



## THE COLLECTIVE SPATIAL KEYWORD QUERY PROCESSING BASED ON USER PREFERENCES

C.Karpagam

Department of Computer Science  
Dr. N. G. P. Arts and Science College  
Coimbatore, India

S. R. Kalaiselvi

Department of Computer Science  
Dr. N. G. P. Arts and Science College  
Coimbatore, India

**Abstract:** Geospatial data is predominantly used in many applications to provide location based services on user queries. This advancement focus on collective spatial keyword query processing (CSKQP) that takes both location and textual description of content. In further to improvise we affix a third component, i.e., user preferences to the generated result set. Existing methods in CSKQP takes user location and user specified keywords to produce effective results. In this paper, we explore a novel approach that incorporates user preferences with user keywords in order to yield accurate results. Based on this model the query in turn returns the user favorite choice has result. We initially study the essential requirements to implement and later focus on a generative algorithm.

**Keywords:** spatial keyword query, user preferences, user priorities, location based services.

### I. INTRODUCTION

Increasing demand of spatial data heads us to develop life changing services such as pre-planned touring, live updates, user alert, and many high-end responsive applications. Mobile and social media gains much attention among users for minute to minute updates. Even service providers become more competitive and researchers focus on improving the methods.

Recently, Tata Teleservices confirmed that it finds huge demand for Location Based Services in India. Tata Teleservices is betting big on the strength of its network across India to carve out its own unique space in the enterprise arena. As Indian mobile market is splurging to make tremendous growth and improvement in retail industry, location-based marketing enables retailers to better influence their customers and engage them.

Users of the web often submit geographical enquiries requesting information about, for example, services relating to retailing, tourist attractions, accommodation, sport, entertainment, transport, public services and cultural heritage [2].

Currently, there are many innovative and advanced methods in SKQP to answer user queries. Collective Spatial Keyword Processing (CSKP) aims to find a group of objects that cover the keywords in a query such that the sum of their spatial distances is achieved at a minimum cost [10]. A spatial keyword query consists of a query area and a set of keywords. The answer is a list of objects ranked according to a combination of their distance to the query area and the relevance of their text description to the query keywords [3]. However, the result is far from meeting user expectations due to lack of ability in understanding user needs. In this paper we focus on the this reason to develop a new method.

In the existing methods constantly we produce the same set of result to all the users. But users exist with different characteristics and preferences. The requirements and liking are different from one to another. For example consider users interest in purchasing books. Each one of them has

their own choice of list. So, when people are with different taste and interest we cannot provide the same set of result to everyone. Thus analysis of user preference plays a major role in query results.

Another example is to find a restaurant according to user inclination. If the user likes to explore new things, we can enlist the result focusing on different cuisine that is unfamiliar and brand new to the user. Some users will be habitual, where they do not prefer many choices like conventional ones. To provide them an optimum result, obviously we should focus on their individual needs. Thus we combine both spatial query processing and user preference technique which would outperform all the existing approaches.

Existing approach formulate user query and determine point of interest based on user location and user specified keywords as shown in Fig. 1. Proposed method is to incorporate user favorite and user first choice while returning the result as depicted in Fig. 2. Therefore this paper, propose a novel method to enhance the result and there is no existing method which combines both the approaches. This development gives prominence to *spatial keyword queries* [5, 6, 8, 10]. A typical such query takes a location and a set of keywords as arguments and returns the single spatial web object that best matches these arguments.

