Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 10, October 2015, pg.318 – 325

SURVEY ARTICLE

A SURVEY ON SEARCHING SPATIO-TEXTUAL TOP K-QUERIES BY REVERSE KEYWORD

T.SUGANIYA[1], M.AKILA RANI[2], K.S.KANNAN[3]

[1]PG STUDENT, [2]ASSISTANT PROFESSOR, [3]ASSISTANT PROFESSOR

[1,2]NPR COLLEGE OF ENGINEERING AND TECHNOLOGY, TAMILNADU, INDIA

[3]MIET COLLEGE OF ENGINEERING AND TECHNOLOGY, TAMILNADU, INDIA

[1] suganiyabe044@gmail.com; [2] akilakamalam@gmail.com; [3] saikannan2012@gmail.com

Abstract— ***Spatio-textual queries recover the greatest similar items with respect to a given site and a keyword set. Existing trainings mostly attention on how to professionally find the top-k result set given a spatio-textual query. However, in many request states, users cannot accurately prepare their keywords and instead favor to select them from some applicant keyword sets. Furthermore, in info looking request's, it is valuable to highpoint the items with the tags under which the items have high grades. Motivated by these applications, suggest a novel query example, namely Reverse Keyword Search for Spatio-Textual top-k Queries (RSTQ). It proceeds the keywords under which a aim item will be a spatio-textual top-k result. To professionally process the new query, devise a novel hybrid index KcR-tree to store and summarize the spatial and textual info of objects. Location-based services have many diverse applications like finding the nearest stores in a particular area, location-based advertisement, analyzing wildlife and traffic movements, location-based gaming etc. So the system proposes Best Algorithm and new method is Spatio-Text Inverter Method and then Multiple Query Processing Methods. It has the ability to perform keyword-augmented nearest neighbor search in time. To improve the effectiveness and efficiency it focuses on the top-k matches for the query, for reasonable values of k.***

Keywords— ***Spatio-Text Inverter; Multiple Query Processing; Top-k matches; Reverse Queries***

I. INTRODUCTION

Spatial data is also known as geospatial data or geographic information.Spatial data is usually stored as coordinates and topology, and is data that can be mapped.A spatial database, or geodatabase is a database that is optimized to store and query data that represents objects defined in a geometric space. Most spatial databases allow representing simple geometric objects such as points, lines and polygons.

Location-based services are the services offered to users based on their locations. Location-based services have many diverse applications like finding the nearest stores in a particular area, location-based advertisement, analyzing wildlife and traffic movements, location-based gaming etc. The advancements in database and mobile technology and rapidly increasing popularity of location-based services results in huge amounts of data being collected in databases. Location- based services have attracted significant attention from the industrial and research community.

R-kNN (Reverse k-Nearest Neighbor) query finds applications in decision support systems where the task is to open a new facility like restaurant in an area such that it will be least influenced by its competitors and attract good business. Another application is a profile based marketing, where a company maintains profiles of its customers and wants to start a new service such that the service is under the influence of maximum number of customers i.e. maximum number of customers are interested in that service. The R-kNN query in decision support system is an example of a monochromatic query as the database objects and the query are of the same type i.e. restaurants. The application of R-kNN in profile based marketing is an example of a bichromatic query as the database objects are customers and the query is service to be started by the company.

Location-based services require the users to report their exact location continuously. A user who doesn't want to send his/her exact location has to stop using the location-based services provided by the service provider. The data collected by such servers can be used to study the user behaviour, visiting patterns, and stalking personal locations. There is a need to find a way such that the user can enjoy the benefits of using location-based services while preserving his/her location privacy. A framework for preserving location privacy of a user and defined three types of Nearest Neighbor queries namely, Public Query over Private Data, Private Query over Public Data and Private Query over Private Data.

Public Query over Private Data signifies that the exact location of the querying user is known to the server but the exact location of database objects is unknown. An example of nearest neighbor Public Query over Private Data is a location based advertisement where a restaurant wants to send its advertisement to customers in its vicinity. In this scenario, the customers want to preserve their privacy and the restaurant (query point) wants to find its nearest customers. An example of Private Query over Public Data is a customer trying to find the nearest restaurant.

An example of Private Query over Private Data is a friend finder application, where a user wants to find his/her nearest buddy. Here, both the querying user and his friends (the database objects) want to hide their exact location from the location-based server. However, our thesis is a privacy preserving evaluation of reverse nearest neighbor query. Privacy preserving reverse nearest neighbor queries also find its applications in many scenarios. An example of Reverse Nearest Neighbor Public query over Private data is a shopping mall trying to find customers which have the shopping mall as one of its k nearest neighbors.

Another example is a pizza store owner trying to find customers who have the pizza store in its k nearest neighbors in order to send them discount coupons. An example of Private Query over Public Data is a customer finding a good location for his home such that he is in the influence of at least k facilities like School, Hospitals etc. Examples of Private query over Private Data include peer to peer applications like file sharing and match fixing application where a match fixer wants to avoid an area where there are other match fixers around him. Another application is location-based gaming, where a gamer wants to be in the influence of at least k friends without revealing their exact location.

