

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 5 日現在

機関番号：12401

研究種目：基盤研究(C)

研究期間：2012～2014

課題番号：24500107

研究課題名(和文) 位置情報サービスの為の各種検索の高速化に関する研究

研究課題名(英文) Study on Efficient Query Algorithms for Location Based Services

研究代表者

大沢 裕 (OHSAWA, Yutaka)

埼玉大学・理工学研究科・教授

研究者番号：50152111

交付決定額(研究期間全体)：(直接経費) 4,000,000円

研究成果の概要(和文)：本研究では、位置情報サービス(LBS)において必要となる道路網距離に基づく各種検索を高速に実行し得るデータ構造とアルゴリズムを提案した。LBSでは、道路網上で2点間の距離を求める演算が重要であり、それを高速に求めることが必要になる。本研究では、この目的でSSMTA*アルゴリズムと、SMPV距離計算アルゴリズムの2種類を提案した。特にSMPV構造は従来方式に比して大幅に少ないデータ量でほぼ同量の処理時間を達成する。更に、それらを逆近接検索(RkNN検索)に適用し、大幅な処理時間の短縮を図った。加えて、長大な処理時間を要する旅行計画問題を実用的な処理時間で解くアルゴリズムを提案した。

研究成果の概要(英文)：In this study, we proposed fast algorithms to perform several kinds of queries in road network distances necessary in location based services (LBS). The basic operation in queries for LBS is the operation to search the road network distance between specified two points on the road network. This study proposed two fast algorithms; one is the single-source multi-target A* (SSMTA*) algorithm, and the other one is the simple materialized path view (SMPV). Especially, SMPV works on the pre-calculated distance data whose amount is only about five times larger than the adjacency list, and the performance is almost the same with conventional encoded path view methods. Applying those algorithms to RkNN query achieves several order of magnitude faster than the conventional algorithms. Furthermore, we proposed fast algorithms to solve the trip planning queries.

研究分野：時空間データベース

キーワード：位置情報サービス 地理情報処理 道路網

1. 研究開始当初の背景

近年、携帯電話やカーナビによる位置に関連した情報サービス(LBS: location based services)の利用が進んでいる。LBSにおいて必要となる検索には、現在位置から目的地への最適経路探索、近隣施設を k 個(任意個)探す k NN 検索、複数の検索点から総移動量が最短となる施設を探索する ANN(aggregate nearest neighbor)検索、複数の検索点から何らかの尺度で有意な施設のみを絞り込む空間的スカイライン検索など、様々ある。更に、現在地から目的地への旅行途中で銀行、レストラン、映画館などいくつかの施設を訪れる際の最適経路を探索する各種旅行計画も重要となる。

LBS の為の各種検索・探索に関する研究は 1990 年代から活発に行われている。しかし、初期の研究は移動に拘束が無いユークリッド距離に基づくものであった。一方、実用的な LBS の為の検索では、人や車の移動は道路網上に拘束される。即ち、道路網上で移動したときに、移動距離や時間が最小となる対象物の検索が重要となる。そのような道路網上で移動を前提とした各種検索方式の研究は 2003 年の Papadias らによるものが最初と思われる。その後、上記の多様な検索が検討されてきた。

ユークリッド距離に基づく検索と道路網上の距離に基づく検索の本質的な違いは、前者は距離計算が単純であり計算を高速に行える一方、後者では道路網上で距離の計算コストが高いことである。道路網上で距離の計算時間は、距離が長くなるにつれ長い処理時間を要する。一方、道路網上で距離を高速に算出するため、前計算(マテリアライズ化)を行う方式や近似距離を用いる方式が用いられるが、これらの方式では一般に道路網上のノード(交差点など)数を N とするとき、前処理により作られるデータ量が $O(N^2)$ 以上に及ぶ。例えば、 N を $1M(10^6)$ とすると、このデータ量は $1T$ となり非現実的なものとなる。また、前処理による方式では検索対象物(POI: point of interest)に動き(出現や消滅)があったとき、再度全域に対して再計算を必要とする問題もある。

申請者らは LBS に関して、双方向探索による高速単純旅行計画アルゴリズム(大沢、藤野、2010)、OSR 探索の高速化(藤井、大沢、他、2010)などの研究を行ってきた。特に OSR 探索では、従来方式の 100 倍程度の高速化を達成した。更に、準備研究として多数の目的地に対する A^* アルゴリズム(SSMTA*)を提案し、ANN 検索の高速化を達成した(Htoo、Ohsawa、et al. 2011)。本研究は、これらの研究により得られた経験と知見を基に行われた。