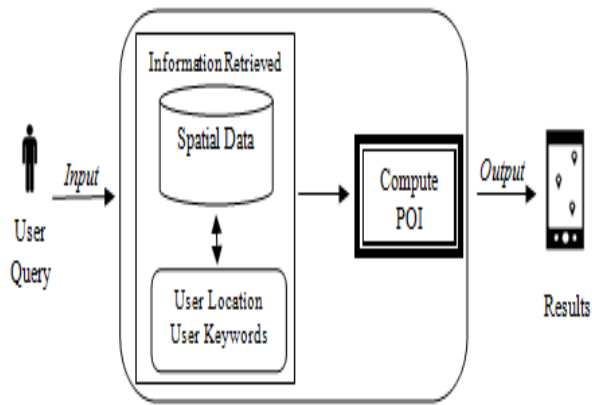


Figure 1. Existing Approach

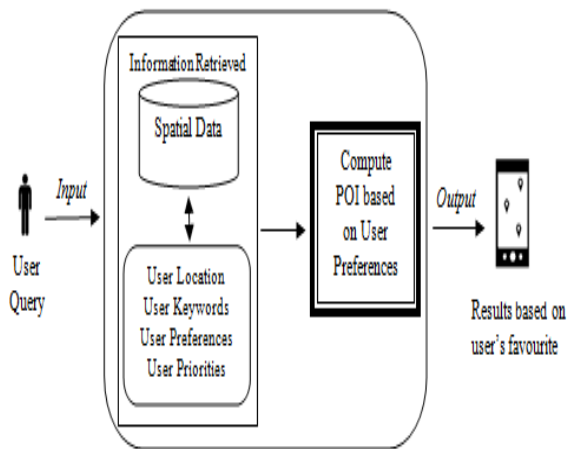


Figure 2. New Approach

By considering the above scenario we propose a novel method in spatial keyword query processing based on user preferences. The input data contains spatial information, current user location, user keywords and their preferences. Through which point of interest (POI) is obtained to the respective user. The following points are highlighted in the paper,

- Requirements to analyze user preference and priorities
- Appropriate method to implement this approach.

Internet privacy involves the right of personal privacy concerning the storing, repurposing, provision to third parties, and displaying of information pertaining to oneself via of the Internet. To conduct experimental analysis, reviewing user history and user's check-in data should also be considered and properly authorized due to privacy issues.

## II. RELATED WORK

### A. Spatial Keyword Queries

Spatial keyword query processing is progressing deep in research with bunch of notable advancements. Spatial indexing of web documents in combination with text indexing is proposed by Subodh Vaid et al. They propose methods to merge text index with results of spatial index in documents. Comparison of them using a pure text index in

conjunction with a separate spatial index of documents and with a pure text index by itself is also presented [2]. An efficient incremental algorithm was presented that uses the IR2 -Tree to answer spatial keyword queries, proposed by I.D.Felipe et al [3]. M-closest keywords (mCK) query addressed by D.Zhang et.al, which is an extension of R-Tree [4]. The efficient processing of continuously moving top-k spatial keyword (MkSK) queries over spatial keyword data propose two algorithms for computing safe zones that guarantee correct results at any time and that aim to optimize the computation on the server as well as the communication between the server and the client [5].

The collective spatial keyword query processing aims to find a set of objects that collectively covers user queried keywords with minimum cost. The problem of collective spatial keyword queries on road networks explored by Yunjun Gao et al. [6], propose algorithms with provable approximation bounds and exact algorithm for efficient processing. It aims to retrieve objects being close to query location and meanwhile close to each other. A simple and popular variant, which is used as a running example, is the 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 [3].

### B. User Preferences

Research shows that travelers make better travel decisions when they are well informed. N. Lathia et al. [7] proposes personalizing travel information to learn implicit preferences and user interaction paradigms can be tuned based on environmental conditions and user habits. J. Feng and Y. Liu [8] address the influencing factors of user background such as cognitive abilities, personality differences and psychological differences that varies from one user to another. Oliver et al. [12] proposed individual travel information for commuters to optimize their daily trip to work. Jie Bao et al. [1] proposed a location-based and preference-aware recommender system that provides location recommendations based on the user's personal preferences learnt from her location history and social opinions mined from the local experts who could share similar interests. Adaptive user interface model states that, people often have different action according to the environment, cognition and personality differences [9]. As the development exceeds in numerous way to obtain precise results, still we focus on improving the user performance. The content providers supply the end user with geo-specific information, a positioning component. Research improves this technique so that users issue the query and it returns with accurate results.

