Best Keyword Cover Search Using Spatial Data

[1] N.Shiva Prasad,[2] G.JoseMary, [3]P.Srinivas Rao
[1]M.Tech ,CSE, Jayamukhi Institute Of Technological Sciences,Warangal,India
[2]Assistant professor,CSE, Jayamukhi Institute Of Technological Sciences,Warangal,India
[3]Associate professor,CSE, Jayamukhi Institute Of Technological Sciences,Warangal,India

Abstract: In recent years, we notice the increasing convenience and value of keyword rating in object analysis for the simpler choice creating. . An interesting drawback known as closest Keywords search is to question objects, referred to as keyword cowl, which along cowl a collection of question keywords and have the minimum repose-objects distance as well because the keyword ranking of objects. The baseline formula is motivated with the help of the strategies of closest key words search that is established on exhaustively combining objects from extraordinary question keywords to get candidate keyword covers. When the amount of question keywords will increase, the performance of the baseline formula drops dramatically so of huge candidate key phrase covers generated. To attack this downside, this work proposes a way additional ascendable formula referred to as keyword nearest neighbor growth (keyword-NNE). In comparison with the baseline formula, keyword-NNE algorithm immensely reduces the amount of candidate keyword covers generated. The in-depth analysis and broad experiments on actual information units have even the prevalence of our key phrase-NE algorithm.

Keywords: Spatial Database, Point of Interests, Keywords, Keyword Rating, Keyword Cover.

I. INTRODUCTION

Data mining is that the methodology of extracting knowledge from a dataset for purchasers to use it in quite heap of reason. The purpose of such knowledge performs an enormous position in keyword looking out. looking out may be a designed activity taking place in data processing. craving for special objects from spatial information has not too some time past sparked enthusiasm among researchers. This excited to boost strategies to retrieve Spatial objects. Spatial objects contain objects related to spatial elements. In different words, Spatial objects involve Spatial data along with line of longitude and latitude of place. Querying such knowledge is named fantabulous key word quilt querying. Search is named quality key word quilt search. Existing method to such knowledge confine mind either minimum inter goal distance and keyword search. As a result new approaches for fine key word cowl search were once developed. Common nearest neighbor search compute nearest neighbor with the help of considering that distance as function. during this context, nearest neighbour search center of attention discovering nearest neighbors the place keywords and spatial knowledge plays a predominant have an impression on. It comes with algorithms to reply such question. (dimension ten &usual)This report may be a example. N electronic copy also will be downloaded from the conference web site. For queries on paper recommendations, please contact the conference publications committee as indicated on the conference computing machine. Knowledge concerning final paper submission is accessible from the convention computing machine.

The content used for querying takes the sort of spatial information. Nice keyword cowl question takes form of keyword phrases or objects. for example, school. Given a spatial information P, that embrace set of components. For a query Q, the place Q belong to line of objects, it search for nearest neighbor within the factor by victimization looking its significance in this knowledge when that participate in nearest neighbor search to accumulate the reply to the question. For higher deciding, proposal of keyword score accustomed be conferred beside its components different than distance. For such search , question can take type of characteristic of objects. It rummage around for nearest neighbour based on a replacement similarity live, named weighted ancient of index score which mix key phrase rating, keyword search and nearest neighbour search.

II. LITERATURE SURVEY

I. D. Felipe, V. Hristidis, and N. Rishe, Keyword search on spatial databases.

Many applications require finding objects closest to a specified location that contains a set of keywords. For example, online yellow pages allow users to specify an address and a set of keywords. In return, the user obtains a list of businesses whose description contains these keywords, ordered by their distance from the specified address. The problems of nearest neighbor search on spatial data and keyword search on text data have been extensively studied separately. However, to the best of our knowledge there is no efficient method to answer spatial keyword queries, that is, queries that specify both a location and a set of keywords. In this work, we present an efficient method to answer top-k spatial keyword queries. To do so, we introduce an indexing structure called IR2-Tree (Information Retrieval R-Tree) which combines an R-Tree with superimposed text signatures. We present algorithms that construct and maintain an IR2-Tree, and use it to answer top-k spatial keyword queries. Our algorithms are experimentally compared to current methods and are shown to have superior performance and excellent scalability.

