観的の近傍 P O I 情報は、行政区域ごとに設けられる請求項 3 から請求項 7 のいずれか 1 項に記載の P O I 系列データ拡張装置。

【請求項 9】

前記選択部は、時間帯に応じて近傍 P O I を選択する請求項 1 から請求項 8 のいずれか 1 項に記載の P O I 系列データ拡張装置。

20

【請求項 10】

P O I 系列データ拡張装置が実行する P O I 系列データ拡張であって、  
 ユーザごと P O I の履歴であるユーザ P O I 系列を取得するユーザ P O I 系列取得ステップと、  
乱数を用いて前記ユーザ P O I 系列から一の P O I を決定し、当該一の P O I の近傍に存在する近傍 P O I を示す近傍 P O I 情報を取得する近傍 P O I 情報取得ステップと、  
 近傍 P O I 情報に含まれる前記近傍 P O I から 1 つの近傍 P O I を選択する選択ステップと、  
 前記選択ステップが選択した近傍 P O I と前記一の P O I とを置換するか、前記選択ステップが選択した近傍 P O I を前記ユーザ P O I 系列に挿入する置換挿入ステップと、  
 前記置換挿入ステップが置換または挿入した前記近傍 P O I を含む前記ユーザ P O I 系列を出力するユーザ P O I 系列出力ステップと、  
 を有する P O I 系列データ拡張方法。

30

【請求項 11】

P O I 系列データ拡張装置のコンピュータに、  
 ユーザごと P O I の履歴であるユーザ P O I 系列を取得するユーザ P O I 系列取得ステップと、  
乱数を用いて前記ユーザ P O I 系列から一の P O I を決定し、当該一の P O I の近傍に存在する近傍 P O I を示す近傍 P O I 情報を取得する近傍 P O I 情報取得ステップと、  
 近傍 P O I 情報に含まれる前記近傍 P O I から 1 つの近傍 P O I を選択する選択ステップと、  
 前記選択ステップが選択した近傍 P O I と前記一の P O I とを置換するか、前記選択ステップが選択した近傍 P O I を前記ユーザ P O I 系列に挿入する置換挿入ステップと、  
 前記置換挿入ステップが置換または挿入した前記近傍 P O I を含む前記ユーザ P O I 系列を出力するユーザ P O I 系列出力ステップと、  
 を実行させるプログラム。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、P O I 系列データ拡張装置、P O I 系列データ拡張方法、およびプログラム

50

に関する。

【背景技術】

【0002】

人の移動パターンには自宅と職場を往復する等の規則性がみられるが、移動パターンは必ずしもいつも同じとは限らない。例えば、普段はめったに寄らない店でも、好奇心などのために通りがかりに寄ることがある。したがって、ユーザ毎のPOI (Point of Interest) 系列には確率的なゆらぎがあると考えられる。

【0003】

こうした位置の履歴に関する技術として、位置履歴データのデータ拡張手法として、緯度経度の正確な正解データとノイズの重畳した生データを用いて、ガウズノイズの分散を推定し、これを用いて正解データにノイズを重畳することでデータ拡張する方法が開示されている(例えば、非特許文献1参照)。この技術では、緯度経度データを対象としており、正解データを必要とする。

10

【0004】

また、位置履歴データのデータ拡張手法として、POI履歴を一定の時間間隔にし、POIの記録がついていない時間帯のPOIを推定する機械学習モデルを学習し、推定して穴埋めすることでデータ拡張する方法が示されている(例えば、非特許文献2参照)。この非特許文献2に開示された技術では、提案手法の比較対象として、2つのPOIを結ぶ線分の中央を基準として、最も近いPOIを選ぶか、周辺で最も人気のあるPOIを選ぶ方法について言及されている。

20

【先行技術文献】

【非特許文献】

【0005】

【文献】Feng, J., Li, Y., Zhao, K., Xu, Z., Xia, T., Zhang, J., Jin, D.: DeepMM: Deep Learning Based Map Matching with Data Augmentation. IEEE Trans. Mob. Comput. 3-6 (2020). <https://doi.org/10.1109/TMC.2020.3043500>.

