

AN IMPROVED APPROACH FOR SUGGESTING BEST TRAVEL PLACES USING SKYLINE QUERY PROCESSING AND SOCIAL MEDIA RECORDS MINING

¹ Ultra Kumar Verma, ² Jasmine Minj

¹ Research Scholar, ² Assistant Professor

¹ Computer Science and Engineering,

¹ Vishwavidyalaya Engineering College, Lakhapur, India

Abstract : Recommender frameworks plan to help clients by choosing and proposing things that might be of significance to them, drawing from vaults that can be self-assertively large. A recommender is the dynamic party in a recommender framework that creates and gives the suggestions to a client; a recommender can be a bit of software as well as a user. Proposals can be verifiable or express: clients may unequivocally make a demand for suggestions, or proposals might be conveyed to them without their particular request. Recommender frameworks encounter numerous issues which reflect dwindled viability. Ongoing examination on recommender frameworks uncovers a thought of using informal organization information to upgrade customary recommender framework with better expectation and enhanced exactness. This paper proposes an enhanced travel recommender framework that first channels the most visited puts based on client remarks and sift through hopeful courses utilizing Skyline handling procedure. At that point the nature of the hopeful courses is enhanced by incorporating time obliged based most brief way calculation. The trial result demonstrates the nature of the recommended spots is extraordinarily upgraded by client time limitations.

Index Terms - Recommender Systems, Skyline Processing, Shortest path algorithm, Time Dependent Shortest Path Algorithm, Query Lattice.

I. INTRODUCTION

A recommender system proposes things like news, music, places, and films etc to customers in the casual association. Using customer tendency, for instance, customer profiles and thing examinations, recommender structure envision things the customer is presumably going to be possessed with. The information over-weight can be reduced by recommender systems by filtering the information in perspective of its congruity. Recommender systems use reasonable figuring to depict needs into some specific course of action of decisions from which the customer can pick. Relational associations are advancing rapidly and awesome arrangements of region based casual networks are in like manner being used. In territory based casual networks the spatial parts of customers are incorporated. It very well may be described as a casual association in which geographic organizations and capacities are joined to engage social components. The overhauls in area administrations brought new open entryways for the recommender structures. Region recommendation and timetable proposition are facilitated with area based administrations. The colossal proportion of data accumulated from area based administrations can be analyzed and used to deal with the zone proposition systems. With a particular ultimate objective to give the proposition, diverse counts and dataset are being used by the structures.

Redone proposition engages the online introduction expansion, suggestion of data in any setup that is pertinent to each and every customer. Modified recommendation structures are organized into five sorts depends upon their approach to manage proposition:

- Content-Based Filtering: prescribe a thing to clients in view of a depiction of the thing and a profile of the client's advantages. The suggestion of a substance construct framework is based with respect to individual data and disregards the commitments of different clients.
- Collaborative Filtering: Collaborative separating method in light of client's history through rating given by the client to a thing as their data source. Communitarian sifting approaches regularly experience the ill effects of three issues: cold start, adaptability, and sparsity.

Recently, there has been much interest in processing skyline queries for various applications that include decision making, personalized services, and search pruning. Skyline queries aim to prune a search space of large numbers of multi-dimensional data items to a small set of interesting items by eliminating items that are dominated by others.

Definition:

Given a search space of D independent dimensions u_1, u_2, \dots, u_d a point p_x is said to dominate another point p_y if the value of p_x is better than all other points in the same dimension. In general, a skyline query reduces the search space S to only the set of skyline points S_{sky} that are of interest to the user.

The nature of movement based recommender frameworks can be exceedingly enhanced by utilizing Skyline handling methods. Significant favorable circumstances are:

- It can be viably connected to huge dataset gathered from informal communication destinations with respect to clients travel inclinations for sifting best client inclinations based on highlights like rate, separations and so on.
- Additionally it can be utilized on area based frameworks keeping in mind the end goal to distinguish the briefest course to a goal or the nearest purpose of enthusiasm among many.
- Another application is disseminated question enhancement. This can be especially valuable in cloud models where information are scattered among servers or for the situation where Quality of (web) administrations is the essential objective.

