

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/329570549>

Tour recommendation and trip planning using location-based social media: a survey

Article in Knowledge and Information Systems · September 2019

DOI: 10.1007/s10115-018-1297-4

CITATIONS

45

READS

3,914

4 authors:



Kwan Hui Lim

University of Melbourne

71 PUBLICATIONS 959 CITATIONS

SEE PROFILE



Jeffrey Chan

RMIT University

86 PUBLICATIONS 1,542 CITATIONS

SEE PROFILE



Shanika Karunasekera

University of Melbourne

167 PUBLICATIONS 2,985 CITATIONS

SEE PROFILE



Christopher Leckie

University of Melbourne

349 PUBLICATIONS 9,234 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Radioactive source localisation [View project](#)



Linear System Cyber Security [View project](#)

Tour Recommendation and Trip Planning using Location-based Social Media: A Survey

Kwan Hui Lim^{*}, Jeffrey Chan^{††}, Shanika Karunasekera[†], Christopher Leckie[†]

^{*}Information Systems Technology and Design Pillar, Singapore University of Technology and Design

[†]School of Computing and Information Systems, The University of Melbourne, Australia

[‡]School of Science, RMIT University, Australia

kwanhui.lim@sutd.edu.sg, jeffrey.chan@rmit.edu.au, {karus, caleckie}@unimelb.edu.au

Abstract. Tourism is both an important industry and popular leisure activity undertaken by millions around the world. One important task for tourists is to plan and schedule tour itineraries that comprise multiple captivating Points-of-Interests (POIs) based on the unique interest preferences of the tourist. The complex task of tour itinerary recommendation is further complicated by the need to incorporate various real-life constraints such as limited time for touring, uncertain traffic conditions, inclement weather, group travel, queuing times and crowdedness. In this survey, we conduct a comprehensive literature review of studies on tour itinerary recommendation and present a general taxonomy for touring-related research. We discuss the entire process of tour itinerary recommendation research including: (i) data collection and types of datasets; (ii) problem formulations and proposed algorithms/systems for individual travellers, groups of tourists and various real-life considerations; (iii) evaluation methodologies for comparing tour itinerary recommendation algorithms; and (iv) future directions and open problems in tour itinerary recommendation research.

Keywords: Tour Recommendation, Itinerary Planning, Group Recommendation, User Interests, Personalization, Location-based Social Networks

1. Introduction

Tourism is a popular leisure activity undertaken by more than 1.18 billion international tourists per annum [97]. Economically, tourism is an important industry,

Received 15 Nov 2017

Revised 12 Jun 2018

Accepted 03 Oct 2018

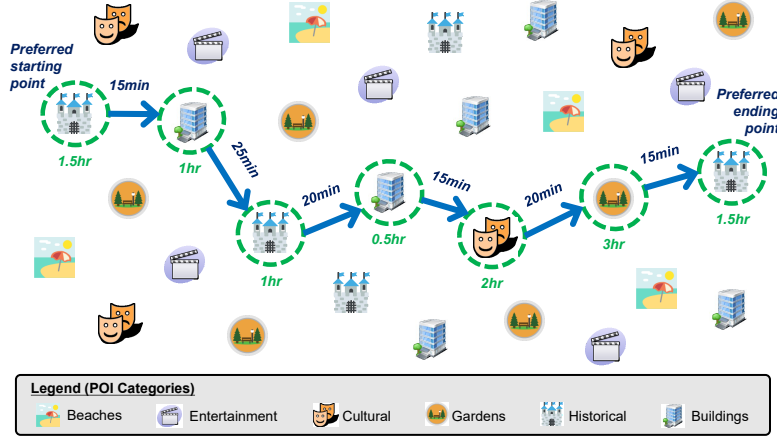


Fig. 1. Example of the tour recommendation problem.

generating more than 284 million jobs and accounting for more than US\$7.2 trillion in revenue annually [105]. Despite its importance and popularity, planning a tour itinerary in a foreign city is both challenging and time-consuming due to the need to identify captivating Places-of-Interest (POIs) and plan visits to these POIs as a connected itinerary. Adding to these challenges are the need to personalize the recommended itinerary according to the interest preferences of tourists, and to schedule the itinerary based on relevant temporal and spatial constraints, such as having limited time to complete the tour and needing to start and end near certain locations (e.g., the tourist's hotel). Figure 1 shows an example of the tour recommendation problem, where there are multiple POIs of different categories, and the tourist has to find an itinerary that optimizes the time taken and number of POIs visited, while satisfying various trip constraints.

Although tourism-related information can be obtained from the Internet and travel guides, these resources simply recommend popular POIs or generic itineraries but otherwise do not address the unique interest preferences of individual tourists or adhere to their various temporal and spatial constraints. Moreover, the large amount of information available increases the challenge of identifying the relevant information for the tourist. One popular alternative is to engage the services of tour agencies, but likewise, these tour agencies normally only recommend standard package tours that may not address the interest preferences or trip constraints of all tourists.

To address these issues, many researchers have studied tour itinerary recommendation problems and proposed various algorithms for solving these problems. These problems originated from the operations research community where the main focus is to schedule an optimal path, where the measure of optimality is typically based on a global metric such as POI popularity, and thus there is no personalization based on unique user interests. With the prevalence of smart phones and location-based social media, there has been an increased emphasis on data-driven approaches to tour itinerary recommendation to better model the interest preferences of tourists, and recommend personalized tour itineraries that satisfy these interest preferences as well as other trip constraints. In this survey paper, we focus on such data-driven tour recommendation research, par-

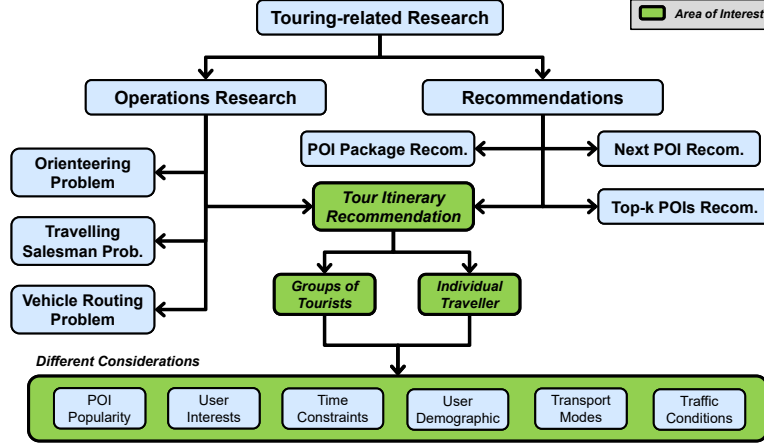


Fig. 2. Taxonomy of touring-related research.

ticularly on the types of data sources used, the problem variants formulated, the algorithms proposed and the evaluation methodology used.

Closely related to the field of tour itinerary recommendation are the fields of next-location prediction/recommendation [76, 6, 41, 91, 66], top-k location recommendation [108, 112, 65, 107, 103, 62] and travel package/region recommendation [8, 7, 95, 82]. Although these fields are related to tour itinerary recommendation, there are distinct differences in terms of the problem studied. Next-location prediction and recommendation aim to identify the next location that a user is likely to visit based on his/her previous trajectories, whereas tour itinerary recommendation aims to recommend multiple POIs or locations in the form of a trajectory. Top-k location recommendation and travel package/region recommendation do fulfill the criterion of recommending multiple POIs as part of a ranked list or travel package, but they do not structure these POIs as a connected itinerary. In contrast, tour itinerary recommendation has the additional challenges of planning an itinerary of connected POIs that appeal to the interest preferences of the users, while adhering to the temporal and spatial constraints in the form of a limited time budget for touring and having to start and end at specific POIs.

