

Title	位置依存型Pub/SubシステムにおけるTop-k検索手法に関する研究
Author(s)	西尾, 俊哉
Citation	大阪大学, 2020, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.18910/76655">https://doi.org/10.18910/76655</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

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

Osaka University

## 論文内容の要旨

氏名 ( 西尾 俊哉 )	
論文題名	位置依存型Pub/SubシステムにおけるTop-k検索手法に関する研究
論文内容の要旨	
<p>インターネット上では様々なデータが生成されており、近年では、飲食店が発信する広告など、位置情報が付加されたデータが大量に生成されている。これらの大量に生成されるデータ群の中から有用なデータをユーザへ配信するには、ユーザが事前に登録したキーワードなどの条件（サブスクリプションと呼ぶ）に合致するデータを配信するパブリッシュ・サブスクライブモデルが有用であり、このモデルを位置情報付きのデータに応用した位置依存型Pub/Subシステムが注目を集めている。また、データが大量に生成される環境では、ユーザに対して有用な上位k個のデータ（Top-kデータと呼ぶ）を検索するTop-k検索を行い、その結果のみを配信することが実用的である。位置依存型Pub/Subシステムでは、データが頻繁に生成および削除され、各サブスクリプション（ユーザ）に対するTop-kデータは時々刻々と変化するため、システムにおいて継続的にTop-k検索を行い、最新のTop-kデータをユーザへ配信することが有用である。本論文では、位置依存型Pub/Subシステムにおいて、情報受信側と情報配信側それぞれに対して、新たな付加価値を付けたサービスを提供するためのTop-k検索手法について論じる。</p> <p>各サブスクリプション（ユーザ）に対して、ユーザ個々のニーズに合わせてTop-kデータを計算することで、情報受信側は自身の要求により合致したデータを受信可能になる。これを実現するためには、ソーシャル関係に基づくユーザの潜在的な興味を考慮したり、ユーザの移動に伴うユーザの現在地の変化を考慮した上でTop-kデータを計算する必要がある。そこで、本論文ではまず、ソーシャル関係を考慮した上での継続的なTop-k検索、およびユーザの移動を考慮した上での継続的なTop-k検索を効率的に行う手法をそれぞれ提案する。</p> <p>一方、あるデータをTop-kデータに含むユーザを検索する逆Top-k検索を行えるサービスを提供することで、情報提供側は自身のデータに興味をもつ潜在顧客を把握することが可能になる。潜在顧客を把握し、それらの統計的な特性を分析することで、情報提供側は、よりユーザのニーズに合わせたデータの発信や新たな顧客の獲得などが可能になる。そこで、本論文では、逆Top-k検索を効率的に行う手法を提案する。</p> <p>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章では、序論として研究の背景と目的について述べる。</p> <p>第2章では、位置、キーワード、およびソーシャル関係に基づく継続的なTop-k検索手法を提案する。提案手法は、データの生成および削除に対して、各サブスクリプションのTop-kデータを高速に更新できる。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すことにより、提案手法の有効性を示す。</p> <p>第3章では、ユーザの移動を考慮した位置およびキーワードに基づく継続的なTop-k検索手法を提案する。提案手法は、ユーザの移動、データの生成および削除に対して、各サブスクリプションのTop-kデータを高速に更新できる。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すことにより、提案手法の有効性を示す。</p> <p>第4章では、位置およびキーワードに基づく逆Top-k検索手法を提案する。提案手法は、逆Top-k検索の解に含まれないユーザに対する処理を途中で中断することで、高速に解を計算できる。また、複数の逆Top-k検索を同時に行うとき、重複する処理を削減し高速に検索を行うため、バッチ処理およびマルチコア環境でバッチ処理を行う手法を提案する。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すことにより、提案手法の有効性を示す。</p> <p>最後に、第5章では、本論文の成果を要約し、今後の研究課題について述べ、本論文のまとめとする。</p>	

## 論文審査の結果の要旨及び担当者

氏 名 ( 西 尾 俊 哉 )			
	(職)	氏 名	
論文審査担当者	主 査	教授	原 隆浩
	副 査	教授	鬼塚 真
	副 査	教授	春本 要 (大阪大学データビリティフロンティア機構)
	副 査	教授	藤原 融
	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、位置情報が付加されたデータが大量に生成されている。大量に生成されるデータ群の中から有用なデータをユーザへ配信するには、ユーザが事前に登録したサブスクリプションに合致するデータを配信するパブリッシュ・サブスクライブモデルが有用であり、このモデルを位置情報付きのデータに応用した位置依存型Pub/Subシステムが注目を集めている。また、データが大量に生成される環境では、ユーザに対して有用な上位k個のデータを検索するTop-k検索を行い、その結果のみを配信することが実用的である。本論文では、位置依存型Pub/Subシステムにおいて、情報受信側のユーザと情報配信側のPoIそれぞれに対して、新たな付加価値を付けたサービスを提供するためのTop-k検索手法の実現を目標としている。具体的には、情報受信側のユーザに対して、ユーザ個々のニーズに合わせたデータを検索できる手法を実現し、情報配信側のPoIに対して、自身の発信したデータに興味を持つユーザである潜在顧客を把握できる手法を実現する。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) ユーザの潜在的な興味といった新たな条件を考慮した上で上位k個のデータを計算することで、ユーザの嗜好に合う結果を出力できると考えられる。そこで、ソーシャル関係に基づくユーザの潜在的な興味を考慮した上で、システムにおいて、継続的にTop-k検索を行う手法を提案している。提案手法では、サブスクリプションを木構造で管理し、木構造に対応するリストを用いて、ソーシャル関係に基づくスコアを管理する。これにより、新たにデータが生成されたとき、上位k個のデータが変化する可能性のあるサブスクリプションを効率的に発見できる。
- (2) ユーザの現在地といった動的に変化する条件を考慮した上で上位k個のデータを計算することで、ユーザは今現在の状況において有用なデータを受信できると考えられる。そこで、ユーザの移動を考慮した上で、システムにおいて、継続的にTop-k検索を行う手法を提案している。この手法では、ユーザの現在地を把握することなくサブスクリプションを管理できる新たなデータ構造を考案しており、このデータ構造に基づいて、新たに生成されたデータにより上位k個のデータが変化する可能性のあるサブスクリプションを効率的に発見できる。
- (3) Pub/Subシステムにおいて、あるデータに対する逆Top-k検索の解は、そのデータに対する潜在顧客とみなすことができる。そこで、Pub/Subシステムにおいて、逆Top-k検索を行う手法を提案している。提案手法では、データ集合およびサブスクリプション集合をキーワードごとに管理することで、逆Top-k検索の解となり得ないサブスクリプションに対する処理を途中で中断でき、高速に解を得ることができる。

以上のように、本論文は位置依存型Pub/SubシステムにおけるTop-k検索手法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。