OUKA
The University of Osaka
Institutional Knowledge Archive

Title	モバイルセンサデータベースにおけるTop-k検索結果の多様化に関する研究
Author(s)	横山, 正浩
Citation	大阪大学, 2018, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/69721
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

様式3

論 文 内 容 の 要 旨

氏 名 （ 横山 正浩 ）	
論文題名	モバイルセンサデータベースにおけるTop-k検索結果の多様化に関する研究

論文内容の要旨

近年，半導体技術の発展により，スマートフォンやウェアラブルデバイスなどの人々が身につけて持ち歩く携帯端末に，音や温度，湿度などの環境情報を生成できるセンサが搭載されつつある．これらのセンサを搭載した携帯端末の普及に伴い，一般ユーザがもつ端末が生成したモバイルセンサデータ（環境情報および位置情報）を提供してもらい，都市部における環境モニタリングなどのサービスに利活用するユーザ参加型センシングが注目されている．端末保持者がセンシング領域を移動しながらデータを生成することで，従来の固定センサネットワークに比べて，より細かい地理的粒度でのデータ収集が可能となる．

収集されたモバイルセンサデータを参照し分析することで，センシング領域の環境情報を調査できる．一般的に，検索を行うユーザは自身の興味に適合する少数のデータにのみ関心がある．そこで，モバイルセンサデータの環境属性値から，ユーザの興味に基づいて算出されるスコアを定義し，よりユーザが関心を示す高いスコアをとる少数のデータを取得することが重要である．さらに，データ間の空間距離を併せて考慮し，地理空間上で広く分散し，かつ高いスコアをとるデータの集合を考える．このようなデータの集合を取得することで，その時々のユーザの興味に基づいた注目すべき領域を，広大なセンシング領域から効果的に割り出せる．高いスコアをとり，かつ地理空間上で分散するデータからなるデータ集合を，本論文では多様集合と呼び，また多様集合を取得することをTop-k検索結果の多様化と呼ぶ．モバイルセンサデータに対するTop-k検索結果の多様化は，広大なセンシング領域の環境情報を調査する方法として有用であるが，大量のモバイルセンサデータに対して多様集合を計算するための単純な手法を用いると，データセット全体を複数回走査する必要があり，計算コストが極めて大きくなってしまう．

そこで本論文では，モバイルセンサデータベースにおいて，ユーザの興味に適合し地理空間上で分散した多様集合を短時間で計算する手法について議論する．本論文は，5章から構成され，各章の内容は次の通りである．まず，第1章で序論を述べ，第2章において，環境属性値の空間的相関性を考慮した，クラスタリングを用いた効率的なTop-k検索結果の多様化手法を提案する．この手法では，クエリが到着する前にオフラインでモバイルセンサデータのクラスタリング処理を行い，生成されたクラスタの性質を利用した効率的なオンラインクエリ処理を行う．提案するオンラインクエリ処理手法では，作成したクラスタ構造の性質を利用し，データセット全体と比較して少数のモバイルセンサデータのみを走査することで，計算時間を短縮する．また，提案手法の性能評価のために行ったシミュレーション実験の結果を示し，その有効性について検証する．

次に第3章において，第2章の手法における環境情報が高次元である場合の問題点に言及し，その問題点を緩和する手法を提案する．第2章におけるクラスタリングでは，モバイルセンサデータの環境属性値ベクトルの類似度に基づいてクラスタリングを行っている．そのため，モバイルセンサデータが高次元である場合，クラスタの数が大幅に増加することでオンラインクエリ処理に要する時間が増加してしまう．そこで，第3章の手法では，階層的クラスタリング処理によりモバイルセンサデータを2階層のクラスタに構造化する．上位クラスタはモバイルセンサデータの空間位置の近接性のみを考慮して構成されるため，その数は環境属性の次元数に依存しない．これにより，クラスタの性質を利用したオンラインクエリ処理における，走査対象のデータの絞り込みのための計算コストを削減でき，第2章の手法よりも計算時間を短縮できる．また，提案手法の性能評価のために行ったシミュレーション実験の結果を示し，その有効性について検証する．