In this survey, we focus on works related to tour itinerary recommendation and the different real-life considerations incorporated into this problem. Figure 2 illustrates a taxonomy of the general area of touring-related research, which is further divided into the sub-areas of operations research and recommendations.

1.1. Related Surveys and Reviews

There exist a variety of survey and review articles that cover different aspects of the tour recommendation problem. In this section, we discuss these related articles and highlight the differences between this paper and the earlier articles.

The tour recommendation problem is closely related to the tourist trip design problem covered in the operations research community, and consequently, there have been various survey papers [88, 45] focusing on the aspects of problem formulation, algorithmic design and the complexity of this problem. Similarly, many tour recommendation problems are based on variants of the Orienteering Prob-

lem (OP), and [101, 48] provide in-depth discussions on the OP. Researchers such as [11] performed a review of tour recommendation systems, focusing on applications and systems aspects such as the types of interface, the system functionalities, the recommendation techniques and the artificial intelligence methods used. Others studied recommendations in general on location-based social networks [5] and the general types of research utilizing Flickr photos [90], with a small portion of their survey covering tourism-related applications. While these articles offer interesting discussions into different aspects of tour recommendation, this paper differs from the earlier articles in the following ways: (i) First, we review tour recommendation research as a holistic problem, covering the whole process from data collection, data pre-processing, tour itinerary recommendation, experimentation and evaluation; and (ii) Second, we provide a comprehensive review of the current state-of-the-art in tour recommendation research.

1.2. Structure and Organization

The rest of the paper is structured as follows. Section 2 discusses the different sources and methods for obtaining tourist visit data. Section 3 describes various tour recommendation algorithms targetted at the individual traveller. Section 4 examines the problem of recommending tours to groups of tourists and examines various algorithms and applications that aim to fulfill this task. Section 5 studies the various methodologies that can be used to evaluate the performance of tour recommendation algorithms. Section 6 summarizes this review paper and discusses research directions for future work.

2. Methods for Retrieving Tourist Visit Data

In all tour recommendation works, one of the initial steps is to identify an appropriate data source that is representative of real-life tourist trajectories. This data source is mainly used to infer the implicit preferences of users and to evaluate the proposed tour recommendation algorithms. Typical data sources are geo-tagged photos or other social media, location-based social networks, or GPS-trajectory traces. In this section, we discuss these three types of datasets with the main focus on geo-tagged photos, which is also the most prevalent data source used in tour recommendation works.

Mining Tourist Trajectories using Geo-tagged Photos. We first discuss the mining of users' past trajectories, and cover optimization-based approaches to itinerary recommendation in the subsequent sections. Choudhury et al. [29] was one of the earliest works to study both itinerary recommendation using an optimization-based approach and the mining of users' past trajectories based on geo-tagged photos (Figure 3). Using geo-tagged Flickr photos, Choudhury et al. construct these past trajectories using the following steps:

1. *Constructing an Ordered Sequence of Relevant Photos.* The entire set of photos are first filtered to remove those that are: (i) not taken in the specific city, based on user tags containing city names; (ii) taken in the specific city but not by a tourist, based on the photo taking time-frame; and (iii) stamped with an inaccurate time taken, based on comparison between taken time and upload time. The remaining photos are then ordered in a temporal sequence.

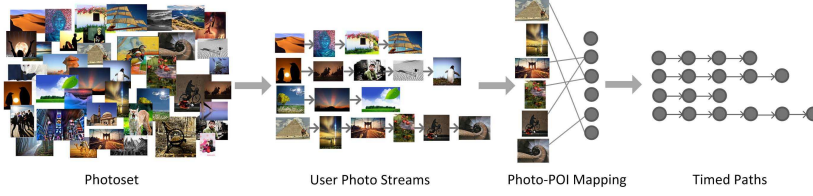


Fig. 3. Construction of tourist trajectories from geo-tagged photos. Retrieved from [29].

2. *Mapping Photos to Popular POIs.* Using a list of POIs and their the latitude/longitude coordinates, the authors map the photos to a POI if either: (i) the locations of the photo and POI differ by $\leq 100m$; or (ii) the trigram set similarity between the photo tags and POI name is above a threshold of 0.3.
3. *Generating Timed Sequences of POI Visits.* After Step 2, the authors then determine POI visit duration and POI-to-POI travel duration based on the photo timestamps, and the sequence of POI visits are divided into smaller sub-sequences if two consecutive photos are taken more than eight hours apart.

Other authors such as [78, 15, 17, 20, 69, 71, 73, 72, 60, 61, 80] also adopted variations of this approach in their tour recommendation works. Similarly, this approach can be easily adapted to other forms of social media with a geo-tagged location such as Tweets or Facebook posts. Apart from mining tourist trajectories, many authors further refine this trajectory mining problem by assigning categories to POIs and using these POI categories to determine the interests of tourists. Many of these approaches are discussed later in Section 3.2.

Mining Tourist Visits on Location-based Social Networks. Another source for obtaining tourist trajectories or visits is from Location-based Social Networks (LBSNs) such as FourSquare or JiePang. LBSNs users are able to follow one another and form friendship links, and are able to explicitly check-in to locations or venues that they have visited. These check-in locations include POIs, restaurants, businesses or general venues, which can be further divided into categories such as Food, Coffee, Nightlife, Fun and Shopping. LBSNs provide another popular source of data and have been used by numerous researchers in their tour recommendation and path planning problems [47, 59, 37, 115, 114, 110].

GPS-based Trajectory Traces. GPS-based trajectory traces are another popular data source commonly used for tour recommendation and path planning problems [117, 109, 27, 119]. These traces are typically recorded based on GPS-enabled devices such as today’s smart phones and dedicated GPS trackers. With the advent of smart phones, such GPS-based trajectory traces are increasingly common but privacy issues prevent such datasets from being publicly shared on a large-scale, unlike datasets based on public geo-tagged photos and LBSNs. Despite the restrictions on GPS-based trajectory traces, these datasets provide a very detailed record of a user’s movement trajectory based on fine-grained GPS locations, in contrast to geo-tagged photos and LBSNs that only record visits to specific POIs or locations.

3. Tour Recommendation for Individual Travellers

In this section, we first review optimization-based approaches to tour recommendation that do not include any user personalization. Thereafter, we discuss

Table 1. Survey of tour recommendation for individual travellers

Paper Reference	Popularity -based	Interest -based	Determines Interest	Constructs Itinerary	Considers Time	Transport / Traffic
[24, 115, 114]	Yes	Yes	Yes	Yes	Yes	Yes
[15, 17, 28, 71, 73, 70, 104, 110, 54]	Yes	Yes	Yes	Yes	Yes	No
[100, 98]	Yes	Yes	No	Yes	Yes	No
[106]	Yes	Yes	No	Yes	No	No
[69]	Yes	Partly	Partly	Yes	Yes	No
[78]	Yes	Partly	Partly	No	No	No
[47]	Yes	Partly	No	Yes	Yes	No
[30, 29]	Yes	No	No	Yes	Yes	No
[38]	Yes	No	No	Yes	Yes	Yes
[44]	Partly	Yes	No	Yes	Yes	Yes
[60, 61]	Partly	Yes	Yes	Yes	Yes	Yes
[79]	Partly	Yes	Yes	Yes	Yes	No
[21]	No	Yes	Partly	Yes	Yes	Yes
[85, 9]	No	Yes	No	Yes	Yes	No
[40]	No	No	No	Yes	Partly	No
[83]	No	No	No	Yes	No	No

data-driven tour recommendation approaches that include personalization based on user interests, traffic conditions, and travelling uncertainty. Table 1 presents an overview of various works on tour recommendation for individual travellers.

