ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, *Vol 9(30), DOI: 10.17485/ijst/2016/v9i30/99075, August 2016* ISSN (Online) : 0974-5645

A Novel Approach for Making Recommendation using Skyline Query based on user Location and Preference

Sanket Shah*, Amit Thakkar and Sonal Rami

Department of Information Technology, CSPIT, CHARUSAT, Anand - 388421, Gujarat, India;
14pgit014@charusat.edu.in, amitthakkar.it@charusat.ac.in, sonalrami.it@charusat.ac.in

Abstract

Objectives: To propose a method to handle large number of user and to improve the accuracy and quality of recommendation system. **Methods/Statistical Analysis:** This paper presents an effective method to identify user location based on his/her preference using Skyline query outline Dominated object. Dominance object suggests that an object falls under good or better in all dimension or good at least one dimension. Skyline query using Recommendation system has increased in recent years. Skyline query using recommendation system mainly used location-based services to find the nearest location, based on user preference. Location-based Services are information services and have a number of uses in social networking. Location-based Service finds the nearest location based on user preferences but not provide location based on similarity and rating. So, the user is not satisfied by the given result. **Findings:** To resolve above problem, the collaborative filtering technique, K-nearest neighbor algorithm and Ranking Scheme being used by us. Using Collaborative filtering technique, we find the similarity and rating of an item. Using K-nearest neighbor approach finds the nearest distance of the similar item and ranking technique being used by us, to choose the most nearest location. In this paper we take temporary dataset and mathematically evaluate our proposed system. **Application/Improvements:** In future, we will develop web tool which identify location and display result on map. We will also check user s' past movement history based on content based recommendation system. Skyline query using recommendation system is use various domain i.e. House Rent/buying, travel and tourism business.

Keywords: Collaborative Filtering Technique, Dominated Object, K-Nearest Neighbor, Recommendation System, Skyline Query

1. Introduction

Data mining[1,2] is outlined as "The nontrivial extraction of implicit, previously unknown, and doubtless helpful data from knowledge" . Data mining typically referred to as knowledge or data discovery that aims whatever knowledge is obtainable that knowledge finds some conclusions within the style of rules. Data mining are often think about as economical and efficient way to discover or to remodel the invisible to visible knowledge.

Now a day, Skyline query and Recommendation System[3] are trendier in database and data mining field. Skyline query is sorting out interesting point based on dominance relationship. It's outlined dominated object. Dominated object means those objects which are good in all dimensions and it is also good in at least one dimension[4-6]. Recommendation Systems are now pervasive in consumers' lives. They aim to assist users to find things that they might wish to purchase or think about supported vast amounts of information collected. Parsing a large quantity of information to predict a user's choice or his or her similarity with an alternative cluster of users is that the core of a recommendation system. Based on this Characteristics skyline query and Recommendation are more fashionable in Multi-Criteria decision-making System.

Skyline query and Recommendation[7,8] for the user is an important sub-topic especially Location based service. Location based mostly services use in real time Geo-Spatial knowledge to produce data, diversion or security.

* *Author for correspondence*

Some services allow customers to "check in" at restaurants, occasional retailers, stores, concerts and various places or events. Location primarily based services use a Smartphone's GPS technology to trace person's location if that person has opted-in to allow to service to do that. After a Smartphone user opt-in, the service can identify his or her location without the require manual entry. Most of existing work on Location based skyline query[7,8], find the nearest location based on user preferences but using this technology not provide location based similarity and rating. So, the user is not satisfying the given result. To solve this Problem, we combine three techniques; First technique is collaborative filtering technique. Using this technique finding the similarity and the rating of an item. The second technique is K-nearest neighbor approach. Using this Approach find the nearest distance of the similar item. The third technique is ranking scheme. Using this Technique find the highest ranked item. So, using these three techniques we recommend the most appropriate user's location based user's preference.

The Rest of the paper is organized as follows. Section 2 describes however skyline query and recommendation system work to assemble using query processing algorithm and find from survey. In this Section 3, we explain system framework. In this section 4, we explain mathematical evaluation of our system. In this section 5 describe conclusion and future work of our proposed system.

