International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

COPY RIGHT

Paper Authors

M Manasa, V Harisha Reddy, G Rithwin Reddy, Mrs K Pushpa Rani, M Kavya Sri

USE THIS BAR CODE TO GET THE PAPER

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

Best Keyword Cover Search

M Manasa
B Tech(CSE)
Email : manu.minnie19@gmail.com
Marri Laxman Reddy Institute of Technology(MLRIT)

V Harisha Reddy
B Tech(CSE)
Email :vedaharishareddy95@gmail.com
Marri Laxman Reddy Institute of Technology(MLRIT)

M Kavya Sri
B Tech(CSE)
Email :kavyas563@gmail.com
Marri Laxman Reddy Institute of Technology(MLRIT)

G Rithwin Reddy
B Tech(CSE)
Email : rithwinreddy1106@gmail.com

Mrs K Pushpa Rani
B Tech(CSE)
Assistant Professor
Email : rani536@gmail.com
Marri Laxman Reddy Institute of Technology(MLRIT)

ABSTRACT: Spatial databases are stores the information about the spatial objects which are associated with the keywords to show the information such as its business/services/features. Very important problem known as closest keywords search is to query objects, called keyword cover. In nearest keyword search, it covers a set of query keywords and minimum distance between objects. From last few years, keyword rating increases its availability and importance in object evaluation for the decision making. This is the main reason for developing this new algorithm called Best keyword cover which is considers inter distance as well as the rating provided by the customers through the online business review sites. Closest keyword searchalgorithm combines the objects from various query keywords to a generate candidate keyword cover. Two algorithms Base-line algorithm and keyword nearest-neighborexpansion algorithms are used to finding best keyword cover. The performance of the closest keyword algorithm drops dramatically, when the number of query keyword increases. The solution of this problem of the existing algorithm, this work proposes generic version called keyword nearest neighbor expansion which reduces the resulted candidate key-word covers.

KEYWORDS: Spatial -database, point of the interests, query keywords, keyword ratings, key-word cover

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

I. INTRODUCTION

An improving number of the applications require the efficient execution of nearest neighbour (-NN) queries constrained by the properties of spatial objects. Due to the popularity of keyword search, particularly on the Internet, many of these applications allow the user to provide a list of keywords that spatial objects (hence forth referred to simply as objects) should contain, in their description or other attribute. 1. For E.g, online yellow pages are allowed system user to specify an address and set of keyword, and send back businesses related description contains these keywords, ordered by their actual distance to specified address location. The spatial, keyword query consists of a query area and the group of keywords. The answer is the list of objects ranked according to a combination of their distance to the query area and the relevance of their text description to query keywords. A simple yet' popular variant, which is a used in our runninge. g, is distance first spatial keyword query, where objects are ranked by distance and keywords are applied as a conjunctive filter to eliminate objects that do not contain them. Which is our running example, displaying a dataset of fictitious hotels with their spatial co-ordinates and a set of descriptive attribute (name, amenities)? An example of a spatial keywords query is "find the closest hotels to the point that contain keywords internet and pool". The top result of this query is the hotel object. Unfortunately, there is nothing efficient support for top k spatial keyword queries, where the prefix of the result list has required. Current applications use ad-hoc combination of the nearest neighbour (NN) and keyword search techniques to tackle a problem. For instance, an R-Tree is used to find out the nearest neighbours and for each neighbour an inverted index is using to check if the query keyword is contained. In this project show that such two phase approaches are inefficient. This project develops two BKC query processing algorithms, baseline & keyword NNE. The base-line algorithm is inspired by the mCK query processing techniques Both the base-line algorithm & keyword-NNE algorithm are supported by the indexing the objects with an R*-tree like index, called KRR*-tree. In the baseline algorithm, the idea is to combine nodes in higher hierarchical levels of

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

KRR*-trees to generate candidate keyword covers. Then, the most promising candidate is assessed in priority by combining their child nodes to generate new candidates. Even though BKC query can be effectively resolved, when the number of query keywords increases, the performance drops dramatically as a result of massive candidate keyword covers generated. To come this analytic drawback, we are developed much climbedkeyword nearest-neighbor expansion (keyword-NNE) algorithms which applies the various strategy. Keyword-NNE selects one of query keyword as a chief query keyword. The objects are associated with principal query keyword has principal objects.

II. RELATED WORK

Given a set of query keywords, an essential task of spatial keywords search is to identify spatial object(s) which are associated with keywords relevant to a set of query keywords and have desirable spatial relationships (e.g., close to each other and/or close to a query location). This problem has unique value in various applications because user requirements are often expressed as multiple keywords. For example, a tourist who plans to visit a city may have particular shopping, dining and accommodation needs. It is desirable that all these needs can be satisfied without long distance traveling. Due to the remarkable value in practice, several variants of spatial keyword search problem have been studied. The works aim to find out the number of separate objects, each of which is the near to a query location and the associated keywords (or called document) are very relevant to a set of query keywords (or called query document).Li et al. the spatial relevance's to the query, They focused on the efficiency issue of geographic document search and proposed an efficient indexing structure, namely, IR-tree, along with a top-k document search algorithm. From an extensive experimentation, IR-tree is demonstrated to outperform the state-of-the-art approaches. At present, they are prototyping a geographic search engine with IR-tree as the score and building a testbed based on IR-tree for future research. They also plan to further enhance the IR-tree index based on various access patterns. Cao et al. describes, that, they propose a new type of query, the LkPT query that retrieves the top-k spatial web objects ranked according to both location proximity and

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

so-called prestige-based relevance that considers both the text relevance of an object to a query and the presence of nearby objects that are relevant to the query].

We develop twobaseline algorithms and propose two new algorithms to process the LkPT query. Results of empirical studies on real data demonstrate the effectiveness of LkPT the query and the efficiency of the new algorithms. They propose two algorithms that compute LkPT queries. Empirical studied with real-world spatial data indicate that LkPT querie are effective in retrieving web objects than the previous issues that does not consider the effects of closest objects; and they can show that a proposed algorithms are scalable and outperform baseline issue significantly. Rocha-Junior et al. describes that they present a new index named Spatial Inverted-Index (S2I) and algorithms (SKA and MKA) to support top-k spatial keyword queries effectual. Similar to an inverted index, S2I maps distinct terms to the set of objects that contains the term. The lists of objects that contain a term are stored differently according to the document frequency of the term. If the term occurs often in the collection, the objects with the term are stored in an aggregated R-tree and can be retrieved in decreasing order of partial-score efficiently. Differently, the objects of infrequent term are stored together in a block in a file. Furthermore, we presentalgorithms to process single-keyword (SKA) queries and multiple keyword (MKA) queries effectual. Then,we view through extensive-experiments that our issues outperform the state of the-art approach in terms of query and update cost.

III. PROPOSED ALGORITHM

A. KEYWORD-NNE:

In previous work, BKC algorithm drops its performance when the number of query keywords is increases. To solve this problem, here developed a more efficient keyword nearest neighbour expansion (keyword-NNE) which uses the different strategy. In this algorithm, one query is considered as a principal query keyword. Those objects are associated with principal query keyword are considered as principal objects. Keyword-NNE computes local best solution

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

for each principal object. BKC algorithm returns the lbkc with having highest evaluation. For each of the principal object, its lbkc can be simply selects few closest and highly rated objects by the viewer/customer. Compared with the baseline-algorithm, the keyword covers significantly reduced. These keyword covers a further processe in keyword-NNE-algorithm that will be optimal, and each keyword candidate covers processed generates very less new candidate keyword are covers.

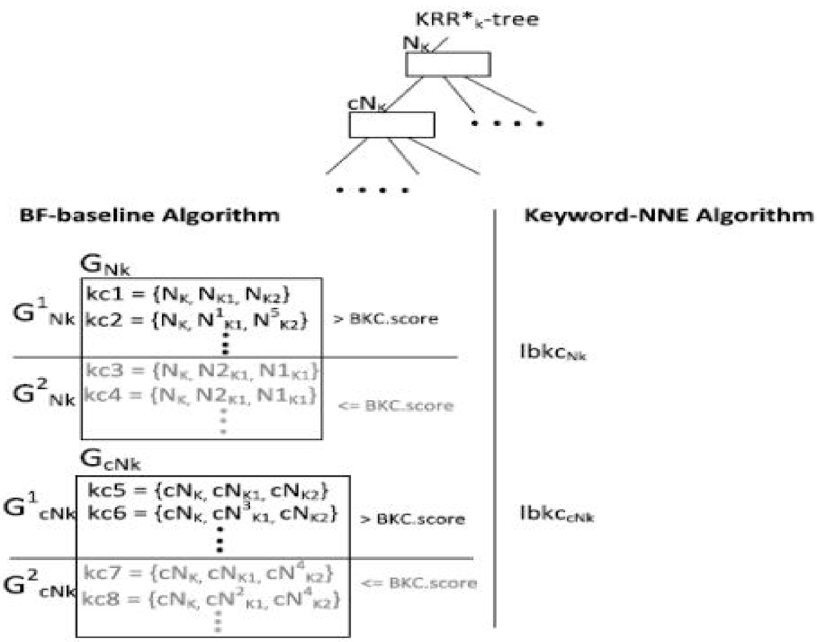

Fig. 1: Baseline& Keyword-(NNE).

B. PRELIMINARY:

In spatial database, each object present in database i.e be associated with either one or multiple keywords. In this object with multiple keywords are directly transformed to multiple objects located at the same location without loss of generality. These objects are in the form of where location of the objects in two dimensional geographical space represented by x and y.

Definition 1 (Diameter): Let O be a set of objects {o1,.......,on}. For oi; oj € O, dist(oi, oj) is the Euclidean distance between oi, oj in the twodimensional geographical space. The diameter of O is Diam(O)=max dist(oi, oj). eq.(1) Each objects has its score with respect to diameter of object and keyword rating of objects in O. Interest of the user may be different in keyword ratings of the objects. Definition 2 (keyword Cover): Let T be a set of keywords {k1, . . . , kn} and O a set of objects {o1, . . . , on} O is a keyword cover of T if one object in O is associated with one and only one keyword in T. Definition 3 (Best Keyword Cover Query): Given a spatial database D and a set of query keywords T, BKC query returns a keyword cover O of T (O subset D) such that O.score ≥ O'.score for any keyword cover O‟ of T (O‟ subset D).In keyword-NNE algorithm, instead of individually processing principal objects are processed in blocks. Suppose k be the principal query keyword. KRR*k-tree used for indexing principal objects. Given principal node Nk in KRR*k-tree, and lbkcNk consider as local keyword cover of Nk, that consists of Nk and other corresponding nodes of Nk in each non-principal query keyword.

IV.SYSTEM ARCHITECTURE

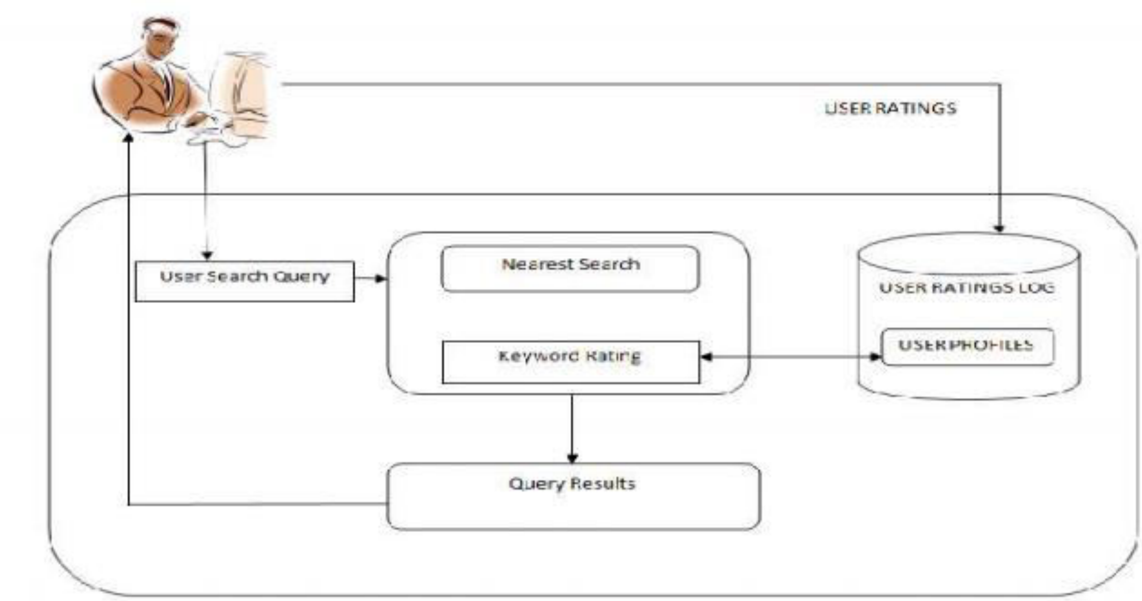

Fig 2. System architecture

- 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, Citysearch, 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.