3.1. Optimization-based Approaches (without Personalization)

Tour recommendation has its roots in the OP and similar variants, where a key feature is that they do not incorporate any personalization for individual users. As a result, the same tour itinerary is recommended to all users, given the same starting/ending POIs and time budget as inputs.

Orienteering Problem (OP). The OP originated from a sport of the same name, where participants visit check-points with pre-determined scores, in an attempt to maximize their total score within a specific time. In recent years, many tour recommendation studies have modelled tour recommendation based on the OP and its many variants. Similarly, there have been many web applications [98, 75] developed based on variants of the OP. We first describe the original OP [96, 101, 48] and how it is applied to the field of tour recommendation.

Many tour recommendation works are focused on individual cities, each of which comprises a set of POIs P . For a tourist visiting a particular city, he/she will have considerations of a certain time or distance budget B , and preferred starting and ending POIs p_1 and p_N , respectively. The budget typically represents the amount of time that a tourist would want to spend on a tour or the distance that he/she is willing to travel. Similarly, the starting and ending POIs reflect the preferences of the tourist to start the tour near a particular point (e.g., the tourist’s hotel) and end the tour at another point (e.g., near a restaurant for dinner). Thus, given the set of POIs P , a budget B , starting POI $p_1 \in P$, destination POI $p_N \in P$, our main goal is to recommend a tour itinerary that maximizes a particular score, while adhering to the constraints of the budget, starting and destination POIs. We formally define this as recommending a tour itinerary $I = (p_1, \dots, p_N)$ that optimizes the following objective:

$$\text{Max} \sum_{i=2}^{N-1} \sum_{j=2}^N x_{i,j} \text{Score}(i) \quad (1)$$

where $x_{i,j} = 1$ if the itinerary involves travelling from POI i to j , and 0 otherwise. Such that:

$$\sum_{j=2}^N x_{1,j} = \sum_{i=1}^{N-1} x_{i,N} = 1 \quad (2)$$

$$\sum_{i=1}^{N-1} x_{i,k} = \sum_{j=2}^N x_{k,j} \leq 1, \quad \forall k = 2, \dots, N-1 \quad (3)$$

$$\sum_{i=1}^{N-1} \sum_{j=2}^N Cost(i,j) x_{i,j} \leq B \quad (4)$$

Equation 1 aims to maximize a certain score that is allocated to POIs in the recommended tour itinerary. This score is typically based on POI popularity, POI alignment to user interests, or some variation of the two. Constraint 2 ensures the tour starts and ends at specific POIs, while Constraint 3 ensures that the recommended tour itinerary comprises POIs connected as a trajectory and no POIs are re-visited. Finally, Constraint 4 ensures that all POIs in the tour itinerary can be visited within the budget B , where the function $Cost(p_x, p_y)$ determines the travelling time or distance between POI p_x and POI p_y .

Itinerary Mining Problem. Based on the OP, [29] proposed the Itinerary Mining Problem (IMP), which aims to find an itinerary that maximizes POI popularity while ensuring that touring time is within a pre-determined budget. They model POI Popularity based on the visit count by distinct tourists, transit times between POIs based on the median transit time by all tourists, and POI visit times based on the 75th percentile of visit time by all tourists. A recursive greedy algorithm [22] is used to solve the IMP, where it tries to estimate the middle node of the itinerary and the associated utility (popularity) gained and cost (time) incurred, then recursively calls itself on both halves of the itinerary.

Tour Recommendation with Specific POI Category Sequence. Gionis et al. [47] approached tour recommendation in a similar fashion as the OP, except that they: (i) consider the POI categories in their tour recommendation; and (ii) recommend tours with a specific visit order over all POI categories, e.g., Cafe \rightarrow Parks \rightarrow Beach. Apart from this constraint of a POI category visit order, the authors also consider variations of this ordering constraint, such as:

- *Partial Ordering of POI Categories.* Instead of a total ordering, this relaxed constraint allows for a partial ordering of POI categories, e.g., Cafe \rightarrow Parks is a partial ordering of Cafe \rightarrow Parks \rightarrow Museum \rightarrow Shopping.
- *Subset Grouping of POI Categories.* Instead of a specific order, this relaxed constraint allows for a subset of the ordering of POI categories, e.g., a visit order would be Cafe *OR* Restaurants \rightarrow Beach *OR* Parks, instead of the specific order Cafe \rightarrow Beach \rightarrow Parks.
- *Skipping of POI Categories.* Instead of having to visit all POI categories at least once, this relaxed constraint allows for one or more POI categories to be skipped, i.e., not visited in the recommended tour.

The authors proposed two schemes for evaluating the utility of each tour, namely: (i) an *additive satisfaction function* on the perceived benefit from visiting

a particular POI, based on either a general measure (e.g., POI popularity) or personalized measure (e.g., personal satisfaction); and (ii) a *coverage satisfaction function* that determines the number of additional, nearby POIs that can be visited during the tour, i.e., POIs within a certain distance from the tour. To solve this tour recommendation problem and the different variations of the relaxed constraints, the authors used a dynamic programming approach.

Tour Recommendation with POI Category Visit Constraints. One possible issue with typical tour recommendations is that it may include excessive visits to the same POI categories (e.g., visiting 10 museums for a tour), resulting in a “sensory overload”. To overcome this issue, Bolzoni et al. [9] proposed the CLuster Itinerary Planning (CLIP) algorithm that aims to recommend tour itineraries with constraints on the maximum number of times that each POI category can be recommended. The proposed CLIP algorithm makes extensive use of clustering and pruning techniques to reduce the search time required to generate a tour itinerary, with the following steps:

1. CLIP first uses agglomerative clustering to group POIs into k clusters using a bottom-up approach, based on the proximity of POIs.
2. To recommend a tour, CLIP generates a path starting from POI S , followed by POI cluster C_1, C_2, \dots, C_N , ending at POI D , with the assumption that the travel costs within a POI cluster are negligible.
3. After Step 2, CLIP selects a subset of individual POIs from each POI cluster, with the aim of maximizing the obtained utility score. This selection problem is modelled as a multi-dimensional knapsack problem.

3.2. Personalization-based Approaches

After discussing optimization-based approaches, we now review data-driven approaches to tour recommendation that include personalization to recommend a customized and unique tour itinerary to each tourist based on their interest preferences. In such personalization-based approaches, the key research challenges are: (i) implicitly inferring the interest preferences of tourists; and (ii) incorporating these interests as part of the recommended tour itinerary.