Skyline questions can likewise be utilized to centre on a subspace of ascribes with a specific end goal to distinguish the horizon on a little subset of the measurements of the dataset that are characterized [6].

Table no 1: Comparison between previous work and problems

Sno	Title	Description	Problem
1	Efficient Keyword-Aware Representative Travel Route Recommendation	Designed a keyword extraction module to classify the POI-related tags. Also designed route construction algorithm for obtaining route candidates. Also used skyline query concepts for filtering best candidates.	The computational cost is increased due to application real time requirements
2	A new time-dependent shortest algorithm for multimodal transportation network	Designed a single source single destination algorithm that takes account the closeness to the target node as heuristic to the destination.	For the dense network the computational time increases.
3	A survey on Skyline processing	This papers surveys different skyline processing techniques with key aspects of each algorithm.	Effort needs to be done for validating with very large datasets and will consider the case of distributed data
4	A Survey on Various Techniques of Recommendation System in Web Mining	Describe the recommendation system related research and then introduces various techniques and approaches used by the recommender system User-based approach, Item-based approach, Hybrid recommendation approaches.	The collaborative recommendation approach suffers from three problems: cold start, scalability, and sparsity.
5	Venue Recommendation Based on User Location	Presented an idea of an android application which will recommend venues to users. This system preferred MobiContext framework, which is used for mobile social networks. The MobiContext is a hybrid cloud-based Bi-Objective Recommendation Framework (BORF). Above framework contains venue selection based on users personal preference and venue closeness on geographical information. Has implemented collaborative filtering	Intended to insert more functionalities in the proposed framework such as contextual data of user that is users' check-ins time interval, users' modified profile, users' hobbies and user's time preferences.

	(CF) before using BORF, because CF will make more ideal choices in real life practical application.	
--	---	--

II. PROBLEM IDENTIFICATION

The real prerequisites recognized in the wake of breaking down the past methodologies are:

- Finding ideal travel course physically is muddled predominantly on the grounds that it's hard to discover suitable focuses between movement courses.
- The recommended POI ought to fulfill greatest inclination of client like the area, stay time and so forth.
- The size of online life dataset is ordinarily tremendous so suitable strategies must be connected to diminish the size and channel just required dataset as per client inclinations.

The existing methodology utilizes DFS for applicant course age. The significant burden of DFS is that It is conceivable that may states keep reoccurring. There is no assurance of finding the objective hub. Once in a while the states may likewise go into unbounded circles.

III. METHODOLOGY

The improved approach is divided into following major modules as mentioned below:

1. Pre-processing on dataset: The dataset is filtered and scores are assigned on the basis of geo and temporal features.
2. Candidate route generation: A route module is constructed by applying Depth First Search Algorithm.
3. Skyline Processing: This technique is applied for reducing the no of candidate routes found in previous step and reduce dominating points. Our approach uses an efficient query lattice technique for skyline computation.
4. Filter shortest route to a tourist spot according to user preferences: The time dependent shortest path algorithm is applied for filtering shortest route to a tourist path as per user time preferences.

3.1 Candidate Route Generation Algorithm:

Information: Raw direction set T;

Yield: New applicant direction set Tc.

- 1: Initialize a stack S;
- 2: Split each course $r \in T$ into (head,tail) subsequences;
- 3: Reconstruct (headSet).
- 4: Procedure Reconstruct (Set).
- 5: foreach (head,tail) Set do
- 6: endFlag = False;
- 7: if S is vacant or $\text{tail.time} > \text{S.pop}().\text{time}$ at that point
- 8: Push head in S;
- 9: Push tail in S;
- 10: else
- 11: Push head in S;
- 12: endFlag = True;
- 13: if endFlag is False at that point
- 14: Reconstruct (tailSet)
- 15: Insert S in Tc;
- 16: Procedure End

