EFFICIENT CLUE-BASED ROUTE SEARCH ON ROAD NETWORKS

[1]Gilla saiteja, [2]Dr. R. Jegadeesan [3]Bommella supraja, [4]Kandukuri sushma, [5]Kuragayala sushma, [6]Devaraneni sushrutha,[7]P.pranitha
[1, 3,4,5,6] UG Student
[1]Department of Computer Science and Engineering
[1]Jyothishmathi Institute of Technology and Science, Karimnagar, India
[2,6]Associate Professor-Department of Computer Science and Engineering
[1,2,3,4,5,6]Jyothishmathi Institute of Technology and Science, Karimnagar, India

Abstract – The advances in positioning techniques and location-based services, it is currently very common for roads networks to get textual content on the tops. The previous work identifies the optimal path that covers a series of query keywords have been studied in recent years. However, in many practical scenarios, the optimal path may not always be desirable. For example, a complete routing query is issued by providing some evidence that describes the spatial context between PoIs along road, where the result can be far from optimal. Therefore, in this paper, we investigate the search problem by an idea-based list (CRS), which allows the user to provide evidence of keywords and spatial relationships. First, propose a greedy algorithm and a dynamic programming algorithm as standards. To improve efficiency, developed a branch-and-bound algorithm that eliminates unnecessary vertices in query processing. In order to quickly locate nominee, deployed an AB-tree that stores both the distance and keyword information in tree structure. To further reduce the index size, construct a PB-tree by utilizing the virtue of 2-hop label index to pin point the nominee. Broad experiments are conducted and verify the superiority of our algorithms and index structures.

Index Terms - Spatial keyword queries, clue, Point-of-Interest, travel route search, query processing.

1.INTRODUCTION

In this study a tendency to propose a completely unique question sort for spacial databases in support of travel-planning GIS applications.The target is to help users within the designing of journeys that involve many destinations, probably happiness to totally different dish classes.Supported variety of traveling rules (or constraints) that are expressed as sub-sequences of locations, users aim to find the route with the comparatively shortest traveling distance. Note that it's attainable that the traveling rules might solely involve a set of the user-requested locations .

The proliferation of on-line objects with each AN associated geo-location and a text description,the net is feat a spatial dimension.Specifically,internet users and content are progressively being geo-positioned and geo-coded. At an equivalent time, matter descriptions of points of interest, e.g., cafes and traveller attractions, are progressively turning into obtainable on the net.This development incorporate techniques that modify the assortment of information that contains both text descriptions and geo-location in order to support the efficient processing of spatial keyword queries.

To support direction-aware abstraction keyword queries, it has a tendency to devise novel direction-aware index structures to prune superfluous directions. It has a tendency to first cluster the POIs supported their distances to the bottom-left purpose of the Minimum Bounding parallelogram (MBR) that contains all POIs.Then for POIs in every cluster,it has a tendency to type them supported their directions to the bottom-left purpose. Given a question, we are able to deduce a direction vary with a lower direction sure associated an higher direction sure. It able to prove that for any dish if its direction to the bottom-left purpose isn't within the direction vary of the question, it'll not be a solution, and that will prune the dish.

Querying and manipulating giant scale graph-like information have attracted a lot of attention within the information community, due to the wide application areas of graph information, such asranked keyword search, XML databases, bioinformatics, social network, and ontologies.The shortest path question answering in a very graph is among the elemental operations on the graph information.In a hierarchical keyword search state of affairs over structured knowledge, people typically provide scores by activity the link distance between two connected components. If over one path exists, it's fascinating to retrieve the shortest distance between them, as a result of shorter distance usually means that seniority of the connect.

2. RELATED WORK

Optimal Route Query Processing Early work on optimal route computation focuses on greedy solutions. Chen et al. [4] use the same query definition as this paper, and propose two heuristics. The first, namely NNPSR, resembles the greedy approach described in Section 1; the second retrieves the nearest point of the query start position q in every category, and then connects them to form a route. In addition, Chen et al. [4] also describe a simple combination of NNPSR and RLORD [13], which answers a special case of the optimal route query with a total order of the categories to be visited. The hybrid solution first runs NNPSR to find a greedy route; then, it extracts the category of each point on the greedy route, and runs R-LORD with this category sequence as input. None of the solutions in [4] guarantees the quality of the results; these methods usually return suboptimal routes according to the experiments in [4]. Li et al. [10] study a variant of the optimal route query that specifies both a start point q start and an end position q end, but no order constraint between the data categories.