1.1 Motivating Example

The fusion of geo-location and preferences makes potential a replacement Skyline queries that takes under consideration each location proximity and user s' preferences. A location-based Skyline[9] is associate extension of Skyline that depends on geographical locations of interest objects. Skyline is more useful to recommend the user location based on his/her preferences. Suppose, we consider one area, user finding out more interesting hotel near to the beach, conference and airport shown in figure1. In this area four hotels are there. Hotel H1, H2, H3, H4.first user map the distance between hotel H2 to beach, airport and conference and also map the distance between Hotel H1 to beach(B), airport(A) and conference(C). Based on mapping we conclude that hotel H1's distance less compare to hotel H2 's distance. So we can say that hotel H1 is dominate hotel H2.next map the distance between hotel H3 to the beach, airport and conference and conclude that choose any hotel either H1 or H3 any two place is near and one place s' distance is large. Similar way map the gap in hotel h4 and supported all mapping realize three most interesting hotel and it's contemplate in skyline. Hotel H1, H3, H4 is contemplate as skyline point.

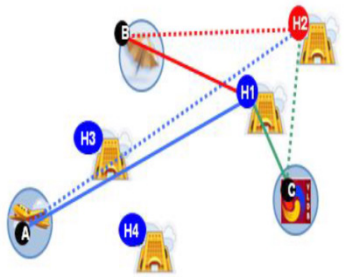

Figure 1. Motivation Example.

2. Related Work

Skyline query[10] is defined those point which are not dominated by any other point. A point is dominated by another point if it is good or better in all dimensions and better in at least one dimension. A query which is used to find the skyline point is called skyline query[11]. Let us consider two object Op and Oq. If one object Op is equal to or better other object Oq in all dimension and if Op is better than Oq at least one dimension, so we can say that Op is dominated Oq. For any comparison operator denoted as > and least one dimension is equal is denoted as ≥. Mathematically describe as below.

$$Op > Oq \Rightarrow \forall i, Opi \geq Oqi \text{ and } \exists j, Opi > Opj$$

Number of researchers design various algorithm to find out dominated object. In this section we represent existing query processing algorithm[12-14]. proposed block nested loop algorithm in 2001.In the Block Nested Loop Algorithm maintain window of an object not yet dominated. In the block Nested Loop algorithm check every new object against the list following three conditions. First condition is if an object is not dominated so prune it. Second condition is if an object not dominated some point in window, so prune this object. If neither, add object in to window. If the window is full, so write an object in to temporary file on disk and note the time stamp. In the next iteration, read the tuples from temporary file. Finally window contains skyline point. Jan Chomicki et al. was proposed sort filter skyline algorithm in 2003.In the sort

filter skyline algorithm use the order list same as input for block nested loop algorithm. In this algorithm sort all object given by entropy function. Remaining process same as Block Nested Loop Algorithm. Marlene Goncalves et al. proposed Location based skyline algorithm in 2012. LOS gets as input a set of spatial objects O, a preference function m defined on r attributes of the objects belonging to O, a given location l, and an R-Tree structure[15]. We use an R-Tree to retrieve the objects in order of closeness using the nearest neighbor search. In this algorithm first compute the skyline point in terms of the function. In this algorithm check next object if it is not belong to the location based skyline object so it is discarded otherwise add in to skyline bucket. Finally bucket contains the skyline point. Spatial skyline with preference Algorithm was proposed by Kazuki Kodama et al. in 2009. They are utilizing Euclidean Distance and consider closer object as skyline object. We assume dominance relationship for each category attribute can obtained from user s' profile. For economical query process, we have a tendency to use the construct of nearest neighbor queries and value more and more whether or not every object belongs to the skyline. If associate degree object isn't dominated by alternative objects in terms of distance and every one the class attribute, it belongs to the skyline. Multi level spatial Skyline with Preference was planned by in 2009. Spatial skyline with preference algorithmic rule presents delicacies supported the dominance relationship. The algorithmic rule works well, however it's the matter that the ensuing skyline might encompass a tiny low range of objects. This happens once one or two objects are very robust compared to alternative objects, or once the amount of category attributes is tiny. For instance, we have a tendency to solely have three skyline objects within the result. A user who needs to match many candidates wouldn't be happy by such a result. As an answer for this drawback, we have a tendency to propose the thought of multi-level skyline queries. We have a tendency to decision the skyline objects obtained by the algorithmic rule within the previous section the first-level skyline objects. If the quantity of the first-level skyline objects is smaller than k, that may be a range given by the user, we have a tendency to calculate the second-level skyline objects.

