

氏名	董于洋
学位の種類	博士(工学)
学位記番号	博甲第8998号
学位授与年月日	平成31年3月25日
学位授与の要件	学位規則第4条第1項該当
審査研究科	システム情報工学研究科

学位論文題目 A Study on Rank-aware Query Processing for Multidimensional Data  
(多次元データに対するランキング問合せ処理に関する研究)

主査	筑波大学 教授	理学博士	北川 博之
副査	筑波大学 教授	博士(工学)	三末 和男
副査	筑波大学 教授	博士(工学)	山本 幹雄
副査	筑波大学 教授	博士(工学)	天笠 俊之
副査	筑波大学 講師	博士(工学)	陳 漢雄
副査	筑波大学 教授	博士(工学)	佐藤 哲司

(図書館情報メディア研究科)

## 論文の要旨

本論文は、多次元データを対象にしたランキング問合せ処理に取り組んだものである。IoT、SNSなどの普及に伴い、多様な多次元データの利用が進みつつある。データオブジェクトを何らかの指標で評価し、評価値の順位にしたがったランキングリストを求めるランキング問合せは、データを活用する上で重要な技術である。本論文では、静的な多次元データとしてマーケティングデータ、動的な多次元データとして位置データ・キーワードデータを具体的な対象とし、それらに関する新たなランキング問合せとその効率的処理について提案している。

マーケティングデータに対しては、顧客属性と商品属性をそれぞれ多次元ベクトルとして表現し、顧客の商品に対する嗜好度合いを両ベクトルの内積値を指標として定量化するモデルについて検討している。はじめに、与えられた複数の商品群をより上位にランク付けするユーザを探すという Aggregate Reverse Rank Query (集約逆ランク検索) を提案している。この問合せを効率的に処理するため、低次元データ向きの bound-and-filter 手法と高次元データ向きの Grid-index 手法を提案し評価を行っている。さらに、この問合せを一般化し、商品群中の各商品に重みを与えることを可能とした Weighted Aggregate Reverse Rank Query (重み付き集約逆ランク検索) を提案し、効率的な問合せ処理手法を提案している。

位置データ・キーワードデータに対しては、位置とキーワードを多次元ベクトル属性としてもつ移動体オブジェクト群とクエリポイント群を想定し、各クエリポイントに対して移動体オブジェクト群を、両者の距離とキーワードの類似度を組み合わせた指標でランキングする問題を対象としている。

具体的には、オブジェクト群の位置とキーワードが時間的に動的に変化することを仮定し、各クエリポイントに対して継続的にトップ k 件のオブジェクトを維持する Continuous Search on Dynamic Spatial Keyword Objects (動的空間・キーワードオブジェクトに対する継続的検索) を提案している。この問合せに対して、オブジェクトとクエリポイントの情報を効率良く管理し、Grid-based index や PCL (Partial Cell List) という解候補バッファを用いる問合せ処理方式を提案し評価している。

## 審査の要旨

### 【批評】

多次元データに対するランキング問合せ処理は、情報検索、データマイニング、レコメンデーションなど、種々の分野に適応可能な重要な技術である。本論文においては、マーケティングデータと位置データ・キーワードデータを対象に、Aggregate Reverse Rank Query、Weighted Aggregate Reverse Rank Query、Continuous Search on Dynamic Spatial Keyword Objects の 3 種類のランキング問合せを新たに定義し、その効率的問合せ処理のための方式を提案している。マーケティングデータを対象とした従来の手法では、単一の商品を上位にランク付けするユーザを探すという問合せを対象としていたが、本論文の Aggregate Reverse Rank Query では、複数商品のバンドル販売などの状況を想定し複数の商品群を対象とした点に新規性が認められる。さらに、それを拡張した Weighted Aggregate Reverse Rank Query においては、商品群中の各商品にウェイトを与えることを可能にしている。本論文では、これらの検索が多次元空間における探索問題に帰着されることに着目し、空間索引を適用し効率的にフィルタリングする手法を提案するとともに、その実験的評価と理論的解析を行っている。さらに、商品群が多次元空間において分散した場合にフィルタリング効率が低下する問題に対して、クラスタリングを適用する手法を示した。さらに、次元数が多い高次元の場合に空間索引の有効性が低下するという問題に対して、grid-index と呼ぶ事前計算を用いる効率化手法を提案している。位置データ・キーワードデータを対象とする Continuous Search on Dynamic Spatial Keyword Objects に関しては、オブジェクトのもつ位置とキーワードの両者が時間と共に動的に変化する状況を扱っている点で、従来研究にない新規性が認められる。この問題に対しては、位置データに注目したグリッド構造の索引と、キーワードの類似度を組み合わせる絞り込みを行う独自の手法を提案している。さらに、オブジェクトの動的な変化に対して、PCL という構造を用いたバッファ管理アルゴリズムを用いて問合せ処理時間を大幅に削減できることを示した。

多次元データに対するランキング問合せに関して、マーケティングデータと位置データ・キーワードデータの検索において必要とされる新たな問題を定式化し、その効率的な処理手法を提案し、実験と理論的解析によりその有効性を示した点で、本論文は新規性と情報工学上の貢献が認められる。

### 【最終試験の結果】

平成 31 年 2 月 8 日、システム情報工学研究科において、学位論文審査委員の全員出席のもと、著者に論文について説明を求め、関連事項につき質疑応答を行った。その結果、学位論文審査委員全員によって、合格と判定された。

### 【結論】

上記の学位論文審査ならびに最終試験の結果に基づき、著者は博士(工学)の学位を受けるに十分な資格を有するものと認める。