Tour Recommendation based on Gender, Age and Race. Cheng et al. [28] aim to recommend tours based on the current location of a user and his/her demographic details such as gender, age and race, which are automatically detected from Flickr photos using a facial detection algorithm [102]. Their tour recommendation then takes two forms, namely:

1. *Recommending Next POI.* Using the user demographic details, the recommender utilizes a Bayesian learning model that also considers the user’s current location and their learned tourist travel model based on travel sequences by other users with similar demographic attributes.
2. *Recommending Tour Itinerary.* They modelled tour recommendation as a shortest path problem from a starting POI to destination POI, while also including N other POIs with POI scores based on popularity and alignment to the user demographic profile. While there is no time or distance budget (like in typical OPs), the authors implemented a penalty function that favours shorter paths.

A later work [26] extended upon [28] by considering the size of the group

in which a user is travelling, i.e., individuals, friends, couples or families. They perform this consideration by using facial recognition techniques to detect the number of faces in a photo, thus identifying the number of travellers in a group.

TripBuilder Algorithm. Brilhante et al. [15, 17] developed the TRIPBUILDER algorithm for planning personalized tour itineraries for tourists based on the Generalized Maximum Coverage problem [31]. TRIPBUILDER aims to plan a tour comprising POIs that maximize tourists' personal interests while adhering to a specific visiting time budget. TRIPBUILDER comprises two steps:

1. *Selection of Sub-trajectories.* As part of the Trip Cover problem, the authors use an approximation algorithm to select a set of sub-trajectories among POIs that best satisfies the tourist interests and is within the specified time constraint.
2. *Joining of Sub-trajectories.* As part of the Trajectory Scheduling Problem, the sub-trajectories found in Step 1 are then joined together to form a complete tour itinerary using a local search algorithm.

The TRIPBUILDER algorithm has also been developed as a web-based application with the same name [16].

TourRecInt Algorithm. The TOURRECINT algorithm [69] aims to recommend tour itineraries with a mandatory-visit POI category c_m . In turn, this mandatory-visit POI category is based on the POI category that the tourist is most interested in, which the author defined as the most frequently visited POI category. TOURRECINT is based on an OP variant with the addition of the mandatory-visit category, which is formally defined as:

$$\sum_{i=1}^{N-1} \sum_{j=2}^N x_{i,j} \delta(Cat_i = c_m) \geq 1, \quad \forall c_m \in C \quad (5)$$

where $\delta(Cat_i = c_m) = 1$ if $Cat_i = c_m$ (POI i is of category c_m), and 0 otherwise. The optimization function and other constraints are the same as the basic OP. Other works have also studied problem variants with must-visit POIs [93].

PersTour Algorithm. The PERSTOUR algorithm [71, 73] recommends tour itineraries with POIs and visit durations tailored to the interest preferences of individual tourists. This personalization is based on both POI popularity and *time-based user interests*, which is a relative measure of user interest in a POI category based on how long a tourist visits a POI compared to the average visit duration by other tourists. Given that S_u is the POI visit history of tourist u , the *time-based user interest* of tourist u in POI category c is formally defined as:

$$Int_u^{Time}(c) = \sum_{p_x \in S_u} \frac{\bar{V}_u(p_x)}{\frac{1}{|T|} \sum_{t \in T} \bar{V}_t(p_x)} \delta(Cat_{p_x} = c), \quad \forall c \in C \quad (6)$$

where $\delta(Cat_{p_x} = c) = 1$ if $Cat_{p_x} = c$, and 0 otherwise.

The function $\bar{V}_t(p_x)$ indicates the average amount of time spent by tourist t at POI p_x , based on all the travel history of tourist t . Thereafter, the PERSTOUR algorithm attempts to recommend tour itineraries similar to that of the OP, with two main differences, namely: (i) PERSTOUR optimizes for POI popularity and

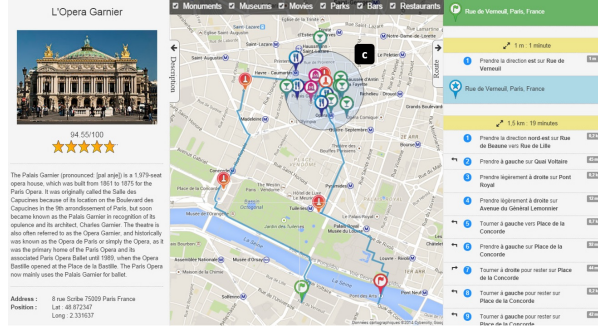


Fig. 4. Example of an itinerary generated by Aurigo, with the Pop Radius feature (blue circle) that allows users to add/delete POIs that are in close proximity. Retrieved from [106].

time-based user interest; and (ii) PERS TOUR uses a time budget based on both travelling time and a personalized POI visit duration based on user interest.

Aurigo System. Aurigo is a recommendation system that recommends personalized itineraries via an End-to-End mode and a Step-by-Step mode [106]. The End-to-End mode, like the OP, aims to recommend tours with specific starting and ending points, while maximizing POI popularity and user interests. POI popularity is determined based on Yelp review counts and ratings, while interest preferences are explicitly provided by users in the form of 1-5 star ratings on each POI category. For the Step-by-Step mode, the user first chooses a starting point then iteratively chooses the next POI to visit until he/she is satisfied with the self-constructed itinerary. The tourist is able to modify the itinerary via a Pop Radius feature (Figure 4), which shows all POIs within a specific radius to a selected POI and allows for fine-tuning the recommended itinerary.

Photo2Trip System. Lu et al. [77] developed the Photo2Trip system that utilizes 20 million geo-tagged photos and 0.2 million travelogues for the main purposes of identifying popular POIs, POI-to-POI path discovery and tour recommendation. More specifically, Photo2Trip achieves these functions by:

1. *Identifying popular attractions.* Photo2Trip used MeanShift clustering to group photos into clusters based on their location. They then picked the top 10% largest clusters and named them based on the nearest POI in the travelogues.
2. *POI-to-POI path discovery.* As a single user may not post all photos of his/her entire trajectory, Photo2Trip combine multiple fragments of photo-to-photo paths from different users into a single POI-to-POI path based on the density of the photo fragments and their actual distance.
3. *Tour recommendation.* Using the list of POIs and paths (from Steps 1 and 2), Photo2Trip then uses Dynamic programming to find an optimal (popular and interesting) tour that can be completed within a specific time budget.

Context-aware Tour Recommendation. Instead of mapping photos to known POIs, Majid et al. [79] infer the location of POIs and their semantic meaning using clustering approaches on geo-tagged photos. Their approach also infers popular travel sequences between POIs and considers the context of the tour recommendation, i.e., time, day and weather. In summary, [79] performs this context-aware tour recommendation in the following steps:

1. *Inferring POI Locations.* The P-DBSCAN algorithm [57] is first used to cluster geo-tagged photos into a set of POI locations in a city. Thereafter, user tags and Google Places data are used to determine the semantic meaning of POI.
2. *Mining frequent travel sequences.* Next, they mapped geo-tagged photos to the discovered POIs to construct travel sequences and used the re-fixSpan algorithm (partially based on [49]) for mining frequent travel patterns.
3. *Determining Weather Conditions.* Using the Wunderground API, the authors then associate each POI visit with the weather conditions (temperature, wind chill, humidity, pressure and wind speed) when the visit took place.
4. *POI and Tour Recommendation.* Finally, they utilized user-based collaborative filtering to determine POI interest scores for users, then used the joint probability of POI interest scores, time and weather to determine the likelihood of including a POI in a recommended list.