X. Cao, G. Cong, C. Jensen, and B. Ooi, Collective spatial keyword querying.

With the proliferation of geo-positioning and geo-tagging, spatial web objects that possess both a geographical location and a textual description are gaining in prevalence, and spatial keyword queries that exploit both location and textual description are gaining in prominence. However, the queries studied so far generally focus on finding individual objects that each satisfy a query rather than finding groups of objects where the objects in a group collectively satisfy a query. We define the problem of retrieving a group of spatial web objects such that the group's keywords cover the query's keywords and such that objects are nearest to the query location and have the lowest inter-object distances. Specifically, we study two variants of this problem, both of which are NP-complete. We devise exact solutions as well as approximate solutions with provable approximation bounds to the problems. We present empirical studies that offer insight into the efficiency and accuracy of the solutions.

G. Cong, C. Jensen, and D. Wu, Efficient retrieval of the top-k most relevant spatial web objects.

The conventional Internet is acquiring a geo-spatial 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 a new kind of top-k query that takes into account both location proximity and text relevancy. To our knowledge, only naive techniques exist that are capable of computing a general web information retrieval query while also taking location into account. This paper proposes a new indexing framework for location-aware top-k text retrieval. The framework leverages the inverted file for text retrieval and the R-tree for spatial proximity querying. Several indexing approaches are explored within the framework. The framework encompasses algorithms that utilize the proposed indexes for computing the top-k query, thus taking into account both text relevancy and location proximity to prune the search space. Results of empirical studies with an implementation of the framework demonstrate that the paper's proposal offers scalability and is capable of excellent performance.

D. Zhang, Y. Chee, A. Mondal, A. Tung, and M. Kitsuregawa, Keyword search in spatial databases: Towards searching by document.

This work addresses a novel spatial keyword query called the m-closest keywords (mCK) query. Given a database of spatial objects, each tuple is associated with some descriptive information represented in the form of keywords. The mCK query aims to find the spatially closest tuples which match m user-specified keywords. Given a set of keywords from a document, mCK query can be very useful in geotagging the document by comparing the keywords to other geotagged documents in a database. To answer mCK queries efficiently, we introduce a new index called the bR*-tree, which is an extension of the R*-tree. Based on bR*-tree, we exploit a priori-based search strategies to effectively reduce the search space. We also propose two monotone constraints, namely the distance mutex and keyword mutex, as our a priori properties to facilitate effective pruning. Our performance study demonstrates that our search strategy is indeed efficient in reducing query response time and demonstrates remarkable scalability in terms of the number of query

p-ISSN: 2348-6848
e-ISSN: 2348-795X
Volume 04 Issue 07
June 2017

keywords which is essential for our main application of searching by document.

III. SYSTEM MODEL

This paper investigates a typical variant of mCKquery, called Best Keyword Cover (BKC) query,which considers inter-objects distance as good askeyword ranking. It's encouraged through the statementof increasing availability and importance ofkey phrase rating in selection making. Thousands oforganizations/offerings/features around the world havebeen rated through users by means of online industryassessment sites similar to Yelp, Citysearch, ZAGAT andDianping, and so on.

This work develops two BKC query processingalgorithms, baseline and key phrase-NNE. Thebaseline algorithm is stimulated with the aid of the mCK queryprocessing ways. Both the baseline algorithmand keyword-NNE algorithm are supported viaindexing the objects with an R*-tree like index,known as KRR*-tree.

We developed much scalable key word nearestneighbor expansion (key phrase-NNE) algorithmwhich applies a further approach. Key word-NNEselects one query keyword as predominant questionkeyword. The objects related to the importantquery keyword are primary objects. For everyfundamental object, the nearby great answer (often called localbest key word duvet lbkc) is computed.