2. 研究の目的

本研究では、LBS で必要となる各種検索を道路網上で距離に基づき高速に算出する

アルゴリズムとデータ構造に関して研究する。

(1) ユークリッド距離に基づく検索で対象を絞り込み、道路網での距離で検証する枠組み(IER)において、検証に SSMTA*(single-source multi-target A^*)を用いることにより各種検索方式の高速化を達成する。

(2) 隣接リストを道路網上の距離で隣接するもの同士でブロック化し、各ブロック毎に道路網距離の簡易マテリアライズ化した構造(SMPV: simple materialized path view)と最適優先探索を用いることにより、各種検索の高速化を図る。

(3) 道路網距離に基づく旅行計画問題の高速化方式を提案する。

3. 研究の方法

(1) IER に基づく高速検索方式

IER(incremental Euclidean restriction)とは、ユークリッド距離で検索対象を絞り込み、その対象物が道路網上で距離で目的にかなったものであるかを確認する計算の枠組である。IER は 2003 年に Papadias らにより提案されて以来、多くの空間検索に適用されてきた。ユークリッド距離での候補点が算出された後、道路網での距離でそれを評価する際に、Dijkstra 法や A^* アルゴリズム(pair-wise A^*)が用いられてきた。しかしこれらの方式では、探索範囲が必要以上に広くなったり、同じ道路網上のノードに対して処理が繰り返される等の問題がある。本研究では、準備研究で検討を進めてきた SSMTA*アルゴリズムを更に改良し適用する。

(2) 隣接リストの効率化

道路網上で最短路を求める際に、道路網グラフの隣接リストの構成法が処理時間に影響する。実用規模の道路網の隣接リストは通常大きなデータ量になることから、適切なサイズのブロック(例えば 8KB)に分割され、必要な部分が LRU バッファに読み込んで参照される。ブロック分けは Peano 順序などにより空間的に近接するもの同士が同じブロックに含まれるように分割される。しかし、従来方式では道路セグメント同士が道路網上で近接するか否かは考慮されてこなかった。本研究では、道路網上で近接関係にあるノード同士をグループ化することにより LRU バッファアクセスの効率化を図り、各種検索時間の短縮を図る。道路網上で近接性に基づくブロック化には、Huang らにより提案された方法、ネットワークボロノイ図の作成と同様に、ランダムに設定された代表点からの並列 Dijkstra 法による領域分割などを試みる。

(3) 動的環境における前計算方式の効率化

道路網上で各種検索を高速に実行するためには、前もって道路網上の任意の 2 ノード間の距離を計算しておく方式(マテリアライズ化)が採られる。しかし、この方式の最大の問題は前計算の結果得られるデータ量が膨大になる点にある。道路網上のノード数

を N とするとき、それは $O(N^2)$ や $O(N^3)$ になる。検索対象 (POI) に対して前計算を行っておく方式も各種提案されているが、POI に出現・消滅の動きがある場合、多くは全域にわたる再計算を必要とする。本研究では検索対象と道路網を分離し、道路網に対しては簡易な前処理を行い、それを IER の枠組みで参照することにより多様な検索が高速に実現でき、検索対象の動きに対して前計算を再度行う必要の無い方式を提案する。

4. 研究成果

(1) SSMTA* アルゴリズムの改良と応用 (雑誌論文、学会発表)

SSMTA* アルゴリズムの基本

車や人は道路網上を移動するため、LBS において道路網上での距離に基づく検索が重要となる。これらの検索で基本となる演算は 2 点間の道路網距離の検索である。例えば kNN 検索を道路網距離で実行する場合に、ユークリッド距離での検索により kNN 候補を得て、それを道路網距離で検証するアプローチ (IER) がよく用いられる。この際に、検索点は固定であるが、ユークリッド距離上での検索により得られる候補点は多く、検索点からそれらの全ての候補点への道路網距離を求める必要がある。

2 点間の道路網距離を高速に求めるためには、従来 A* アルゴリズムが用いられてきた。しかし、その場合にも多数の候補点に対して繰り返し道路網距離の検索が繰り返されるため、検索処理の多くの時間を道路網距離を求めることに費やされていた。そこで、1 つの検索点 q に対して、多数の目的地集合 D を設定し、それぞれの目的地への道路網距離を並行的に求める方式 (SSMTA* アルゴリズム) を提案した。

