

氏 名	Aye Thida Hlaing
博士の専攻分野の名称	博士（工学）
学位記号番号	博理工甲第981号
学位授与年月日	平成27年3月24日
学位授与の条件	学位規則第4条第1項該当
学位論文題目	Road Network Distance Based Efficient Algorithms Suitable in Location Based Services (位置情報サービスに適した道路網距離に基づく高効率アルゴリズム)
論文審査委員	委員長 教授 大澤 裕 委員 教授 島村 徹也 委員 准教授 小林 貴訓 委員 准教授 山田 敏規

論文の内容の要旨

This thesis describes the studying of efficient algorithms for real-time monitoring, shortest-path finding and Reverse k nearest neighbor (R- k NN) query applying on road network distances for Location Based Services (LBS). In recent time, the role of location-based services is rising as the increasing numbers of users are requesting the location-based information. The main idea of LBS is to provide the service that depends on the positional information which is associated with the user, most importantly, the user's current location. The service may also be dependent on other information, such as personal preferences and interests of the user. For example, finding the cheapest hotel within 20 km, where the nearest gas station is how long it will take to go to Italian restaurant, are some specific user's queries in location-based services. Also a service may inform its users about traffic jams, weather situation, the position of emergency vehicles, hazardous materials or public transportation. Moreover, in other related services like delivery management, trip planning, bus routes queries for Urban planning, checking movement of terrorists in emergency situations for crime prevention, etc., require LBS services based on spatial information. For all these requirements, studying on efficient algorithms for LBS services has been demanded. In this thesis, some background knowledge and related works are described in chapter (1) and chapter (2).

Various route queries in road networks have gained significant research interests due to advances in GIS and mobile computing such as car navigation system. The main challenge of processing such a query is how to efficiently monitor the moving object and how to retrieve the route rapidly. In recent times, much work has been conducted on a real time monitoring of moving objects (cars and humans). Some studies adopted to get high accuracy tracking of moving objects with less communication between moving objects and server. However, these studies assumed the moving object is moving in Euclidean space or in fixed routes. This thesis proposed a real time monitoring method aiming for both thick client and thin client. The proposed algorithms are described in chapter (3). The main differences between thick client and thin client are their own computational ability and data storage capacity. In thick client (e.g. car navigation

system), it still requires at least periodic connection to a server but it can perform many functions, such as predict and compute the route without that connection. In thin client (e.g. mobile phone user), it depends on the server to fulfill its requirements, however, it is convenient to use. Using the “frequently used routes” (FUR) information, we offer high scalability monitoring system with empirical comparisons of the proposed method and the conventional dead-reckoning on road network (DRRN) method for both thick client and thin client. Because conventional real-time monitoring does not use knowledge about routes, scalability remains low and the accuracy of tracking is also low. On the other hand, using FURs extracted from historical trajectories decreases the communication cost and achieves highly accurate monitoring. According to the experiment results, our proposed method outperforms the conventional method DRRN by factors from 2 to 4 in thick client and 4 to 6 in thin client.

Despite the importance of spatial networks in real applications, most of the spatial query methods focus on Euclidean distance, where the distance between two objects is determined solely by their relative position in space. However, in practice, objects can usually move only on a pre-defined set of routes as specified by the real road network (road, railway, river etc.). Thus, the important measure is the road network distance i.e., the length of the shortest trajectory connecting two objects rather than their Euclidean distance. Only using Euclidean distance on spatial network databases (SNDB) is scarce and too restrictive for emerging applications such as mobile computing and location-based queries. The main difference between using Euclidean distance and road-network distance is based on their calculation costs. As considering, the Euclidean distance between two arbitrary points can be computed easily, however, in the road network distance, it takes longer processing time. For example, if the river or mountain lies between two points, there is totally difference between the Euclidean distance and the road-network distance, and the costs of the distance calculation as well. In the conventional methods, to promptly acquire the distance between two terminal points for spatial queries in LBS applications, we can use well-known shortest path finding algorithms, Dijkstra's algorithm and A* algorithm. However, due to the complexity of some queries, for example, RNN (reverse nearest neighbor) queries, skyline queries, k-NN queries and several kind of TPR (trip planning route) queries, these two algorithms take very long processing time and will cause high calculation cost. These algorithms refer an adjacency list to find the neighboring nodes to a currently noticed node. When two specified points (s and d) are located on a long distance, they need much repetitive processing (*node-expansions*).