【0006】

【文献】Li, Y., Luo, Y., Zhang, Z., Sadiq, S., Cui, P.: Context-aware attention-based data augmentation for POI recommendation. Proc. - 2019 IEEE 35th Int. Conf. Data Eng. Work. ICDEW 2019. 177-184 (2019). <https://doi.org/10.1109/ICDEW.2019.00-14>

30

【発明の概要】

【発明が解決しようとする課題】

【0007】

上述した非特許文献2に開示された技術では、機械学習モデルを学習する必要があることから事前準備のコストが大きく、さらに比較手法は決定的であるため、人の移動パターンにみられる確率的なゆらぎを表現できない。

本発明は、上記問題を解決すべくなされたもので、低コストで妥当なゆらぎをPOI系列に与えることができるPOI系列データ拡張装置、POI系列データ拡張方法、およびプログラムを提供することにある。

40

【課題を解決するための手段】

【0008】

(1) 本発明の一態様は、ユーザごとPOIの履歴であるユーザPOI系列を取得するユーザPOI系列取得部と、乱数を用いて前記ユーザPOI系列から一のPOIを決定し、当該一のPOIの近傍に存在する近傍POIを示す近傍POI情報を取得する近傍POI情報取得部と、近傍POI情報に含まれる前記近傍POIから1つの近傍POIを選択する選択部と、前記選択部が選択した近傍POIと前記一のPOIとを置換するか、前記選択部が選択した近傍POIを前記ユーザPOI系列に挿入する置換挿入部と、前記置換挿入部が置換または挿入した前記近傍POIを含む前記ユーザPOI系列を出力するユーザPOI系列出力部と、を備えるPOI系列データ拡張装置である。

50

(2) 本発明の一態様は、上記(1)に記載のPOI系列データ拡張装置において、前記近傍POI情報には、複数のPOIどうしの物理的距離に基づいて、前記近傍POIが定められている客観的の近傍POI情報が含まれる。

(3) 本発明の一態様は、上記(2)に記載のPOI系列データ拡張装置において、前記近傍POI情報には、ユーザの行動に基づく複数のPOIどうしの関連の程度に基づいて、前記近傍POIが定められている主観的の近傍POI情報が含まれる。

(4) 本発明の一態様は、上記(3)に記載のPOI系列データ拡張装置において、前記選択部は、前記客観的の近傍POI情報と、前記主観的の近傍POI情報とを所定の方法で選択することにより、前記客観的の近傍POI情報または前記主観的の近傍POI情報のいずれかから前記近傍POIを選択する。

10

(5) 本発明の一態様は、上記(4)に記載のPOI系列データ拡張装置において、前記所定の方法とは、ユーザの行動に基づく複数のPOIどうしの関連の程度を示す遷移確率に基づいて選択する方法である。

(6) 本発明の一態様は、上記(5)に記載のPOI系列データ拡張装置において、前記遷移確率に基づいて選択する方法は、前記遷移確率の有効度合いが大きいほど、前記主観的の近傍POI情報が選択される確率が大きくなる方法である。

(7) 本発明の一態様は、上記(4)に記載のPOI系列データ拡張装置において、前記所定の方法とは、無作為に選択する方法である。

(8) 本発明の一態様は、上記(3)から(7)のいずれかに記載のPOI系列データ拡張装置において、前記主観的の近傍POI情報は、行政区域ごとに設けられる。

20

(9) 本発明の一態様は、上記(1)から(8)のいずれかに記載のPOI系列データ拡張装置において、前記選択部は、時間帯に応じて近傍POIを選択する。

#### 【0009】

(10) 本発明の一態様は、POI系列データ拡張装置が実行するPOI系列データ拡張であって、ユーザごとPOIの履歴であるユーザPOI系列を取得するユーザPOI系列取得ステップと、乱数を用いて前記ユーザPOI系列から一のPOIを決定し、当該一のPOIの近傍に存在する近傍POIを示す近傍POI情報を取得する近傍POI情報取得ステップと、近傍POI情報に含まれる前記近傍POIから1つの近傍POIを選択する選択ステップと、前記選択ステップが選択した近傍POIと前記一のPOIとを置換するか、前記選択ステップが選択した近傍POIを前記ユーザPOI系列に挿入する置換挿入ステップと、前記置換挿入ステップが置換または挿入した前記近傍POIを含む前記ユーザPOI系列を出力するユーザPOI系列出力ステップと、を有するPOI系列データ拡張方法である。

30

#### 【0010】

(11) 本発明の一態様は、POI系列データ拡張装置のコンピュータに、ユーザごとPOIの履歴であるユーザPOI系列を取得するユーザPOI系列取得ステップと、乱数を用いて前記ユーザPOI系列から一のPOIを決定し、当該一のPOIの近傍に存在する近傍POIを示す近傍POI情報を取得する近傍POI情報取得ステップと、近傍POI情報に含まれる前記近傍POIから1つの近傍POIを選択する選択ステップと、前記選択ステップが選択した近傍POIと前記一のPOIとを置換するか、前記選択ステップが選択した近傍POIを前記ユーザPOI系列に挿入する置換挿入ステップと、前記置換挿入ステップが置換または挿入した前記近傍POIを含む前記ユーザPOI系列を出力するユーザPOI系列出力ステップと、を実行させるプログラムである。