II. LITERATURE SURVEY

2.1 Top-K Spatial Keyword Queries on Road Networks

The popularization of GPS-enabled devices there is an increasing interest for location-based queries. In this context, one interesting problem is processing top-k spatial keyword queries. Given a set of objects with a textual description (e.g., menu of a restaurant), a query location (latitude and longitude), and a set of query keywords, a top-k spatial keyword query returns the k best objects ranked in terms of both distance to the query location and textual relevance to the query keywords.

Top-k spatial keyword queries return the k most excellent spatio-textual objects ranked in terms of both spatial nearness to the query site and textual consequence to the query keywords. Even though the wide range of location-based applications that can promote from these queries, the current approaches for processing top-k spatial keyword queries are limited to the Euclidean space. A top-k spatial keyword query on road networks returns the k best objects in terms of both 1) direct path to the query site, and 2) textual consequence to the query keywords

In top-k spatial keyword queries on road networks both direct path and textual consequence are considered. For example, for the query "Home fries" point as q.l, the spatio-textual item point as p6 may appear better ranked than point as p7 because the description of p6 ("Egan Boliden bar & snack bar") is more textually related to the query keywords than the description of p7 ("Choco coffee shop"), and p6 is only slightly more distant to q.l than p7. The peak object conversely, is p10 because it is very close to q.l and is also related to the query keywords. Note that p11 is not returned as an effect of this query, since none of the conditions in the report of p11 appear in the query keywords. Top-k spatial keyword queries on road networks can be in use by location-based applications to provide a more accurate and reasonable result. However, dealing out these queries is costly, since it requires computing several shortest paths.

To the greatest of information, dealing out top-k spatial keyword queries on road networks has never been planned before. They formalize the concepts of this new query type and explain how to rank objects allowing for both the network distance and the textual consequence. Then, present an enhanced approach that employs inverted files to index the description of the spatio-textual objects lying on the road networks and, therefore, can process queries more professionally.

2.2. Spatial Keyword Query Processing: An Experimental Evaluation.

Geo-textual indices play an important role in spatial keyword querying. The existing geo-textual indices have not been compared systematically under the same experimental framework. This makes it difficult to determine which indexing technique best supports specific functionality

With the proliferation of online objects with both an associated geo-location and a text description, the web is acquiring a spatial dimension. Specifically, web users and content are increasingly being geo-positioned and geo-coded. At the same time, textual descriptions of points of interest, e.g., cafes and tourist attractions are increasingly becoming available on the web. Spatial keyword queries are being supported in real-life applications, such as Google Maps where points of interest can be retrieved, Foursquare where geo-tagged documents can be retrieved, and Twitter where tweets can be retrieved. Spatial keyword querying is also receiving increasing interest in the research community where a range of techniques have been proposed for efficiently processing spatial keyword queries.

Three types of spatial keyword queries:

a) **Boolean kNN Query**: “Retrieve the k objects nearest to the user’s current location (represented by a point) such that each object’s text description contains the keywords tasty, pizza, and cappuccino.”

b) **Top-k kNN Query**: “Retrieve the k objects with the highest ranking scores, measured as a combination of their distance to the query location (a point) and the relevance of their text description to the query keywords tasty, pizza, and cappuccino.”

c) **Boolean Range Query**: “Retrieve all objects whose text description contains the keywords tasty, pizza, and cappuccino and whose location is within 10 km of the query location.”

2.3. Efficient Retrieval of the Top-K Most Relevant Spatial Web Object

The conventional Internet is acquiring a geospatial dimension. Web documents are being geo-tagged and geo-referenced objects such as points of interest are being associated with descriptive text documents. The resulting fusion of geo-location and documents enables new kinds of queries that take into account both location proximity and text relevancy. The framework encompasses algorithms that utilize the proposed indexes for computing location-aware as well as region-aware top-*k* text retrieval queries, thus taking into account both text relevancy and spatial proximity to prune the search space.

Driven in part by the emergence of the mobile Internet, the conventional Internet is acquiring a geo-spatial dimension. On the one hand, many (geo-referenced) points of interest—e.g., stores, tourist attractions, hotels, entertainment services, public transport, and public services—are being associated with descriptive text documents. On the other hand, web documents are increasingly being geo-tagged.

It has found that about one fifth of web search queries are geographical and have local intent, as determined by the presence of geographical terms such as place names and postal codes . Indeed commercial search engines have started to provide location based services, such as map services, local search, and local advertisements. For example, Google Maps supports location-aware text retrieval queries. Additional examples of location-based services include online yellow pages.

An example query may request a "good micro-brewery that serves pizza" and that is close to the user's hotel. The answer to such a top-k query is a list of k objects ranked according to a ranking function that combines their distances to the query location and the relevance of their textual descriptions to the query phrase. The LkT (Location-aware top-k Text retrieval) query is different from the query that retrieves relevant documents within a geographic range.