Tour Recommendation with Time-variant Interests. Instead of the OP, Yu et al. [110] proposed a tour recommendation problem with a starting POI and touring time budget, but with no consideration of a specific destination POI. One key difference between this work and others is how Yu et al. proposed the idea of time-variant interest preferences, e.g., visit tourist attractions in the morning and have lunch at a restaurant at noon. This work uses the following steps:

1. *Modelling User Interest Preferences.* Interest preferences are modelled based on six time periods throughout the day (except sleeping time from midnight to 8am), and interest levels are based on visit frequency to POI categories.
2. *Modelling POI Scores.* POI scores are derived from a combination of POI popularity (based on the number of visits to that POI in a specific month) and the POI rating (as assigned by JiePang users to that POI).
3. *POI Recommendations.* This next step involves identifying POIs that are interesting to the user and near the specific starting location, then ranking them using user-based collaborative filtering [116].
4. *Construction of Tour Itinerary.* The final step includes constructing a tree rooted starting at a specific POI and subsequent levels based on a list of top-N POIs [119] for each time period. The recommended tour itinerary is determined based on a tree traversal, where the POI-to-POI transition probability is based on user interests, POI popularity, touring time and POI-to-POI distances.

Tour Recommendation based on Time and Seasons. Compared to other tour recommendation systems, Jiang et al. [54] proposed a system that considers interest preferences, POI admission costs, POI opening hours and the visiting seasons, which they automatically obtain from geo-tagged photos and travelogue websites. Their tour recommendation system comprises the following steps:

1. *Extracting POI Statistics from Travelogues and Photos.* The authors utilize tags and description of travelogue articles to determine the various categories, admission cost and opening hours of POIs. Photo timestamps are also used to determine the visiting distribution at the POIs during the different seasons.
2. *Determining User Interest, Cost, Time and Season Preferences.* Using users' posted photos as travel sequences, the authors determine their interest preferences based on the associated tags, and cost, time and season preferences based on photo timestamps.

3. *Tour Itinerary Recommendation.* The recommended tour is personalized to individual users based on popular routes, which are filtered to match the interest, cost, time and season preferences of individual user. Using these popular routes, the authors then replace POIs with those that better match user preferences based on a variant of collaborative filtering [64, 118].

3.3. Consideration for Situational Awareness

The consideration of user interest preferences (in the previous section) is an attempt to make tour recommendations more personalized and there are various works that incorporate other real-life considerations. Other practical considerations that raise novel optimization challenges include incorporating forms of situational awareness such as multiple modes of transport, considering traffic conditions, POI crowdedness and queuing times and including uncertainty in travelling times, which we discuss next.

Recommending Tours with Consideration for Traffic Conditions. TripPlanner [24] is a traffic-aware route planning system that recommends personalized routes comprising a set of must-visit POIs while considering the traffic conditions at different POIs and times based on Foursquare check-ins and taxi GPS traces. TripPlanner operates based on three main steps:

1. *Generating a Dynamic POI Network Model.* TripPlanner uses FourSquare to determine the popularity, category, location, opening hours and visit duration at POIs. Similarly, taxi GPS traces are used to derive the travelling time between POIs based on a time-dependent traffic condition.
2. *Searching for Possible Routes.* Based on user-specified starting/destination POIs, must-visit POIs and touring time, TripPlanner searches for valid routes that satisfy these constraints. If no valid routes can be found, iteratively suggests removals from the must-see POIs, until a valid route can be found.
3. *Augmenting Routes with Preferred POIs.* If the route found in Step 2 has unused time budget, TripPlanner augments the original route with additional POIs to maximize user satisfaction based on their interest preferences.

There have also been applications that consider time-varying travelling times between POIs based on the traffic conditions as well as transport modes at the time of POI departure, such as the eCOMPASS tourist tour planner [44].

In the field of Operations Research, there are also route planning works that consider multiple transport modes and uncertain travelling times [36, 13, 12]. While these works present interesting results, they differ from our tour recommendation problem as they are mainly concerned with finding the shortest path between a starting and ending location. Similarly, researchers such as [67, 68] incorporate traffic flow predictions into their route planning problem, enabling them to recommend routes that avoid traffic congestion and hazards in advance.

Recommending Tours based on Interests and Different Transport Types.

Kurashima et al. [60, 61] proposed a method for tour recommendation that considers the current location of the user, his/her interest preferences, available time for touring and available means of transport. The authors use a combined topic and Markov model to recommend POIs that are based on a user's interest and current location. They used Probabilistic Latent Semantic Analysis (PLSA) [51]

as the topic model, which considers the interests of a user and models the probability of this user visiting a POI p given a travel history h , that is:

$$P(p|h) = \sum_{z \in Z} P(z|h)P(p|z) \quad (7)$$

where Z is the set of topics for the POIs, $P(z|h)$ is the probability of a user being interested in topic z , and $P(p|z)$ is the probability that POI p is selected from topic z .

For the Markov model, the authors employ a first-order Markov model, where the probability of visiting a POI p_t depends on a previous POI visit p_{t-1} . This is formally defined as:

$$P(p_t|p_{t-1}) = \frac{N(p_{t-1}, p_t)}{N(p_{t-1})} \quad (8)$$

where $N(p_{t-1}, p_t)$ is the frequency that POI p_t is visited after a prior visit to POI p_{t-1} , and $N(p_{t-1})$ is the total visit count at POI p_{t-1} .

The authors then combine the topic and Markov models as one single model, using the following formula:

$$P(p_t|p_{t-1}, h) = \frac{P(p_t|p_{t-1})}{C(p_{t-1}|h)} \frac{P(p_t|h)}{P(p_t)} \quad (9)$$

where $C(p_{t-1}|h)$ is a normalization factor based on unigram rescaling [46]. This combined topic and Markov model then recommends the next POI to visit based on the user's current location and his/her interest preferences. To recommend a tour itinerary, the authors use a best-first search algorithm to select tour itineraries with the highest probability and adhere to the available touring time.

Recommending Tours with Transport-cost Awareness. Unlike other works that consider transportation based on traffic conditions or transport modes, [38] included the consideration for the transport cost associated with a recommended tour. The authors utilized geo-tagged photos to determine the popularity of POIs, visit durations, opening hours and appropriate visual scenes to display at different times of the day. They modelled their tour recommendation problem based on a variant of the NP-hard Vehicle Routing Problem with Time Window [14]. In this problem, their main aim is to visit the largest number of popular POIs and ensure that the routes taken are minimal and smooth (i.e., no long detours), while adhering to the available touring time and transportation cost budget.

Recommending Tours based on Interests, POI Opening Hours and Travel Time Uncertainty. Zhang et al. [115, 114] studied tour recommendation with the goal of recommending personalized itineraries based on the interest preferences of users and available touring time, while considering opening hours of POIs and uncertainty in travelling time. Their work involves the following:

- *Modelling User Interest Preferences.* The authors modelled interest preferences based on user ratings on POI features, instead of specific POIs, using their proposed feature-centric collaborative filtering approach [115].
- *Modelling Travel Time Uncertainty.* Uncertainty in travelling time is represented as a random variable associated with a particular probability distribution. This work aims to recommend tours with a high *completion probability*, i.e., the likelihood of completing an itinerary within the time budget.