40

#### 【発明の効果】

#### 【0011】

本発明によれば、低コストで妥当なゆらぎをPOI系列に与えることができるPOI系列データ拡張装置、POI系列データ拡張方法、およびプログラムを提供できる。

#### 【図面の簡単な説明】

#### 【0012】

【図1】本発明の実施形態のPOI系列データ拡張装置の一例を示す図である。

50

【図 2】置換例を示す図である。

【図 3】挿入例を示す図である。

【図 4】P O I 系列データ拡張処理の流れを示すフローチャートを示す図である。

【図 5】ポイント候補 N 1 ~ N j を示す図である。

【図 6】近傍情報取得部が取得したポイント候補を示す図である。

【図 7】遷移確率例を示す図である。

【図 8】近傍情報取得部が取得したポイント候補を示す図である。

【図 9】学習部を備えた P O I 系列データ拡張装置の構成例を示す図である。

【発明を実施するための形態】

【0013】

本実施形態の P O I 系列データ拡張装置、P O I 系列データ拡張方法、およびプログラムを、図面を参照しつつ説明する。以下で説明する実施形態は一例に過ぎず、本発明が適用される実施形態は、以下の実施形態に限られない。

なお、実施形態を説明するための全図において、同一の機能を有するものは同一符号を用い、繰り返しの説明は省略する。本実施形態における P O I は店舗、公共施設、駅、教育施設といったユーザの訪問の対象となる場所を示したカテゴリカルデータであるものとする。また、P O I には緯度経度といった P O I 間の相対的な距離の遠近を比較可能な情報が付与されているものとする。P O I 系列は、例えばスマートフォンなどのモバイル端末のアプリケーション等によって取得される。取得された P O I 系列は、例えば通信会社が管理するサーバなどにユーザ単位で記憶される。

【0014】

図 1 は、本発明の実施形態の P O I 系列データ拡張装置の一例を示す図である。本実施形態の P O I 系列データ拡張装置 100 は、ユーザごと P O I の履歴であるユーザ P O I 系列が入力され、入力されたユーザ P O I 系列にゆらぎを与えたゆらぎ済みユーザ P O I 系列を出力する。入力されるユーザ P O I 系列は、本実施形態では他の装置から取得される。P O I 系列データ拡張装置 100 がユーザ P O I 系列を記憶しておき、その中からユーザ P O I 系列を取得してもよいし、ユーザ P O I 系列が記憶された媒体（フラッシュメモリや光ディスクなど）から取得してもよい。

【0015】

P O I 系列データ拡張装置 100 は、系列取得部 110 と、近傍情報取得部 120 と、選択部 130 と、置換挿入部 140 と、出力部 150 とを備える。系列取得部 110 は、入力されたユーザ P O I 系列を取得する。近傍情報取得部 120 は、ユーザ P O I 系列に含まれる一の P O I の近傍に存在する近傍 P O I を示す近傍 P O I 情報を取得する。選択部 130 は、近傍 P O I 情報に含まれる近傍 P O I から 1 つの近傍 P O I を選択する。

【0016】

置換挿入部 140 は、選択部 130 が選択した近傍 P O I と一の P O I とを置換するか、P O I 選択部が選択した近傍 P O I をユーザ P O I 系列に挿入する。なお、挿入位置の詳細については後述する。

【0017】

これにより、ゆらぎが与えられたユーザ P O I 系列が生成される。出力部 150 は、置換挿入部 140 が置換または挿入した近傍 P O I を含むユーザ P O I 系列を出力する。

【0018】

上記構成において、系列取得部 110 は、ユーザ P O I 系列取得部の一例である。近傍情報取得部 120 は、近傍 P O I 情報取得部の一例である。出力部 150 は、ユーザ P O I 系列出力部の一例である。

【0019】

次に、置換挿入部 140 の置換と挿入例について説明する。以下の説明において、「系列」とはユーザ P O I 系列を示すものとする。また、「ポイント」とは、系列に含まれる P O I を示すものとする。

【0020】

10

20

30

40

50

図 2 は、置換例を示す図である。図 2 において、入力系列は、ポイント P 1、P 2、P 3 の順に接続されているものとする。なお、ポイント P 2 は、ユーザ P O I 系列に含まれる「一の P O I」に対応する。

