ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

A Novel Trip Planner Using Effective Indexing Structure

M.K Chandrasekharan[1], S Shiva Shankar [2]

P.G. Scholar, Department of CSE, Maharaja Institute of Technology, Coimbatore, India[1].

Assistant Professor, Department of CSE, Maharaja Institute of Technology, Coimbatore, India[2].

ABSTRACT: The administration of transportation frameworks has ended up progressively imperative in numerous genuine applications such as area based administrations, production network administration, movement control, et cetera. These applications normally include questions over spatial street systems with powerfully changing and confused activity conditions. In this paper, we model such a system by a probabilistic time-dependent graph (PT-Graph), whose edges are connected with unverifiable postponement capacities. We propose a valuable inquiry in the PT- Graph, in particular an Trip planner query (TPQ), which recovers excursion arranges that cross a set of inquiry focuses in PT- Graph, having the base voyaging time with high certainty. To handle the proficiency issue, we display the pruning systems time interim pruning and probabilistic pruning to viably discount bogus alerts of trek arrangements. Besides, we outline a pre computation method in view of the expense model and develop a list structure over the pre computed information to empower the pruning by means of the file. We coordinate our proposed pruning techniques into a productive question system to answer TPQs. Through far reaching tests, we exhibit the proficiency and adequacy of our TPQ question noting methodology.

KEYWORDS: probabilistic time-dependent graph, trip planner query, indexing

I. INTRODUCTION

Transportation techniques have played an essential position in real functions such because the site visitors manipulate vicinity-founded services (LBS), travel planning, and geographical knowledge management. One ordinary illustration of such programs is the European site visitors Message Channel (TMC) , which has been operated in many European international locations, North the united states, and Australia. With the growing interest in the administration of transportation programs, recently, the spatial avenue community has bought so much attention from the database group,. In particular, a spatial road network can also be modeled by way of a huge graph in a 2-dimensional geographical area , whose edges correspond to road segments, and are related to weights concerning the site visitors information (eg; street-community distance, velocity of cars, or the lengthen time). Over such street networks, a huge spectrum of functional issues were broadly studied, together with variety queries, k-nearest neighbor (kNN) queries, reverse nearest neighbor queries, shortest course queries, multi-source skyline queries, and so on.

In this paper, we will investigate a useful query in spatial road networks, namely the trip planner query (TPQ), which is helpful for the decision making by travelers. Specifically, travelers may want to visit several cities of interest, and stay at each place for a period of time. Due to the complicated situations, the traffic conditions of road networks may highly depend on specific time slots. For example, during the “rush hour”, traffic jams may occur on some busy roads, while at night there might be few vehicles on roads. Thus, on such time-dependent road networks, a trip planner needs to design an appropriate sequence of driving among cities, as well as the departure time from each city, such that traffic jams can be avoided and the minimum traveling time can be achieved.

II. RELATED WORK

Spatial Road Networks: In the literature, a spatial road network is usually modeled by a set of road segments (lines, or edges) that are connected at their ending points (vertices), which can be also considered as a graph. The distance

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

between any two points in road networks is given by their shortest path distance (rather than Euclidean distance). The most popular algorithm for answering the shortest path query is Dijkstra's algorithm. Later, several variants have been proposed to heuristically prune the search space or materialize some paths. Huang et al. studied the shortest path search with constraints. Many other queries have been studied in spatial road networks, including range queries, k-nearest neighbor (kNN) queries, reverse nearest neighbor queries, and multi-source skyline queries For probabilistic graphs, previous works modeled the (RDF) graph containing edges associated with existence probabilities (i.e., edges either exist or do not exist in the graph). Moreover, some other works [12], [24],

[30] model the probabilistic graph by the graphical model,Bayesian network [20], whose labels of vertices connecting through edges are dependent, and represented by conditional probability tables (CPTs). In contrast, our work considers the spatial road network, whose edge weights are probabilistic and time-dependent, that is, associated with UDFs w.r.t. departure time from vertices. Furthermore, Hua and Pei [13] studied path queries on road networks with uncertain traffic data, which retrieve paths between 2 vertices having total weights less than a threshold with high probability. The path query differs from our TPQ query which involves multiple query points, and road networks with time-dependent traffic on edges. Thus, due to different graph models/query types, techniques proposed by prior works cannot be used for our TPQ problem..

