iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** II **Month of publication:** February 2017

DOI: http://doi.org/10.22214/ijraset.2017.2041

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Analysis on Location Based Nearest Keyword Search

Miss Rachana V. Kurhekar[1], Prof. R. R. Shelke[2]
ME (CSE), HVPM COET, Amt

Abstract: It is common that the objects in a spatial database (e.g., restaurants) are associated with keywords to indicate their businesses/services. An exciting problem known as Closest Keywords search is to query objects, called nearest keyword search , which together cover a set of query keywords and have the minimum inter-objects distance. Observation is the increasing availability and rank of keyword rating in object calculation for the better decision making. This inspires us to study a generic version of Closest Keywords search called Best Keyword Cover which considers inter-objects distance and the keyword rating of objects. The baseline algorithm is inspired by the methods of Closest Keywords search which is based on fully combining objects from different query keywords to generate candidate keyword covers. When the number of query keywords increases, the performance of the baseline algorithm falls melodramatically as a result of massive candidate keyword covers generated. To recover this weakness, this work proposes a much more scalable algorithm called keyword nearest neighbor expansion (keyword-NNE). keyword-NNE algorithm meaningfully reduces the number of candidate keyword covers generated. The in-depth analysis and general experiments on real data sets have correct the advantage of our keyword-NNE algorithm.
Keywords — Spatial database, Point of Interests, Keywords, Keyword Rating, Keyword Cover

I. INTRODUCTION

In a spatial database, each tuple represents a spatial object which is associated with keywords to indicate the information such as its businesses. Given a set of query keywords, an essential task of spatial keywords search is to identify spatial objects which are associated with keywords relevant to a set of query keywords, and have desirable spatial relationships (e.g., near to a query locality). This problem has unique value in various applications because users' requirements are expressed as multiple keywords. For example, a tourist who plans to visit a city may have particular shopping and accommodation needs. It is required that all these needs can be fulfilled without long distance traveling. . The works [4][8-10] aim to find a number of individual objects, each of which is close to a query location and the associated keywords (or called document) are very relevant to a set of query keywords (or called query document). The document similarity is applied to measure the relevance between two sets of keywords. An increasing number of applications need the effective execution of nearest neighbour (NN) queries controlled by the properties of the spatial objects. Due to the reputation of keyword search, particularly on the Internet, no of applications allow the user to make available a list of keywords that the spatial objects (henceforth referred to simply as objects) should contain, in their description or other attribute. As one more example, real lands web sites allow users to search for properties with specific keywords in their description and rank them according to their distance from a detailed location. We call such queries spatial keyword queries. A spatial keyword query be made up of a query area and a set of keywords. The response is a list of objects ranked according to a mixture of their distance to the query area and the significance of their text description to the query keywords.

II. LITERATURE REVIEW & RELATED WORK

This issue has remarkable esteem in different applications since clients' prerequisites are often communicated as various keywords For instance, a traveller who arrangements to visit a city may have specific shopping, feasting and convenience needs. It is attractive that every one of these necessities can be achieved without long separation travelling. Because of the amazing quality practically speaking, a few variations of spatial keyword search issue have been examined. The works mean to detect various individual protests, each of which is close a query location and the related keywords (or called document) are very important to a set of query keywords. 1.IR Tree : An efficient index for geographic document search [1] From This Paper we Discussed- In this paper, we propose an actual record, called IR-tree, that composed with a top-k document search algorithm inspires four of note tasks in file searches, to be detailed, 1) spatial filtering, 2) textual filtering, 3) relevance computation, and 4) document ranking in a entirely coordinated mode. What's more, IR-tree permits searches to hold diverse weights on textual and spatial relevance of documents at the runtime and in this way cooks for a wide variety of utilizations. An arrangement of full examinations over an wide variety of situations has been focussed and the experimental comes about show that IR-tree beats the cutting edge line of attack for geographic file searches.

A. Retrieving top-k prestige-based relevant spatial web objects [2] From This Paper we Discussed:

The location-aware keyword query returns ranked objects that are almost a query location and that have printed portrayals that match query keywords. This query occurs certainly in many sorts of useful and conventional web administrations and applications, e.g., Maps administrations. Previous work considers the possible significances of such a query as being independent when ranking them. All the same, a relevant outcome question with adjacent objects that are similarly applicable to the query is likely to be perfect over an significant protest short of important close-by objects. The paper suggests the idea of prestige-based significance to catch both the printed significance of a question a query and the effects of close-by objects. Established on this, additional sort of query, the Location-aware top- k Prestige-based Text recovery (LkPT) query, is not compulsory that recovers the top-k spatial web objects categorized by prestige-based significance and location closeness. We suggest two calculations that process LkPT questions. Exact analyses with open spatial information display that LkPT inquiries are more exciting in recovering web objects than a previous approach that does not consider the effects of adjacent objects; and they prove that the proposed calculations are adjustable and out Performa standard approach necessarily.