【 0 0 2 1 】

入力系列に対し、置換挿入部 1 4 0 により、ポイント P 2 がポイント P A に置換されたものとする。この場合、P O I 系列データ拡張装置 1 0 0 は、ポイント P 1、P A、P 3 の順に接続された系列を出力する。これにより、ユーザがポイント P 1 の次にポイント P A を経由してポイント P 3 に移動するといったゆらぎが与えられた系列が得られる。

【 0 0 2 2 】

図 3 は、挿入例を示す図である。図 3 において、入力系列は、ポイント P 1、P 2、P 3 の順に接続されているものとする。なお、ポイント P 2 は、ユーザ P O I 系列に含まれる「一の P O I」に対応する。

10

【 0 0 2 3 】

図 3 に示されるように、置換挿入部 1 4 0 は、対象ポイント P 2 の前または後ろに P B を挿入する。例えば、ポイント P B がポイント P 2 の後ろに挿入されたものとする。この場合、P O I 系列データ拡張装置 1 0 0 は、ポイント P 1、P 2、P B、P 3 の順に接続された系列を出力する。これにより、ユーザがポイント P 2 の次にポイント P B を経由してポイント P 3 に移動するといったゆらぎが与えられた系列が得られる。

【 0 0 2 4 】

以下の説明において、図 2 のように置換するポイント P A や、図 3 のように挿入するポイント P B を「近傍ポイント」と表現する。

20

【 0 0 2 5 】

本実施形態では、近傍ポイントの選択方法として 3 つの選択方法がある。1 つめは、客観的的近傍選択方法である。2 つめは、主観的的近傍選択方法である。3 つめは、客観的的近傍選択方法と主観的的近傍選択方法とを使い分けるハイブリッド選択方法である。これらの選択方法について順に説明する。なお、以下の説明において、上述した「一の P O I」を「対象ポイント」と表現する。

【 0 0 2 6 】

選択方法の説明に先立ち、いずれの選択方法の場合であっても共通する全体的な処理の流れについて説明する。図 4 は、P O I 系列データ拡張装置 1 0 0 により実行される P O I 系列データ拡張処理の流れを示すフローチャートを示す図である。図 4 に示されるフローチャートでは、1 つの系列に存在する対象ポイント数は n 個としている。例えば、系列が、始点 P 0、P 1、P 2、...、P n、終点 P n+1 のように構成された系列のうちの、P 1、P 2、...、P n の n 個のポイントを対象ポイントとすること想定している。この場合は、始点と終点は動かさないようにゆらぎを与える処理となる。本実施形態では、始点と終点は動かさない例について説明するが、始点と終点の少なくとも一方を対象ポイントとしてもよい。

30

【 0 0 2 7 】

図 4 において、P O I 系列データ拡張装置 1 0 0 の系列取得部 1 1 0 は、系列を取得する（ステップ S 1 0 1）。P O I 系列データ拡張装置 1 0 0 は、ループカウンタ k を 1 で初期化する（ステップ S 1 0 2）。P O I 系列データ拡張装置 1 0 0 は、k 番目のポイントを置換挿入の対象ポイントとするか否かを判定する（ステップ S 1 0 3）。

40

【 0 0 2 8 】

このステップ S 1 0 3 は、系列の全ポイントを置換挿入の対象ポイントとしないようにするための処理である。全ポイントを置換挿入の対象ポイントとすると、入力系列から大幅に変更された系列が出力されることとなるため、例えば乱数を用いて低確率（例えば 1 0 % 以下の確率）で置換挿入の対象ポイントとするような判定が行われる。

【 0 0 2 9 】

k 番目のポイントを置換挿入の対象ポイントとしないと判定された場合には（ステップ S 1 0 3 : N O）、ステップ S 1 1 3 に処理が進む。k 番目のポイントを置換挿入の対象

50

ポイントとすると判定された場合には（ステップS103：YES）、POI系列データ拡張装置100は、対象ポイントと置換するか、対象ポイントの次に挿入するかのいずれかを判定する。ここでは、例えば乱数を用いて置換するか、挿入するかが判定される。

#### 【0030】

ステップS104において、置換すると判定された場合には、近傍情報取得部120は、k番目のポイントの近傍POI情報を取得する（ステップS105）。選択部130は、近傍POI情報に含まれる近傍ポイントから1つの近傍ポイントを選択する（ステップS106）。置換挿入部140は、選択部130が選択した近傍ポイントと対象ポイントとを置換して（ステップS107）、ステップS111に進む。

#### 【0031】

ステップS104において、挿入すると判定された場合には、近傍情報取得部120は、k番目のポイントの近傍POI情報を取得する（ステップS108）。選択部130は、近傍POI情報に含まれる近傍ポイントから1つの近傍ポイントを選択する（ステップS109）。置換挿入部140は、選択部130が選択した近傍ポイントを対象ポイントの前または後ろのポイントとして系列に挿入して（ステップS110）、ステップS111に進む。