3.2 Travel Routes Generation Using Skyline Processing Module.

The thought is to consolidate the join with settled circle and sort-blend calculations. Notwithstanding, every connection needs to go to different occasions keeping in mind the end goal to process the horizon for each join esteem, and after that the worldwide horizon. What's more, each information question must be checked at any rate once. A cross section based way to deal with single-connection horizon calculation is presented in [6], with the point of proposing an information dispersion autonomous calculation. A cross section structure is utilized to answer horizon questions over measurements with low cardinality areas, or those that can be mapped to low-cardinality spaces, (for example, Price, that can be mapped to value ranges). The guideline is to compose every one of the qualities mixes into a cross section in view of the strength relationship, and afterward to recover those that (a) are available in the info informational index, and (b) are not reachable by the predominance relationship from another component of the grid, additionally having a place with the informational index. Rather than getting to individual tuples in a database table, our methodology depends on the meaning of horizon as the association of the responses to an arrangement of questions. In doing as such, our essential calculation abstains from getting to the table specifically: access to the table is through inquiries, thus free of how the table is actualized or how its tuples are listed. Given a social table R, our methodology depends on what we call the inquiry cross section of R; and our essential calculation develops the horizon set as the association of the responses to a subset of questions from that grid - subsequently without specifically getting to the table R.

3.3 Time Dependent Shortest Path Algorithm on the skyline generated travel routes.

The aim of the algorithm is find the closest node which leads to the target node with least cost.

Step 1: Constraint Parameters

Threshold distance d as mean value from all vertices to the virtual path given by equation below:

$$v_i \in V, d = \sum_i^n \text{dist}(V_i, (st))/n \dots \dots \dots (1)$$

Δd as elementary step distance as mean value from all vertices to its neighbours.

$$v_i \in V, \Delta d_i = x_i/k \dots \dots \dots (2)$$

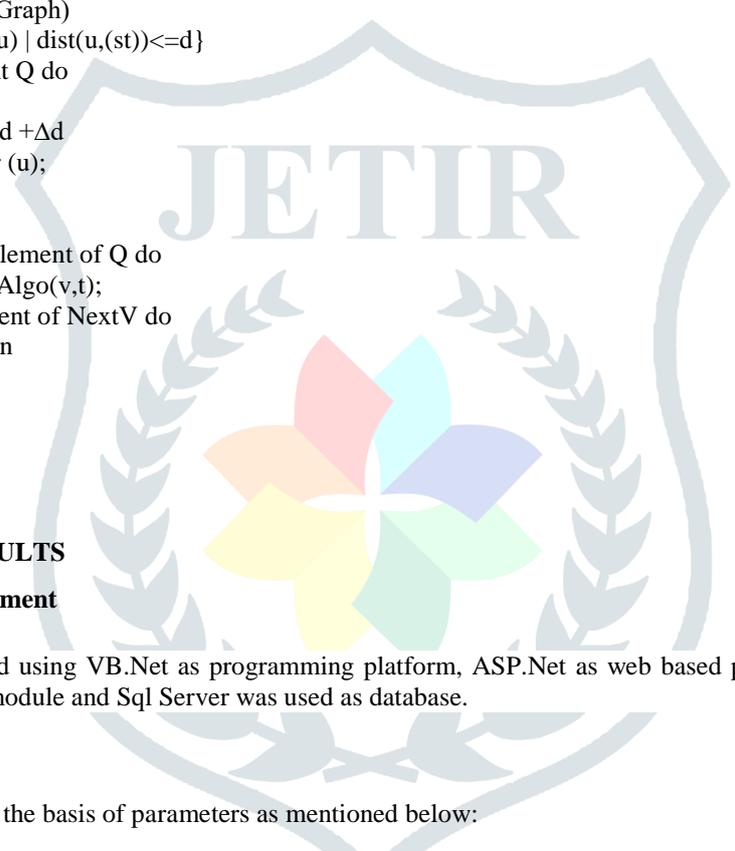
$$\Delta d = \sum_1^n \Delta d_i /n \dots \dots \dots (3)$$

Step 2: Algorithm implementation

Input: Skyline filtered candidate routes as virtual path

Output: Shortest route on user time constraints.