2.1 Finding from Survey

After analyzing of all query processing technique we summarized following important point.

- Block nested loop algorithm provide skyline object as an output but not provide in a sorted manner.
- Sort filter skyline algorithm provides skyline object in a sorted manner as an output.SFS overcome BNL algorithm and also reduce searching time.
- Spatial skyline with preference[16] recommends item such as a hotel to the mobile user taking in considers his current location and preferences.
- Spatial Skyline[17] with preferences based mostly query processing is work well in a little number of the object. However, user wish to match many candidates wouldn't be happy with such a result. As an answer for this drawback, authors have projected the concept of multi-level skyline query.
- Location based skyline algorithm find user location based on his/her preferences. If the dataset contains non-duplicated data and has a large number of attributes in the Skyline criteria, LOS performs better than SFS. But if the dataset is duplicated LOS performance becomes poorer compare to SFS.

After analysis of all techniques, we realize, location is based on user preference but not find the location based on similarity and ranking wise so we will try to propose new technique to find location based user preference and also check similarity and suggest most nearest location. To do this we combine three techniques collaborative filtering in recommendation system, k-neighbor method and ranking scheme.

3. System Frame Work

We proposed our model based on user location and his/her preference. The system consist K-nearest Neighbor Algorithm, Item Based Collaborative Filtering technique and Ranking Scheme. We In the Location based skyline algorithm search result based on user preference but in not search result ranking and similarity wise. So we use two approach K-nearest neighbors and Item Based collaborative techniques[18-20]. Using Item based Collaborative Technique find out similar item based on his preference and we also find the rating on item. Using K- nearest neighbor algorithm[21-23] find out item nearby location. Using Ranking Scheme we select most closet item based on rank.

3.1 Description of Proposed System

We design our proposed model in to two phases. In the

first phase we use the Collaborative Filtering Concept and Second phase we use the K- nearest neighbor algorithm. The step of our proposed algorithm described as below.

- User Search an item on particular location and based on preference
- Using Geo- Locate user track the user location.
- Find out the similar item based on his location. In this step find the similarity of item using below given "Eq. (1)".

$$\cos\left(\overrightarrow{t1}\,\overrightarrow{t2}\right) = \frac{\overrightarrow{t1} \cdot \overrightarrow{t2}}{\left\|\overrightarrow{t1}\right\| \left\|\overrightarrow{t2}\right\|} \quad (1)$$

- Find out the most rated similar item. In this step find rating of similar item using below given "Eq. (2)".

$$p(u,i) = \bar{r}_u + \frac{\sum u' \in N\, s(u,u')\left(r_{u,i} - \bar{r}_u'\right)}{\left(\sum[u' \in N\,|s(u,u')|\right)} \quad (2)$$

- Finding the nearest distance on similar item by using below "Eq. (3)".

$$D = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \quad (3)$$

- Sort the distance and determine nearest neighbors.
- Display Result

3.2 Architecture of Proposed System

Using Our Proposed system user can search item to nearest location based on his/her preference. Figure 2 show our proposed model for location based item search.

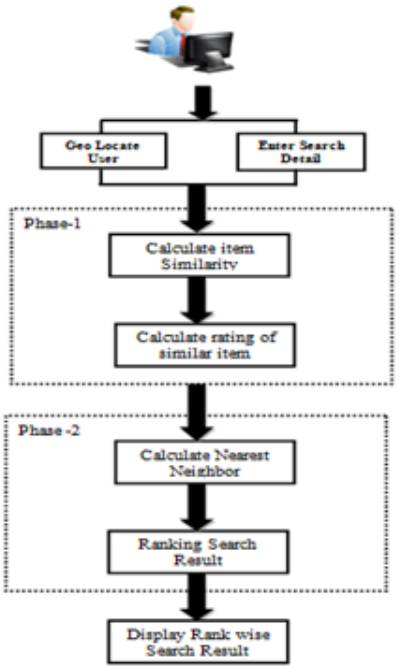

Figure 2. Architecture of Proposed System.