#### 【0032】

POI系列データ拡張装置100は、 $k = n$ か否かを判定する（ステップS111）。 $k = n$ の場合には（ステップS111：YES）、全ての対象ポイントに対して処理を行ったこととなるため、出力部150は、系列を出力して（ステップS112）、処理を終了する。 $k < n$ の場合には（ステップS111：NO）、未処理の対象ポイントが存在しているため、POI系列データ拡張装置100は、kを増分して（ステップS113）、処理を終了する。

#### 【0033】

以上説明したPOI系列データ拡張処理により、POI系列データ拡張装置100は、ゆらぎを与えた系列を出力することができる。このPOI系列データ拡張処理を踏まえ、以下、各選択方法について説明する。以下の説明において、近傍ポイントの候補を「ポイント候補」と表現する。

#### 【0034】

[客観的近傍選択方法]

客観的近傍選択方法とは、複数のPOIどうしの物理的距離に基づく方法である。具体的に図5を用いて説明する。図5は、対象ポイントPのポイント候補 $N_1 \sim N_j$ を示す図である。客観的近傍選択方法では、対象ポイントPから距離d以内に位置する（すなわち半径d内に位置する）ポイントをポイント候補とする。

#### 【0035】

近傍情報取得部120は、これらのポイント候補を含む近傍情報を取得する。図6は、近傍情報取得部120が取得したポイント候補を示す図である。選択部130は、ポイント候補 $N_1 \sim N_j$ の中から、近傍ポイントを選択する。距離dは、例えば10000メートルなどと固定してもよいが、個々のユーザの標準的な移動距離（例えば、系列で連続する2つのPOI間の距離の平均値）に応じてユーザごとに設定されてもよい。

#### 【0036】

近傍情報取得部120は、定期的（例えば1日に1回など）、POIに付与された緯度経度から、例えばヒュニベの公式を用いてポイント候補を取得する。選択部130は、ポイント候補の中から、非確率的な基準（無作為）で近傍ポイントを選択する。なお、ポイント候補は、他の装置で用意されたものを近傍情報取得部120が取得してもよい。

#### 【0037】

客観的近傍選択方法によれば、対象ポイントからの距離を基準として近傍ポイントが選択されるので、機械学習モデルを学習する場合と比較して非常に低コストで妥当なゆらぎをPOI系列に与えることができる。

#### 【0038】

10

20

30

40

50

## [主観的近傍選択方法]

主観的近傍選択方法とは、ユーザの行動に基づく複数のPOIどうしの関連の程度に基づく方法である。具体的に実際の系列を集計して、例えばポイントPから他のポイント $N_k$  ( $k = 1 \sim r$ )への遷移確率を求めて置く。この遷移確率がPOIどうしの関連の程度を示す。図7は、遷移確率例を示す図である。図7において、ポイントPから $N_k$ への遷移確率は $P_k$ で示されている。遷移確率の算出方法としては、Pから任意のポイントに遷移している総数をQとし、そのうちのPから $N_k$ へ遷移している総数を $S(k)$ とした場合、 $P_k = S(k) / Q$ とする方法が挙げられる。言い換えると、系列に存在するPOIの前後をペアとして、ペアの出現回数をカウントし、全ペアの出現回数の合計で除算することで遷移確率が算出される。

10

## 【0039】

近傍情報取得部120は、遷移確率の大きい方から順にR位までの $N_k$ をポイント候補とし、これらを含む近傍情報を取得する。図8は、近傍情報取得部120が取得したポイント候補を示す図である。図8では、一例として $R = 20$ とした遷移確率が大きい上位20のポイント候補 $N'1 \sim N'20$ を示している。

## 【0040】

近傍情報取得部120は、例えば所定時刻が到来したタイミングや、POIが更新されたタイミングなど、遷移確率を算出してポイント候補を取得する。選択部130は、ポイント候補の中から、遷移確率に応じて、または非確率的な基準(無作為)で近傍ポイントを選択する。なお、ポイント候補は、他の装置で用意されたものを近傍情報取得部120が取得してもよい。

20

## 【0041】

主観的近傍選択方法によれば、実際のPOIに基づいて近傍ポイントを選択するという点において現実的なゆらぎをPOI系列に与えることができる。

## 【0042】

## [ハイブリッド選択方法]

ハイブリッド選択方法は、客観的近傍選択方法において取得されたポイント候補(「第1ポイント候補」とする)と、主観的近傍選択方法において取得されたポイント候補(「第2ポイント候補」とする)とを所定の基準で選択する方法である。