この方式の基本は、A* アルゴリズムと同様にヒューリスティクスに基づく距離を設け、最短路が存在する可能性が高い領域から探索を開始し、徐々に探索範囲を拡大していくものである。このヒューリスティクス距離としてはユークリッド距離を用いた。探索は優先順位付キュー (PQ) を用いた最適優先探索で実行する。PQ は次に示すレコードを管理し、 C 値の昇順にレコードを取り出すものとする。

$\langle C, dr, n \rangle$

ここで、 n は現在の注目ノード、 dr は q から n への道路網距離、 C は n に最近接する D 中の要素へのユークリッド距離と dr との和である。探索中に訪れる各ノードでは D 中の要素の内、 n に最近接する要素を求める必要があるが、この計算はユークリッド距離で行うため、高速に計算できる。

探索がある目的地に到達したとき (即ち、1 つの目的地への道路網距離が求まった時) PQ の全てのレコードに対して C 値の再計算を行う。この再計算は既に最短路距離が求まった要素を D 中から除外し、まだ距離が求まっていない D 中の要素に対して最短のヒューリ

スティクス距離を再計算により求めるものである。この計算は、PQ 中の全ての要素に対して実行されるが、主記憶中での処理であるため高速実行可能である。

SSMTA* アルゴリズムの応用

SSMTA* アルゴリズムを ANN (aggregate nearest neighbor) 検索に応用した。ANN 検索は、複数の検索点集合 Q 、対象のデータ集合 P 、ある集約関数 f 、が与えられたとき、全ての Q からの距離が f の条件の下で最適な P 中の要素を求める検索である。例えば、 f として距離の総和 (SUM) が与えられたときは、 Q 中の全ての要素からの距離の総和が最小となる P 中の要素を求める。

この検索では、 Q 中の各々の点から P 中の候補要素への道路網距離の検索を繰り返し実行する必要がある。この処理にかかる時間が全体の処理時間の大部分を占める。そこで、本研究では SSMTA* アルゴリズムをこの道路網距離での検証部分に適用することにより、大幅な検索時間の短縮を図った。

(2) 道路網距離の簡易マテリアライズ化方式とその応用 (雑誌論文、学会発表、)

基本方式

道路網距離の算出方式としては、前節で扱った隣接リストのみを参照しつつ Dijkstra 法や A* アルゴリズムにより探索する方法の他、道路網上の距離を予め算出しておくマテリアライズ化 (以下 MPV: materialized path view) 方式が各種提案されている。最も単純な MPV 方式は対象とする道路網全体に対して、あらゆるノード間の距離を算出しておくものである。この方式では、2 点間の道路網距離を求めるためには距離表を一度参照すればよいことから高速に求められるが、データ量が膨大になる欠点がある。そこで、道路網を多数の小単位に分割し、更に階層的に上位層での距離表を置く方式、即ち階層的 MPV 方式 (HMPV) 方式が提案された。しかし、この方式によってもデータ量が膨大になるという問題は解消しなかった。

本研究では、HMPV 方式と同様に道路網を多数のサブグラフに分割し、サブグラフ間にまたがる 2 点間の距離の算出には最適優先探索を用いることにより、距離表のデータ量が少なく、かつ HMPV 方式に比して同程度の処理時間を達成し得る方式を提案した。表 1 は HEPV と提案方式である SMPV とのデータ量の比較を示している。

表 1 . SMPV と HEPV とのデータ量比較

道路網	ノード数	SMPV (MB)	HEPV (MB)
MapA	16,284	6.7	30.1
MapB	109,373	11.3	367.1
MapC	465,245	65.8	8,287.6

この表に見られるように、SMPV 方式は HEPV 方式に比して大幅なデータ量削減を達成し

ている。一方、2点間の最短経路距離を求める演算では2点間の距離が50km以下の場合、両者の処理時間の差はほとんど認められなかった。一方、2点間の距離が大きい場合にはSMPV方式の処理時間が増大することが予想されるが、本研究で対象としている応用は位置情報サービス(LBS)であるため、多くの検索が50km程度の範囲内で行われると予想される。従って、SMPV方式は道路網距離に基づくLBSでの検索に適した方式であると考えられる。