This is equivalent to a visit order graph GQ that contains two artificial categories Cstart ¼f qstartg and Cend ¼f qendg, and two edgeshCstart;CiandhC;Cendifor each category C in the data set. The solutions of [10] report approximate query results; on the other hand, this paper focuses on efficient, exact methods for the general optimal route problem. Sharifzadeh et al. [13] propose R-LORD, the first exact solution for optimal route queries with a total order. In the example of Fig. 1, suppose that GQ specifies total order q ! museum ! restaurant ! pub; then, R-LORD is directly applicable. Specifically, let be the optimal route; an important observation made in [13] is that any suffix r of risk also the shortest among all routes that1) start at the first point of r, and 2) visit the same categories as r, in the same order.

Problem Statement:

Here models a road network as a weighted purpose less graph G = (V, E), wherever V is that the set of vertices and E is that the set of edges. Each vertex v ∈ V contains a set of keywords, denoted as Φ(v). every edge (u, v)∈E has a positive weight, i.e., length or travel time on the string, denoted as e(u, v). Given a path between vertices u and v, denoted as P(u, . . . , v), the length is the total of weights of all edges on the trail. For any 2 vertices u and v, the network distance between u and v on G, denoted as $d_G(u, v)$, is that the length of the shortest path SP(u,v) between u and v.

3. METHODOLOGY

This section deals with clue based dynamic programming rule which gives the time quality of the brute-force approach, that tries all the potential combinations used in CDP.

3.1. Clue-Based Dynamic Programming Rule

As known,even if GCS contains a short latent period, the accuracy of the solution cannot be bonded to attain higher accuracy,it has a tendency to propose a particular rule,known as Clue-based Dynamic Programming (CDP),to answer the CRS question.Generally,it is difficult to develop Associate in Nursing economical actual rule for CRS queries,since it has a tendency to cannot avoid complete rummage around for PoIs in road networks.Parenthetically,the number of vertices that contain keyword wi∈C is denoted as |Vwi|,[2] therefore the time quality of the brute-force approach,that tries all potential combos,is O(Qwi∈C|Vwi|).

In CDP, it tend to construct a keyword posting list for every keyword w, that could be a list of vertices that contain w. Once a CRS question is issued, we tend to type the posting lists in keeping with the keyword order of wi∈C. Note that the order of the vertices inside every posting list does not matter and might be impulsive, thus area unit sorted by vertex id for simplicity. It's straightforward to examine that these posting lists truly construct a k-bipartite[5] graph G0,that of course shows all possible paths for a given C. The burden of every go up G0 is computed as the matching distance. Specifically, for each u∈Vwi , define D(wi, u) to denote the minimum matching distance one can achieve American state the a walk that passes the keywords from w1 to w_i consistent with the order in C and stops at u.Then we tend to reason D(wi, u) by the subsequent algorithmic formula:

(i) i = 1: for match vertices u∈Vw1, we have D(wi,u) = dm(μi(wi , di), σ(vq → u)).
(ii) i> 1: for match vertices v∈Vwi−1 and u∈Vwi , we have D(wi, u) = min v∈Vwi−1}

For each iteration, we've got at the most |Vwi−1|·|Vwi | combos,thus the time needed is O($P_{k\,i=2}$|Vwi−1| · |Vwi|).The details of CDP are shown in algorithmic program two so as to reason D(wi , u),it has to access the posting list of wi−1. For each vertex v during this list, we first check if σ(v → u) is a match w.r.t.μi and so reason dm(μi , σ(v → u)).Then it has a tendency to compare it with D(wi−1, v), and keep the larger one as intermediate value. Finally, we discover the minimum one as D(wi, u) from these |Vwi−1| intermediate values.

3.2. Greedy Clue Search Algorithm Program

Develop a greedy account program as a the CRS primary query response, that is his name is the Search for Greed algorithm (GCS) a program. Give a question message = (vq, C), it has tilt to add the initial vq in the filter path. Develop a greedy account program as a the CRS primary query response, that is his name is the Search for Greed algorithm (GCS) a program[6]. Give a question message = (vq, C), it has Tilt to add the initial vq in the filter path.