## 【0043】

ハイブリッド選択方法では、近傍情報取得部120は、客観的近傍選択方法において取得されたポイント候補(「第1ポイント候補」とする)と、主観的近傍選択方法において取得されたポイント候補(「第2ポイント候補」とする)とを予め用意しておく。

30

## 【0044】

その上で、ハイブリッド選択方法には、選択部130の選択方法として3種類の方法(、)がある。選択方法は、所定の基準を、非確率的な基準とする方法である。具体的に、選択部130が、第1ポイント候補および第2ポイント候補のいずれか一方を無作為(非確率的な基準)に選択し、選択したポイント候補の中から近傍ポイントを選択する方法である。

## 【0045】

選択方法は、所定の基準を、確率的な基準とする方法である。具体的に、選択方法は、第2ポイント候補を取得する際に用いられた遷移確率を用いる方法である。遷移確率には確率的に意味のあるもの(有意)とないものがある。そこで、POIの遷移パターンの全てが等確率で起こると仮定した場合の遷移確率と比較して、着目している遷移パターンが有意であるかどうかを、有意水準を0.05等とした比の差の検定により判定する。

40

## 【0046】

判定の結果、有意なものと判定されたうちの、遷移確率が上位(10位まで)を改めて第2ポイント候補とする。選択部130は、第2ポイント候補の中から近傍ポイントを選択する。その一方で、集計対象の系列が少ない場合などでは、遷移確率が10位までとなるものがない場合があり得る。このような場合には、選択部130は、第1ポイント候補

50

の中から近傍ポイントを選択する。選択方法 によれば、実際の P O I に基づいて近傍ポイントを選択することも可能であるとともに、有意ではない遷移確率の場合には、第 1 ポイント候補から選択されるので、低コストで妥当なゆらぎを P O I 系列に与えることができる。

#### 【 0 0 4 7 】

選択方法 は、第 2 ポイント候補を取得する際に用いられた遷移確率を用いる方法である。選択部 1 3 0 が、第 1 ポイント候補および第 2 ポイント候補のいずれか一方を確率にもとづき選択する方法である。例えば、第 2 ポイント候補を集計する系列に、対象ポイントが十分多く存在すれば、少ない場合と比較して第 2 ポイント候補はより有効と考えられる。

10

#### 【 0 0 4 8 】

そこで、集計する系列における対象ポイントの十分さの指標として、 $e ( 0 \leq e \leq 1 )$  を設ける。e の算出方法は、対象ポイントに対応する第 2 ポイント候補の数を M とし、第 2 ポイント候補の数の最大値を N としたとき、 $e = M / N$  とする。この e が大きいほど、第 2 ポイント候補の方が第 1 ポイント候補と比較して、より有効と考えられる。すなわち、関連の程度の有効度合いが大きいと考えられる。なお、「第 2 ポイント候補の数の最大値」とは、例えば上述したように上位 1 0 位を第 2 ポイント候補とする場合には、 $N = 1 0$  となる。したがって、第 2 ポイント候補の数が 1 0 位まで存在する対象ポイントでの指標 e は 1 となる。

#### 【 0 0 4 9 】

選択部 1 3 0 は、確率 e で第 2 ポイント候補から近傍ポイントを選択し、確率  $1 - e$  で第 1 ポイント候補から近傍ポイントを選択する。このようにすることで、低コストで妥当なゆらぎを P O I 系列に与えることができる。

20

#### 【 0 0 5 0 】

以上説明したように、本実施形態に係る P O I 系列データ拡張装置 1 0 0 は、妥当なゆらぎを与えた P O I 系列を出力する。そこで、P O I 系列データ拡張装置 1 0 0 は、妥当なゆらぎを与えた P O I 系列を学習対象として学習することにより、ユーザ P O I 系列に基づくユーザの位置情報を出力する学習部を備えてもよい。

#### 【 0 0 5 1 】

図 9 は、学習部を備えた P O I 系列データ拡張装置 2 0 0 の構成例を示す図である。図 9 に示される構成は、図 1 に示される構成に加え、ユーザ P O I 系列記憶部 1 6 0 と学習部 1 7 0 とを備える。

30

#### 【 0 0 5 2 】

図 9 において、出力部 1 5 0 は、ユーザ P O I 系列記憶部 1 6 0 に出力した系列を次々と追記していく。学習部 1 7 0 は、ユーザ P O I 系列記憶部 1 6 0 に記憶された系列を参照して学習を行う。ユーザ P O I 系列記憶部 1 6 0 には、ユーザごとに P O I 系列（本実施形態では、緯度と経度で示される P O I の系列）が記憶される。