B. Efficient retrieval of the top-k most relevant spatial web objects [3] From This Paper we Discussed:

The customary Internet is make safe a geo-spatial dimension. Web information are being geo-labeled, and geo referenced protests, for case in point, purposes of intrigue are being associated with attractive content records. The following grouping of geo-location and reports allows additional kind of top-k query that takes into record both location vicinity and content implication. To our information, just local systems occur that is fit for recording a general web information recovery query while as well taking location into record. This paper put forward another collection framework for location aware top-k content recovery. The framework impacts the disappointed document for content recovery and the R-tree for spatial nearness querying. Rare collation methodologies are studied inside the framework. The framework encloses calculations that use the future records for imagining the top-k query, therefore taking into record both content reputation and location nearness to crop the inquiry space. Significances of experimental analyses with an performance of the framework display that the paper's proposal offers flexibility and is equipped for excellent performance.

C. Keyword search on spatial databases:

In this paper, mostly attention on finding top-k Nearest Neighbors, in this way each node has to match the entire querying keywords. As this way cup tie the entire query to every node, it does not reflect the density of data objects in the spatial space. When no of queries rises then it hints to minor the efficiency and quickness. They present an efficient way to response top-k spatial keyword queries. This work has the next contributions: 1) the problematic of top-k spatial keyword search is defined. 2) The IR2-Tree is projected as an efficient indexing structure to collection spatial and textual data for a set of objects. There are efficient algorithms are used to keep the IR2-tree, that is, insertion and remove objects. 3) An efficient incremental algorithm is existing to response top-k spatial keyword queries by means of the IR2-Tree. Its presentation is projected and likened to the current methods. Actual datasets are used in our trials that display the significant enhancement in performance times.

Disadvantages: -

Every node has to cup tie with querying keyword. So it affects on presentation also it develops time consuming and make the most of searching space. IR2-tree has certain weaknesses.

D. Locating mapped resources in web 2.0," in Proc. IEEE 26th Int. Conf. Data [5] From This Paper we Discussed:

Mapping invention are increasing Web 2.0 applications in which info objects, for example, locations, pictures and soundtracks from many sources are shared and set separately in a guide exploiting APIs that are satisfied by web based mapping schedules, for case in point, Google Maps. These objects are generally linked with an organization of labels holding the put in semantic and an organization of organizes showing their geographic locations. Expected web asset penetrating organizations are not workable in such a domain because of the time off of the gazetteer context in the labels. Rather, a greater option approach is to find a protest by tag organizing. In some situation, the number of labels linked with each question is repeatedly little, creating it difficult for a objection catch the whole semantics in the query objects. In this paper, we focus on the key use of result geographical effects and

Volume 5 Issue II, February 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

suggest a capable tag driven query making procedure. In detail, we malicious to treasure trove an organization of neighbouring co-located objects which composed cup tie the query labels. Assumed the technique that there could be general number of data objects and labels, we build up a creative search control that can scale up as far as the number of objects and labels. Advance, to agreement that the results are valid, we similarly advise a geographical situation delicate geo-tf-idf putting component. Our studies on manufactured data set exhibition its usefulness while the tests using the open data set check its practicality.

E. Processing Spatial-Keyword (SK) Queries in Geographic Information Retrieval (GIR) Systems.[9] From This Paper we Discussed **:**

Location based data stored in GIS database. These data objects of such databases have in cooperation spatial and textual descriptions. This paper suggests a framework for GIR system and attention on indexing policies that can procedure spatial keyword query. The next offerings in this paper: 1) It gives framework for query processing in Geo- graphic Data Recovery (GIR) Systems. 2) Grow a original indexing structure called KR*-tree that detentions the combined distribution of keywords in space and expressively progresses performance over prevailing index structures. 3) This way have shown research on real GIS datasets viewing the effectiveness of our methods paralleled to the existing explanations. It make known to two index structures to collection spatial and textual data.

Separate index for spatial and text features:

Benefits: -

Easy of keeping two separate files.

Performance restricted access lies in the no of candidate object produced during the filtering stage.