The LkT query poses new challenges for both existing spatial database and existing information retrieval techniques that have been developed separately. The research in spatial databases mainly focuses on highly structured, map-based geometric data and their attributes. . First, they introduce a new type of location-aware top-k text retrieval queries, LkT queries, that returns objects ranked according to a linear interpolation function that combines normalized location proximity and text relevancy. Second, to efficiently process the query, a new indexing framework that integrates location indexing and text indexing, and then develops an IR-tree and an associated algorithm for processing the LkT query.

2.4. Reverse Top-K Queries

Rank-aware query processing has become essential for many applications that return to the user only the top-k objects based on the individual user's preferences. Top-k queries have been mainly studied from the perspective of the user, focusing primarily on efficient query processing.

It formally defines reverse top-k queries and introduces two versions of the query, namely monochromatic and bichromatic. They first provide a geometric interpretation of the monochromatic reverse top-k query in the solution space that helps to understand the reverse top-k query conceptually. Such a query, if computed in a straightforward manner, requires evaluating a top-k query for each user preference in the database, which is prohibitively expensive even for moderate datasets. Top-k queries retrieve only the k objects that best match the user preferences, thus avoiding huge and overwhelming result sets. Therefore it is very important for a manufacturer that its products are returned in the highest ranked positions for as many different user preferences as possible.

Reverse top-k queries differ from reverse nearest neighbor (RNN) queries. An RNN query retrieves the set of points having the query point as their nearest neighbor and there exists a monochromatic and a bichromatic version. In contrast to RNN queries, the reverse top-k query q finds the distance functions (in terms of weights) for which q would qualify as a k-nearest neighbor of the point positioned at the origin of the data space. Reverse skyline queries aim at identifying customers that are interested in a product, based on the dominance relationship.

2.5. Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concern

Context - aware computing often involves tracking peoples' location. Many studies and applications highlight the importance of keeping people's location information private. They discuss two types of location based services; location-tracking services that are based on other parties tracking the user's location and position-aware services that rely on the device's knowledge of its own location. To find that even though the perceived usefulness of the two different types of services is the same, location tracking services generate more concern for privacy than position -aware services. They conclude that development emphasis should be given to position- aware services but that location -tracking services have a potential for success if users are given a simple option for turning the location -tracking off.

Context-aware computing describes applications, often implemented for mobile devices that adapt to environmental sensor information. Many of these applications rely solely on location information as their context and although some researchers claim that too much attention is given to this type of sensor information , it is predicted that location -based services will be the most common form of context - aware computing. As mobile telephony becomes increasingly common as a handheld computing platform, location -tracking of mobile phones enables location -based services to spread outside closed environments. Location -tracking of customers by mobile telephony providers via GSM and later GPS enabled services generates a need for addressing privacy issues in relation to the building of location based technologies and services.

Such as mobile telephony service providers, whereas position-aware services are based on the device's own knowledge of its position. Examples of the two services include locating friends or family members and updating the time when entering a new time zone, respectively. By studying peoples' concern for and use of location -based services , it is a goal to provide insight about the degree to which location based services are considered to be intrusive on users' privacy and whether location-tracking or position-aware services generate more concern.

Although some studies of users' privacy and identification preferences exist, none of them focus solely on concerns about location -tracking or positioning. The study find that their concern for privacy depends on what types of information they are asked to give up, but also on the application's usefulness to the user. Another study compares general privacy concerns in different situations to the inquirer of the information and finds that inquirer is a greater determinant for what people want to reveal, than situation.

3. CONCLUSION

The system devise a cross key to store the spatial and textual information of objects to accelerate the processing of RSTQ (Reverse Spatio-Textual Queries). For region-based RSTQ, the system has a reduction-based technique to avoid enumerating an infinite number of query points. The system proposes Spatio-Text Inverter algorithm (STI). It has an ability to perform keyword augmented in a time and aims at promoting the ranking of the target object with the least query modification. Ranking is also an important one to study approximate algorithms that compute quickly a good estimation of the outcome set. In first phase work is ranking process is only based on Hotel Advertisement and then check user search query. To find target object and then choosing nearest distance. Both are based on how much of count. It can be enhanced to customer rating concept and then to eliminate the fake links in future work.

REFERENCES

[1] K. Norvag, J. B. Rocha-Junior. "Top-k spatial keyword queries on road networks". In EDBT, pp. 168-179, 2012.

[2] L. Chen, G. Cong, C. S. Jensen, and D. Wu. "Spatial Keyword Query Processing: An Experimental Evaluation". PVLDB, 6(3): 217-228, 2013.

[3]G. Cong, C. S. Jensen, D. Wu. "Efficient retrieval of the top-k most relevant spatial web objects". PVLDB, 2(1):337-348, 2009.

[4] C. Doulkeridis, Y. Kotidis, and K. Norvag, A. Vlachou. "Reverse top-k queries". In ICDE, pp. 365-376, 2010.

[5] Anind Dey, Louise Barkuus. "Location-Based Services for Mobile Telephony: a Study of Users' Privacy Concerns".-2003

