

ユーザの嗜好を考慮する空間的スカイライン問合せ

兒玉 一樹[†] 飯島 裕一^{††} 石川 佳治[‡][†] 名古屋大学 工学部 電気・電子情報工学科 情報工学コース ^{††} 名古屋大学大学院情報科学研究科[‡] 名古屋大学 情報連携基盤センター

1 まえがき

モバイル機器の発展、位置情報技術の進展により、地図を用いた空間的なアプリケーションが近年盛んになっている。一方、近年では携帯機器の記憶容量の増大に伴い、小規模なデータベースを携帯機器内で稼働させ、その上位にさまざまなアプリケーションを構築できるようになっている。また、モバイル空間に限らず、さまざまな領域において、ユーザの興味を反映した適応的な情報の選択・提供技術が重要となってきており、データベース中のデータのランク付け、それによる情報の推薦の研究も盛んに行われている。

このような背景から、本研究では、携帯機器中にデータベースを格納し、モバイルユーザの状況（位置や嗜好に関する情報）に応じて適応的な情報の提供を実現するためのシステムの開発を進めている [1]。本稿では、スカイライン問合せの概念を発展させた適応的な情報提供手法について述べる。

2 システムの流れ

2.1 概要

ユーザの状況と嗜好から、適切な情報を提供できるようにするのが本システムの目的である。とりわけ、位置情報（ユーザの位置や移動状況）を考慮した情報提供を行うものとし、GPS などで取得した位置情報に応じて情報提供を行う。なお、ユーザの目的地や速度などはひとまずは考慮しないものとする。

各ユーザがどのような興味を有しているかという情報は、プロフィールとしてシステムに登録され、プロフィールに応じた情報提供が行われる。たとえば、レストランを探しているユーザは、「中華料理よりはイタリア料理が好み」といった情報を、プロフィールに登録しておく。システムには、そのユーザの「中華料理」、「イタリア料理」というカテゴリ間の選好関係が登録される。情報提供システムは、ユーザの現在の位置を考慮して周辺の情報（この場合はレストラン）を提供するが、その際にユーザのプロフィール中の選好情報を考慮して、候補の絞り込みを行う。絞り込まれた結果は、ユーザの興味を反映した形でランク付けされ、上位 k 件（ k の値はユーザにより指定される）が提供される。表示領域が限られている携帯端末においては、特に、この絞り込みとランク付けの手法が重要となる。

本システムは携帯機器上で稼働する組込み型の RDBMS 上での実現を図っている。本システムでは、空間問合せ機能を持つ、日立の Entier^{*}を用いている。Entier は、標準的な SQL に加え、空間データ型や空間問合せの機能を支援している。本システムはこの機能を活用する。

2.2 レストラン検索の例

システムの流れをレストランを検索する例を挙げて述べる。本研究では、ユーザに嗜好の情報を入力してもらい、その情報からデータの選択とランク付けを行う手法を考えている。

図 1 のようなレストランのデータベースを考える。このようなデータベースが、モバイルユーザが保持する携帯機器上の RDBMS に収められていると想定する。図 2 はレストランのデータを座標上にプロットしたものである。

レストラン	座標	種類	価格
a	(3, 8)	中華	低
b	(7, 5)	フレンチ	中
c	(7, 7)	和食	低
d	(5, 1)	イタリアン	高
e	(3, 4)	和食	中
f	(4, 8)	フレンチ	低
g	(5, 6)	中華	高
h	(1, 3)	イタリアン	低
i	(6, 2)	和食	中
j	(9, 3)	中華	高

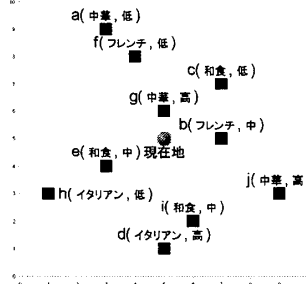


図 1: レストランデータ

図 2: 位置関係

例として、ユーザが以下のような嗜好を持っているとする。

- 現在いる地点からより近いレストランの方が好ましい
- 高級な料理のほうがよい (価格: 高 > 中 > 低)
- イタリアン > 和食 ≒ フレンチ > 中華という好みの順番がある: できればイタリアンで、中華はあまり好きではない、和食とフレンチは同程度

このようなユーザの選好情報がプロフィールとしてシステムに登録される。この場合、情報提供システムは距離、種類、価格の 3 点から候補を絞ることになる。距離、種類、価格の 3 点でほかに良い条件のレストランがあるならば、そのレストランを候補から外す。この例では、次章で述べる問合せ処理を行うと、b, d, g がユーザに提示される。

Spatial Skyline Queries That Considers User Preferences

[†] Kazuki Kodama, Department of Information Engineering, School of Engineering, Nagoya University^{††} Yuichi Iijima, Graduate School of Information Science, Nagoya University[‡] Yoshiharu Ishikawa, Information Technology Center, Nagoya University^{*} <http://www.hitachi.co.jp/Prod/comp/soft1/Entier/>