Drawbacks: -

If spatial filtering is done first, various objects may untruth within a query is spatial level, but very rare of them are significant to query keywords. This rises the disk right of entry cost by producing a huge no of candidate objects. The following stage of keyword filtering becomes luxurious.

F. Keyword Nearest Neighbor Expansion (Keyword-NNE):

Using the baseline algorithm, BKC query can be effectively resolved. However, it is based on fully methods have been explored; it has been detected that the performance dewdrops dramatically, when the no of query keywords rises, because of the fast rise of candidate keyword covers generated. This motivates us to progress a changed algorithm called keyword nearest neighbor expansion (keyword-NNE). We attention on a certain query keyword, called principal query keyword. The objects related with the principal query keyword are called principal objects.[8]

III. EXISTING SYSTEM

Keyword cover query answered by two main algorithms known as Baseline algorithm and keyword-NNE. At first the spatial keyword search uses altered technologies. But baseline algorithm gives greatest results by considering the location, query keywords and inter object distance. Baseline uses m-ck way. Top down strategy is used to look the index. The method uses R*tree index to right of entry. We call it as KRR*-tree. It joints the nodes at top level in order to obtain a candidate keyword cover. Auspicious candidate or new candidate can be gained by joining child nodes; will outcome in fresh candidate keyword cover. Take up that KRR* tree is created for each keyword. Set of query keyword given as TK={kw1,kw2…kwn}. candidate keyword cover is completed by rake through the child nodes of KRR*-tree.

Let the candidate keyword cover is Okc={Mk1,Mk2…Mkn}every Mki is a node in KRR*-tree. O.score=score (M, N)

M=max (distance (Mki, Mkj))

Mki, Mkj є Okc

N=min(Mk.rating)

Mk є Okc Mk

rating means maximum value for object under Mk in keyword rating. distance (Mki, Mkj) means min Euclidian distanced given by xd and yd dimensions. Keyword_cover query can be simply soluble using the algorithm baseline. But the algorithm is not accessible, means when the given query will have more no of keywords, the processing develops hard because it makes huge no of

Volume 5 Issue II, February 2017
ISSN: 2321-9653

candidate keywords.

IV. ANALYSIS OF PROBLEM

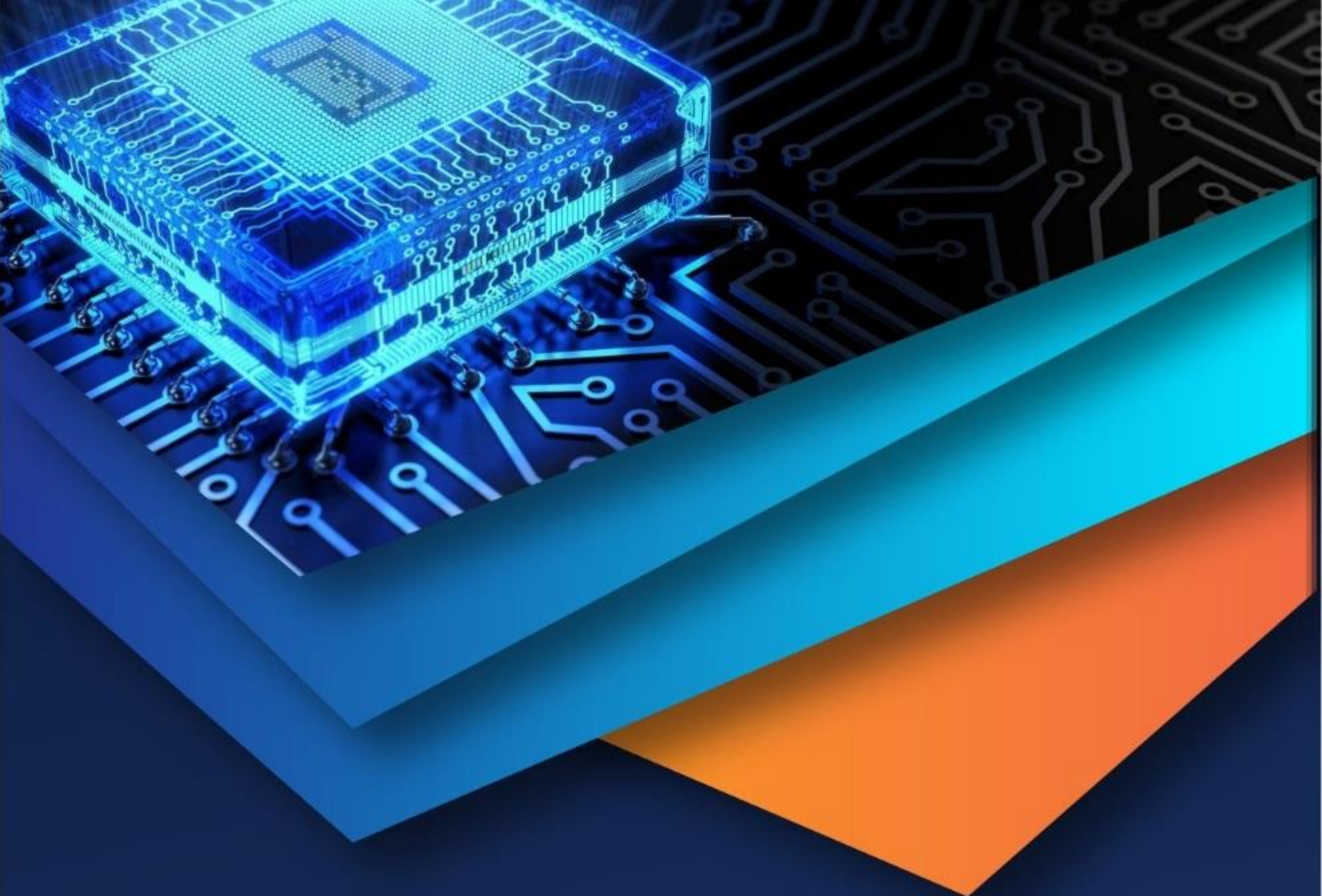
This test shows the impact of the performance. Is an application detailed parameter to equilibrium the weight of keyword rating and the diameter in the score function. Compared to m, the impact of the performance is limited. At what time _ = 1, BKC query is degraded to mKC query where the distance among objects is the sole factor and keyword rating is overlooked. When _ changes from 1 to 0, more weight is taken to keyword rating. An interesting statement is that with the decrease of _ the number of keyword covers generated in both the baseline algorithm and keyword-NNE algorithm shows a constant trend of slight decrease. The reason behind is that KRR*-tree has a keyword rating dimension. Objects close to each other geographically may have very different ratings and thus they are in different nodes of KRR*-tree. If more weight is assigned to keyword ratings, KRR*-tree tends to have more pruning power by distinguishing the objects close to each other but with different keyword ratings. As a outcome, a smaller amount candidate keyword covers are produced.