1. $d = \text{Calculate_theshold}(\text{Graph}, \text{Source}, \text{Destination})$.
 2. $\Delta d = \text{Calculate_Step}(\text{Graph})$
 3. $Q = \{v \in \text{Neighbours}(u) \mid \text{dist}(u, (st)) \leq d\}$
 4. While t not an element Q do
 5. If $Q \neq \text{null}$ then
 6. If $d < d_max$ then $d = d + \Delta d$
 7. Else $w = \text{predessoror}(u)$;
 8. $d = d + \Delta d$
 9. $\text{path} = \text{path} / u$
 10. Else for each v is an element of Q do
 11. $\text{NextV} = \text{shortestPathAlgo}(v, t)$;
 12. Foreach W is an element of NextV do
 13. If $(\text{dist}(w, (st)) \leq d)$ then
 14. $Q = Q \cup \{w\}$;
 15. $\text{addtoPath}(w)$
 16. $Q = Q - \{v\}$
- Return path



IV. EXPERIMENTAL RESULTS

4.1 Platforms used for experiment

The experiment was conducted using VB.Net as programming platform, ASP.Net as web based platform for developing a web based route recommendation module and Sql Server was used as database.

4.2 Performance Measure

Performance was measured on the basis of parameters as mentioned below:

1. **Route Prediction Accuracy** :We used nearly 1000 photos as dataset collected from social media for training and testing the model. For each dataset, the test data were created by collecting the last travel sequence of the top-10 percent of users in the most recent 30 percent time periods. The training dataset consisted of the set of travel sequences excluding the testing data part. The route prediction table is as mentioned below:

Table 2: Prediction accuracy comparison

Title	Feature
Geo-Social Influenced Routes (GSI)	Only consider the geo social influence score of [7]. The route consists of POIs visited by geo-social influential users in the social network.
Keyword-Aware Skyline Travel Route (KSTR).	KSTR [8] outputs full Skyline routes based on both POI and user factors.
Keyword-Aware Representative Travel Route.	The existing approach outputs optimal representative Skyline routes [1].
Hybrid Skyline Based Route Recommendation Combined with Time Dependent Shortest Path algorithm.	Our hybrid approach combines the effective shortest path algorithm for route travel for skyline filtered routes based on

user time preferences.

Efficiency: The efficiency was calculated on the basis of response time for skyline computational approaches. The existing approach uses R-Tree based skyline approach which is more time consuming as compared to lattice based approach. The performance table is as mentioned below:

Table 3: Response time comparison

	Existing R-Tree based approach	Proposed Lattice based approach
Response Time	0.265147393	0.145713425

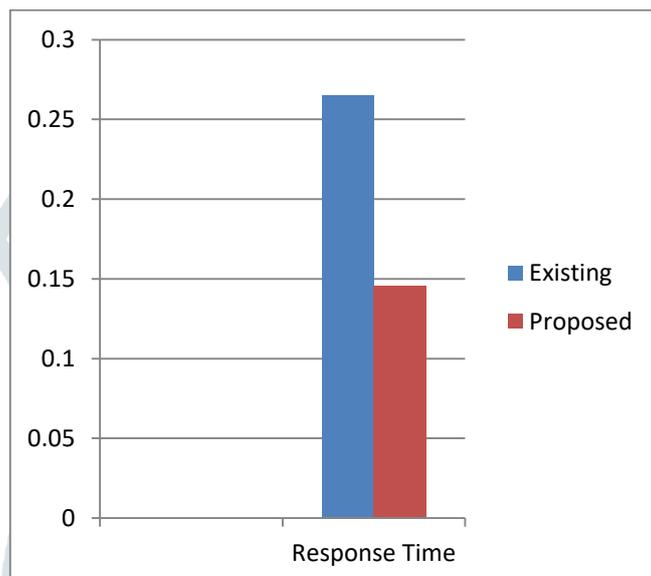


Figure 1: Response time comparison between exiting skyline and proposed skyline.