第4章では，ユーザの興味に基づいた注目すべき領域のリアルタイムの追跡を目的とし，モバイルセンサストリーム環境における継続的なTop-k検索結果の多様化手法を提案する．この手法では，新たに追加，削除されるモバイルセンサデータを短時間で構造化して管理するために，格子グリッドベースのデータ構造を用いる．モバイルセンサデータの位置情報から割当て先の空間セルを決定し，さらに空間セルの中で環境属性値に基づいて格納先のデータリストを決定する．格子グリッドベースのデータ構造の空間セルの性質を利用し，モニタリングの対象のデータセット全体と比較して少数のモバイルセンサデータのみを走査することで，計算時間を短縮する．また，提案手法の性能評価のために行ったシミュレーション実験の結果を示し，その有効性について検証する．

最後に，第5章では，本論文の成果を要約した後，今後の研究課題について述べる．

様式 7

論文審査の結果の要旨及び担当者

氏　名　（　横山　正浩　）		
論文審査担当者	（職）	氏　名
	主 査　教授	原　隆浩
	副 査　教授	藤原　融
	副 査　教授	鬼塚　真
	副 査　教授	松下　康之
	副 査　教授	下條　真司

論文審査の結果の要旨

近年、スマートフォンなどのセンサを搭載した携帯端末が取得したモバイルセンサデータを、その端末のユーザに提供してもらい、都市部における環境モニタリングなどのサービスに利活用するユーザ参加型センシングが注目されている。センシング領域の環境情報を調査することを目的とした場合、ユーザは気温や騒音といった環境属性値が周辺に比べて極めて大きい領域に特に関心があるため、ユーザが関心を示す少数のデータを取得することが重要である。また、ユーザが関心を示す領域を広大なセンシング領域から効果的に割り出すために、データ間の空間距離を併せて考慮する必要がある。このようなユーザの興味に適合し、かつ地理空間上で分散するk個のデータからなるデータ集合を、本論文では多様集合と表現し、また多様集合を取得することをTop-k検索結果の多様化と表現している。大量のモバイルセンサデータに対する多様集合の計算コストは極めて大きいため、効率的なTop-k検索結果の多様化手法が求められる。また、環境情報の高次元化に伴う計算効率の低下にも対応する必要がある。さらに、時々刻々と新しいデータが到着するモバイルセンサストリーム環境において、多様集合をリアルタイムにモニタリングする要求も考えられる。本論文では、これらの課題のそれぞれについて考慮し、モバイルセンサデータベースにおいて計算コストが小さいTop-k検索結果の多様化処理技術の実現を目標としている。本論文の主要な研究成果を要約すると次の通りである。

(1) モバイルセンサデータベースにおいて、環境属性値の空間的相関性を考慮したクラスタリングを用いたTop-k検索結果の多様化手法を提案している。この手法では、クエリが到着する前にオフラインで、空間位置の近接性および環境属性値ベクトルの類似度に基づいてデータをクラスタリングする。提案するオンラインクエリ処理手法では、作成したクラスタ構造の性質を利用し、データセット全体と比較して少数のデータを含むクラスタのみを走査することで、既存手法で得られる多様集合と同一の検索結果を短時間で取得できる。

(2) モバイルセンサデータの環境属性値ベクトルが高次元である場合、(1)で提案した手法ではクラスタの数が大幅に増加することでオンラインクエリ処理に要する時間が増加してしまう。そこで、階層的クラスタリングを利用した、オンラインクエリ処理時間の増加を緩和する手法を提案している。この手法では、データの空間位置の近接性のみを考慮した、環境属性値ベクトルの次元数に非依存なクラスタを上位階層として作成し、(1)で提案した手法における走査するデータの絞り込みを高速化する。

(3) モバイルセンサストリーム環境における、効率的な多様集合のモニタリング手法を提案している。この手法では、新たに追加・削除されるデータを短時間で構造化して管理するために、格子グリッドベースのデータ構造を用いる。データの位置情報から割当て先の空間セルを決定し、さらに空間セルの中で環境属性値に基づいて格納先のデータリストを決定する。このようにして作成された空間セルの性質を利用し、モニタリングの対象のデータセット全体と比較して少数のデータのみを走査することで、計算時間を短縮できる。

以上のように、本論文はモバイルセンサデータベースにおけるTop-k検索結果の多様化に関する先駆的な研究として、情報科学に寄与するところが大きい。よって本論文は博士（情報科学）の学位論文として価値のあるものと認める。