Among them, the lbkc with the very best evaluationis the answer of BKC question. Given a majorobject, its lbkc can also be identified by using readilyretrieving a few local and enormously rated objects in each and every non-predominant question key phrase (two-4objects in normal as illustrated in experiments).

A. Indexing Keyword Ratings

A single tree structure is used to index objects ofextraordinary key terms. The one tree may also be elevated with one more dimension to index key phrase rating. A single tree constitution fits the situation that almost all key words are query key words. For the above mentioned illustration, all keywords, i.e., “resort”, “restaurant” and “bar”, are query keywords. Nevertheless, it is more universal that most effective a small fraction of key terms are query keyword phrases. For illustration in the experiments, only not up to 5% keywords are query key words. In this situation, a single tree is bad to approximate the spatial relationship between objects of few distinct keywords. Consequently, a couple of KRR*- tree are used on this work, every for one keyword.1 The KRR*-tree for key word ki is denoted as KRR*kitree. Given an object, the ranking of a related keyword is usually the imply of ratings given by a number of patrons for an interval of time. The alternate does happen however slowly. Despite the fact that dramatic alternate happens, the KRR*-tree is up to date in the general approach of R*-tree replace.

B. Keyword nearest Neighbor Expansion

Using the baseline algorithm, BKC query can beeffectively resolved. However, it is based onexhaustively combining objects (or their MBRs). Eventhough pruning techniques have been explored, it hasbeen observed that the performance dropsdramatically, when the number of query keywordsincreases, because of the fast increase of candidatekeyword covers generated. This motivates us todevelop a different algorithm called keyword nearestneighbor expansion. We focus on a particular querykeyword, called principal query keyword. The objectsassociated with the principal query keyword are calledprincipal objects. The goal of the interface is toprovide point of interest information (static anddynamic ones) with, at least, a location, somemandatory’s attributes and optional details(description…). In order to provide that information,the component that implements the interface uses themap database information to locate and display pointof interest (POI) or to select POI as route waypoint andfavorite. This component not only provides searchfunctionalities for the local database but also a way toconnect external search engine to this component andenhance the search criteria and the list of results It alsoproposes a solution to get custom POIs (not part of thelocal map database) or to dynamically update contentand description of local POI.

Using the baseline algorithm, BKC query can also be effortlessly resolved. Nevertheless, it's established on exhaustively combining objects (or their MBRs). Although cutting tactics were explored, it has been determined that the efficiency drops dramatically, when the number of query key terms increases, considering of the quick increase of candidate keyword covers generated. This motivates us to strengthen an extra algorithm referred to as key word nearest neighbor growth. We focus on a specific query keyword, known as predominant query keyword. The objects associated with the foremost query key word are known as important objects. The purpose of the interface is to furnish factor of interest expertise (static and dynamic ones) with, at least, a place, some necessary's attributes and not obligatory small print (description…). With a view to provide that understanding, the factor that implements the interface uses the map database information to locate and show point of interest (POI) or to select POI as route waypoint and favorite. This element not handiest provides search functionalities for the local database but also a technique to join outside search engine to this element and increase the search criteria and the list of results It also proposes a solution to get custom POIs (now not part of the neighborhood map database) or to dynamically replace content and description of local POI.

This is achieved by specifying and providing interfacesto:

- Select POIs from one of their attributes (e.g.,Category, Name,…)
- Retrieve POI attributes (e.g., Location andDescription)
- Get dynamic content for a given POI.
- Add custom POI to the map display
- Import new POIs and POIs categories fromlocal file.

C. LBKC Computation

Given a spatial database, each object may be related to one or a couple of key words. Without loss of generality, the article with a couple of key terms are converted to more than one objects placed on the identical area, every with a specified single keyword. When extra processing a candidate key phrase cover, key word-NNE algorithm most commonly generates much less new candidate key phrase covers compared to BFbaseline algorithm. In the grounds that the number of candidate key phrase covers extra processed in key phrase-NNE algorithm is optimal the quantity of key word covers generated in BF-baseline algorithm is much more than that in keyword- NNE algorithm. In flip, we conclude that the number of key word covers generated in baseline algorithm is rather more than that in keyword NNE algorithm. This conclusion is unbiased of the principal question key word for the reason that the analysis does no longer practice any constraint on the choice strategy of foremost query key word.