III. SYSTEM MODEL

The system proposed here will investigate a useful query in spatial road networks, namely the trip planner query(TPQ), which is helpful for the decision making by travelers. Specifically, travelers may want to visit several cities of interest, and stay at each place for a period of time. Due to the complicated situations, the traffic conditions of road networks may highly depend on specific time slots. For example, during the "rush hour", traffic jams may occur on some busy roads, while at night there might be few vehicles on roads. Thus, on such time-dependent road networks, a trip planner needs to design an appropriate sequence of driving among cities, as well as the departure time from each city, such that traffic jams can be avoided and the minimum traveling time can be achieved.

Although many works have been proposed for various queries on road networks, most of them assume that the exact traffic information is available at hand. However, this assumption is too idealistic. Usually, due to many unknown factors such as the sudden change of weather, road closure for the construction, or traffic accidents, it is difficult, if not impossible, to obtain traffic information in time. We can only predict or infer the future traffic information (e.g., the possible speeds of vehicles) from historical data (e.g., statistics of traffic data at the same time of a day in history). The resulting predictions are thus imprecise and uncertain. To represent such a spatial road network, we use the (time-dependent) probabilistic model to capture the uncertainty of its traffic data, and formulate a probabilistic time-dependent graph (PT-Graph) with uncertain edge weights. In turn, our TPQ problem is to retrieve the best traveling plans over such a PT-Graph, which minimizes the total traveling time with high confidence.

To our best knowledge, no prior works have studied this TPQ problem in the PT-Graph. While a relevant query retrieves the shortest path between source and destination in a certain time-dependent graph, our TPQ problem considers multiple ($\geqslant 2$) places (not only 2), uses the PT-Graph with the probabilistic model (rather than a certain graph), and has a different goal of minimizing the traveling time on road networks only (instead of the total time that includes the staying time at vertices). Therefore, our problem is more challenging and cannot directly borrow existing methods.

Moreover, some works considered queries in uncertain graphs whose edges exist with some probabilities. They are, however, different from our PT-Graph where edges always exist, but the traveling time of vehicles (i.e., weight) on each edge (road segment) is uncertain and time-dependent. In addition, these works studied different problems such as *k*NN or reachability queries. Thus, their proposed techniques (w.r.t. distinct data models and query types) cannot solve our TPQ problem directly. Due to the data uncertainty in the PT-Graph, our TPQ problem can be solved by first conducting queries in each possible world of the PT-Graph, and then combining the query results from all the possible worlds, where each possible world is a materialized instance of road network (PT-Graph) with fixed traffic conditions

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

that can appear in reality. However, since the number of possible worlds can be exponential, this method is inefficient, and thus challenging to obtain the best trip plans efficiently from a PT-Graph.