4. Mathematical Evaluation

In this Section we take temporary dataset and evaluate Cosine similarity, rating and evaluated nearest location.

4.1 Measure Cosine Similarity

In this Example predict the rating of User C based On User A and User B Given Rating on this Movie. To do this thing we use cosine similarity equation, using these equations decide the User C similarity of this movie. We also find out the rating. Here, we take temporary dataset shown in Table1.

Table 1. Movie Rating Database

User	Batman Begins	Alice in Wonderland	Dumb and Dumber	Equilibrium
A	4	?	3	5
B	?	5	4	?
C	5	4	2	?
D	2	4	?	3
E	3	4	5	?

- We Measure Cosine similarity between User A and User B are using "Eq. (1)".

  Cos (a, c) = 1.8868

  Cos (d, c) = 2.3255

4.2 Calculate Rating

Consider the ratings matrix in table. We want to find User C's prediction for Equilibrium (pC,e) using "Eq. (2)" with the following configuration.

- Neighborhood size of 2.
- Weighted average with mean offset.

Users C mean rating is 3.667. There are only two users who have rated Equilibrium, and therefore only two candidate users for the neighborhood.

$$= 3.667 + \frac{1.88 \cdot (5 - 4) + 2.32 \cdot (2 - 3)}{0.832 + 0.515} = 4.667$$

4.3 Calculation of Nearest Location

Consider User Position (3, 7). Suppose User Want to see the movie Equilibrium so he search Multiplex nearby his location. The Multiplex Name and Its Location information is retrieved in Dataset. It is shown in Table 2.

Table 2. Multiplex Database

Multiplex Name	Location
Cine Max	(7,7)
Big Cinema	(7,4)
PVR	(3,4)
Cine Pride	(1,4)

Calculate the distance between the query-instance and the entire training sample using "Eq. (3)" and it shown in Table3.

Table 3. Calculate Distance Using "Eq. (3)"

Multiplex Name	Location	Distance
Cine Max	(7,7)	4
Big Cinema	(7,4)	5
PVR	(3,4)	3
Cine Pride	(1,4)	3.60

After Calculate the distance, give the rank in ascending order shown in Table 4.and find the closet distance. After calculate the rank we choose the PVR multiplex because it s' rank is 1.

Table 4. Calculate Rank of Multiplex

Multiplex Name	Location	Distance	Rank
Cine Max	(7,7)	4	3
Big Cinema	(7,4)	5	4
PVR	(3,4)	3	1
Cine Pride	(1,4)	3.60	2

5. Conclusion and Future Work

In this paper we propose a novel approach for making recommendation using skyline query based on user location and preference. The proposal is in the Implementation stage. Using this technique, we will build a web tool. It identifies location based user preferences. In this tool user enter location and their preferences. Our tool track the user position and find the item according to location and display those item which is near to the user location. The filter choice will facilitate effective decision making, where the number of object that are part of the result is greatly reduced in most cases. To build our tool used K- nearest neighbor algorithm and use collaborative filtering technique. In the collaborative filtering technique find the similar item based on user preference and using K-nearest neighbor algorithm find the item which near to user location. After the K- Nearest neighbor algorithm and collaborative filtering technique we use ranking algorithm and find the top most similar result. We also check user past movement history based on content based filtering technique.