Several methods based on materialized path view (MPV) have also been proposed for the fast road network distance computation. They retrieve the distance by looking up a pre-computed distance table. However, this MPV has the following problems: (1) Usually, a road network contains a large amount of nodes, and the data size of the MPV is proportional to the square of the number of nodes. Therefore, the data amount of the distance table becomes huge for a large size of the road network. (2) Very long processing time is necessary to construct MPV table, because the distance must be calculated over all combinations of node pairs. Therefore, the processing time increases rapidly in accordance with the length of the shortest path. For all these reasons, we proposed efficient shortest path finding algorithm presented in chapter (4), based on simple materialized-path-view (SMPV) structure constructed only on partitioned subgraphs to compute this road network distance which is very useful for LBS application. Due to the sizes of the real road network maps, apply this method to three variations SPFLM, SPFMM and SPFFM. A road network is partitioned into subgraphs, and the distance materialization is performed only in the subgraphs. Therefore, the amount of pre-computed data is greatly reduced. According to the experimental results, our proposed method can reduce the used data

amount comparing to the existing hierarchical network distance materialized methods; HEPV and HiTi, and it also can also reduce the overall processing time drastically in all maps.

Using partitioned subgraphs, we studied an efficient algorithm for Reverse k -Nearest Neighbors (R- k NN) query in road network distance, and it was described in chapter (5). A Reverse k -Nearest-Neighbor (R k NN) query finds the data points that take the query point as one of their k nearest neighbors. This type of query is required in a wide variety of applications, including decision support, facility management, taxi allocation, location-based services, advertisement distribution, and games. For example, in a battle field, a backup army unit may issue a RNN query to monitor other units for which is the closest army unit to seek help from the backup army unit. According to our knowledge, the existing methods for R- k NN queries required to find k NN search on every visited node. This causes a large number of node expansions and the processing time increases simultaneously. Therefore, we proposed a fast R- k NN search algorithm based on a simple materialized path view (SMPV) structure and we adopted incremental Euclidean restriction strategy for k NN queries which is the main function in R- k NN queries. Referring to the experimental results, the proposed method outperforms the existing method (Eager algorithm) for finding R k NNs on real road network especially when the data points are sparsely distributed or k value is large. In addition, the proposed method shows stable characteristic and low processing time even the Eager algorithm increases the processing time in low POI density.

The conclusions of our thesis and the future research are contributed in the last chapter (6).

論文の審査結果の要旨

当学位論文審査委員会は当該論文の発表会を平成 27 年 1 月 28 日に公開で開催した。発表後質疑応答に基づく詳細な審査を行った。以下に審査結果を要約する。

本博士論文では、位置情報サービス (LBS: location based services) において必要となる、道路網距離の高速算出可能なデータ構造と検索方式、そのデータ構造の逆近接検索 (RkNN: reverse k nearest neighbor) への応用、移動体の効率的な位置把握方式についての提案をまとめたものである。

まず、道路網距離の基本的な算出方式には Dijkstra 法があり、更に効率化方式として A* アルゴリズムが用いられてきた。しかし、これらの方式は程度の差はあるが、距離を求めようとする 2 点間の距離が大きくなるにつれ、長い処理時間を必要とする。一方、LBS で必要となる各種検索においては、2 点間の道路網距離を多数回求めることが必要となる。そこで、更なる高速化方式として、2 点間の距離を予め算出しておく距離のマテリアライズ化方式 (MPV: materialized path view) が提案された。しかし、この MPV では膨大なデータ量の距離表を保持しなければならないという問題があった。本論文では、道路網グラフを多数のサブグラフに分割し、そのサブグラフ内での距離表を作成すること、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、大幅なデータ量の削減と、MPV に匹敵する処理スピードを達成する方式 (SMPV: simple materialized path view) を提案している。

次に、この SMPV 構造を用いて、逆近接検索問題 (RkNN) を道路網距離で高速に求める方式を提案している。RkNN 検索とは、データオブジェクトの集合を P とするとき、検索点を kNN に含む全てのデータ点 ($\in P$) を求める検索である。この演算は、データ点数を N としたとき、基本的には $O(N^2)$ の処理量を必要とする。そこで、これを道路網距離で求めようとするとき、長い処理時間を必要とする。そこで、本論文では SMPV 構造上で、RkNN を求める方式を提案している。SMPV の境界ノードのみでの検証処理を実行することにより、大幅な処理時間の短縮を実現している。また、提案方式ではデータオブジェクトの分布密度によらず一定の処理時間を達成可能であるという良い特性も示している。

