

# Optimal Route Search using top-k Spatial Keyword Query

Priti Panmand<sup>1</sup>, Prof. Pradnya Kasture<sup>2</sup>

M. E Student, RMDSSOE Pune, Maharashtra, India<sup>1</sup>

Professor, RMDSSOE Pune, Maharashtra, India<sup>2</sup>

**ABSTRACT:** With advances in geo-arranging progresses and geo-region organizations, there are a rapidly creating measure of spatiotextual things accumulated in various applications, for instance, region based organizations and casual groups, in which a question is portrayed by its spatial range and a course of action of watchwords (terms). Consequently, the examination of spatial catchphrase look which explores both range and abstract depiction of the articles has pulled in amazing thought from the business affiliations and research gatherings. In the paper, the structure focus two essential issues in the spatial watchword request: best k spatial catchphrase look (TOPK-SK), and amass beat k spatial watchword look for (BTOPK-SK). Given a course of action of spatio-printed objects, a request zone and a game plan of question catchphrases, the TOPK-SK recoups the closest k challenges each of which contains all watchwords in the question. BTOPK-SK is the bunch treatment of sets of TOPK-SK request. Additionally, to oversee BTOPK-SK, this system diagram another figuring perspective which fragment the investigation into social affairs in light of both spatial region and the printed congruity between inquiries. This system exhibit that the IL-Quadtree procedure can in like manner profitably reinforce BTOPK-SK. Thorough trials on real and fabricated data unmistakably displays the profitability of our methodologies. Feedback and proposition structure is executed that will be helpful to new customer. As shown by feedback new customer can pick the result decision pointedly.

**KEYWORDS:** Spatial, Keyword, Batch

## I. INTRODUCTION

With the extending inevitability of the geo-arranging technologies and geo-region organizations, there are tremendous measures of spatio-scholarly articles available in various applications. For instance, in the adjacent chase advantage, online expert asset (e.g., business index) gives the region information and furthermore short depictions of the associations (e.g., lodgings, restaurants). In the GPS course system, POI (reason for interest) is a geographically secured pushpin that some person may find important or captivating, which is regularly clarified with surface information (e.g., depictions and customers' reviews). Plus, in various relational association organizations (e.g., Facebook, Flickr), innumerable named photographs are accumulated normal, which can be geo-marked by customers; GPS engaged phones or cameras with an inborn GPS authority (e.g., Panasonic Lumix DMC-TZ10). These exchanged photo outlines are ordinarily associated with different substance marks. Consequently, starting late unique spatial watchword request models and frameworks have ascended with the true objective that customers can satisfactorily mishandle both spatial and printed information of these spatio-artistic articles. In the paper, this system investigate the issue of coordinating top k spatial catchphrase look for (TOPK-SK) [1], [2]; that is, given a course of action of spatio-artistic articles, a question region question and answer set of watchwords, we hope to recoup the k closest challenges each of which contains all watchwords in the request. The top k spatial watchword interest is enter in spatial catchphrase request and has an extensive variety of employments.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

## II. LITERATURE SURVEY

### 1. Inverted Linear Quadtree: Efficient Top K Spatial Keyword Search.

Author: Chengyuan Zhang, Ying Zhang, Wenjie Zhang, Xuemin Lin.

**Description:** With advances in geo-arranging developments and geo-region organizations, there are a rapidly creating measure of spatiotextual things assembled in various applications, for instance, region based organizations and casual associations, in which a question is depicted by its spatial region and a game plan of watchwords (terms). Along these lines, the examination of spatial catchphrase look which explores both range and artistic delineation of the articles has pulled in exceptional thought from the business affiliations and research gatherings. In the paper, structure think two key issues in the spatial catchphrase questions: beat k spatial watchword look for (TOPK-SK), and group best k spatial watchword look (BTOPK-SK). Given a course of action of spatio-scholarly articles, a request zone and a game plan of question catchphrases, the TOPK-SK recoups the closest k challenges each of which contains all watchwords in the request. BTOPK-SK is the gatherings get ready of sets of TOPK-SK request. In perspective of the annoyed document and the direct Quadtree, this system propose a novel rundown structure, called turned around straight quadtree (IL-Quadtree), which is carefully expected to manhandle both spatial and watchword based pruning frameworks to effectively lessen the chase space.

### 2 Hybrid Index Structures for Location-based Web Search

Authors: Y. Zhou, X. Xie, C. Wang, Y. Gong, and W.-Y. Ma.