IV. CONCLUSION

The baseline algorithmgenerates an outsized variety of candidate keyword coverswhich ends up in dramatic production drop once morequery keywords ar given. The projected keywordNNE algorithmic rule applies a distinct process strategy,i.e., looking out native best answer for every object in acertain question keyword. As a consequence, the numberof candidate keyword covers generated is significantlyreduced. The analysis reveals that the quantity ofcandidate keyword covers which require to be furtherprocessed in.

REFERENCES

[1] Ke Deng; Xin Li; Jiaheng Lu; Xiaofang Zhou," Best Keyword Cover Search" Knowledge and Data Engineering,IEEE Transactions onYear: 2015.

[2] XinCaoy, Gao Cong, Christian S. Jensenz, ‟ Retrieving Top-k Prestige Based Relevant Spatial Web Objects‟, Proceedings of the VLDBEndowment, Vol. 3, No.1 , 2010.

[3] Ian De Felipe, Vagelis Hristidis, Naphtali Rishe, Keyword Search on Spatial Databases, in Proc. IEEE 24th Int. Conf. Data Eng., 2008, pp.656–665.

[4] Z. Li, K. C. Lee, B. Zheng, W.-C. Lee, D. Lee, and X. Wang,‟ IRTree: An efficient index for geographic document search‟, IEEE Trans. Knowl.Data Eng., vol. 99, no. 4, pp. 585–599, Apr. 2010.

[5] G. Cong, C. Jensen, and D. Wu, Efficient retrieval of the top-k most relevant spatial web objects, Proc. VLDB Endowment, vol. 2, no. 1, pp.337–348, Aug. 2009.

[6] D.Amutha Priya, Dr. T.Manigandan, Fast Accurate Mining on Spatial Database Using Keywords, International Journal for Research in AppliedScience & Engineering Technology (IJRASET), Volume 3, Special Issue-1, May 2015.

[7] Ruicheng Zhong, Ju Fan, Guoliang Li, Kian-Lee Tan and Lizhu Zhou, „Location-Aware Instant Search‟ CIKM‟12, October 29–November 2,2012, Maui, HI, USA

[8] X .cao, G Cong, C Jensen "COLLECTIVE SPATIAL KEYWORD QUERYING IN Proc.ACM SIGMOD Int. Conf manage. Data , 011 ,pp.373-384.

[9] G. R. Hjaltason and H. Samet, "Distance browsing in spatial databases," ACM Trans. Database Syst., vol. 24, no. 2, pp. 256–318, 1999.

[10] Z. Li, K. C. Lee, B. Zheng, W.-C. Lee, D. Lee, and X. Wang, "IRTree: An efficient index for geographic document search," IEEETrans. Knowl. Data Eng., vol. 99, no. 4, pp. 585–599, Apr. 2010.

[11] N. Mamoulis and D. Papadias, "Multiway spatial joins," ACMTrans. Database Syst., vol. 26, no. 4, pp. 424–475, 2001.

[12] D. Papadias, N. Mamoulis, and B. Delis, "Algorithms for queryingby spatial structure," in Proc. Int. Conf. Very Large Data Bases, 1998,pp. 546–557

[13] D. Zhang, B. Ooi, and A. Tung, "Locating mapped resources in web 2.0," in Proc. IEEE 26th Int. Conf. Data Eng., 2010, pp. 521–532.

[14] D. Zhang, Y. Chee, A. Mondal, A. Tung, and M. Kitsuregawa, "Keyword search in spatial databases: Towards searching by document," in Proc. IEEE Int. Conf. Data Eng., 2009, pp. 688–699.