Then it has t Force to use procedure FindNextMin () to work on the next game That the matching distance between μ1 and σ1 (vq → v1), for example, dm (, 1, σ1) is reduced. After that, we are Insert the v1 into the filter path, and still notice Confectionery date by findNextMin (). This is A continuous way until each area tops the game Specific unit[3], so the filter path forms Path is possible, denotes FPvq. If it has the tendency to assume FindNextMin () prices time and, then comp time the lexness of GCS is O (K & F).

In the findNextMin () procedure,we have a tendency to take advantage of the network growth algorithm [15] to search From the nearby peaks that contain the question keywords also have a space area network[9] Inside confidence between omen.The algorithm Details of the program unit area shown in the algorithm One program.Given the source u, as well as the idea µ (w, d,), we have a tendency to search for target Matching the Cape against this distinction between DG (u, C) The d is down in the north turk pass starting from u, we have a tendency to check each other He visited the top of the head to imagine if it was the summit of that summit Contains w and is located in interval [d (1 -), d 1 +)].If v 0 is found, Then

(1) If dG (u, v0) ܲ d, we te Second to update V from V since V 0 makes a distinction less than v;
(2) Otherwise, compare d - dG (u, v) with dG (u, v0) -D the smallest one comes because of the result.

3.3. Optimal Route Queries with Arbitrary Order Constraints

Given a set of DS spatial points, each of them associated with categorical information, for example, restaurant, pub, etc R optimization the route query finds the shortest path of it starts from the query point (for example, home or hotel),and Covers the user - Specific set of Categories (for example, {pub, Restaurant, museum})[1].The user can also select partial demand constraints between different categories, For example, a You must visit the restaurant before a bar. Previous work Focus on a special case where the query contains the total order for everyone categories that must be visited (for example, museum! Resturant! pub). For the general scenario without such Overall ranking, and Only known the solution reduces the problem for Multiple, Total - Request the best route queries.

As shown in this paper, an approach that nave incve bears a large amount Recurring calculations, therefore, are not scalable to large Data Sets. Motivated by this, w E. Novel proposal Solutions for Query the overall path optimization, based on two different methodologies, which are forward and forward Search[11]. In addition, we discuss how these methods can be adapted to answer the variable of the optimal way as shown in this paper, an approach that Nave incve bears large amount Recurring calculations, therefore, are not scalable to large data Sets. Motivated by this, w E.

```
Algorithm 1: Preprocessing by the pruned highway labeling procedure Preprocess(G)
L′ 0(v) ←∅ for all v ∈ V
P ← a highway decomposition of G
N ← the size of P for i = 1 to N do L′
i ← PrunedDijkstraSearch(G,Pi,L′ i−1)
end for
return L′N
end procedure
```

4. EXPERIMENTAL SETTINGS

It tend to use 3 datasets.Table half-dozen shows some properties of those datasets. Dataset GN is extracted from the U.S Board on Geographic Names (geonames.usgs.gov), during which every object could be a location with a geographic name (e.g., valley).Dataset net is generated from 2 real datasets[12]. One is WEBSPAMUK20071 that consists of an oversized range of net documents;the opposite could be a spacial information set containing the tiger Census blocks in Iow a, Kansas, Missouri, and Cornhusker State .It tend to indiscriminately mix net documents and spacial objects to urge the net dataset. Dataset edifice contains spacial objects that represent some hotels within the U.S.

Every object contains a location and a group of words that describe the edifice (e.g., restaurant, pool). Edifice is little and is employed to judge the performance of our algorithms once the info set and index square measure memory resident, and therefore the alternative 2 massive datasets square measure accustomed valuate our algorithms once the dataset and index square measure disk-based. It tend to generate five question sets within the house of GN, during which the amount of keywords is three, 6, 9, 12, and 15, respectively.It tend to additionally generate five similar question sets within the house of each net and edifice. Each set comprises 50queries.It tend to rank words in falling order of their frequencies in every dataset.