**Description:** There is progressively business and research excitement for region based web look, i.e. finding web content whose subject is related to a particular place or zone. In this sort of interest, zone information should be recorded and furthermore message information. Nevertheless, the record of routine substance web searcher is set-masterminded, while range information is two-dimensional and in Euclidean space. This brings new research issues on the most ideal approach to capably address the territory properties of site pages and how to merge two sorts of documents. In this paper, we propose to use a blend document structure, which facilitates changed records and R\*-trees, to deal with both printed and zone careful request.

### 3. Keyword Search on Spatial Databases

Author: I. D. Felipe, V. Hristidis, and N. Rische,

**Description:** Various applications require finding objects closest to a predefined range that contains a course of action of catchphrases. For example, online professional resource allows customers to decide an address and a course of action of catchphrases. Thus, the customer procures an once-over of associations whose depiction contains these watchwords, asked for by their division from the predefined address. The issues of nearest neighbor look for on spatial data and catchphrase look on substance data have been extensively considered autonomously. Regardless, to the best of our knowledge there is no powerful system to answer spatial catchphrase request, that is, inquiries that decide both a territory and a game plan of watchwords. In this work, this wander demonstrate a gainful method to answer best k spatial catchphrase request.

### 4. Batch Query Processing for Web Search Engines

Author: S. Ding, J. Attenberg, R. A. Baeza-Yates, and T. Suel,

**Description:** Gigantic web look instruments are at present get ready billions of request each day. A substantial bit of these inquiries are instinctive in nature, requiring a response in parts of a minute. Nevertheless, there are in like manner different basic circumstances where tremendous gatherings of request are submitted for various web mining and structure headway errands that don't require a brisk response. Given the significant cost of executing interest request more than billions of site pages, it is a trademark thing to inquire as to whether such clusters of inquiries can be more gainfully executed than savvy questions. In this paper, the structure goad and discuss the issue of cluster question taking care of in web searchers, perceive basic segments for upgrading the execution of such request, and give a preliminary trial evaluation of the proposed strategies.

### 5. Text vs. Space: Efficient Geo-Search Query Processing

Author: M. Christoforaki, J. He, C. Dimopoulos, A. Markowetz, and T. Suel.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

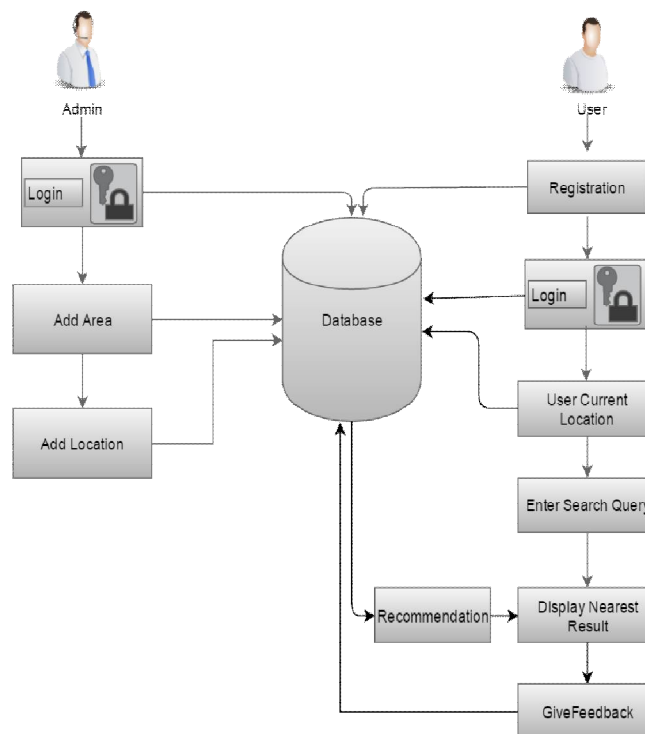
**Description:** Many web look organizations allow customers to oblige content inquiries to a geographic zone (e.g., yoga classes close Santa Monica). Key outlines join neighborhood web crawlers, for instance, Google Local and territory based output organizations for cutting edge cell phones. A couple explore packs have concentrated the capable execution of inquiries mixing substance and geology; their systems when in doubt solidify modified records with a spatial get to strategy, for instance, a R-tree or space-filling twist. Many web look organizations allow customers to force content inquiries to a geographic range (e.g., yoga classes close Santa Monica). Basic delineations join neighborhood web crawlers, for instance, Google Local and territory based search organizations for cutting edge cells. A couple explore groups have concentrated the capable execution of inquiries mixing substance and geography; their strategies usually join changed records with a spatial get to strategy, for instance, a R-tree or space-filling twist.