3 ユーザの嗜好を反映する問合せ処理

3.1 スカイライン問合せ

本研究では、ユーザの嗜好を反映した情報提供を実現するためにスカイライン問合せに着目している。スカイライン問合せは、大量のデータの中からユーザの興味に合う少数のデータを選択するために有効な手法であり、空間データベースへの問合せに適用した研究も見られる [2, 3]。直感的には、スカイラインオブジェクトとは、ある属性の組合せに着目したとき、データベース中の他のどのオブジェクトよりも優れているようなオブジェクトのことを指す。データベース中のすべてのスカイラインオブジェクトをスカイラインと呼び、これを見つめる問合せをスカイライン問合せという。

本研究で提案するスカイライン問合せのアルゴリズムを 2.2 節で挙げた例を用いて述べる。

まず、前処理として図 1 のデータを「種類」と「価格」(カテゴリ属性と呼ぶ)について射影する。このとき、選好関係から「和食 ≒ フレンチ」なので、「和食」と「フレンチ」を一体化する。前処理で得られたリレーションを図 3 に示す。これをカテゴリ表と呼ぶ。

種類	価格
イタリアン	高
イタリアン	低
和食/フレンチ	中
和食/フレンチ	低
中華	高
中華	低

図 3: カテゴリ表

これは、データベース中に現れるカテゴリ属性の全ての組合せのパターンである。

本システムでは下位の RDBMS が最近傍問合せ機能を提供することを想定している。現在地から最近傍のレストランを RDBMS から検索すると g が得られる。 g (中華, 高) は、距離の面で他のどのオブジェクトよりも優れているのでスカイラインに属する。つまり、この時点のスカイラインは $SQ = \{g\}$ となる。ここで、カテゴリ表から、現在のスカイラインに対して劣っている(支配されている)エントリを削除する。この例の場合、レストラン g (中華, 高) に支配されるパターンは(中華, 高)と(中華, 低)なのでこれらを削除する(図 4)。支配されたパターンをもつレストランは、検索されてもスカイラインには属さない。

ここで、次に近いレストランを検索する。この場合、 b が検索される。 b の組合せ(フレンチ, 中)は、カテゴリ表に存在している。これは、(フレンチ, 中)が支配対象でないことを意味している。つまり、 b はスカイラインに属するため、 $SQ = \{g, b\}$ となる。そして、同様に b (フレンチ, 中)に支配されるパターンを削除する(図 4)。

次に近いレストランは e である。 e の(和食, 中)はカテゴリ表に存在しないため、スカイラインには属さない。同様の処理を続けると、 c, f, i もスカイラインには含まれない。その後、 d を検索すると、組合せが表に存在するのでスカイラインとなる($SQ = \{g, b, d\}$)。 d

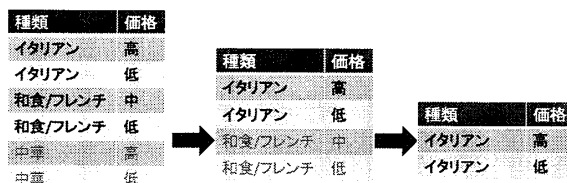


図 4: カテゴリ表の変化

に支配されるパターンを削除するとカテゴリ表が空になる。カテゴリ表が空になったときにアルゴリズムは終了し、 SQ の要素がスカイラインとして出力される。

3.2 スカイラインのランク付け

前述の例では 3 件のレストランがスカイラインとして提示された。しかし、ユーザはたとえば 5 件以上提示してほしい、と要求することもある。

そこで、全データから既にスカイラインとなったレストランを除いたデータ集合で再びスカイラインを計算するアプローチをとる。前節の例を用いて説明する。ユーザは 5 件以上候補を要求しているとする。全データで計算したとき、 g, d, b の 3 件が出力した。まだ 5 件でないので、スカイラインとなったこの 3 点を除いた残りのデータで再びスカイラインを計算する。手順は 3.1 節の例と同じである。計算過程は省略するが、結果的に e, h, j がスカイラインとなる。この処理によって合計 6 件提示することができる。最初に得られた g, d, b をランク 1、次点として得られた e, h, j をランク 2 のスカイラインと呼ぶ。さらに必要な場合は、これらのレストランも除いた残りのデータで再度スカイラインを計算する。

ユーザにとって最も有力な候補は最初に求めたスカイライン g, d, b だが、これらが不満なとき(前に行ったことがあるから別のところがいい等)は、 e, h, j が次点として挙げられる。

4 まとめと今後の課題

本稿では、モバイル空間でユーザの状況と嗜好に反映した情報提供の手法として、スカイライン問合せについて述べた。今後、システムの実装を行い、仮想データのシミュレーションや、実データを用いてモバイル上で実験をする予定である。

謝辞

本研究の一部は科学研究費補助金(19024037, 19300027, 18200005)による。

参考文献

- [1] 石川佳治, 兒玉 一樹, 飯島 裕一: 「ユーザの嗜好と位置に応じた情報選択・提供のためのデータベース問合せ手法」, 情報学ワークショップ, 2008 年.
- [2] Baihua Zheng, Ken C. K. Lee, Wang-Chien Lee: "Location-Dependent Skyline Query", MDM 2008.
- [3] Raymond Chi-Wing Wong, Ada Wai-Chee Fu, Jian Pei, Yip Sing Ho, Tai Wong, Yubao Liu: "Efficient Skyline Querying with Variable User Preference on Nominal Attributes", VLDB 2008.