#### 【 0 0 5 3 】

学習部 1 7 0 が、P O I 訪問予測モデルを構築する場合には、P O I 系列データ拡張装置によってゆらぎが与えられる前の系列で構築する場合と同様に行えばよい。例えば、P O I 訪問予測モデルが L S T M と全結合層と s o f t m a x 関数からなるニューラルネットワークである場合には、P O I 系列データのうち最新の P O I を除く部分を古いものから順に入力し、出力と 1 h o t エンコーディングされた最新の P O I とから交差エントロピー損失を算出し、誤差逆伝播法によりモデルパラメータを更新すればよい。

40

#### 【 0 0 5 4 】

第 1 ポイント候補または第 2 ポイント候補から選択される近傍ポイントで置換したり、または挿入する操作を変換と見なすと、本実施形態に係る P O I 系列データ拡張装置 2 0 0 を用いて学習させた機械学習モデルは、系列の確率的なゆらぎに対して頑健となる。したがって、P O I 系列データ拡張装置 2 0 0 によれば、系列を入力とした P O I 訪問予測モデルの性能が向上する。

50

## 【 0 0 5 5 】

なお、図9では、学習部はP O I系列データ拡張装置200に設けられているが、これに限るものではない。P O I系列データ拡張装置200とは異なる装置に設けられてもよい。例えば、P O I系列データ拡張装置100の通信可能な装置に設けられもよい。

## 【 0 0 5 6 】

以上説明した実施形態において、第2ポイント候補は、行政区域（都道府県や市町村）ごとに設けられてもよい。これは、ユーザの行動に基づく複数のP O Iどうしの関連の程度は行政区域ごとに異なると考えられるためである。このようにすることで、P O I系列データ拡張装置は、より実態に即した近傍ポイントを用いてゆらぎが与えられた系列を出力することができる。

10

## 【 0 0 5 7 】

また、選択部130は、時間帯に応じて近傍ポイントを選択してもよい。これは、例えば酒類を提供する飲食店は主に夜間にしか開店してなく、またデパートなどは深夜には開店していないことから、ある時間帯によっては訪問する可能性が0%となるポイントが存在する。さらに、ユーザによっては、ある時間帯は訪問しないが、それ以外の時間帯には訪問するようなP O Iも存在する。そこで、選択部130は、時間帯に応じて近傍ポイントを選択してもよい。このようにすることで、P O I系列データ拡張装置は、より実態に即した近傍ポイントを用いてゆらぎが与えられた系列を出力することができる。

## 【 0 0 5 8 】

なお、これにより、低コストで妥当なゆらぎをP O I系列に与えることができるP O I系列データ拡張装置、P O I系列データ拡張方法、およびプログラムを提供できることから、国連が主導する持続可能な開発目標（SDGs）の目標9「レジリエントなインフラを整備し、持続可能な産業化を推進するとともに、イノベーションの拡大を図る」に貢献することが可能となる。

20

## 【 0 0 5 9 】

以上、実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更、組合せを行うことができる。これら実施形態は、発明の範囲や要旨に含まれると同時に、特許請求の範囲に記載された発明とその均等の範囲に含まれるものである。

30

なお、上述したP O I系列データ拡張装置100、200は、コンピュータで実現するようにしてもよい。その場合、各機能ブロックの機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録する。この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、C P Uが実行することで実現してもよい。ここでいう「コンピュータシステム」とは、O S（O p e r a t i n g S y s t e m）や周辺機器などのハードウェアを含むものとする。

また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、R O M、C D - R O Mなどの可搬媒体のことをいう。また、「コンピュータ読み取り可能な記録媒体」は、コンピュータシステムに内蔵されるハードディスクなどの記憶装置を含む。

40

さらに「コンピュータ読み取り可能な記録媒体」とは、短時間の間、動的にプログラムを保持するものを含んでいてもよい。短時間の間、動的にプログラムを保持するものは、例えば、インターネットなどのネットワークや電話回線などの通信回線を介してプログラムを送信する場合の通信線である。

また、「コンピュータ読み取り可能な記録媒体」には、サーバやクライアントとなるコンピュータシステム内部の揮発性メモリのように、一定時間プログラムを保持しているものも含んでもよい。また上記プログラムは、前述した機能の一部を実現するためのものであってもよい。また、上記プログラムは、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるものであってもよい。また、上記プログラムは、プログラマブルロジックデバイスを用いて実現されるものであってもよい