6. CONCLUSION

Creating systems for proficiently give customized proposals has picked up a considerable measure of force with ubiquity of online life mining. The present paper gives a talk about online networking mining, horizon handling and recommender systems. The current methodology centers on proposing a productive half breed procedure by joining horizon calculations and briefest way calculation in view of client time inclinations. The future execution of the work depends on actualizing time subordinate briefest way calculation. The future work aims to integrate more keyword aware techniques using advanced big data algorithms like Map Reduce.

REFERENCES

- [1] Y.Ting Wen, J.Yeo, W.Peng, "Efficient Keyword-Aware Representative Travel Route Recommendation", Ieee Transactions On Knowledge And Data Engineering, Vol. 29, No. 8, August 2017.
- [2] A.Idri, M.Oukarfi, A.Baulmakoul, K.Zeitouni & A.Masri, "A new time-dependent shortest path algorithm for multimodal transportation network", The 8th International Conference on Ambient Systems, Networks and Technologies, 2017.
- [3] C.Kalyvas & T.Tzouramanis, "A survey on skyline processing" arXiv: 1704.01788, 2017.
- [4] Y.G. patel & V.P. Patel, A Survey on Various Techniques of Recommendation System in Web Mining, IJEDR, 2015
- [5] A.Kute, M.Kathale, P.Choudhari & P.Tapkir, "Venue Recommendation Based on User Location", International Journal of Innovative Research in Computer and Communication Engineering, October 2016.
- [6] M. Morse, J. Patel, H. V. Jagadish. Efficient Skyline Computation over Low Cardinality Domains. In Proc. of VLDB 2007, pp. 267–278, 2007.
- [7] Y.-T. Wen, P.-R. Lei, W.-C. Peng, and X.-F. Zhou, "Exploring social influence on location-based social networks," in Proc. IEEE Int. Conf. Data Mining, 2014, pp. 1043–1048.
- [8] Y.-T. Wen, K.-J. Cho, W.-C. Peng, J. Yeo, and S.-W. Hwang, "KSTR: Keyword-aware skyline travel route recommendation," in Proc. IEEE Int. Conf. Data Mining, 2015, pp. 449–458.
- [9] S.Shah, A.Thakkar & Sonal Rami, "A Novel Approach for Making Recommendation using Skyline Query based on user Location and Preference", Indian Journal of Science and Technology, Vol 9(30), DOI: 10.17485/ijst/2016/v9i30/99075, August 2016.
- [10] J.Bao, Y.Zheng, D.Wilkie & M.Mokbel, "Recommendations in Location-based Social Networks: A Survey", Geoinformatica (2015) 19:525565 DOI 10.1007/s10707-014-0220-8.

- [11] Jian-Guo Liu, Michael Z. Q. Chen, Jianchi Chen, Fei Deng, Hai-Tao Zhang, Zi-Ke Zhang, And Tao Zhou.” Recent Advances In Personal Recommender Systems”,International Journal Of Information And Systems Sciences, 2009.
- [12] Y.G. patel&V.P.Patel,”A Survey on various techniques of Recommendation System on Web Mining”, 2015 IJEDR | Volume 3, Issue 4 | ISSN: 2321-9939
- [13] Y.Wen, J.Yeo &W.Peng ,” An Efficient Keyword-Aware Representative Travel Route Recommendation”, IEEE Transactions on Knowledge and Data Engineering (Volume: 29, Issue: 8, Aug. 1 2017)
- [14] K.Pachpande , M.Mohanan ,D.Zodage , H.Thakare4 &K.Jadhav,” Recommendation and suggestion System For Easy Tourism”, Volume 4, Issue 11, November -2017
- [15] J.Li, Y.Yang, &W.Liu,” Exploring Personalized Travel Route Using POIs”, International Journal of Computer Theory and Engineering, Vol. 7, No. 2, April 2015.
- [16] J.Zhang ,H. Kawasaki &Y.Kawai,” A Tourist Route Search System Based on Web Information and the Visibility of Scenic Sights”, Second International Symposium on Universal Communication,IEEE 2008.

