International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 04, April 2017
www.ijarse.com

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

BEST KEYWORD COVER SEARCH

Balaprasad Bhojane[1], Lakhan Chavan[2] ,Shubham Jadhav[3]

Sanket Kutwal[4]

[1,2,3,4]*Computer Department, RajarshiShahu College of Engineering, (India)*

ABSTRACT

Spatial databases are stores the data about the spatial items which are related with the Keyword to demonstrate the data, for example, its business/administrations/highlights. Important issue known as nearest watchwords pursuit is to question objects, called catchphrase cover. In closest watchword look, it covers an arrangement of question Keyword and least separation between articles. From most recent couple of years, catchphrase rating builds its accessibility and significance in protest assessment for the basic leadership. This is the fundamental purpose behind building up this new calculation called Best catchphrase cover which is considers bury separate and also the rating given by the clients through the online business survey destinations. Nearest catchphrase seek calculation consolidates the articles from different question watchwords to a create hopeful watchword cover. Two calculations k-implies grouping and catchphrase closest neighbor development calculations are accustomed to discovering best watchword cover. K-implies grouping calculations are utilized to discover the similitude of various classes. The execution of the nearest watchword calculation drops drastically, when the quantity of question catchphrase increments. It is common that the objects in a spatial database (e.g., restaurants/hotels) are associated with keyword(s) to indicate their businesses/services/features. An interesting problem known as Closest Keywords search is to query objects, called keyword cover, which together cover a set of query keywords and have the minimum inter-objects distance. In recent years, we observe the increasing availability and importance of keyword rating in object evaluation for the better decision making. This motivates us to investigate a generic version of Closest Keywords search called Best Keyword Cover which considers inter-objects distances well as the keyword rating of objects. The baseline algorithm is inspired by the methods of Closest Keywords search which is based on exhaustively combining objects from different query keywords to generate candidate keyword covers. When the number of query keywords increases, the performance of the baseline algorithm drops dramatically as a result of massive candidate keyword covers generated. To attack this drawback, this work proposes a much more scalable algorithm called keyword nearest neighbor expansion (keyword-NNE).

Keyword- Spatial database, Point of Interests, Keywords, Keyword Rating, Keyword Cover

I. INTRODUCTION

Now a days, utilization of portable processing increments. Enlivened by the portable registering, the spatial watchwords look issue has pulled in much consideration as of late in view of area based administrations and wide accessibility of broad advanced maps and satellite symbolism. So the quantity of clients utilizing the area based administrations has been additionally expanded to vast broaden. Spatial articles demonstrates the data, for example, its Business/administrations/highlights which are related to keyword(s). In spatial database, each tuple

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

speaks to a spatial question. The primary thought behind the spatial catchphrases pursuit is to distinguish spatial object(s) which are related with watchwords pertinent to an arrangement of inquiry catchphrases which are near each other as well as near the question area. This issue has special incentive in different applications in light of the fact that users‟ prerequisites are frequently communicated as various watchwords. In existing, spatial catchphrase seek issue have been examined on account of the estimation of the unique watchword look practically speaking.. In this venture k-implies grouping calculations is utilized to discover the watchword. An increasing number of applications require the efficient execution of nearest neighbor (NN) queries constrained by the properties of the 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 the spatial objects (henceforth referred to simply as objects) should contain, in their description or other attribute. For example, online yellow pages allow users to specify an address and a set of keywords, and return businesses whose description contains these keywords, ordered by their distance to the specified address location. As another example, real estate web sites allow users to search for properties with specific keywords in their description and rank them according to their distance from a specified location. We call such queries spatial keyword queries.