50

。プログラマブルロジックデバイスは、例えば、FPGA (Field Programmable Gate Array) である。

【符号の説明】

【0060】

100、200...POI系列データ拡張装置、110...系列取得部、120...近傍情報取得部、130...選択部、140...置換挿入部、150...出力部、160...系列記憶部、170...学習部

10

20

30

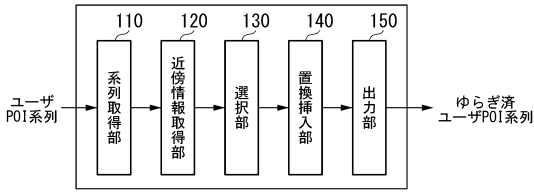
40

50

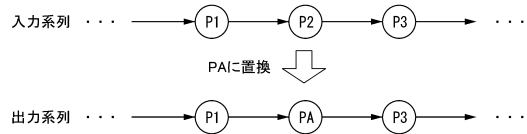
【図面】

【図 1】

100

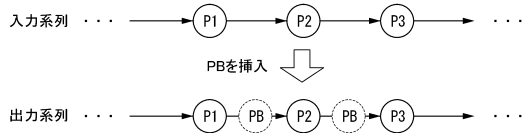


【図 2】

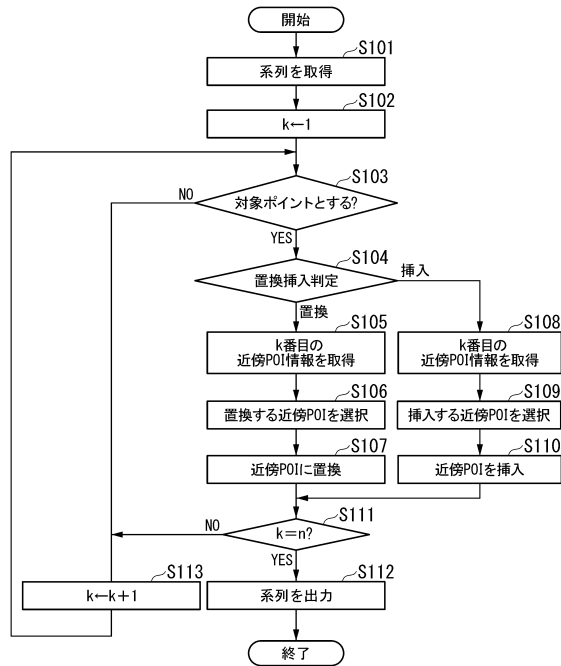


10

【図 3】



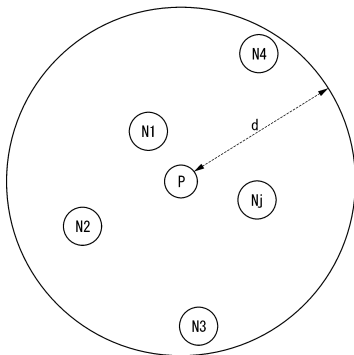
【図 4】



20

30

【図 5】



【図 6】

Pのポイント候補	N1	N2	...	Nj
----------	----	----	-----	----

40

50

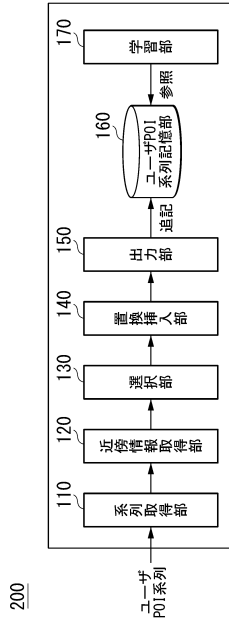
【図7】

	N1	N2	...	Nr
Pからの遷移確率	P1	P2	...	Pr

【図8】

Pの候補ポイント	N' 1	N' 2	...	N' 20
----------	------	------	-----	-------

【図9】



10

20

30

40

50

## フロントページの続き

(72)発明者 春田 秀一郎

埼玉県ふじみ野市大原二丁目1番15号 株式会社KDDI総合研究所内

審査官 高 木 真顕

(56)参考文献

特開2021-056761(JP,A)

特開2016-161339(JP,A)

特開2015-095083(JP,A)

特開2014-112301(JP,A)

特開2017-106779(JP,A)

米国特許出願公開第2015/0285652(US,A1)

(58)調査した分野 (Int.Cl., DB名)

G08G 1/00 - 99/00

G01C 21/00 - 21/36

G06Q 10/00 - 10/30

G06Q 30/00 - 30/08

G06Q 50/00 - 50/20

G06Q 50/26 - 99/00

G16Z 99/00