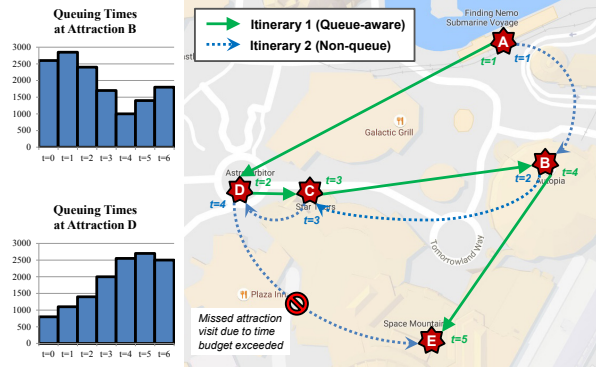


Fig. 5. Example of queue-aware tour itinerary recommendation. Retrieved from [70].

- *Modelling POI Opening Hours.* The authors consider POI opening hours by implementing a POI availability constraint, and adopt a tour recommendation approach by iteratively adding a POI to an itinerary if it satisfies this constraint, along with the total time budget constraint.

Personalized Tours with Queuing Time Awareness. In tourist attractions like theme parks, the queuing times are an important considerations when recommending a tour itinerary. Figure 5 illustrates an example where Itinerary 2 results in incomplete visits due to excessive queuing times, while Itinerary 1 is able to complete all visits due to considering for queuing times. [70] studied this queue-aware tour recommendation problem based on an OP variant with time-dependent queuing times, and travel costs comprising queue times, ride durations and travel times. To solve this problem, the authors proposed the PERSQ algorithm, adapted from Monte Carlo Tree Search [33, 18], which involves:

- **Selection.** PERSQ initializes at a starting POI s and iteratively *selects* a next POI n to expand based on a strategy of *exploring* unvisited POIs and *exploiting* POIs with high reward, where the reward is based on a heuristic of most popular/interesting POIs with the shortest distance/queuing-time.
- **Expansion.** If selected POI n is not the destination POI or exceeds the total time budget, expand POI n and randomly select one of the unvisited POIs.
- **Simulation.** Steps 1 and 2 are then iteratively *simulated* until either the total time budget is exhausted or the specified destination POI is reached.
- **Back-propagation.** During this step, the current itinerary is the set of all POIs visited and the reward for this itinerary is calculated then *back-propagated* to all the visited POI during this iteration.

Steps 1 to 4 are considered one iteration of PERSQ, and PERSQ is then repeated for either a fixed number of iterations or for a fixed time period. At the end of these multiple iterations, there will be multiple itineraries being explored and the recommended itinerary is the one with the highest reward.

Personalized Crowd-aware Tour Recommendation. Wang et al. [104] studied personalized crowd-aware tour recommendation, with the main objective of planning itineraries that optimize the conflicting objectives of POI popularity, user interests and POI crowdedness. The authors determine POI popularity

based on past visit frequency, user interests using user-based collaborative filtering [113], and POI crowdedness from a pedestrian sensor dataset. Using a variant of the Ant Colony Optimization algorithm [34], the authors proposed the PERSCT algorithm for solving this problem. PERSCT utilizes a number of agents and works in two main steps:

1. **Route Initialization.** Agents start at a specific POI and iteratively visit next POIs until reaching the destination POI. The selection of next POI favours those with higher profits, nearer to the destination, and recently selected.
2. **Route Update.** After the previous step and all agents have generated itineraries, the itinerary with the highest profit (based on POI popularity, user interests and POI crowdedness) is recorded before being de-emphasized over time.

Steps 1 and 2 are repeated a fixed number of times and PERSCT will lean towards itineraries with high profits that were recently used by agents. Eventually, the recommended tour itinerary is the one with the highest profit based on the joint objectives of POI popularity, user interests and POI crowdedness.

3.4. Other Approaches

While user personalization and real-life traffic/transport considerations are important, there are tourists that emphasize different aspects of travelling, such as recommending routes that may not be the most popular but are the most scenic or safest. In the following sections, we present some of these works.

Recommending beautiful, quiet and happy routes. Instead of popular or personalized routes, [83] recommended routes that are emotionally pleasing, i.e., beautiful, quiet and happy. They use crowd-sourcing to identify photos that are beautiful, quiet or happy, with photos corresponding to specific locations with their perceived scores. Thereafter, they recommend tours as follows:

1. Given a starting point s and destination d , use Eppstein’s algorithm [35] to recommend the M shortest routes from s to d , where M is set to be arbitrarily large (e.g., 1 million) such that it covers all possible routes.
2. Given $k \leq M$, calculate the average beauty, quietness or happiness rank for each of the top- k routes and record the route with the best rank. Instead of all M routes, the intuition is to iteratively explore a smaller set of k routes.
3. Repeat Step 2 for the next k routes and identify the route with the best rank. This step repeats until the improvement is less than a threshold ϵ , and the last route is recommended as the most beautiful, quietest or happiest route.

Quercia et al. also utilized the user tags of photos to determine how beautiful, quiet and happy each photo is, based on the Linguistic Inquiry Word Count dictionary [81]. Similarly, the ScenicPlanner system [23] also used Flickr photos and FourSquare check-ins to determine scenic scores for individual road segments as part of a larger scenic route planning problem. Other researchers have also focused on other aspects of non-touristic tour recommendations, such as [40, 39, 55] who used crime statistics for recommending short but safe paths.

Random Walks with Restart. Also using geo-tagged photos, Lucchese et al. [78] recommended POI visits using random walks on a graph-based representation of past tourist trajectories. This algorithm comprises the following steps:

1. The authors first construct an itinerary graph $G = (P, E, W)$, where P is the set of all POIs, E is the set of edges representing co-visits to two POIs, and W is the edge weight based on unique visitors of that POI pair.
2. Itinerary graph G is then transformed into an itinerary transition matrix, with the transition probability between POIs. The authors then use the Random Walk with Restart algorithm [94] to compute the steady-state probability distribution for the set of POIs previously visited by the tourist.
3. Using the itinerary transition matrix, tourist POI visit history and its steady-state probability distribution, they calculate the scores of the unvisited POIs based on the product of entries in the steady-state probability distribution. Finally, the algorithm then recommends the top- k POIs with the highest scores, which can be constructed as a connected itinerary.

3.5. Web and Mobile-based Applications

In this section, we examine various web and mobile-based applications for recommending personalized tours. Unlike the personalization-based approaches that infer interest preferences (previous section), these applications employ user interfaces to explicitly solicit interest preferences from tourists before using these interest preferences to recommend personalized tour itineraries.

City Trip Planner. This system was proposed by Vansteenwegen et al. [100] for recommending personalized tours in five cities in Flanders, Belgium (Antwerp, Bruges, Ghent, Leuven and Mechlin) based on user-provided interest preferences. The City Trip Planner works in the following steps: (i) soliciting trip constraints such as the trip duration, starting/ending locations and break timings; (ii) estimating user interests using the Vector Space Model [4, 89] on the user-provided interest levels in various POI categories; (iii) recommending tours using a Greedy Randomised Adaptive Search Procedure [99]; and (iv) altering the recommended tour based on user feedback, i.e., removing specific POIs.

myVisitPlanner^{GR}. *myVisitPlanner^{GR}* is a web-based application that is targeted at recommending touristic activities to visitors of Northern Greece [85]. There are three main steps to using the *myVisitPlanner^{GR}* system, namely: (i) getting users to provide their demographics details and trip constraints; (ii) suggesting activities using a hybrid recommender based on a variant of collaborative filtering on user-rated activities; and (iii) recommending an itinerary of suggested activities using the scheduling engine of *SelfPlanner* [84, 86]. This system uses an ontological approach to represent activities where there are multiple hierarchical levels. Activity providers have the flexibility to describe their activity at a higher, more general hierarchical level or a lower, more specific level. In addition, the system keeps track of the ratings that users assign to the various activities and uses these ratings as input to their hybrid recommender system.