最後に、移動体の効率的な位置モニタリング方式を提案している。多数の人や車などの移動体位置をサーバが捕捉しようとするとき、(1) 移動体が一定時間間隔ごとにサーバに位置を報告する、または (2) 移動体が一定距離移動するごとにサーバに位置を報告する、の内いずれかの方式がとられてきた。しかし、1 台のサーバで多数の移動体の位置を捕捉しようとするとき、可能な限り通信量や通信回数を減らすことが求められる。そこで、サーバと移動体が同じアルゴリズムで将来の移動位置を予測し、その予測が外れたときに移動体が現在位置を報告する方式がとられている。従来方式は、自律航法 (dead-reckoning) と呼ばれるもので、移動体と同じ速度で道路に沿って移動するという仮定による位置予測が行われてきた。しかし、この方式では予測が外れることが多く、より性能のよい予測方式が求められてきた。本論文では、個々の移動体の過去の経路を記録保持し、その履歴から将来の移動位置を予測する方式を提案している。また、これを実現するためのデータ構造を提案している。実際の移動データを用いた実験により、提案方式は従来方式に比して数分の一の通信量で位置の捕捉が可能であることが示されている。

本博士論文は、6 章から構成されている。

第 1 章は序論であり、LBS の現状と問題点が述べられ、本論文の貢献点、3～5 章で行われる実験で用いられる道路地図データについて説明されている。

第 2 章では、本論文に関連した従来研究についてまとめられている。道路網上で与えられた 2 点間の最短経路探索方式として、Dijkstra 法、A* アルゴリズム、事前に計算された距離表を用いる方式について、利点

と欠点が示されている。次に、移動体の位置捕捉方式の従来研究がまとめられている。更に、LBS で必要となる各種検索方式に関する従来研究が示され、最後にユークリッド距離での検索において広く用いられている空間索引構造 (R木) が述べられている。

第3章では、新しい移動体の位置捕捉方式が提案されている。この方式は、前述のように個々の移動体の過去の移動履歴から「よく通るルート: FUR (frequent used route)」を作成し、それをサーバとクライアントが共有することにより適切な近未来位置を予測する方式が述べられている。提案方式では、移動体側が道路地図を保有するモデル (リッチクライアント) と道路地図を保有せず必要に応じてサーバから得るモデル (シンクライアント) に分け、それぞれのモデルにおいて有効な方式が提案されている。

第4章では、道路網距離の簡易マテリアライズ化方式 (SMPV) について述べられている。SMPV での道路網を多数のサブグラフへ分割する方式、サブグラフ内で作成される各種距離表、最適優先探索による最短経路算出方式が説明されている。サブグラフ内の各ノード間の距離表として、3種類の距離表をあげ、それらの一部または全部を用いた最短経路探索アルゴリズムが示されている。SMPV 方式は、従来の HMPV (階層的 MPV) 方式と同等の速度で最短経路の検索ができ、かつデータ量は地図の規模によるが数桁少なく抑えられている。また、SMPV を kNN 検索に適用した場合に、A* アルゴリズムに比して 100 倍程度高速な検索を行える結果が示されている。

第5章は SMPV データ構造上での高速 RkNN 検索方式が提案されている。道路網上での RkNN 検索アルゴリズムとしては、Yiu らにより Eager アルゴリズムが提案されている。また、その後提案されたアルゴリズムのほとんどは、Eager アルゴリズムの考え方を基礎としている。本章では、Eager アルゴリズムを基本方式として、それを SMPV 上で実行するための工夫について述べられている。Eager アルゴリズムでは、データオブジェクトの密度が高い場合には高速実行が可能であるが、道路網上での存在密度が低くなるにつれ急速に実行時間が長くなる傾向が見られる。一方、提案方式ではデータオブジェクトの存在密度には依存せずほぼ一定の実行時間で検索可能である。この性質は、LBS への応用を考えたとき、常に一定時間で検索できることを示しており、良好な性質である。

第6章は結論の章であり、本論文の貢献点をまとめ、今後の課題が述べられている。

以上の研究成果は4編の国際会議プロシーディングス論文として公表されている。また、そのうち2編が第1著者である。

以上の結果を総合し、同審査委員会は本学位論文が博士(工学)の学位授与に値する十分な研究内容を持つものであると判断した。