SMPVの応用

SMPV方式を用いることにより、道路網上の2点間の距離が高速に得られる。そこで、SMPV方式を道路網距離でのRkNN(逆近接)検索に応用した。RkNN検索は、データ点集合Pとある検索点q($q \in P$)が与えられたとき、qをkNNに含むP中の要素を求める検索である。P中の要素数をNとしたとき、RkNN検索は $O(N^2)$ の処理量を必要とする検索であり、特に道路網距離でのRkNNを実行するためには多くの処理時間を必要とする。本課題の目的は、この処理時間を大幅に短縮することである。

道路網上でRkNN検索の基本は以下の通りである。まず、検索点qからDijkstra法と同様に、道路網上で探索範囲を拡大していく。道路網上の各ノードnにおいて、kNN検索を行い、検索点qがkNNの結果に含まれる時、そのノードを経由し隣接するノードに探索範囲を拡大する。一方、qがkNNの結果に含まれない時、ノードnを経由する領域の探索を枝刈りする。この方式は道路網上でのRkNN検索で通常用いられるものであるが、データ点の存在密度が低い場合に膨大な数のノードに対して上記の処理が繰り返されることになる。

本方式では、SMPV構造の境界ノードにおいてのみ上記の処理を行うことと、IERの枠組みでのkNN検索を実行することにより処理時間の大幅な短縮を達成している。SMPVを適用することにより、各部分グラフの境界ノードでのみkNN探索を行えばよいことから、探索回数が大幅に減少する。また、IERの枠組みを用い、かつ道路網距離の検証においてもSMPV構造を用いることができることから、

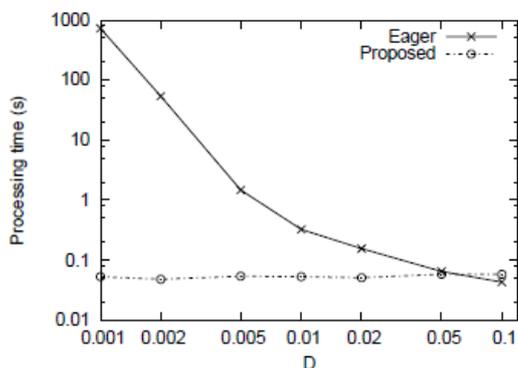


図1 SMPVを用いたRkNN検索時間

kNN探索に要する時間も減少する。

図1は、SMPVを用いたRkNN検索の処理時間を示している。なお、ここで用いた道路網は、表1のMapAである。比較対象は、道路網上でRkNN検索の為にYiuらにより提案されたEagerアルゴリズムである。横軸はデータ集合の存在密度を示している。この図に見られるように、Eagerアルゴリズムはデータの存在密度が低い時、長い処理時間を要している。これは道路網上で広い範囲の探索が必要なためである。一方、提案方式ではデータの存在密度に依存せずほぼ一定の処理時間を示している。データの存在密度が極端に高い時($D=0.1$)には、Eagerアルゴリズムの方が処理時間が若干短くなっている。これはRkNN探索を行う範囲が非常に小さくなるためであり、その場合には従来方式でも高速検索が可能なが分かる。ここで示した結果は他の道路網グラフを用いた場合にもほぼ同量の結果が得られることを確認している。

(3) 旅行計画(雑誌論文、学会発表、)

旅行計画とは、現在地(s)と最終目的地(e)及び旅行の途中で経由するm個の施設のカテゴリ($C_i, 0 \leq i \leq m$)が与えられたとき、sからeに到達する途中で C_i の中の要素を1つずつ訪れる際の最短経路を求める検索である。従来、途中で経由するカテゴリの訪問順序に制約が与えられる場合と、訪問順序が自由な場合など、多くの研究が行われて来た。しかし、それらの多くの研究はユークリッド距離に基づくのもであり、道路網距離を対象とした研究は少ない。また、道路網上での演算方式では多大な処理時間を要していた。

本研究では、この旅行計画を道路網距離で解く高速アルゴリズムを2種類提案した。1つはINEに基づくものであり、他の1つはIERに基づくものである。

INEに基づく検索では、現在までに求まっているデータ点のカテゴリまでの道路網距離と、そのカテゴリのデータ点からeまでのユークリッド距離の和が最小のものから順次探索するものである。この方式では、データ点の存在密度が高い場合に高速な探索が行える。一方、データ点の存在密度が低い場合には、道路網上で広範囲を探索する必要があることから、探索時間が長くなる欠点を有している。