SAMAP System. The authors of [21] proposed the *SAMAP* system for recommending and planning a personalized daily itinerary that considers his/her user profile and available touring time. *SAMAP* was designed for mobile devices such as smart phones and operates as a multi-agent system, comprising the following agents: (i) first, an *interface agent* that solicits interests, trip preferences and personal information from the tourist; (ii) next, a *user modelling agent* builds a

Table 2. Survey of Tour Recommendation for Groups of Tourists

Paper Reference	Popularity-based	Interest-based	Determines Interest	Constructs Itinerary	Considers Time	Transport/Traffic	Tour Guides
[72]	Yes	Yes	Yes	Yes	Yes	No	Yes
[26]	Yes	Yes	Yes	Yes	Yes	No	No
[2, 87]	Yes	Yes	Yes	Yes	Yes	No	No
[42, 43]	Yes	Yes	Partly	Yes	Yes	No	No
[3]	No	Yes	No	Yes	Yes	Yes	No
[53]	No	Yes	No	No	No	No	No

model of the user to pass to the next agent; (iii) a *case-based reasoning agent* then identifies a set of POIs that are aligned to the user’s interest, based on the preferences of similar users; and (iv) finally, a *planning agent* schedules an itinerary comprising a subset of the earlier POIs that maximizes the tourist’s utility score, while accounting for POI opening hours and transport modes between POIs.

Thus far, we have examined tour itinerary recommendation for individual travellers and covered optimization-based approaches, personalization-based approaches, and web and mobile-based applications. In real-life, people frequently travel in groups of varying sizes, such as couples, family and friends, and we discuss such works in the next section.

4. Tour Recommendation for Groups of Tourists

Tour recommendation research typically focus on the single traveller, as seen in Section 3. In real-life however, tours frequently involve multiple travellers such as couples, friends or families, which are challenging due to the need to appeal to multiple travellers within the same group. In the next section, we examine some early efforts on resolving this group tour recommendation problem. Table 2 presents a broad overview of the various group tour recommendation works.

Group Tours with Tour Guides. Lim et al. introduced the Group Tour Recommendation (GROUPTOURREC) problem, where the main aim is to recommend tours that satisfy groups of tourists with diverse interest preferences [72]. They solve the GROUPTOURREC problem by decomposing the problem into more manageable sub-problems of tourist grouping, POI recommendation and tour guide assignment. Given the set of tourists $T = \{t_1, \dots, t_l\}$, tour guides $U = \{u_1, \dots, u_m\}$, tour groups $G = \{g_1, \dots, g_m\}$, and POIs $P = \{p_1, \dots, p_n\}$, the main goal of GROUPTOURREC is to optimize the following function:

$$\begin{aligned}
Max \alpha \sum_{g \in G} \sum_{t \in T} \sum_{p \in P} x_{t,g} y_{g,p} & \left(\eta Int_t(Cat_p) + (1 - \eta) Pop(p) \right) \\
& + (1 - \alpha) \sum_{g \in G} \sum_{u \in U} \sum_{p \in P} z_{u,g} y_{g,p} Ept(u, p)
\end{aligned} \tag{10}$$

where $x_{t,g} = 1$ if tourist t is assigned to group g , $y_{g,p} = 1$ if group g is recommended POI p , $z_{u,g} = 1$ if tour guide u is assigned to group g , and 0 otherwise.

In short, the main objectives of Equation 10 are to find optimal values for: (i) tourist allocation to tour group, i.e., $x_{t,g}$; (ii) POI recommendation to tour group, i.e., $y_{g,p}$; and (iii) tour guide assignment to tour group, i.e., $z_{u,g}$. GROUPTOURREC is divided into more manageable sub-problems and solved by: (i) using

k-means to cluster users into tour groups based on their interest preferences; (ii) using an OP variant for recommending itineraries to tour groups based on their group interest; and (iii) using integer programming to assign tour guides to tour groups based on the guide expertise and recommended itineraries.

Group Tours with Pre-assigned Groups. Anagnostopoulos et al. [2] also proposed similar group tour recommendation problems, focusing on recommending tour itineraries that best satisfy a group of tourists, which is pre-determined in advance. These are termed the TOURGROUPSUM, TOURGROUPMIN and TOURGROUPFAIR problems, which differ based on the objective function to be optimized. Instead of solving these problems as an Integer Linear Program like [72], they utilized greedy heuristics and Ant Colony Optimization to solve their group tour recommendation problems. Another difference is that [2] did not consider the assignment of tour guides to lead each tour group, but they consider multiple forms of optimization objectives to maximize the overall group interest (TOURGROUPSUM), interest of the least satisfied user (TOURGROUPMIN) and fairness among all members of a group (TOURGROUPFAIR).

e-Tourism System. E-TOURISM [42, 43] is a system that aims to recommend interesting activities (including POI visits) to either individuals or groups of tourists. There are several main steps in using the *e-Tourism* system, namely:

1. *Provide Tourist Profile and Groupings.* The individual traveller first provides his/her demographic details and interest preferences, and additionally, groups of tourists need to explicitly state the members of their group.
2. *Recommendation to Individuals and Groups.* E-TOURISM recommends a list of activities or POIs to individual tourists, using approaches based on demographics [19], content [19], general likes-based filtering [50] or a hybrid. For groups of tourists, a group interest preference is calculated using aggregation or intersection techniques on the preferences of individual tourists in that group.
3. *Tourist Feedback.* Each tourist is able to rate individual items in the recommended itineraries, which is used to improve future recommendations.

One unique characteristic of *e-Tourism* is their representation of tourist interest preferences, which are based on a hierarchical taxonomy of features instead of explicit POI categories. This approach provides a general representation of items and allows the system to be easily generalized to other application domains.

Intrigue System. Interactive TouRist Information GUIdE (or INTRIGUE) is a web and mobile based system that aims to recommend tours to both individuals and groups of tourists [3]. For a group of tourists, the usage scenario for *Intrigue* is that a specific (lead) tourist will use the system to: (i) indicate the number of tourists in that tour group; (ii) manually specify the sub-group that each tourist belongs to, based on the tourist demographics (e.g., age and background) and interest preferences; and (iii) enter details about each sub-group via a registration form (an optional step). The main idea is that a large tour group could be divided into smaller homogeneous sub-groups, e.g., a particular sub-group could be defined by the characteristics of ages ranging from 30 to 40 years old, backgrounds in engineering, and interests in architecture and museums. The tour recommendation occurs in an iterative fashion that involves the tourist indicating their trip constraints and explicitly adding recommended POIs, before INTRIGUE schedules an itinerary based on selected POIs and trip constraints.

Travel Decision Forum. The *Travel Decision Forum* [53] was one of the earlier applications that focused on group recommendations in the context of tour planning. This system utilizes group-oriented interfaces and virtual agents to enhance mutual awareness among group members for soliciting interest preferences to resolving conflicted preferences. Some key features of this application are:

- *Solicitation of Interest Preferences.* Interest preferences are solicited in the form of user ratings and importance, which are viewable by all group members.
- *Generation of Proposal.* Solutions are generated using either average/median rating, random choice or a non-manipulable, joint-rating mechanism [32].
- *Discussion of Proposal.* Each user can either accept the proposed solution or discuss with other users to come to a mutually-agreeable solution.