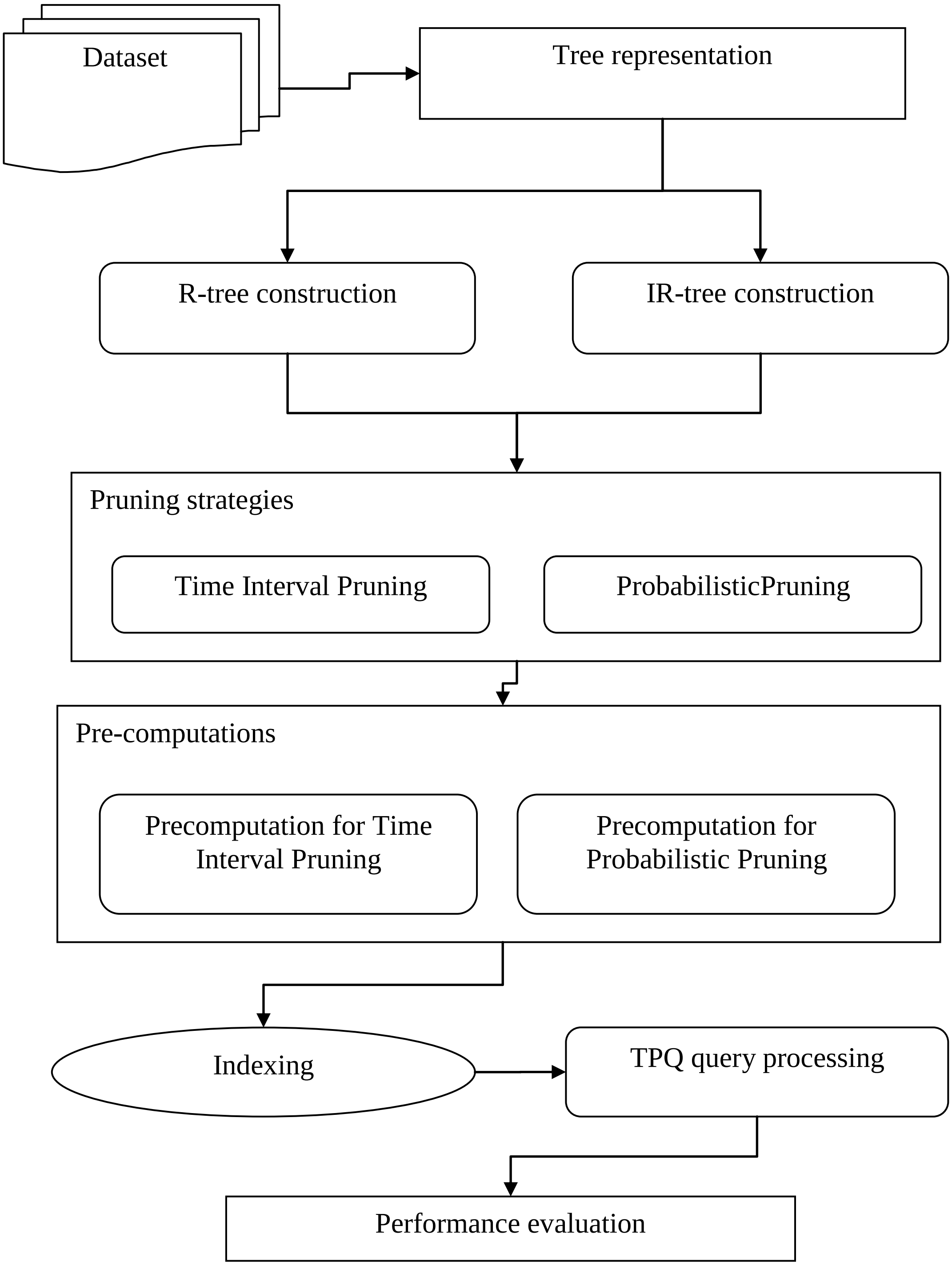

Fig 1: System architecture

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

1. R-tree construction

R-trees are data structures designed to store multi-dimensional objects. An essential problem with multi-dimensional objects (or rather, any object with dimension two or larger), is that there exists no good ordering on such objects. This makes the design of searchable multi-dimensional data structures difficult. R-trees have in common that they attempt to order data elements by using Minimum Bounding Rectangles (MBRs).

2. Pruning strategies

In this module, we present two pruning strategies, time interval pruning and probabilistic pruning. Specifically, the time interval pruning utilizes lower/upper bounds of the traveling times for trip plans to filter out false alarms; the probabilistic pruning considers probabilistic distributions to prune those trip plans with low confidences.

3. Pre-Computations

Our basic idea of pre-computations is as follows. We select m vertices from PT-Graph G as pivots, denoted as piv1, piv2, . . . , and pivm. Then, our goal is to estimate bounds of the traveling time between any 2 vertices vi and vj by using these m pivots. Unfortunately, traditional approaches that use pivots to derive the pruning bounds are not applicable to the traveling time on edges of road networks. This is because the triangle inequality may not hold for the traveling time (due to reasons such as traffic jam on road networks). That is, given 3 vertices x, y, and z in PT-Graph G, we may have: $T(x \rightarrow y) + T(y \rightarrow z) < T(x \rightarrow z)$, where $T(a \rightarrow b)$ is the traveling time of the shortest path from vertex a to vertex b. As a result, we cannot directly use the triangle inequality technique to derive the traveling time interval via pivots.