### III. PROPOSED SYSTEM

Inventor propose a sentiment analysis approach for the feedback sentiment analysis play very important role for product review, feedback, Business intelligence as well as in decision making

- The user will give the feedback on the search query
- Sentiment analysis is the task of identifying positive and negative opinions, emotions.
- sentiment analysis will be done as positive and negativefeedback given by customer by using Sentiment classification algorithm
- Recommendation for nearest location

### IV. SYSTEM ARCHITECTURE



This system propose a novel record structure, particularly IL-Quadtree, to create the spatio-printed objects. A capable figuring is delivered to support the top k spatial catchphrase look by abusing the IL-Quadtree. This wander furthermore propose a portion based procedure to update the practicality of the sign of straight quadtree. To support a great deal of spatial watchword request, this system propose a BTOPK-SK estimation and what's more a request total computation to

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

redesign the execution of the structure. This wander look at how to satisfactorily deal with the cluster beat k spatial watchword request. The system exhibits the troubles and motivations, and Section develops a computation to gainfully portion the request to groups. This system shows a capable computation to support cluster examine for IL-Quadtree in perspective of the aggregate of the imprints from noteworthy watchwords.

1. Go to the Web Application
2. Admin Log in
3. Add the location and the related keyword
4. Get the Geographical location of the that location
5. Store in spatial database with Description
6. If new user then he First Register and then Log in
7. Search the location and keyword on the site
8. Get the Result of that location with nearest places
9. Then the user will give the feedback on the search query.

## V. MATHEMATICAL MODEL

Let S is the Whole System Consist of

$S = \{s, e, X, Y, F_{me}\}$

s=Start of program

1. Login with webpage
2. Search the location.

e=end of program.

1. Display the result.

X= Input of the program

1. Location which user want to search

Y= Output of program

Output will be keyword search.

$F_{me} = \{F1, F2, F3, F4\}$

F1= TOPK-SK (Q, q, k)

Input :LQ : the IL-Quadtree, q : the *spatio-textual* query

k : number of objects returned,

Output :R : TOPK-SK query result.

F2= : BTOPK-SK (Q, k)

Input :k : number of objects returned, Q : the query group,

Output :R : top-k results of each query in Q

F3= Partition algorithm.

Input: Accept there are n protests in the dataset

Output: the points of interest of the Grouping calculation

F4: Grouping Algorithm

Input: z-arrange qualities are initially computed in view of the spatial areas of the inquiries.

Output: Final keyword

Step1: User will enter the inquiry.

Step2: After entering inquiry the accompanying operations will be performed.

Step3: Spatial catchphrase seek which investigates both area and literary depiction of the articles to be hunt.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

Step4: Based on the altered list and the direct quadtree, this propose a novel file structure, called transformed straight quadtree (IL-Quadtree), which is deliberately intended to abuse both spatial and watchword based pruning procedures to successfully lessen the inquiry space.

Step5: A proficient calculation is then created to handle beat k spatial watchword seek. To further upgrade the sifting capacity of the mark of direct quadtree, this propose a parcel based technique. What's more, to manage BTOPK-SK, this venture plan another figuring worldview which segment the inquiries into gatherings in view of both spatial nearness and the printed importance between questions.

Step6: System demonstrate that the IL-Quadtree procedure can likewise proficiently bolster BTOPK-SK.

Step7: Classification.

Step8: Final yield improved classifier and its execution marker.

Algorithms used in this project.

## Algorithm 1: TOPK-SK (Q, q, k)

Input : LQ : the IL-Quadtree, q : the *spatio-textual* query

k : number of objects returned,

Output : R : TOPK-SK query result.

```
1 R =  $\emptyset$ ; H =  $\emptyset$ ;  $\lambda = \infty$  ;
2 for each LQi where  $i \in I(q.T)$  do
3 Push root node of LQi into H ;
4 while H  $\neq \emptyset$  do
5 e  $\leftarrow$  the quadtree node popped from H;
6 if e is a black leaf node then
7 Suppose e is from LQi;
8 for each LQj where  $j \neq i$  and  $j \in I(q.T)$  do
9 CheckSignature(e, LQj) ;
10 if e passed the signature test then
11 Load objects contained by e;
12 for each object o contained by e do
13 ohit := ohit + 1 ;
14 if ohit = 1 then
15 Compute  $\delta(o, q)$  ;
16 if  $\delta(o, q) > \lambda$  then
17 Break;
18 else
19 Update  $\lambda$  and R ;
20 else if e is a non-leaf node then
for each child node e'
21 of e do
if e is not a white leaf node and
22  $\delta_{\min}(e', q) \leq \lambda$  then
Push e' 23 into H;
24 return R
```