## III. ANALYZING USER PREFERENCES AND PRIORITIES

User preferences are pieces of information that you store persistently to determine their interest. One possible way to perceive user information is to question every detail, but finding answer is not feasible due to frustration of user in answering n number of queries. And also this kind of analysis is not likely to achieve because huge amount of data need to be categorized on every object. So, we propose to determine the result through history of user visits to a particular location. If a user visits a particular destination or

search specific places frequently then obviously it is understood that user interest lies over there. This approach is also applied in Google to infer user interest and demographic categories by the type of pages that user is visiting.

Based on the history of user visits, we rank the objects. When the user enters the same query again, the determined result gives the object with highest ranking. At the end it returns the result as user favorite choice. The main notations used in this paper are enlisted in Table I. A particular user search and visit to restaurants is shown in Table II. It contains name of the restaurant, type of food preferred, frequency of visits, and rank of objects. The table detail clearly depicts the personal preferences of the user, non vegetarian lover, mostly prefer fast food, and that too from different sellers. Consider the user issue a query to find restaurant again. Now the system process to find a subway restaurant in nearby user location hence it has the highest ranking. Frequency of visit ( $f$ ) is calculated using Equation

1, through which rank ( $r_i$ ) of each object is determined. User priorities were recognized through order of keyword entered by the user  $p_i$ . Now we determine the objects based on  $p_i$  and  $r_i$  values respectively.

$$f = \left(\frac{n}{T}\right) \times 100 \tag{1}$$

Table I. Symbols and Description

Notation	Description
f	Frequency of Visit
n	Number of Visits
T	Total number of entries
q	Query
$r_i$	The rank of i-th object
$p_i$	Priority of i-th object
l	User current location

Table II. Restaurants Visited

Name	Type		Number of Visits (n)	Frequency of Visit (f)	Rank
	Type1	Type3			
McDonald's	NonVeg	Fast Food	6	25%	2
Subway	NonVeg	Fast Food	9	38%	1
Pizza Hut	NonVeg	Fast Food	5	21%	3
Dunkin Donuts	NonVeg	Fast Food	3	13%	4
Persian Darbar	NonVeg	Persian Cuisine	1	4%	5

#### IV. APPROPRIATE METHOD

This section presents the generative algorithm that includes the concepts of SKQP and user preferences. Given a set of objects in a road network, we aim to match the user queried keyword with the objects according to the individual preference of user. Initially, the process has been divided in to two phases. Phase I is to determine user preferences and priorities with the user supplied query. Phase II will obtain the input from Phase I and the resultant will be point of interest which focuses on user’s favorite.

##### A. Phase I Information Retrieval

**Input:** user query  $q$

**Output:** User location, keywords, preferences and priorities.

- *Step 1 - Determine geospatial information of the user.* Given a user query  $q$ , determine the geo-location  $l$  of the user.
- *Step 2 - Determine the keywords from user query  $q$ .* Partition the user query in to set of keywords, consider them as objects to determine user preferences.

- *Step 3 - For each keyword in the  $q$ .* Prioritize values as  $p_1, p_2, \dots, p_n$  as per the user entered order.
- *Step 4 - For each  $p_1, p_2, \dots, p_n$  analyze the history of visits.* Determine user preference by ranking the set of objects as  $r_1, r_2, \dots, r_n$ . Frequency of user visits  $f$  obtained from Equation (1) is used in deriving the preferable choice of users. Finally, by utilizing information from a user’s location history, the system aims to predict user interests.

The aim of the above procedure is to obtain user’s favorite choice based on the previous history of visits that assist in increasing the relevance to users’ choice. It takes user query as input and generates four kinds of measures for further analysis. They are

- User current location  $l$
- Keywords are retrieved based on user query
- User priorities are determined by order of keywords.
- User preferences by history of visits

The generated output from Phase I act as an input to Phase II with set of parameter values. Phase I act as a preliminary stage to retrieve fundamental key values, which allows us to perform complex computation in the next phase.

### B. Phase II: To determine user's favourite results

**Input:** User location, keywords, preferences and priorities

**Output:** Result set based on user preferences

- *Step 1* – Consider user priorities  $p_1, p_2, \dots, p_n$  and user preferences  $r_1, r_2, \dots, r_n$  as inputs. Correspond each  $p_i$  with respective  $r_i$  value for each object.
- *Step 2* – Compute POI for objects in a given dataset using an iterative approach. Derive a table of values from the route map where the user interest might lie on.
- *Step 3* – Final result is consolidated to obtain user first and best choice as result.