4.Indexing

In this module, we illustrate how to index the PT-Graph, as well as the pre-computed data, to facilitate the pruning. Specifically, we create a hierarchical tree index I over PT Graph G. On the leaf level of the tree, we store data of vertices vi ∈ V(G), including their spatial locations, the pre computed distance vectors to pivots, that is, (dist(vi, piv1), dist(vi, piv2), . . . , dist(vi, pivm)), and pointers pointing to vertex vi in PT-Graph G (represented by an adjacency list), in which functions of edges such as UDF(·), Min-UDF(·), Ma-UDF(·), and (1-β)-UDF(·) are stored. Then, we construct a hierarchical tree structure I, by recursively grouping vertices according to spatial locations of vertices/nodes, until finally a root is obtained. In particular, each intermediate node ep of the index I contains aggregated information such as spatial locations of its underlying vertices and distance vectors of its children. That is, each node ep contains a minimum bounding rectangles (MBR) that minimally bounds spatial locations of all vertices under this node. Here, our grouping criteria for the index construction are similar to the one in R∗-tree, that is, grouping vertices based on their closeness in the 2-dimensional spatial space.

5. IR-Tree construction

In our proposed system, we present IR-tree, an efficient index that provides the following required functions for geographic document search and ranking: 1) spatial filtering: all the spatially irrelevant documents have to be filtered out as early as possible to shrink the search space; 2) textual filtering: all the textually irrelevant documents have to be discarded as early as possible to cut down the search cost; and 3) relevance computation and ranking: since only the top-k documents are returned and k is expected to be much smaller than the total number of relevant documents, it is desirable to have an incremental search process that integrates the computation of the joint relevance and document ranking seamlessly so that the search process can stop as soon as the top-k documents are identified. In addition, IR-tree is designed by taking into account the storage and access overheads since a document set is very large in terms of numbers of documents and their words.

IV.**EXPERIMENT RESULT**

In order to support efficient geographic document search, the IR-tree clusters a set of documents into disjointed subsets of documents and abstracts them in various granularities. By doing so, it enables the pruning of those (textually or spatially) irrelevant subsets. The efficiency of IR-tree depends on its pruning power, which, in turn, is highly related to the effectiveness of document clustering and the search algorithms. Our IR-tree clusters spatially close documents

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

together and carries textual information in its nodes. These designs distinguish our IR-tree from other hybrid indexes. IR-tree associates each leaf entry with an inverted file and associates a document summary that provides textual information of documents with each node so that the tf and idf values of the document words can be estimated at nodes without examining individual documents.

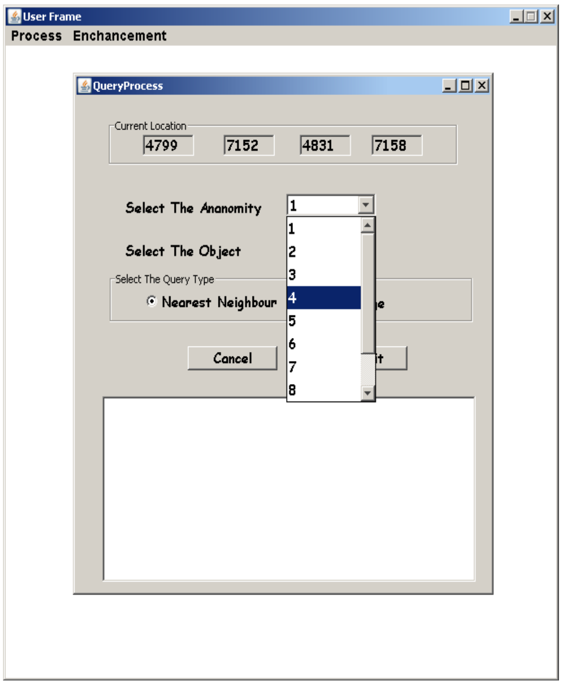

Fig 2:Query input

The system will provide high search efficiency and also minimizes input/output cost. Thus it will enhance the overall reliability of the system.