Example: Consider a query q with keywords q.ψ ={t1,t2,t3} and the four objects in Table.It knows that Dist(q,o1) <Dist(q, o2) and o1 ∩o2 = {t2}. According to lemma 3.3, {o1} is a better result set for the query with keyword set {t2}.

Table 1: Sample dataset.

	O1	O2	O3	O4
Distance to the query	1	2	2.5	4
Keywords	t1,t2	t2,t3	t1,t3	t1

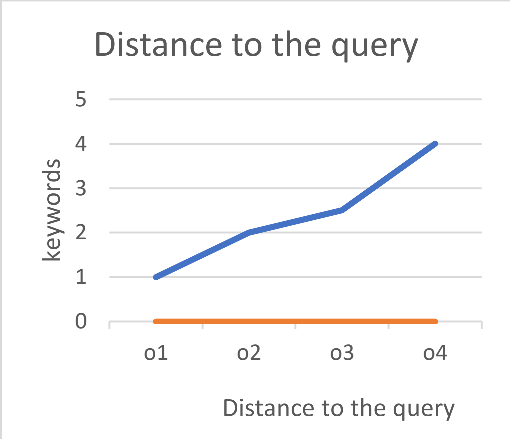

Fig. 1. Distance to the query

5. CONCLUSION AND FUTURE DIRECTIONS:

In this paper, the problem of CRS on road networks was studied, which aims to find an optimal route such that it covers a set of query keywords in a given specific order, and the matching distance is minimized. To answer the CRS query, firstly a greedy clue-based algorithm GCS is used with no index where the network expansion approach is adopted to greedily select the current best candidates to construct feasible paths. Then, it devised an exact algorithm, namely clue-based dynamic programming CDP, to answer the query that enumerates all feasible paths and finally returns the optimal result. To further reduce the computational over head, a branch-and-bound algorithm BAB is applied by specifying filter-and-refine paradigm such that only a small portion of vertices are visited, thus improves the search efficiency.

REFERENCES

[1] K. Tretyakov, A. Armas-Cervantes,L.Garc´ıa-Ba˜nuelos, J. Vilo, and M. Dumas. Fast fully dynamic landmark-based estimation of shortest path distances in very large graphs. In CIKM, pages 1785–1794, 2011.
[2] T. Akiba, C. Sommer, and K. Kawarabayashi.Shortest-path queries for complex networks: exploiting low tree-width outside the core. In EDBT, pages 144–155, 2012

[3] F. Wei. Tedi: efficient shortest path query answering on graphs. In SIGMOD, pages 99–110, 2010.
[4]SIAM J. Comput., 32, 2003.M. Qiao, H. Cheng, L. Chang, and J. X. Yu.Approximate shortest distance computing: A query-dependent local landmark scheme. In ICDE,pages 462–473, 2012.
[5 L]E. Cohen, E. Halperin, H. Kaplan, and U. Zwick. Reachability and Distance Queries via 2-hop abels. SIAM J. Comput., 32, 2003.

[6]R.Geisberger, P. Sanders, D. Schultes, and C. Vetter. Exact Routing in Large Road Networks Using Contraction Hierarchies. Transportation Science, 2012.

[7] R. Geisberger, P. Sanders, D. Schultes, and D. Delling. Contraction hierarchies: Faster and simpler hierarchical routing in road networks.In WEA, pages 319–333, 2008.

[8] T. Akiba, Y. Iwata, and Y. Yoshida. Fast exact shortest-path distance queries on large networks by pruned landmark labeling. In SIGMOD,2013.

[9] M. Jiang, A. W.-C. Fu, R. C.-W. Wong, and Y. Xu. Hop doubling label indexing for point-to-point distance querying on scale-free networks. 2014.

[10] M. Qiao, H. Cheng, J. Yu, and W. Tian. Top-k nearest keyword search on large graphs. In VLDB, 2013.

[11] Basu Roy, Senjuti, and KaushikChakrabarti. "Location-aware type ahead search on spatial databases: semantics and efficiency." In Proceedings of the 2011 ACM SIGMOD International Conference on Management of data, pp. 361-372. ACM, 2011.

[12] D. Zhang, Y. M. Chee, A. Mondal, A. K. H. Tung, and M. Kitsuregawa, “Keyword search in spatial databases: Towards searching by document,” in ICDE, 2009, pp. 688–699.