一方、IERに基づく方式では、ユークリッド距離での探索により旅行計画路の候補を得て、それを道路網距離で検証することにより最短旅行計画路を得るものである。この方式により、データの存在密度が低い場合には、INE方式に比して高速な探索を行える。逆に、データ点の存在密度が高い場合には、ユークリッド距離での探索時に経路長の似た経路が多数生成され、それらを道路網距離で検証する必要があることからデータ処理時間が

増大する傾向にある。

この研究の発展として、次の2つが今後の研究テーマとしてあげられる。1つは、INEとIERを組み合わせることによる最適化である。上述のように、それぞれの戦略には相反する有利な状況と不利な状況がある。そこで、データ点の存在状況に応じてこれらの方式を切り替えて適用することにより、処理時間の面から最適な探索が可能である。他の1つは、連続旅行計画への適用である。ある現在点から最適な旅行計画路を得た場合に、その経路路に従って旅行を行うものと考えられるが、工事中や渋滞、または運転者の好みにより提案された旅行計画路から外れる場合がある。その時に、最初に旅行計画路を探索した時のデータ(PQなどの内容)を用いて、2回目以降の探索を高速化することができる。このような課題への対応は今後の課題である。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計3件)

Aye Thida Hlaing, Tin Nilar Win, Htoo Htoo, Yutaka Ohsawa, RkNN Query on Road Network Distances, Journal of Information Processing, Vol .23, No. 2, 査読有、2015、pp . 163-170 (DOI:10.2197/ipsjip.23.163)

Htoo Htoo, Yutaka Ohsawa, Noboru Sonehara, Masao Sakauchi, Incremental Single- source Multi-target A* Algorithm for LBS Based on Road Network Distance, IEICE Trans, Inf. & Syst, Vol .E96-D, No. 5, 査読有、2013、pp. 1043-1052 (DOI:10.1587/transinf.E96.D.1043)

大沢 裕、トウ トウ、曾根原登、坂内雅夫、道路網上での旅行計画の為にインクリメンタル検索方式、日本データベース学会論文誌、Vol . 11, No . 2, 査読有、2012、pp . 1-6

http://dbsj.org/journal/dbsj_journal/dbsj_journal_vol11_no2/

[学会発表](計8件)

Aye Thida Hlaing, Htoo Htoo, Yutaka Ohsawa, Efficient Reverse kNN Query Algorithm on Road Network Distances Using Partitioned Subgraphs, SeCoGIS2014, LNCS8623, 査読有、pp . 212-217, Atlanta (USA)2014年10月27日-29日

Yutaka Ohsawa, Aye Thida Hlaing, Light Materialized Path View for Location Based Services, ICCA2014, 査読有、pp . 210-216, Yangon (Myanmar) 2014年2月17, 18日

Yutaka Ohsawa, Aye Su Yee Win, Trip Route Planning Qoute Query in Road Network Distance, ICEEHE2013, 査読有、TL-2-3, Yangon(Myanmar) 2013年12月21日, 22日

Aye Thida Hlaing, Htoo Htoo, Yutaka Ohsawa, Noboru Sonehara, Masao Sakauchi, Shortest Path Finder with Light Materialized Path View for Location Based Services, WAM2013, LNCS7923, 査読有、pp . 229-234, Beidaihe (China)2013年6月14日-16日

Yutaka Ohsawa, Htoo Htoo, Noboru Sonehara, Masao Sakauchi, Sequenced Route Query in Road Network Distance Based on Incremental Euclidean Restriction, DEXA2012, LNCS7446, 査読有、pp .484-491, Viena (Austria)2012年9月3日-6日

Htoo Htoo, Yutaka Ohsawa, Noboru Sonehara, Masao Sakauchi, Aggregate Nearest Neighbor Search Methods Using SSMTA* Algorithm on Road-Network, ADBIS2012, LNCS7503, 査読有、pp . 181-194, Poznan(Poland) 2012年9月18日-20日

Htoo Htoo, Yutaka Ohsawa, Noboru Sonehara, Optimal Sequenced Route Query Algorithm Using Visited POI Graph, WAIM2012, LNCS7418, 査読有、2012、pp . 198-209, Harbin(china)2012年8月18日-20日

[図書](計0件)

[産業財産権]

出願状況(計0件)

取得状況(計0件)

[その他]

ホームページ等

<http://www.mm.ics.saitama-u.ac.jp>

6. 研究組織

(1)研究代表者

大沢 裕(OHSAWA, Yutaka)

埼玉大学・大学院理工学研究科・教授

研究者番号：50152111

(2)研究分担者

なし

(3)連携研究者

なし