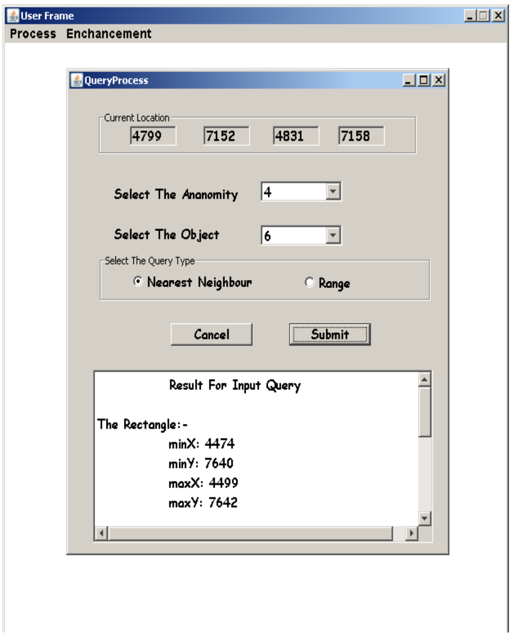

Fig 3: Query processing

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 5, May 2015

V. CONCLUSION AND FUTURE WORK

In the existing system, investigate a useful and important problem, called the trip planner query (TPQ), in the probabilistic and time-dependent spatial road network. Specifically, a TPQ query helps travelers find those trip plans, which visit several places of interest (staying at each place for some period of time), and have the minimum traveling time on road networks with high confidence. Although, in this work Search complexity is highest and also Accuracy rate of the system is lower. In order to overcome these problems, we are proposing the efficient index such that IR – tree. The basic indexing architecture is the combination of a spatio- textual index such as IR – tree and the road network architecture. In top-K spatial keyword queries over a road network, the distance between the objects and query location is constrained by a road network. A top-k spatial keyword query returns the k best objects ranked in terms of both the distance to the query location and textual relevance to the query keywords. This system Provides efficient indexing approach as well as minimize I/O costs for high search efficiency. Experimental result shows that our proposed system is well effective than the existing system.

Recent years have witnessed an increased interest in recommender systems. Despite significant progress in this field, there still remain numerous avenues to explore. Indeed, our future plan is to provide a study of exploiting online travel information for personalized travel package recommendation. A critical challenge along this line is to address the unique characteristics of travel data, which distinguish travel packages from traditional items for recommendation.

REFERENCES

[1]. Dalvi N. and Suciu D (2007) "Efficient query evaluation on probabilistic databases," VLDB J., vol. 16, no. 4, pp. 523–544.
[2]. Deng K., Zhou.X, and Shen H.T (2007) "Multi-source skyline queryprocessing in road networks," in Proc. IEEE 23rd ICDE, Istanbul,Turkey, pp. 796–805.
[3]. B. Ding.B, Yu.J.X, and Qin.L (2008) "Finding time-dependent shortestpaths over large graphs," in Proc. 11th EDBT, Nantes, France, pp. 205–216.
[4]. Hua.M and Pei.J (2010) "Probabilistic path queries in road networks:Traffic uncertainty aware path selection," in Proc. 13th EDBT,Lausanne, Switzerland, pp. 347–358.
[5]. Huang.H and Liu.C (2009) "Query evaluation on probabilistic RDF databases," in Proc. 10th WISE, Poznan, Poland, pp. 307–320.
[6]. Jampani.R*et al*(2008), "MCDB: A montecarlo approach to managing uncertain data," in *Proc. SIGMOD*, Vancouver, BC, Canada,pp. 687–700.
[7]. Jeung.H, Yiu.M.L, Zhou.X, and Jensen.C.J (2010) "Path prediction and predictive range querying in road network databases," VLDB 4, pp. 585–602.
[8]. Jin.R, Liu.L, Ding.B, and Wang .H(Jun. 2011) "Distance-constraint reachability computation in uncertain graphs," in Proc. VLDB, pp. 551–562.
[9]. Li.J, Saha.B, and Deshpande.A (2009) "A unified approach to ranking in probabilistic databases," in Proc. VLDB, Secaucus, NJ, USA, pp. 502–513.
[10].

BIOGRAPHY

M.K Chandrasekharan is pursuing Master of Engineering in Computer Science and Engineering in Maharaja Institute of Technology, Coimbatore, India.

S.Shiva shankar is assistant Professor in Department of Computer Science, Maharaja Institute of Technology, Coimbatore, India.