The above procedure rank user's favorite choice and enlist them as the final result set. The proposed method is a concise approach to solve the stated problem.

## V. COMPARATIVE STUDY

Many papers show how to resolve location information using shortest path techniques at minimum cost. Countless advancement was made to retrieve accurate results at a minimum span of time and complexity. To project the difference between the others, our model objective is to present user favorite results. By comparing accurate result with user favorite one, the user prefers the latter.

## VI. CONCLUSION

Therefore in this paper we focus on how to enhance mobile location based services based on user preferences to meet user needs. The system aims in predicting the user choice that he likes to visit by analyzing the past history of visits. The user will be satisfied with appropriate results returned by the system. We plan to improve the routing system with exact user choice involving component study of existing methods and samples of route data in different regions, as well as deploy the prototype to solicit feedback from a large set of users.

Further we plan to design a system and experimental results will be compared with existing ones. The main question to be addressed is the accuracy in data results. Since the accuracy of data is dependent on analyzing user preferences in an efficient manner. Some important evaluations are to be made to test the efficiency of algorithms.

## VII. REFERENCES

- [1] J. Bao, Y. Zheng, M.F. Mokbel, "Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data", in Proc. 20<sup>th</sup> International Conference on Advances in Geographic Information Systems, pp. 199-208.
- [2] S. Vaid, C. B. Jones, H. Joho, and M. Sanderson, "Spatio-textual indexing for geographical search on the web," in Proc. SSTD, 2005, pp. 218–235.
- [3] I. D. Felipe, V. Hristidis, and N. Risse, "Keyword search on spatial databases," in Proc. IEEE ICDE, 2008, pp. 656–665.
- [4] D. Zhang, Y. M. Chee, A. Mondal, A. Tung, and M. Kitsuregawa, "Keyword search in spatial databases: Towards searching by document," in Proc. IEEE ICDE, 2009, pp. 688–699.
- [5] D. Wu, M. L. Yiu, C. S. Jensen, and G. Cong, "Efficient continuously moving top-k spatial keyword query processing," in Proc. IEEE ICDE, 2011, pp. 541–552.
- [6] Y. Gao, J. Zhao, B. Zheng, and G. Chen, "Efficient Collective Spatial Keyword Query Processing on Road Networks, IEEE Transactions on Intelligent Transportation System, Vol. 17. No. 2 2016 pp. 469 - 479.
- [7] N. Lathia, L. Capra, D. Magliocchetti, F. D. Vigili, G. Conti, R. D. Amicis, T. Arentze, J. Zhang, D. Cali, V. Alexa, "Personalizing Mobile Travel information Services" Proc Social and Behavioral Sciences, 2012, pp 1195 – 1204
- [8] J. Feng and Y. Liu, "Context-Aware User Interface Framework for Mobile GIS", International Journal of Multimedia and Ubiquitous Engineering, 2014, pp. 253-262
- [9] J. Feng and Y. Liu, "Intelligent Context-Aware and Adaptive Interface for Mobile LBS", Hindawi Publishing Corporation, 2015.
- [10] B. John Oommen, Anis Yazidi and Ole-Christoffer Granmo, "An Adaptive Approach to Learning the Preferences of Users in a Social Network Using Weak Estimators," Journal of Information Processing Systems, vol. 8, no. 2, pp. 191~212, 2012. DOI: 10.3745/JIPS.2012.8.2.191.
- [11] X. Cao, G. Cong, C. S. Jensen, and B. C. Ooi, "Collective spatial keyword querying," in Proc. ACM SIGMOD, 2011, pp. 373–384.
- [12] O. Roider, C. Rudloff, M. Ray, "An Individual Travel Information System for Optimizing Mode and Route Choice behaviour of Commuters", in Proc. Social and Behavioral Sciences 48, 2012, pp. 1948-1957.