V. SIMULATION RESULTS

The simulation studies involve the deterministic develops two BKC queryprocessing algorithmsbase-lineand keyword-NNE. The baseline algorithm is inspired by the mCK query processing methods. Both the baseline algorithm and keyword-NNE algorithm are supported by indexing the objects with an R*-tree like index, called KRR*-tree. The developed much scalable keyword nearest neighbor expansion (keyword-NNE) algorithm which applies a different strategy. The following screenshot show the result of project.

1. Home page:

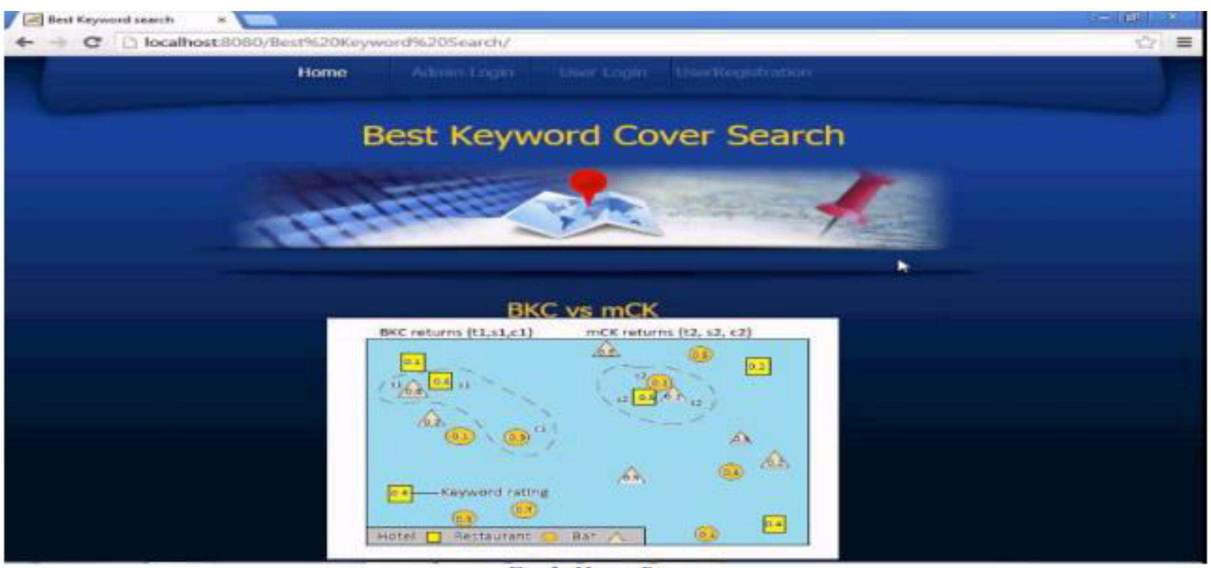

Fig 3. Home Page

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

2. Search result:

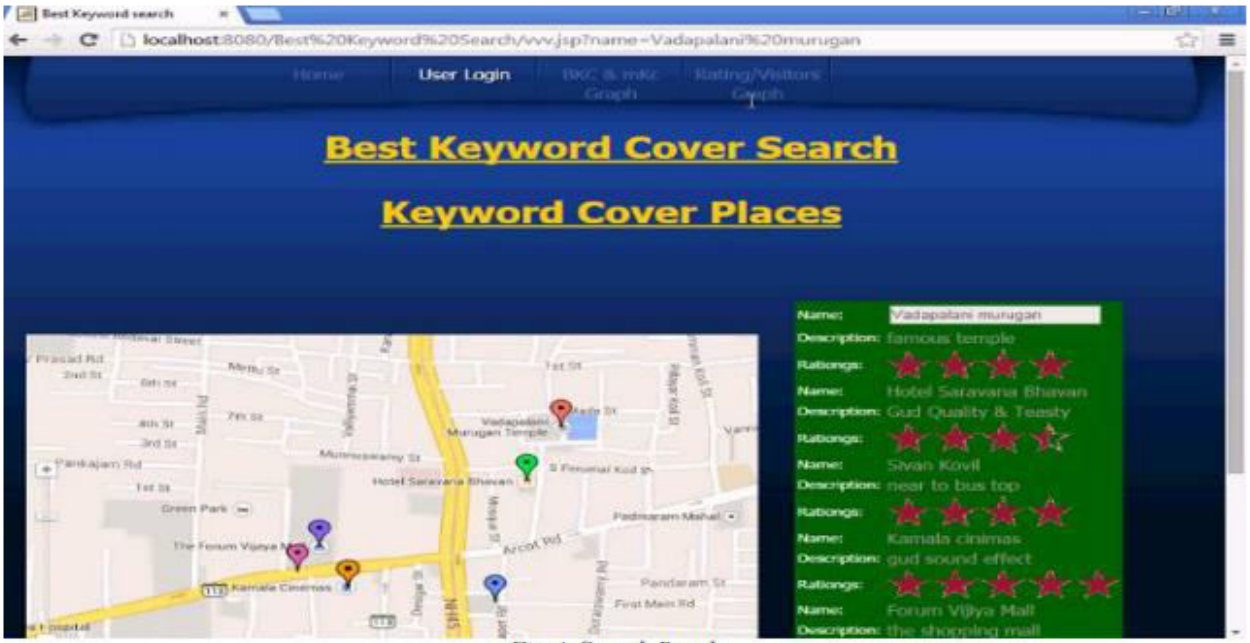

Fig 4. Search Result

3. Keywords Ratings:

Fig 5. Keyword Ratings

4. Keywords NNE Graph:

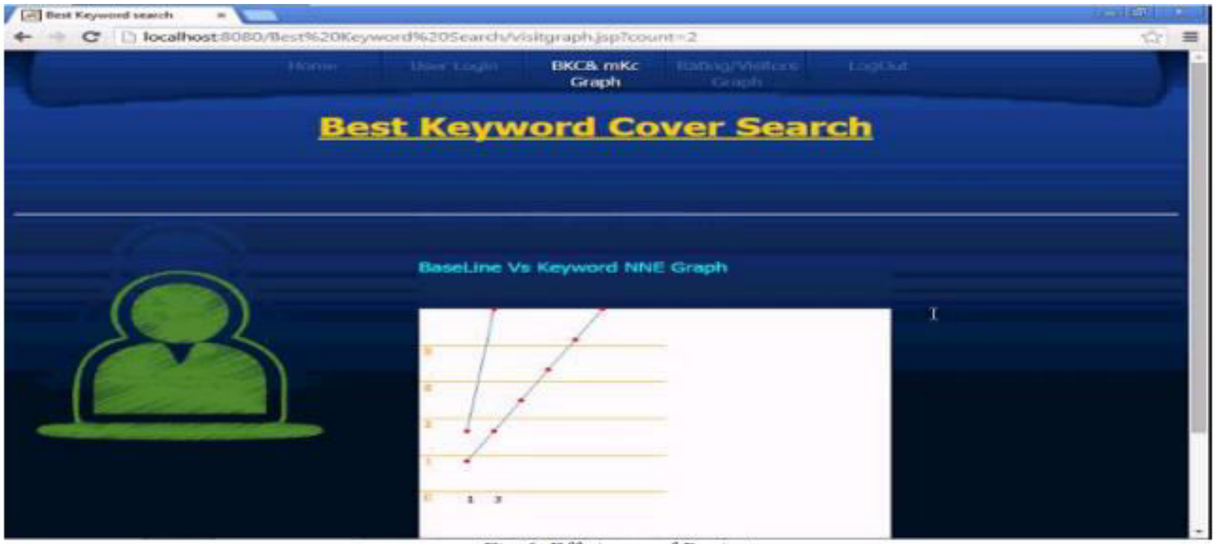

Fig 6. Efficiency of Project

VI.CONCLUSION AND FUTURE WORK

The proposed system in this paper provides a flexible approach and a very sensible decision making than the existing approach. The bKC query provides the result on the basis of not only the inter object distance but also with the keyword rating of that object. The keyword rating of the object is provided by the user on his personal experience while using the system. So as the keyword rating is important in decision making this approach gives the optimized result than the mCK query given in existing approach. The KNN algorithm provides optimized approach for the system in which the generated candidate cover set is minimized. The future work with this system is adding the concept of personalized search. The personalized search is gaining popularity due to its benefits. So the use of personalized search will increase the flexibility of the system. The future work is to provide the methods which automatically provide the methods for detecting the keyword rating than provided by the user.

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

REFERENCES

[1] 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

[2] 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.

[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 Applied Science & Engineering Technology (IJRASET), Volume 3, Special Issue-1, May 2015.\

[7] 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.

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

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

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