An Efficient and Robustness of Travel Route Recommendation Framework

Yenumala Sankara Rao[1], Popuri Hanuma[2]
[1]associate Professor, Dept of MCA
[2]Dept of MCA
[1, 2] St. Mary's Group Of Institutions, Guntur, Andhra Pradesh, India

Abstract- *with the notoriety of online networking (e.g., Face book and Flicker), clients can without much of a stretch offer their registration records and photographs amid their excursions. In perspective of the gigantic number of client verifiable portability records in online networking, we expect to find travel encounters to encourage trip arranging. When arranging a trek, clients dependably have particular inclinations with respect to their outings. Rather than confining clients to restricted inquiry choices, for example, areas, exercises, or eras, we consider subjective content portrayals as keywords about customized prerequisites. In addition, a various and delegate set of prescribed travel routes is required. Earlier works have expounded on mining and positioning existing routes from registration information. To address the issue for programmed trip association, we assert that more highlights of Places of Interest (POIs) ought to be extricated. In this manner, in this paper, we propose a productive Keyword-aware Representative Travel Route structure that utilizations information extraction from clients' verifiable portability records and social associations. Unequivocally, we have planned a keyword extraction module to group the POI-related labels, for successful coordinating with inquiry keywords. We have additionally outlined a route remaking calculation to develop route hopefuls that satisfy the necessities. To give befitting inquiry comes about, we investigate Representative Skyline ideas, that is, the Skyline routes which best depict the exchange offs among various POI highlights. To assess the adequacy and productivity of the proposed calculations, we have led broad investigations on genuine area based informal organization datasets, and the examination comes about demonstrate that our techniques do undoubtedly show great execution contrasted with cutting edge works.*

Keywords- Location-based social network, text mining, travel route recommendation

I. INTRODUCTION

LOCATION-BASED social network (LBSN) services allow users to perform check-in and share their check-in data with their friends. In particular, when a user is traveling, the check-in data are in fact a travel route with some photos and tag information. As a result, a massive number of routes are generated, which play an essential role in many well-established research areas, such as mobility prediction, urban planning and traffic management. In this paper, we focus on trip planning and intend to discover travel experiences from shared data in location-based social networks. To facilitate trip planning, the prior works in [1], [2], [3], [4], provide an interface in which a user could submit the query region and the total travel time. In contrast, we consider a scenario where users specify their preferences with keywords. For example, when planning a trip in Sydney, one would have “Opera House”. As such, we extend the input of trip planning by exploring possible keywords issued by users.

However, the query results of existing travel route recommendation services usually rank the routes simply by the popularity or the number of uploads of routes. For such ranking, the existing works [6], [7], and [8] derive a scoring function, where each route will have one score according to its features (e.g., the number of Places of Interest, the popularity of places). Usually, the query results will have similar routes. Recently, [9] aimed to retrieve a greater diversity of routes based on the travel factors considered. As high scoring routes are often too similar to each other, this work considers the diversity of results by exploiting Skyline query.

In this paper, we develop a Keyword-aware Representative Travel Route (KRTR) framework to retrieve several recommended routes where keyword means the personalized requirements that users have for the trip. The route dataset could be built from the collection of low-sampling check-in records.

Definition 1 (Travel route). Given a set of check-in points recorded as a series of travel routes, each check-in point represents a POI p and the user’s checked-in time t. The check-in records were grouped by individual users and ordered by the creation time. Each user could have a list of travel routes fT g ¼ fT_0; T_1; . . .g, where T_0 ¼ ðp_0; t_0Þ; ðp_1; t_1Þ; . . . ; ðp_i; t_i Þ, T_1 ¼ ð$p_{iþ1}$; $t_{iþ1}$Þ; ð$p_{iþ2}$; $t_{iþ2}$Þ; . . . and $t_{iþ1}$ t_i is greater than a route-split threshold. We set the route-split threshold to one day in this paper.

Consider the example illustrated in Fig. 1, the related route information of which is stored in Table 1. For ease of illustration, each POI is associated with one keyword (though our model can support multiple keywords) and

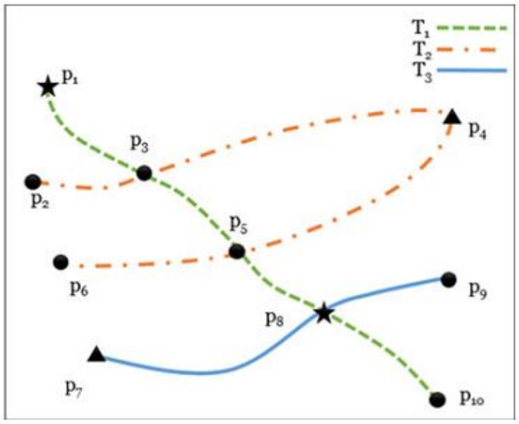


fig. 1. Keyword-aware travel routes query running example

Two-dimensional score vector (each dimension represents the rank of a feature). Assume a tourist plans a date with a set of keywords [“Whisky” “Sydney Cove” “Sunset”]. First, we can find that these keywords vary in their semantic meaning: “Sydney Cove” is a geographical region; “Sunset” is related to a specific time period (evening) and locations such as beach; “Whisky” is the attribute of POI.

We argue that knowing semantics is important, as some query keywords do not need to be matched in the POI key-word. For example, p_9, even though its name does not include “Whiskey”, is a good match, as it is an important attribute of Bar POIs. Similarly, “Sydney Cove” is not mentioned, but based on the location of Opera House, p_8 matches the requirement. As a result, T_3 matches all the requirements, which could not be supported by existing simple keyword-based matches. In this example, the key-word “Sunset” can be easily matched. Although the other two words are not stored in the database, we want to correspond them to Drinking whisky at a bar and Opera House in Sydney Cove. Finally, T_3 matches all the requirements. Mean-while, there is still a possibility that no existing route is in accordance with the query keywords. For this challenge, we propose a candidate route generation algorithm to increase the number of routes. For instance, a travel sequence T ^{0} ¼ fp_1 ! p_3 ! p_4 ! p_5 ! p_8 ! p_9g, which is aggregated from the route segments of T_1 to T_3, also matches all the key-words specified.

Additionally, we have mentioned that the final results may have similar characteristics and be monotonous due to the fact that all of the factors are aggregated into one score for each travel route. Consequently, the system will retrieve the top-k routes with the highest score as the results. Users may not understand the characteristic of these routes through the final single score (e.g., Which one has the most interesting landmarks? Which one is well-connected to the place I want to go?) so it may be hard to choose a route from the final results. Furthermore, users need to pre-define the weight for each factor, although it is hard to select a suitable weight in most cases. Since travel route recommendation has to take several factors into consideration to emphasize the unique travel factors of travel routes, we borrowed the concept of Distance-based Representative Skyline [10] to retrieve travel routes. Distance-based Representative Sky-line search on the travel routes also includes a small number k of skyline routes that best describe the full optimal (Skyline) results in terms of the features derived. Consider an example in Fig. 1, where the score vector of POIs represents the attractiveness score and the visiting time information. To compute the average POI score of T_1, T_2 and T_3, we get the final score values (0.1, 0.34), (0.15, 0.44), and (0.18, 0.3) respectively. For example, with k ¼ 3, the skyline points in Fig. 2 can be divided into three subsets {T_4}, {T_2; T_5; T_6} and {T_3; T_8}. Our representative skyline travel route solution will report fT_2; T_3; T_4g.

TABLE 1 Example of Route Dataset

Tid	Uid	Pid	keyword	time	POI score vector
T_1	u_1	p_1	Opera House	10:00	(0.04, 0.2)
T_1	u_1	p_3	Bar	12:00	(0.25, 0.2)
T_1	u_1	p_5	Bar	15:30	(0.2, 0.8)
T_1	u_1	p_8	Opera House	17:30	(0.04, 0.3)
T_1	u_1	p_{10}	Bar	19:00	(0.04, 0.2)
T_2	u_2	p_2	Bar	10:30	(0.02, 0.2)
T_2	u_2	p_3	Bar	12:30	(0.25, 0.2)
T_2	u_2	p_4	Sunset	17:00	(0.05, 0.2)
T_2	u_2	p_5	Bar	19:00	(0.2, 0.8)
T_2	u_2	p_6	Bar	19:30	(0.25, 0.8)
T_3	u_3	p_7	Sunset	18:30	(0.4, 0.8)
T_3	u_3	p_8	Opera House	19:30	(0.04, 0.3)
T_3	u_3	p_9	Bar	20:00	(0.1, 0.1)

This paper builds on and significantly improves the KSTR framework [9] of recommending a diverse set of travel routes based on several score features mined from social media. KSTR then constructs travel routes from dif-ferent route segments. Specifically, we extend KSTR to consider representative and approximate results under an optional k limit in Section 5. Additionally, resources including passive check-ins such as GPS-tagged photos are discussed in Section 6. This addition would enable KRTR to consider a larger input including active and passive check-ins with high efficiency and scalability.

The contributions of this paper are summarized as follows:

We propose a KRTR framework in which users are able to issue a set of keywords and a query region, and for which query results contain diverse trip routes.

Check-in information is mined from passive check-ins to enrich the input data. GPS-tagged photos are larger in scale than foursquare check-ins. This mining thus improves the coverage of the input data.

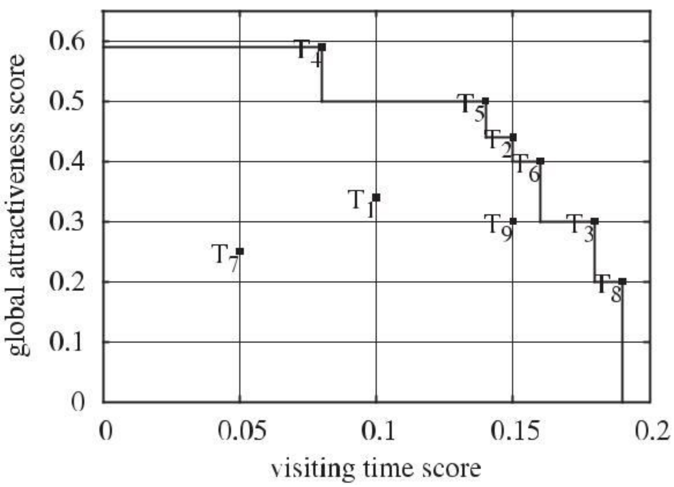


Fig. 2. An extended example of skyline travel routes built by Table 1.

TABLE 2 Symbols and Notations

Notation	Definition
[1]	location as Point-of-Interest (POI)
t	low-sampling route as travel sequence
w	tag that describes a POI
n	the number of routes in the dataset
K	a set of query keywords
G	geo-specificity score of a tag w
T	temporal-specificity score of a tag w
AT	attribute score of a tag w
D	a set of d-dimensional featured routes
m	the number of routes in D
S	the full skyline of D
k	maximum number of the returned travel routes
R	the returned representative skyline travel routes

We propose a route reconstruction method to partition routes into segments by considering spatial and temporal features.

Representative Skyline query for travel route search is adopted to combine the multi-dimensional measurements of routes, which increases the diversity of the recommended results. Moreover, a greedy method is designed for the efficiency of the online application.

To evaluate our proposed framework, we conducted experiments on real LBSN and photo datasets. The experiments show that KRTR is able to retrieve travel routes that are of interest to users.\

II. LITERATURE SURVEY

[1]**Efficient Keyword-Aware Representative Travel Route Recommendation Yu-**Ting Went, Jinyoung Yeo, Wen-Chih Peng and Seung-Won Hwang

Description: —With the popularity of social media (e.g., Facebook and Flicker), users can easily share their check-in records and photos during their trips. In view of the huge number of user historical mobility records in social media, we aim to discover travel experiences to facilitate trip planning. When planning a trip, users always have specific preferences regarding their trips. Instead of restricting users to limited query options such as locations, activities, or time periods, we consider arbitrary text descriptions as keywords about personalized requirements. Moreover, a diverse and representative set of recommended travel routes is needed. Prior works have elaborated on mining and ranking existing routes from check-in data. To meet the need for automatic trip organization, we claim that more features of Places of Interest (POIs) should be extracted. Therefore, in this paper, we propose an efficient Keyword-aware Representative Travel Route framework that uses knowledge extraction from users' historical mobility records and social interactions. Explicitly, we have designed a keyword extraction module to classify the POI-related tags, for effective matching with query keywords. We have further designed a route reconstruction algorithm to construct route candidates that fulfill the requirements. To provide befitting query results, we explore Representative Skyline concepts, that is, the Skyline routes which best describe the trade-offs among different POI features. To evaluate the effectiveness and efficiency of the proposed algorithms, we have conducted extensive experiments on real location-based social network datasets, and the experiment results show that our methods do indeed demonstrate good performance compared to state-of-the-art works.

[2]**Mining interesting locations and travel sequences from GPS trajectories**
Authors: Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma

Description: The increasing availability of GPS-enabled devices is changing the way people interact with the Web, and brings us a large amount of GPS trajectories representing people's location histories. In this paper, based on multiple users' GPS trajectories, we aim to mine interesting locations and classical travel sequences in a given geospatial region. Here, interesting locations mean the culturally important places, such as Tiananmen Square in Beijing, and frequented public areas, like shopping malls and restaurants, etc. Such information can help users understand surrounding locations, and would enable travel recommendation. In this work, we first model multiple individuals' location histories with a treebased hierarchical graph (TBHG). Second, based on the TBHG, we propose a HITS (Hypertext Induced Topic Search) based inference model, which regards an individual's access on a location as a directed link from the user to that location. This model infers the interest of a location by taking into account the following three factors. 1) The interest of a location depends on not only the number of users visiting this location but also these users' travel experiences. 2) Users' travel experiences and location interests have a mutual reinforcement relationship. 3) The interest of a location and the travel experience of a user are relative values and are region-related. Third, we mine the classical travel sequences among locations considering the interests of these locations and users' travel experiences. We evaluated our system using a large GPS dataset collected by 107 users over a period of one year in the real world. As a result, our HITS-based inference model outperformed baseline approaches like rank-by-count and rank-by frequency. Meanwhile, when considering the users' travel experiences and location interests, we achieved a better performance beyond baselines, such as rank-bycount and rank-by-interest, etc.

[3]Exploiting geographical influence for collaborative point-of-interest recommendation Authors: M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee

Description: In this paper, we aim to provide a point-of-interests (POI) recommendation service for the rapid growing location-based social networks (LBSNs), e.g., Foursquare, Whrrl, etc. Our idea is to explore user preference, social influence and geographical influence for POI recommendations. In addition to deriving user preference based on user-based collaborative filtering and exploring social influence from friends, we put a special emphasis on geographical influence due to the spatial clustering phenomenon exhibited in user check-in activities of LBSNs. We argue that the geographical influence among POIs plays an important role in user check-in behaviors and model it by power law distribution. Accordingly, we develop a collaborative recommendation algorithm based on geographical influence based on naive Bayesian. Furthermore, we propose a unified POI recommendation framework, which fuses user preference to a POI with social influence and geographical influence. Finally, we conduct a comprehensive performance evaluation over two large-scale datasets collected from Foursquare and Whrrl. Experimental results with these real datasets show that the unified collaborative recommendation approach significantly outperforms a wide spectrum of alternative recommendation approaches.

[4]Exploring social influence on location-based social networks
Authors: Y.-T. Wen, P.-R. Lei, W.-C. Peng, and X.-F. Zhou

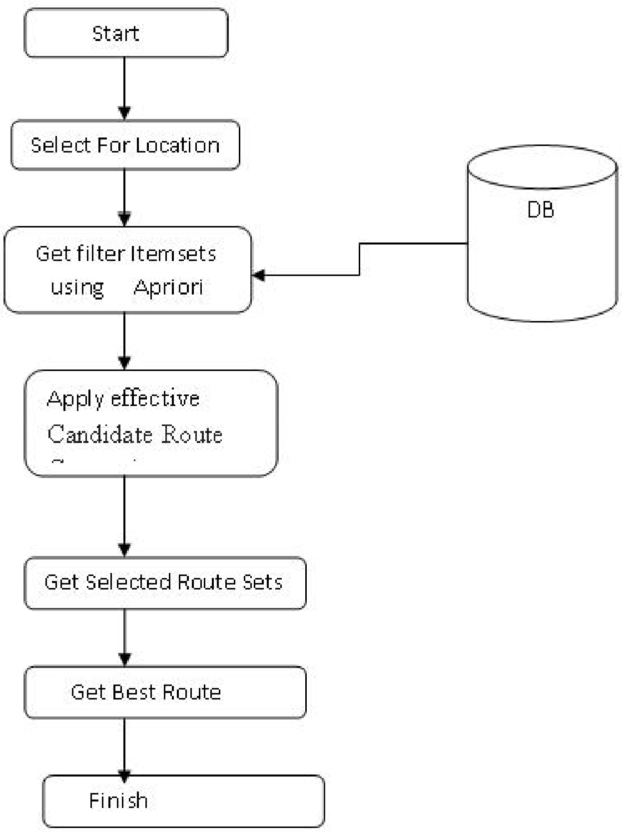
Description: In recent years, with the popularization of mobile network, the location-based service (LBS) has made great strides, becoming an efficient marketing instrument for enterprises. For the retail business, good selections of store and appropriate marketing techniques are critical to increasing the profit. However, it is difficult to select the retail store because there are numerous considerations and the analysis was short of metadata in the past. Therefore, this study uses LBS, and provides a recommendation method for retail store selection by analyzing the relationship between the user track and point-of-interest (POI). This study uses regional relevance analysis and human mobility construction to establish the feature values of retail store recommendation. This study proposes (1) architecture of the data model available for retail store recommendation by influential layers of LBS; (2) System-based solution for recommendation of retail stores, adopts the influential factors with specified data in LBS and filtered by industrial types; (3) Industry density, area categories and region/ industry clustering methods of POIs. Uses KDE and KMeans to calculate the effect of regional functionality on the retail store selection, similarity is used to calculate the industry category relation, and consumption capacity is considered to state.

III. RELATED WORK

Problem Definition: We propose an efficient Keyword-aware Representative Travel Route framework that uses knowledge extraction from users' historical mobility records and social interactions. Explicitly, we have designed a keyword extraction module to classify the POI-related tags, for effective matching with query keywords. To provide befitting query results, we explore Representative Skyline concepts, that is, the Skyline routes which best describe the trade-offs among different POI features. The experiment results show that our methods do indeed demonstrate good performance compared to state-of-the-art works.

IV. PROPOSED SYSTEM

We propose an efficient Keyword-aware Representative Travel Route framework that is knowledge extraction from users' historical mobility records and social interactions. The experiment results show that our method do indeed demonstrate good performance compared to state of the art work.



V. CONCLUSION

These travel routes are identified with all or halfway client inclination keywords to enhance the productivity of succession mining and gathering clients in view of their area histories or grouping areas regarding individuals' visits are potential works. We propose a novel keyword extraction module to distinguish the semantic significance and match the estimation of routes

REFERENCES

[1] Yu-Ting Wen, Jinyoung Yeo, Wen-Chih Peng, Member, IEEE, and Seung-Won Hwang ‖ Efficient Keyword-Aware Representative Travel Route Recommendation IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 29, NO. 8, AUGUST 2017

[2] H.-P. Hsieh and C.-T. Li, ―Mining and planning time-aware routes from check-in data,‖ in Proc. 23rd ACM Int. Conf.
Conf. Inf. Knowl. Manage., 2014, pp. 481–490.

[3] V. S. Tseng, E. H.-C. Lu, and C.-H. Huang, ―Mining temporal mobile sequential patterns in location-based service environments,‖ in Proc. Int. Conf. Parallel Distrib. Syst., 2007, pp. 1–8.

[4] W. T. Hsu, Y. T. Wen, L. Y. Wei, and W. C. Peng, ―Skyline travel routes: Exploring skyline for trip planning,‖ in Proc.
IEEE 15th Int. Conf. Mobile Data Manage., 2014, pp. 31–36.

[5] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, ―Mining interesting locations and travel sequences from GPS trajectories,‖ in Proc. 18th Int. Conf. World Wide Web, 2009, pp. 791–800.

[6] Y. Tao, L. Ding, X. Lin, and J. Pei, ―Distance-based representative skyline,‖ in Proc. IEEE 25th Int. Conf. Data Eng.,
2009, pp. 892–903.

[7] T. Lee, Z. Wang, H. Wang, and S.-W. Hwang, ―Attribute extraction and scoring: A probabilistic approach,‖ in Proc.
IEEE 29th Int. Conf. Data Eng., 2013, pp. 194–205.

[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. 449458.

[9] R. Agrawal, T. Imielinski, and A. Swami, ―Mining Association Rule between Sets of Items in Large Databases‖, in
Proceedings of the ACM SIGMOD Conference on Management of Data, Washington, D.C., pp. 207-216, May 1993.