## Algorithm2: BTOPK-SK (Q, k)

Input : k : number of objects returned, Q : the query group,

Output : R : top-k results of each query in Q

```
1 R :=  $\emptyset$ ; H =  $\emptyset$ ;  $\lambda_{\max} = \infty$  ; Q.T =  $\emptyset$ ;
```

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

```

2 e ← new a root node;
3 for each subquery qi ∈ Q do
4 Ri ← new max heap;
5 initialize Ri with k null object with distance ∞;
6 let qi.λ be the maximum distance in Ri; qi.λ ← ∞ ;
7 e.Q ← e.Q ∪ qi;
8 for each LQt where t ∈ I(qi.T) do
9 if t ∈ e.T then
10 e.T ← e.T ∪ t ;
11 update e by the root node of LQt;
12 Push e into H, λmax ← ∞ ;
13 while H ≠ ∅ do
14 e ← the node popped from H;
15 if emin ≥ λmax then
16 break;
Q' ← {qi ∈ e.Q | δmin(qi, e) ≤ qi.λ ∧ qi.T ∈ e.T };
if Q' is ∅ then
19 continue the while-loop;
20 for each t ∈ e.T do
if the object list of e'
21 of t is not empty then
split the objects into child entries e'
22 ; for each child entries e'
23 of e do
if |Q'.T ∩ e'.T| ≥ Q.μ and δmin(Q', e') < Q'
24 .λ then
Q* ← {qi ∈ Q' | δmin(qi, e') ≤ qi.λ ∧ qi.T ∈ e'
25 .T }; if Q* is not ∅ then
if e'.level ≥ w'
27 then Node-process(Q*,e',R);
29 else
30 update e' by Q*
31 Push e' into H;
32 return R

```

### Algorithm3:

#### Partition Algorithm.

Accept there are n protests in the dataset, this framework expect to parcel these items into \_ bunches/groups so that the quantity of false hits can be altogether diminished. As it is infeasible to specify every single conceivable segment to accomplish the negligible false hit costs

### Algorithm4:

Grouping Algorithm. In light of the meaning of spatiotextual positioning score, Algorithm 3 outlines the points of interest of the Grouping calculation. The z-arrange qualities are initially computed in view of the spatial areas of the inquiries. Next, in view of this z-arrange values, k focuses are consistently chosen from these questions as the seeding focus purposes of the gatherings. At that point, the rest of the questions will be allocated to the most comparative gatherings in view of their spatio-literary positioning score esteem in the k implies grouping calculation worldview.

### Proposed Algorithm: Sentiment Analysis

Algorithm 3 Sentiment Analysis:

\_ Get Terms - Reduce each review to the list of words

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

\_ Filtering - Remove unnecessary words that will not add value for sentiment analysis

- is, among, but, and, it, that

\_ Base Word - Convert all inflections to their root word

- fry, fries, fried -> fry

- going, go, went, goes -> go

- movies, movie -> movie

- MakeFeatures - Use the words thus extracted from a review as features to indicate the positiveness or negativeness of that review

\_ Classifier - Train a classifier to predict positivity

Comment Analyzer

Input: Preprocessed comment

Output: Comment categorized as positive negative or neutral.

overallPol: Polarity of the whole comment.

sentPol: Polarity of the sentence within the comment.

POS: Part of Speech of the word.

BEGIN

1. overallPol = 0

2. For each sentence in comment

3. f

4. sentPol = 0

5. For each word in sentence

6. f

7. If Polarity[word] == Positive

8. f

9. If POS[word] == Adjective

10. sentPol = sentPol + 4

11. Else if POS[word] == Adverb

12. sentPol = sentPol + 3

13. Else if POS[word] == Verb

14. sentPol = sentPol + 2

15. Else:

16. sentPol = sentPol + 1

17. g

18. Else if Polarity[word] == Negative

19. f

20. If POS[word] == Adjective

21. sentPol = sentPol - 4

22. Else if POS[word] == Adverb

23. sentPol = sentPol - 3

24. Else if POS[word] == Verb

25. sentPol = sentPol - 2

26. Else

27. sentPol = sentPol - 1

28. g

29. Else if word is a Negation word

30. sentPol = - sentPol

31. g

32. overallPol = overallPol + sentPol

33. g

34. If overallPol > 0



# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

- 35. Return Positive Comment
- 36. Else if overallPol < 0
- 37. Return Negative Comment
- 38. Else
- 39. Return Neutral Comment

OUTPUT: The yield will be the reaction to the client question.