6. References

1. Fayyad U, Piatetsky-shapiro G, Smyth P. From data mining to knowledge discovery. American Association for Artificial Intelligence. 1996; 17(3):1–18.
2. Roddick JF, Spiliopoulou M. A bibliography of temporal, spatial and spatio-temporal data mining research. ACM SIGKDD Explorations Newsletter. 1999 Jun; 1(1):34–8.
3. Wang H, Xing C. An approach to online recommendation of products with high price-performance ratios based on a customized price-dominance relationship. J Softw. 2013 Dec; 8(12):1–6.
4. Borzsony S, Kossmann D, Stocker K. The Skyline operator. Proceedings 17th Int Conf Data Eng. 2001. p. 421–30.
5. Chomicki J, Godfrey P, Gryz J, Liang D. Skyline with presorting. Proceedings - Int Conf Data Eng. 2003 Mar. p. 717–9.
6. El Maarry K, Lo C, Balke W. Crowdsourcing for Query Processing on Web Data : A Case Study on the Skyline Operator. Journal of Computing and Information Technology. 2015 Mar; 23(1):43–60.
7. Kodama K. Skyline Queries based on user Location and Preferences for Making Location-based Recommendations Preference-based Query Processing Multi-Level Skyline Query, DB Group Nagoya University, 2009 Nov.
8. Goncalves M, Torres D, Perera G. Making recommendations using location-based skyline queries. Proceedings - Int Work Database Expert Syst Appl DEXA. 2012; 111–5.
9. Li Y, Qu W, Li Z, Xu Y, Ji C, Wu J. Skyline query based on user preference with mapreduce. 2014 IEEE 12th Int Conf Dependable, Auton Secur Comput. 2014 Aug. p. 153–8.
10. Ma Z, Xu Y, Sheng L, Li L. QBHSQ: A quad-tree based algorithm for high-dimension skyline query. 3rd Int Symp Intell Inf Technol Appl IITA. 2009 Nov; 1:593–6.
11. Lin Y, Wang ET, Chen ALP. Finding Targets with the Nearest Favor Neighbor and Farthest Disfavor Neighbor by a Skyline Query. SAC '14 Proceedings of the 29th Annual ACM Symposium on Applied Computing. 2014; 821–6.
12. Zhao L, Yang Y-Y, Zhou X. Continuous probabilistic subspace skyline query processing using grid projections. J Comput Sci Technol. 2014 Mar; 29(2):332–44.
13. Lim J, Lee Y, Bok K, Yoo J. A continuous reverse skyline query processing for moving objects. 2014 Int Conf Big Data Smart Comput BIGCOMP. 2014 Jan. p. 66–71
14. Papadias D, Fu G, Seeger B. An Optimal and Progressive Algorithm for Skyline Queries, SIGMOD '03 Proceedings of the 2003 ACM SIGMOD International Conference on Management of data. 2003. p. 467–78.
15. Guttman A. R-Trees: A Dynamic Index Structure For Spa-

tial Searching. 84 Proc 1984 ACM SIGMOD Int Conf Manag data. 1984 Jun. p. 47–57.
16. Hou C-C, Chang C-K, Chen Y-C, Su H-Y, Hsu Y-L. Finding similar users in social networks by using the neural-based skyline region. J Comput. 2015; 10(5):292–9.
17. Sharifzadeh M, Shahabi C. The Spatial S kyline Queries. VLDB '06 Proceedings of the 32nd International Conference on Very Large Data Bases. 2006. p. 751–62.
18. Bilge A, Kaleli C. A multi-criteria item-based collaborative filtering framework. 2014 11th Int Jt Conf Comput Sci Softw Eng Human Factors Comput Sci Softw Eng - e-Science High Perform Comput eHPC, JCSSE. 2014 May. p. 18–22.
19. Huang YHY. An item based collaborative filtering using item clustering prediction. 2009 ISECS Int Colloq Comput Commun Control Manag. 2009 Aug; 4:54–6.
20. Ying JCB, Shi N, Tseng VS, Tsai HW, Cheng KH, Lin SC. Preference-aware community detection for item recommendation. Proceedings - 2013 Conf Technol Appl Artif Intell TAAI 2013. 2013 Dec. p. 49–54
21. Kumar T. Solution of Linear and Non Linear Regression Problem by K Nearest Neighbour Approach: By using Three Sigma Rule. 2015 IEEE Int Conf Comput Intell Commun Technol. 2015 Feb. p. 197–201.
22. Yi X, Paulet R, Bertino E, Varadharajan V. Practical k nearest neighbor queries with location privacy. Icde'14. 2014 Mar -Apr; 640–51.
23. Xu W, Miranker DP, Mao R, Ramakrishnan S. Anytime K-nearest neighbor search for database applications. Proceedings - First Int Work Similarity Search Appl. SISAP 2008. 2008; 139–48.