V. CONCLUSION

Compared to the mCK query, BKC query provides an additional dimension to support more sensible decision making. The introduced baseline algorithm is motivated by the methods for processing mCK query. The baseline algorithm creates a large number of candidate keyword covers which leads to dramatic performance fall 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 ascertain query keyword. As a significance, the no of candidate keyword covers produced is knowingly reduced.

REFERENCES

[1] Li, K. C. Lee, B. Zheng, W.-C. Lee, D. Lee, and X. Wang, “IR Tree: 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] 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: Emantics 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.
[6] Ashlesh S. Patole; Shripadrao Biradar “ A Survey on Best Keyword Cover Search “ IJIRCCE Vol. 3, Issue 11, November 2015 ISSN(Online): 2320-9801 ISSN (Print): 2320-9798.
[7] Ke Deng; Xin Li; Jiaheng Lu; Xiaofang Zhou,” Best Keyword Cover Search” Knowledge and Data Engineering, IEEE Transactions on Year: 2015.
[8] Ian De Felipe, Vagelis Hristidis, and Naphtali Rishe. “Keyword Search on Spatial Databases”. In: ICDE. 2008, pp. 656–665.
[9] Ramaswamy Hariharan et al. “Processing Spatial- Keyword (SK) Queries in Geographic Information Retrieval (GIR) Systems”. In: Proceedings of the 19th International Conference on Scientific and Statistical Database Management. 2007, pp. 16–23.
[10] Jo˜ao B. Rocha-Junior et al. “Efficient processing of top-k spatial keyword queries”. In: Proceedings of the 12th international conference on Advances in spatial and temporal databases. 2011, pp. 205–222.
[11] R. R. Shelke, Dr. R. V. Dharaskar, Dr. V. M. Thakare , “Client Side Mobile User’s Consumptive Behaviour Based On Data Mining Techniques” Proc. of National Conference on Advances in Information and Communication Technology , Madhav Institute of Technology and Science ,Gwalior
[12] R. R. Shelke, Dr. R. V. Dharaskar, Dr. V. M. Thakare, “Analysis of User Movement Behaviour Pattern Mining Approaches for Mobile Environments”, International Journal of Advanced Research in Computer Science, Volume 5, No. 4, April 2014, ISSN No. 0976-5697.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY