K-NNE ALGORITHM FOR BEST KEYWORD COVER SEARCH

Dhawale Reshma Ramdas[1], Thete Madhuri Dhondibhau[2], Badhekar Pratima Malhari[3], Prof. Rathod R. R.[4]

[1] *B.E., Dept. of Computer, SGOI, COE, Belhe, SPPU, Pune, Maharashtra, India*
[2] *B.E., Dept. of Computer, SGOI, COE, Belhe, SPPU, Pune, Maharashtra, India*
[3] *B.E., Dept. of Computer, SGOI, COE, Belhe, SPPU, Pune, Maharashtra, India*
[4] *Assistant Professor., Dept. of Computer, SGOI, COE, Belhe, SPPU, Pune, Maharashtra, India*

ABSTRACT

Spatial databases stores the knowledge regarding the spatial posts which are related with the keywords featuring the knowledge like their business/administrations/ highlights Crucial situation called nearest keyword search is obviously to inquiry items called keyword cover In nearest keyword search it covers a couple of query keywords and minimal range between objects From latest few years keyword rating builds their convenience and significance in protest examination for the decision making This is actually the principal purpose for building the new algorithm called Best keyword cover which thinks inter range and also the ratings gave by the clients through the web business review destinations Best keyword search algorithm mixes the items from various query keywords to a create candidate keyword cover Two algorithms Base line algorithm and keyword nearest neighbor expansion algorithms are used to find best keyword cover The efficiency of the closest keyword algorithm declines drastically when just how several query keyword increases The right solution is of this problem of the existing algorithm this function proposes common variation called keyword nearest neighbor expansion which decreases the resulted candidate keyword covers

Keyword : *- Spatial database, keyword rating, inter object distance, keywords, keyword cover, POIs, etc….*

1. INTRODUCTION

An enhancing quantity of the purposes require the effective delivery of nearest neighbor NN queries obliged by the qualities of spatial objects Because of the popularity of keyword search especially on the Net a significant quantity of these applications allow the client to give a set of keywords that spatial 0bject must include in their information and other attribute The spatial keyword query comprises of a query area and the band of keywords The correct response is the list of things located with a blend of their separation to the question area and the pertinence of their content interpretation to question keywords A basic yet' conventional alternative which is an applied as a part of our working e.g. is split up first spatial keyword query where things are placed by separation and keywords are linked as a conjunctive station to destroy protests that do not include them An incident of a spatial keyword query is "find the nearest accommodations to the level that have keywords internet and pool" The best consequence of this question is the hotel object Unfortunately there is nothing efficient support for top k spatial keyword queries where in actuality the prefix of the end result list has expected Traction purposes use especially appointed mix of the nearest neighbor NN and keyword search procedures to deal with a concern Like an R Tree is applied to find out the best neighbors and for every single neighbor a developed file is applying to test if the query keyword is included In that undertaking show that such two stage techniques are inefficient This work produces two BKC query processing algorithm baseline and keyword NNE The baseline algorithm is enlivened by the mCK query processing algorithm Both the baseline algorithm and keyword NNE algorithm are upheld by the purchasing the articles by having an R* Tree like list called KRR* Tree In the baseline algorithm the thought is always to combine hubs in higher progressive levels of KRR* tree to create candidate keyword covers At that point the most stimulating opponent is considered in require by joining their tyke hubs to create new applicants Despite the fact that BKC query could be successfully resolved when the total amount of question keywords grows the delivery falls considerably due to gigantic candidate keyword covers made Ahead that diagnostic downside we're made significantly climbed keyword nearest neighbor expansion (keyword NNE) algorithm which applies the various

methods Keyword NNE decides among query keyword as a central query keyword The objects are connected with principal query keyword has principal object

2. LITERATURE REVIEW

JuFan, Guoliang Li, Ruicheng Zhong, Kian-Lee Tan ,Lizhu Zhou, Location Aware Instant Search[1]. : Now a days, mobile users widely uses the Location-Based Services(LBS). Location-based systems can work efficient if user enters the complete keyword, otherwise it had been shown incomplete output. It difficult to enter the complete keyword on mobile devices for get the proper relative result. In order to avoid this problem, proposed system studied about the location aware search. It could be returned search answers as the user enters in queries letter by letter. In this paper, the key challenge is to provide the relevant answers speedily. Author runs on the new index structure, prefix-region tree (called PR-tree), that can be support to provide the speedily results to the users. PR-Tree is really a tree-based index structure which seamlessly integrates the textual description and spatial information to index the spatial data. Using the PRTree ,authors develop efficient algorithms to guide single prefix queries and multi keyword queries.

Gao Congy, Christian S.Jensenz, Xin Caoy, Retrieving Top-k Prestige Based Relevant Spatial Web Objects[2]: The location-aware keyword query returns ranked objects that are near a question location and which have textual descriptions that match query keywords. There are lots of mobile applications and traditional services uses this kind of query, e.g. Yellow pages and Maps services. In previous work, ranked query returns independent potential results. Ranking is very important in decision making. However, a relevant result object with nearby objects that are also highly relevant to the query is likely to be preferable over a relevant object without relevant nearby objects. The paper proposes the thought of prestige-based relevance to recapture both textual relevance of an item to a question and the consequences of nearby objects. Based on this, a fresh form of query, the Location-aware top-k Prestige-based Text retrieval (LkPT) query, is proposed that retrieves the top k spatial web objects ranked according to both prestige-based relevance and location proximity. We propose two algorithms that compute LkPT queries. Empirical studies with real-world spatial data demonstrate that LkPT queries are more efficient in retrieving web objects when compared to a previous approach that will not consider the consequences of nearby objects; and they reveal that the proposed algorithms are scalable and outperform a baseline approach significantly. De Felipe, Vagelis Hristidis, Naphtali Rishe, Keyword Search on Spatial Databases. There are plenty of applications that finds the objects nearest to the specified location which contains a set of keywords. Yellow pages required address and a set of keywords to get the results. Yellow pages returns a listing of business/features/service whose description contains entered keywords, ordered by their inter-object distance from the specified location. In this paper, author studied problems of nearest neighbor search on location data and keyword search on text data separately. There's no any method that returns response to the best for spatial and keyword queries which relates to the same. In this paper, author proposed an efficient algorithm that returns top-k spatial keyword queries. Proposed system introduces indexing structure called Information retrieval R-Tree which will be combination of R-tree with super imposed text signatures. Algorithm returns the clear answer from IR2 tree which construct and maintain by the algorithm to the keyword queries. Proposed algorithms are superior performance and excellent scalability to the prior work experimentally.

Ken C.K. LEE, Baihua ZHENG, Zhisheng LI, Wang-Chien LEE, Dik Lun LEE,IR-tree[3]:An Efficient Index for Geographic Document Search Geographic search engine returns documents which are very close textually and spatially to the query keywords. Retrieved documents are ranked according to their joint textual and spatial relevance to the entered query. Existing indexing scheme inefficient in answering spatial queries because of lacking in index simultaneously handle both the textual and location aspect.

Ramakrishnan Srikant and Rakesh Agrawal. Fast algorithms for mining association rules in large databases[4]: We consider the situation of discovering association rules between items in a large database of sales transactions. We present two new algorithms for solving thii problem that are fundamentally distinctive from the known algorithms. Empirical evaluation shows these algorithms outperform the known algorithms by factors ranging from three for small problems to significantly more than an order of magnitude for big problems. Weal so show how the very best features of the two proposed algorithms may be combined in to a hybrid algorithm, called Apriori Hybrid. Scale-up experiments show that Apriori Hybrid scales linearly with the amount of transactions. Apriori Hybrid even offers excellent scale-up properties regarding the transaction size and the amount of items in the database

3. PROBLEM STATEMENT

Whenever the total amount of query keyword increases the actual performance falls drastically as a result of large candidate keyword generated. To beat this kind of critical issue proposed system design significantly scalable Keyword nearest neighbor expansion (KNNE) algorithm which applies a good alternative strategy.

4. EXISTING SYSTEM

Some existing works focus on finding specific objects by specifying problem consisting of a problem location and a couple of query keywords. . Each recovered object is related to keywords strongly related the query keywords and is near the query location. Existing system focus on finding specific objects by specifying a problem consisting of a problem location and a couple of query keywords.

4.1 DRAWBACKS OF EXISTING SYSTEM

- Whenever query keywords raises, the efficiency falls significantly consequently of significant candidate keyword covers generated.
- The inverted index at each node identifies a pseudo report that represents the keywords underneath the node. Thus, to manage to validate in the event a node is strongly related some query keywords, the inverted index is accessed at each node to judge the matching relating to the query keywords and the pseudo-document connected with the node.

5. PROPOSED WORK

5.1 Keyword-NNE ALGORITHM:

In prior function, BKC algorithm falls its effectiveness when query keyword is increases. To solve this problem, here developed a more effective keyword nearest neighbor expansion (Keyword-NNE) which employs the different strategies. For the reason that algorithm, one issue is recognized as a principal query keyword. Those ideas are connected with principal query keyword are thought as principal objects. Keyword-NNE computes local best solution for each principal object. BKC algorithm earnings the lbkc with having highest evaluation. For all the principal objects, its lbkc might be only selects several closest and exceptionally placed objests by the viewer/customer. Weighed against the baseline-algorithm, the keyword covers considerably reduced. These keyword covers yet another strategy in keyword-NNE-algorithm which is perfect, and each keyword selected covers processed and generates very less number of keyword covers

6. GOALS AND OBJECTIVES

The goal is to rank the techniques, therefore we only record here on the binary comparisons that permitted people to ascertain the ordering of the four strategies (excluding unnecessary comparisons).Our recent goals are to allow specific queries, and to rank document effects with the aim of maximizing the coverage of all the in the spatial database, while minimizing redundancy in a prospect of the best keyword search. A keyword cover of keyword that is the phrase connected to that particular keyword, and cover keyword. is known as to be the best keyword for these arch find's important search and ranking, without interrupting the discussion flow, therefore ensuring the functionality of our system. Later on, this will be tried with human people of the system within real life meetings.

7. SYSTEM ARCHITECTURE

This project introduced the generic version of the mCK query, called Best Keyword Cover (BKC) query, which considers inter-objects distance as well as keyword ratings. It is inspired by the observation of the improving availability and very good of keyword rating in decision making. Number of businesses/services/features are the world have be rating by users through online business review sites such as Yelp, City search, ZAGAT and Dianping, etc.

This work can be introduced two BKC query processing algorithms, base-line and keyword-(NNE). The baseline algorithm is a inspire by the mCK query processing technique. Both the base-line algorithm and keyword-(NNE) algorithm are supporting by indexing the objects with an R-tree index, called as KRR*-tree.

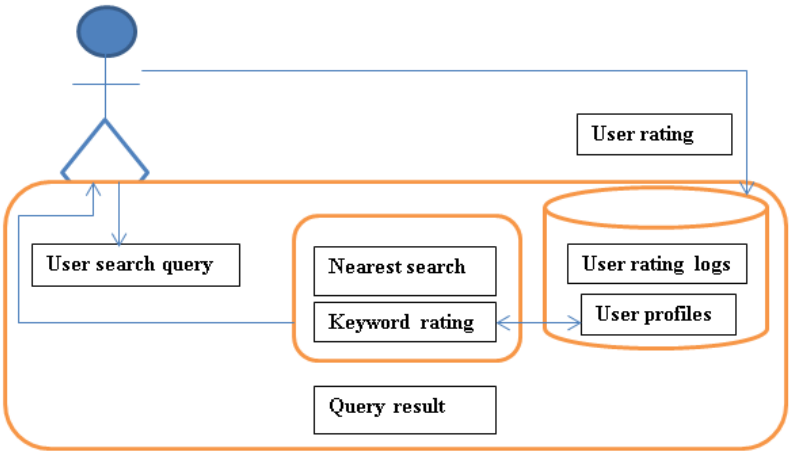

Fig -1: System Architecture

8. RESULT

It ranks the result of user query on the basis of three evidence that is rating, review, rank. So it will help to retrieve exact or relevant search for user query. The Result of System is that optimal results of searching mainly you search a place of any city then nearby All location Such as Temple, Logging, Hospital, Hotel. The main Aim of this system is that Find shortest Distance between Two Location. As keyword increases it is not affect on search of data. The relevant data as per requirement of user can be generate. The list generated which is dynamic in nature which shows most visited object first. So user knows which is best in list

9. CONCLUSIONS

The proposed program in this report provides a flexible approach and a really smart decision making than the existing approach. The bKC query gives the result on the basis of not merely the inter object range but in addition with the keyword rating of that object. The keyword rating of the object is provided by the user on his personal knowledge when using the system. Therefore because the keyword rating is important in decision making this method gives the improved result compared to mCK query provided in present approach. The K-NNE algorithm gives improved approach for the system in that the produced candidate cover set is minimized. The future work with this method is introducing the idea of individualized search. The individualized search is increasing reputation due to its benefits. Therefore the usage of individualized search increases the flexibleness of the system. The future perform is to provide the methods which immediately provide the methods for finding the keyword standing than provided by the user.

10. ACKNOWLEDGEMENT

We owe a great many thanks to a great many people who helped and supported us during our project work. Our deepest thanks to the Guide of the project Prof. R. R. Rathod for guiding and correcting various documents with attention and care. He has taken pain to go through the project and make necessary correction as and when needed. We would like to thank our HOD Prof. M. R. Shimpi for providing us with a platform on which we could conduct research extensively on a topic of our choice.

11. REFERENCES

[1] Ruicheng Zhong, JuFan, Guoliang Li, Kian-Lee Tan, Lizhu Zhou, Location Aware Instant Search. (IJCSIT) Vol. 5 (2) , 2014.

[2] Xin Caoy, Gao Congy,Christian S.Jensenz ,Retrieving Top-k Prestige Based Relevant Spatial Web Objects., IJIRCCE/ijircce.2015.

[3] Zhisheng LI, Ken C.K. LEE, Baihua ZHENG, Wang-Chien LEE, Dik Lun LEE,IR-tree: An Efficient Index for Geographic Document Search,(IJSRD /Vol. 3/2015. Zhisheng LI, Ken C.K. LEE, Baihua ZHENG, Wang-Chien LEE, Dik Lun LEE, IR-tree.

[4] Rakesh Agrawal and Ramakrishnan Srikant.Fast algorithms for mining association rules in large databases. In: VLDB. 1994, pp. 487-499., (IJSRD /Vol. 3/2015.

[5] T.Brinkhoff, H. Kriegel, and B. Seeger. Efficient processing of spatial joins using R-trees. In: SIGMOD (1993), pp. 237-246., (IJSRD /Vol. 3/2015.

[6] 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.

[7] Mehjbeen N. Tamboli, Prof. Suvarna Pawar," Advanced Best Keyword Cover Search" IJIR, Vol. 5, Issue 11, November 2016

[8] 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.

[9] 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

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

[11] Ke Deng, Xin Li, Jiaheng Lu, and Xiaofang Zhou, ‟ Best Keyword Cover Search‟, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 27, NO. 1, JANUARY 2015.

[12] X .cao, G Cong, C Jensen "COLLECTIVE SPATIAL KEYWORD QUERYING‟ IN Proc.ACM SIGMOD Int. Conf manage. Data , 2011 ,pp. 373-384.

[13] J .fan, G Li, L. Zhou, S Chen, and J. Hu Seal: "SPATIO-TEXTUAL SIMILARITY SEARCH ". PVLDB, 5(9): 824-835, 2012.

[14] W Huang ,G .Li,K-L. tan and J Feng. "EFFICIENT SAFE-REGION CONSTRUCTION FOR MOVING TOP-K SPATIAL KEYWORD QUERIES" in CIKM 2012

[15] Xin Caoy, Gao Cong, Christian S. Jensenz, ‟ Retrieving Top-k Prestige Based Relevant Spatial Web Objects‟, Proceedings of the VLDB Endowment, Vol. 3, No.1 , 2010.

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

[17] N. Beckmann, H. Kriegel, R. Schneider, and B. Seeger. The R*-tree: An efficient and robust access method for points and rectangles. In Proc. of ACM Management of Data (SIGMOD), pages 322–331, 1990.

[18] G. Bhalotia, A. Hulgeri, C. Nakhe, S. Chakrabarti, and S. Sudarshan. Keyword searching and browsing in databases using banks. In Proc. of International Conference on Data Engineering (ICDE), pages 431–440, 2002.

[19] X. Cao, L. Chen, G. Cong, C. S. Jensen, Q. Qu, A. Skovsgaard, D. Wu, and M. L. Yiu. Spatial keyword querying. In ER, pages 16–29, 2012.

[20] Y.-Y. Chen, T. Suel, and A. Markowetz. Efficient query processing in geographic web search engines. In Proc. of ACM Management of Data (SIGMOD), pages 277–288, 2006.

[21] D. Felipe, V. Hristidis, and N. Rishe. Keyword search on spatial databases. In Proc. of International Conference on Data Engineering (ICDE), pages 656–665, 2008