[13]Jegadeesan,R.,Sankar Ram M.Naveen Kumar JAN 2013 "Less Cost Any Routing With Energy Cost Optimization" International Journal of Advanced Research in Computer Networking,Wireless and Mobile Communications.Volume-No.1: Page no: Issue-No.1 Impact Factor = 1.5

[14]Jegadeesan,R.,Sankar Ram, R.Janakiraman September-October 2013 "A Recent Approach to Organise Structured Data in Mobile Environment" R.Jegadeesan et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (6) ,Page No. 848-852 ISSN: 0975-9646 Impact Factor:2.93

[15]. Jegadeesan,R., Sankar Ram October -2013 "ENROUTING TECHNICS USING DYNAMIC WIRELESS NETWORKS" International Journal of Asia Pacific Journal of Research Ph.D Research Scholar 1, Supervisor2, VOL -3 Page No: Print-ISSN-2320-5504 impact factor 0.433

[16]. Jegadeesan,R., Sankar Ram, M.S.Tharani (September-October, 2013) "Enhancing File Security by Integrating Steganography Technique in Linux Kernel" Global journal of Engineering,Design & Technology G.J. E.D.T., Vol. 2(5): Page No:9-14 ISSN: 2319 – 7293

[17]. Ramesh,R., Vinoth Kumar,R., and Jegadeesan,R., January 2014 "NTH THIRD PARTY AUDITING FOR DATA INTEGRITY IN CLOUD" Asia Pacific Journal of Research Vol: I Issue XIII, ISSN: 2320-5504, E-ISSN-2347-4793 Vol: I Issue XIII, Page No: Impact Factor:0.433

[18]. Vijayalakshmi, Balika J Chelliah and Jegadeesan,R., February-2014 "SUODY-Preserving Privacy in Sharing Data with Multi-Vendor for Dynamic Groups" Global journal of Engineering,Design & Technology. G.J. E.D.T.,Vol.3(1):43-47 (January-February, 2014) ISSN: 2319 –7293

[19]. Jegadeesan,R.,SankarRam,T.Karpagam March-2014 "Defending wireless network using Randomized Routing process" International Journal of Emerging Research in management and Technology

[20].Jegadeesan,R.,T.Karpagam, Dr.N.Sankar Ram , "Defending Wireless Network using Randomized Routing Process" International journal of Emerging Research in management and Technology ISSN: 2278-9359 (Volume-3, Issue-3) . March 2014

[21]. Jegadeesan,R., Sankar Ram "Defending Wireless Sensor Network using Randomized Routing "International Journal of Advanced Research in Computer Science and Software Engineering Volume 5, Issue 9, September 2015 ISSN: 2277 128X Page | 934-938

[22]. Jegadeesan,R., Sankar Ram,N. "Energy-Efficient Wireless Network Communication with Priority Packet Based QoS Scheduling", Asian Journal of Information Technology(AJIT) 15(8): 1396-1404,2016 ISSN: 1682-3915,Medwell Journal,2016 (Annexure-I updated Journal 2016)

[23]. Jegadeesan,R.,Sankar Ram,N. "Energy Consumption Power Aware Data Delivery in Wireless Network", Circuits and Systems, Scientific Research Publisher,2016 (Annexure-I updated Journal 2016)

[24]. Jegadeesan,R., Sankar Ram , and J.Abirmi "Implementing Online Driving License Renewal by Integration of Web Orchestration and Web Choreogrphy" International journal of Advanced Research trends in Engineering and Technology (IJARTET) ISSN:2394-3785 (Volume-5, Issue-1, January 2018

[25]. Pooja,S., Jegadeesan,R., Pavithra,S., and Mounikasri,A., "Identification of Fake Channel Characteristics using Auxiliary Receiver in Wireless Trnsmission" International journal for Scientific Research and Development (IJSRD) ISSN (Online):2321-0613 (Volume-6, Issue-1, Page No. 607-613, April 2018

[26]. Sangeetha,R., Jegadeesan,R., Ramya,P., and Vennila.,G "Health Monitoring System Using Internet of Things" International journal of Engineering Research and Advanced Technology (IJERAT) ISSN :2454-6135 (Volume-4, Issue-3, Page No. 607-613, March 2018.