1.1 Proposed System

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

1.2 Proposed system Architecture

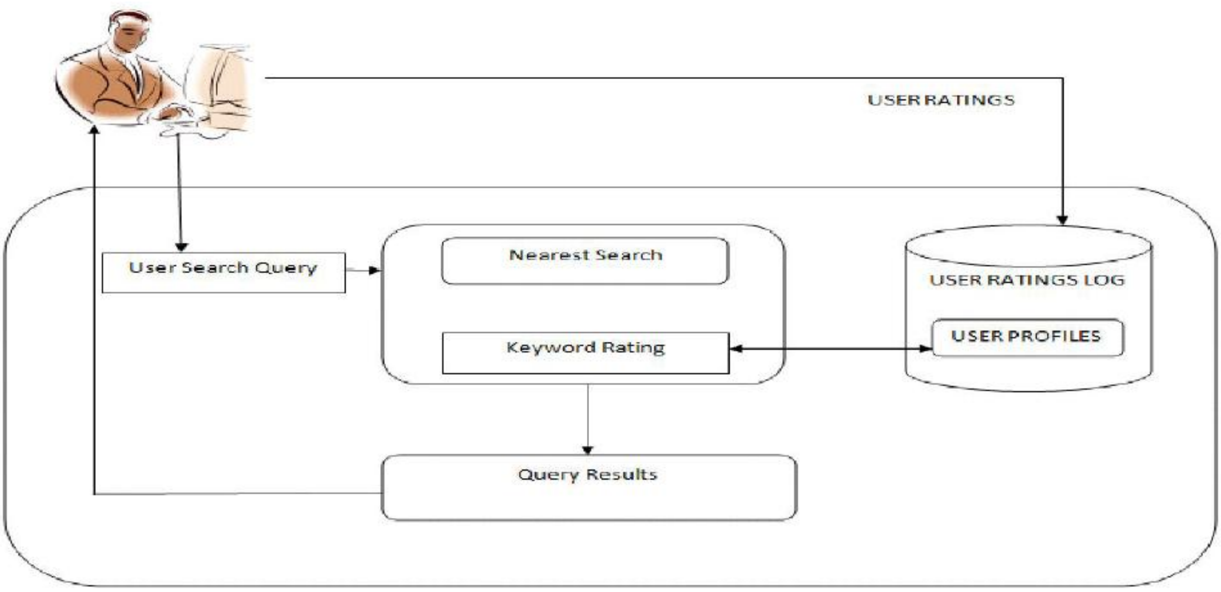

Fig 1. System architecture

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 04, April 2017
www.ijarse.com

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

II. ALGORITHMS

2.1. 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 neighbor 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 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 k-means clustering, the keyword covers significantly reduced. These keyword covers a further processes in keyword-NNE-algorithm that will be optimal, and each keyword candidate covers processed generates very less new candidate keyword are covers.

2.2. K-MEANS

Let X = fx1, x2, x3,..,xn g be the set of data points and

V = fv1, v2, vcg be the set of centers.

1) Randomly select ‘c’ cluster centers.

2) Calculate the distance between each data point and cluster centers.

3) Assign the data point to the cluster canter whose distance from the cluster centre is minimum of all the clustercentres.

4) Recalculate the new cluster canter using:

$Vi = (1/Ci)\ \Sigma cj=1\ Xi$

Where, Ci represents the number of data points in ith cluster.

5) Recalculate the distance between each data point and new obtained cluster centres.

6) If no data point was reassigned then stop, otherwise repeat from step 3).

III. FIGURES AND TABLES

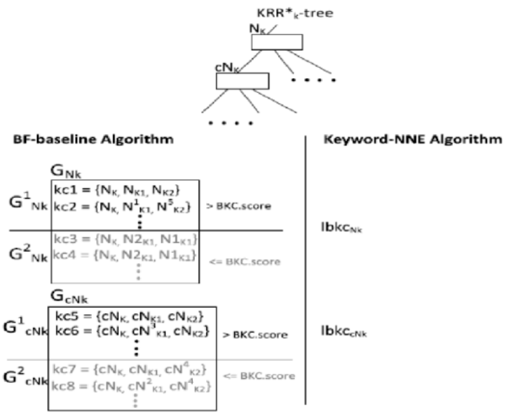

Fig 2: Algorithms

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

IV. CONCLUSION

Compared to the most relevant mCK query, BKC query provides an additional dimension to support more sensible decision making. The introduced baseline algorithm is inspired by the methods for processing mCK query. The baseline algorithm generates a large number of candidate keyword covers which leads to dramatic performance drop when more query keywords are given. The proposed keyword-NNE algorithm applies a different processing strategy, i.e., searching local best solution for each object in a certain query keyword. As a consequence, the number of candidate keyword covers generated is significantly reduced. The analysis reveals that the number of candidate keyword covers which need to be further processed in keyword-NNE algorithm is optimal and processing each keyword candidate cover typically generates much less

new candidate keyword covers in keyword-NNE algorithm than in the baseline algorithm.

V. ACKNOWLEDGEMENTS

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. N.V.Kimmatkar** for guiding and correcting various documents with attention and care. She has taken pain to go through the project and make necessary correction as and when needed.

We would like to thank our HOD **Prof. S. V. Keddar** for providing us with a platform on which we could conduct research extensively on a topic of our choice.

We express our thanks to the Principal **Dr.R.K.Jain**, for extending his support.

We would also thank our College and our faculty members without whom this project would have been a distant reality. We also extend our heartfelt thanks to our family and well-wishers.

REFERENCES

Papers

[1.] 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.

[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]** *Sentiment Crawling: Extremist Content collection through a sentiment analysis guided web-crawler*

[3.] 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

[4.] 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. 361–37.

[5.] .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.