



## Keyword NNE for the Fast Keyword Cover Search

Author

**Veena Kumari Crasta**

Dept of Computer Science, Sahyadri College of Engineering, Bantwal, D.K

Email: [vinaroshan@gmail.com](mailto:vinaroshan@gmail.com)

### ABSTRACT

*Mobile computing is very wide popular and so as its applications. Along with it, accessing position based services of the objects have increased the use of spatial object search. Spatial database contains each object as group of keywords associated with it. Keyword rating plays important role in deciding about the objects importance. Considering all these promoted a design that uses inter object distance and service rating of an object called closet keyword search, which improves decision making process. The knne algorithm which can be applied with additional dimension and which can consider more number of keywords in the query is used here. Baseline algorithm can do same task. Baseline is based on grouping concept, objects from various keywords form intermediate candidate keywords. Algorithm shows poor performance as keywords are increased. K\_nne is used which generates less number of candidate keyword covers, which resolves the problem and improves the performance*

**Keywords**—Spatial database, point of interests, keywords, keyword rating, keyword cover

### 1. INTRODUCTION

People are much attracted by search engines and do the day to day work by relying on it. It's obvious that spatial objects locating and accessing their service is wide popular. Men plans for easy world and get the things done faster. There are people who plan to visit a place or city and would like to visit the famous places around. Without wasting much time and without travelling for longer distance it is possible to cover all the famous places around a city by using the application that uses keyword-nne.in spatial database each object is stored along with keywords they associated e.g. services/features. The importance of spatial keyword search is to find spatial objects which are very near to query location and keywords and will have spatial relationships. Spatial keyword searching technique gaining popularity because of digital map app is easily available and also satellite imaginaries.

Two main algorithms can be used to solve BKC search. Baseline algorithm for best keyword covers search and k-nne. Baseline algorithm uses the concept of MCK. These algorithms use R-tree like indexing technique. The baseline algorithm allows to group objects according to the levels they take in the tree. Higher hierarchal level objects group together to have a candidate keyword cover. The child nodes are combined to form new candidate keyword cover. Even though the BKC problem is solvable using this technique it gives poor performance when number of keywords in the query are more. This issue can be handled in K-NNE.

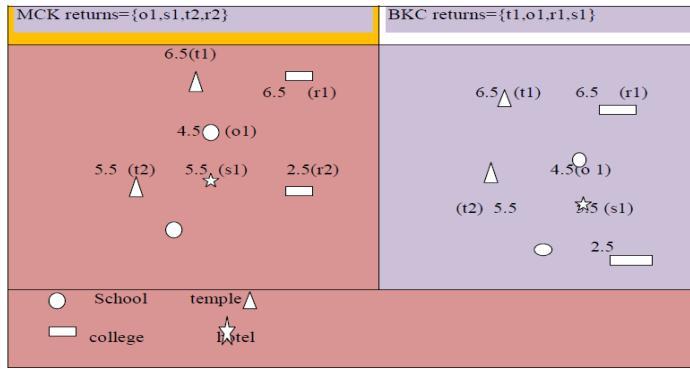


Figure 1 BCK VS MCK

2. LITERATURE SURVEY

X.Cao and others [2], [4], [6], [9] presented that an object is selected based on the keyword in query and its location. As we know that no objects in the database can have all query keywords associated with it, it is difficult to locate an object that is true for every keyword in a query.

D.Zhang, B.Ooi and others discussed in the paper [11] regarding selecting objects that together cover every keyword mentioned in query. The problem here is the objects satisfying this condition should have geometric relationships.

The authors X.Cao and others in the research paper [3] use a method to select spatial objects that satisfy 3 following conditions.

1. Selects an object if it covers every keyword.
2. Select the object with small inter object distance.
3. Select the object with close to query location.

The problem discussed by the authors of [9], [10] will retrieve the objects with low inter object distance and associated with query keywords. What we can observe here is that space is not a constraint for searching.

In the research papers, [4] and [9] the authors G. Cong and et al use an access method called inverted index. This method checks whether a node is equal to a set of query keywords or not. Zannig and Chee proposed an access method called hybrid indexing that is bR\*-tree and bit map is used to index so that m-closest keyword query results in objects that match m keywords. It uses previous knowledge for searching. It reduces space

required for searching and proposes important constraints like distance and keyword mutex. This helps in pruning

3. EXISTING SYSTEM

Keyword cover query solved by two main algorithms known as Baseline algorithm and keyword-NNE. Initially the spatial keyword search uses different technologies. But the baseline algorithm gives best results by considering the location, query keywords and inter object distance. Baseline uses m-ck method. Top down strategy is used to browse the index. The technique uses R\*-tree index to access. We call it as KRR\*-tree. It joins the nodes at top level in hierarchy to obtain a candidate keyword cover. Promising candidate or new candidate can be obtained by combining child nodes; will result in new candidate keyword cover.

Assume that KRR\* tree is constructed for every keyword. Set of query keywords given as  $TK = \{kw1, kw2, \dots, kwn\}$ . Candidate keyword cover is done by combining the child nodes of KRR\*-tree. Let the candidate keyword cover is  $Okc = \{Mk1, Mk2, \dots, Mkn\}$  each  $Mki$  is a node in KRR\*-tree.

$$O.score = score(M, N)$$

$$M = \max(\text{distance}(Mki, Mkj))$$

$$Mki, Mkj \in Okc$$

$$N = \min(Mk.rating)$$

$$Mk \in Okc$$

$Mk.rating$  means maximum value for object under  $Mk$  in keyword rating.  $\text{distance}(Mki, Mkj)$  means minimum Euclidean distance given by  $xd$  and  $yd$  dimensions.