## VI. RESULTS AND DISCUSSIONS

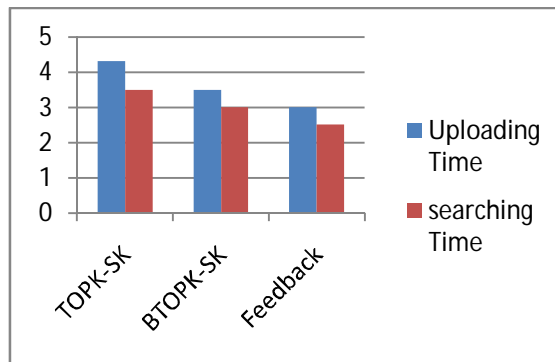


Figure : Graph of Proposed System

| Techniques | Uploading Time | searching Time |
|------------|----------------|----------------|
| TOPK-SK    | 4.3            | 3.5            |
| BTOPK-SK   | 3.5            | 3              |
| Feedback   | 3              | 2.5            |

## VII. CONCLUSION

The issue of top k spatial catchphrase chase is fundamental due to the growing measure of spatio-printed objects accumulated in an extensive variety of uses. In the paper, this wander propose a novel document structure, to be particular IL-Quad tree, to deal with the spation printed objects. A powerful estimation is created to reinforce the top k spatial watchword look for by misusing the IL-Quadtree. This system furthermore propose a fragment based technique to enhance the sufficiency of the characteristic of straight quadtree. To support a great deal of spatial catchphrase request, this structure propose a BTOPK-SK figuring and what's more a request accumulate count to update the execution of the system. The recommendation system will help customer to pick fitting result. Likewise, every customer can give their contribution to that particular result.

## ACKNOWLEDGMENT

It is my privilege to acknowledge with deep sense of gratitude to my guide Prof. Pradnya Kasture for her kind cooperation, valuable suggestions and capable guidance. Also thanks the analysts and also distributors for making their assets accessible. I additionally appreciative to commentator for their significant recommendations.



# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 6, June 2017

## REFERENCES

- [1] Chengyuan Zhang, Ying Zhang, Wenjie Zhang, Xuemin Lin "Inverted Linear Quadtree: Efficient Top K Spatial Keyword Search", 2016
- [2] Y. Zhou, X. Xie, C. Wang, Y. Gong, and W.-Y. Ma, Hybrid index structures for location-based web search, in CIKM, 2005.
- [3] I. D. Felipe, V. Hristidis, and N. Rishe, Keyword search on spatial databases, in ICDE, 2008.
- [4] S. Ding, J. Attenberg, R. A. Baeza-Yates, and T. Suel, Batch query processing for web search engines, in Proceedings of the Forth International Conference on Web Search and Web Data Mining, WSDM 2011, Hong Kong, China, February 9-12, 2011, 2011, pp. 137146. [Online]. Available: <http://doi.acm.org/10.1145/1935826.1935858>
- [5] D. Wu, M. L. Yiu, G. Cong, and C. S. Jensen, Joint top k spatial keyword query processing, TKDE, 2011.
- [6] G. R. Hjaltason and H. Samet, Distance browsing in spatial databases. TODS, vol. 24, no. 2, pp. 265318, 1999.
- [7] M. Christoforaki, J. He, C. Dimopoulos, A. Markowetz, and T. Suel, Text vs. space: efficient geo-search query processing, in CIKM, 2011.
- [8] J. B. Rocha-Junior, O. Gkorgkas, S. Jonassen, and K. Nrvag, Efficient processing of top-k spatial keyword queries, in SSTD, 2011.
- [9] R. Hariharan, B. Hore, C. Li, and S. Mehrotra, Processing spatial- keyword (sk) queries in geographic information retrieval (gir) systems, in SSDBM, 2007.
- [10] G. Cong, C. S. Jensen, and D. Wu, Efficient retrieval of the top-k most relevant spatial web objects, PVLDB, vol. 2, no. 1, 2009.
- [11] D. Wu, G. Cong, and C. S. Jensen, A framework for efficient spatial web object retrieval, VLDB J., 2012.