Keyword\_cover query can be easily solvable using the algorithm baseline. But the algorithm is not scalable, means when the given query will have more number of keywords, the processing becomes difficult because it generates large number of candidate keywords.

#### 4. PROPOSED SYSTEM

Even though the baseline will provide solution for the problem, it supports few dimensions to an object. More characteristics of an object in query drop the performance in the existing method. This encourages having the keyword-nearest neighbor algorithm for bkc problem.

Keyword-NNE considers the keywords specified in a query, location, keyword rating and distance between objects, to solve bkc. Considering the keyword rating will help in correct decision making.

##### Outline of Keyword-NNE:-

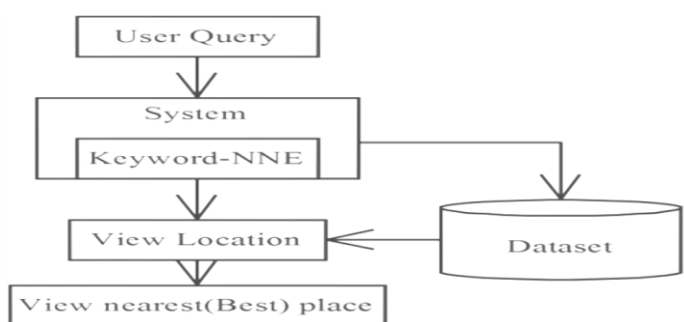
Step 1. Select a keyword from query as the principal\_query\_keyword;

Step 2. Select the objects associated with principal\_query\_keyword as principal\_objects.

Step 3. For every principal\_objects, local\_best\_solution is computed.

Step 4. Identify the Global\_best\_solution from step3;

Step 4. Return the solution as Global\_best\_solution



**Figure 2** Structre of the proposed syyetm

##### Advantages of the proposed method

Computes the local best solution for the principal objects and Later among the local best solution it selects global best solution. So the candidate keywords obtained are less. Thus increases the performances. Keyword rating is also taken into consideration so as to make the decision correctly.

#### 5. RESULTS AND ANALYSIS

In this Technique spatial objects are displayed based on the inter object distance location and the

rating also taken into consideration. So the application of the project is fulfilled as the keyword covers displays the location and also the ratings, which helps the user in better decision making..When the query keywords are many KNN gives better result compared to baseline algorithm. Since the baseline algorithm generates more candidate keyword covers, KNN shows better results.

#### 6. CONCLUSIONS

It is possible in keyword-NNE to add additional dimensions, which gives more description about the object, and also helps in decision making. Whereas Baseline method's performance reduces as the number of dimensions increased (keyword). It generates large number of candidate keyword cover. The proposed technique searches for the local best solution, given the principal object. This reduces the generation of new candidate keyword covers from the existing one. After the a deep analysis it is shown that processing of each candidate keyword cover will generate more number of new candidate keyword cover in baseline algorithm, which is comparatively less in keyword-NNE. Thus keyword -NNE gives best result for bkc query.

#### REFERENCES

1. R. Agrawal and R. Srikant, "Fast algorithms for mining association rules in large databases," in Proc. 20th Int. Conf. Very Large Data Bases, 1994, pp. 487–499.
2. X. Cao, G. Cong, and C. Jensen, "Retrieving top-k prestige-based relevant spatial web objects," Proc. VLDB Endowment, vol. 3, nos.1/2, pp. 373–384, Sep. 2010.
3. X. Cao, G. Cong, C. Jensen, and B. Ooi, "Collective spatial keyword querying," in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2011, pp. 373–384.
4. 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.
5. R. Fagin, A. Lotem, and M. Naor, “Optimal aggregation algorithms for middleware,” J. Comput. Syst. Sci., vol. 66, pp. 614–656,2003.D. Felipe, V. Hristidis, and N. Rische, “Keyword search on spatial databases,” in Proc. IEEE 24th Int. Conf. Data Eng.,2008, pp. 656–665.
  6. I D. Felipe, V. Hristidis, and N. Rische, “Keyword search onspatial databases,” In Proc. IEEE 24th Int. Conf. Data Eng.,2008, pp. 656–665.
  7. G. R. Hjaltason and H. Samet, “Distance browsing in spatial databases,”ACM Trans. Database Syst., vol. 24, no. 2, pp. 256–318, 1999.
  8. 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.
  9. J. M. Ponte and W. B. Croft, “A language modeling approach to information retrieval,” in Proc. 21st Annu. Int. ACM SIGIR Conf.
  10. S. B. Roy and K. Chakrabarti, “Location-aware type ahead search on spatial databases: Semantics and efficiency,” in Proc. ACM SIGMOD Int. Conf. Manage. Data, 2011, pp. 361–372
  11. 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
  12. D. Zhang, Y. Chee, A. Mondal, A. Tung, and M. Kitsuregawa, “Keyword search in the spatial databases: Towards searching by document,” in Proc. IEEE Int. Conf.Data Eng., 2009, pp. 688–69