While *Travel Decision Forum* is used for tourism purposes, its main purpose is to solicit and de-conflict group interest preferences using interaction-based techniques. Additional work is required to translate these group interests into the recommendation of relevant POIs as a group tour itinerary.

Top-k group recommendations. Closely related to the group tour recommendation problem are the problems of top-k recommendations for groups and group recommendation, where the main objective is to recommend a ranked list or set of items that are of relevance to a group of users. These problems typically focus on retail items such as films, songs or books, with [87] recommending sets of POIs as items but not in the form of an itinerary. More specifically, [87] examined the problem of group formation such that the constructed group comprises members who are more likely to prefer a top-k item recommendation. Others have studied a related problem where the group members are pre-determined, and the objective is to model a collective group preference and/or recommend items to each group. For example, [52] proposed an algorithm for representing group preferences as a set of high level features that are not biased towards specific individuals, using collective deep belief networks and dual-wing restricted Boltzmann machines. Others like [1] tried to derive a group consensus score that maximizes item satisfaction for all members of a group, while minimizing the level of disagreement among members of the group. Yuan et al. [111] proposed a probabilistic model for group recommendations that accounts for group members with different levels of influence and how user preferences change when acting as an individual compared to being a member of a group.

While these works are targetted at groups of users, their main objective is to recommend either a ranked list of items or a set of items, which are typically retail items and merchandise. Although these works can be adapted to tourism by treating individual POIs as items, they do not recommend these POIs as a connected itinerary nor do they consider the various spatial and temporal constraints that are associated with tour planning. For a more comprehensive review on group recommendation research, we refer the interested reader to [10].

5. Evaluation Strategies

A key process in tour recommendation research is the evaluation of the recommended tour itineraries, namely how well they satisfy the requirements of the individual tourists. However, there are many interpretations of these requirements, thus leading to a variety of evaluation strategies used by the various works in

this area. In this section, we aim to highlight the various forms of evaluation strategies, and discuss the advantages and disadvantages for each of them.

5.1. Real-life Evaluations

We classify an evaluation strategy as a real-life evaluation if the recommended itineraries are compared against the real-life travel history of a tourist. As discussed in Section 2, these real-life travel histories can be obtained from various sources, namely: (i) geo-tagged photos; and (ii) location-based check-ins. For both of these sources, the real-life visits of tourists can be narrowed down to individual POIs, which allow a researcher to compare the POIs in a recommended itinerary against the real-life POI visits by tourists. To facilitate such a comparison, various Information Retrieval (IR) based metrics are used, such as:

1. **Precision.** The proportion of recommended POIs that were also visited by the tourist in real-life.
2. **Recall.** The proportion of POIs visited by the tourist in real-life that were also recommended.
3. **F1-score.** The harmonic mean of both precision and recall.

Other definitions of Precision, Recall and F1-score include variants to measure how well the categories of POIs in the recommended tour reflect the POI categories that were visited in real-life, i.e., how well the recommendations match real-life user preferences [15, 17]. Similarly, Chen et al. [25] used variants that account for POI visit orders in real-life, while others [71, 73] used root-mean-square error to determine the variations between recommend POI visit durations and real-life visits.

5.2. Heuristic-based Evaluations

In cases where the real-life travel histories of tourists are not available, or as a supplemental analysis to the IR-based metrics introduced in Section 5.1, heuristic-based metrics are often used to evaluate the effectiveness of the recommended itineraries. Some examples of heuristic-based metrics are:

1. **Total POIs Recommended.** The total number of POIs recommended to a tourist as part of an itinerary.
2. **POI Popularity.** The summation of popularity scores of all POIs recommended to a tourist, i.e., total tour popularity.
3. **Tourist Interests.** The summation of interest alignment scores of all POIs recommended to a tourist, i.e., total tour interest.

Apart from using the summation of POI popularity and tourist interests, possible variations include using other statistical values based on the average, median, minimum, maximum or other percentile/quartile values. Particularly for group tour recommendation, the minimum and maximum tour interest values would respectively reflect the least and most satisfied tourist in a group. Novelty is another important aspect of recommendations, thus another measure would be based on how new and interesting recommended POIs are.

5.3. Crowd-based Evaluations and User Studies

In contrast to the quantitative measures used in the real-life and heuristic-based evaluations (covered in the previous sections), an alternative evaluation methodology is to utilize qualitative measures, which are typically smaller in scale but more detailed in scope. Examples of such qualitative measures are:

- **User Studies.** This evaluation involves a small number of experiment participants using the proposed tour recommendation system and other baseline systems, before answering a survey based on subjective criteria, such as the usability of the system. For example, [106] utilized user studies of 10 participants comparing their Aurigo system against the baseline Google Maps.
- **Crowd-based Evaluations.** Crowdsourcing services, like Amazon Mechanical Turk (AMT), are another popular platform for evaluation purposes with more focus on the recommendation results, i.e., the tour itineraries. In contrast, user studies focus on the user experience in using the tour recommendation system. For example, [30, 29] used AMT workers to evaluate their recommended tour itineraries against baseline itineraries by tour agencies.

One key advantage of these evaluations is that they provide us with information regarding active user volumes and user feedback, thus serving as real-life indicators of how successful a specific system or algorithm is.

5.4. Online Controlled Experiments

Online companies such as Google, Microsoft and Yahoo! frequently use online controlled experiments to evaluate the effects of website user interface and design changes on live users in a real-life setting [92, 58, 74, 63]. Such experiments involve showing a specific design or algorithmic variant to a user group, while an alternative variant is shown to another user group. Similarly, tour recommendation systems can utilize online controlled experiments on a smaller subset of users before introducing new features. The evaluation of these features could include:

- **Design-based Variants.** The evaluation of changes to a website user interface, which could include the way recommended tours are displayed or the design of a form to solicit user information.
- **Algorithm-based Variants.** The evaluation of changes to the underlying tour recommendation algorithms, e.g., comparing a Naive Bayes recommender against a popularity-based recommender such as in [56].

More importantly, online experiments serve as a form of implicit user feedback, which allows us to determine the interest level of users based on their interactions with the tour recommendation systems.

6. Conclusion and Future Directions

We have provided a comprehensive review of the literature in the area of tour itinerary recommendation and highlighted the key differences between tour itinerary

recommendation and the related areas of Operations Research, next-location prediction, top-k location recommendation and travel package/region recommendation. We developed a taxonomy to describe general touring-related research, with a detailed breakdown of tour itinerary recommendations based on various real-life considerations such as POI popularity, user interests, time constraints, user demographics, transport modes and traffic conditions. In addition to reviewing a large selection of tour itinerary recommendation problems and solutions, we also discuss the various types of datasets (geo-tagged social media, location-based social networks, and GPS trajectory traces) and evaluation methodologies (real-life and heuristic-based metrics, user studies and online experiments) that can be employed in tour itinerary recommendation research.

Based on our survey, we observed a trend of tour recommendations that originated from optimization approaches and moving towards personalized and context-aware approaches with the prevalence of big social data. Although tour itinerary recommendation has been well-studied in recent years, there still remain interesting research directions to explore. Moving forward, we highlight future directions that consider various new context and personalization, such as:

- ***Consideration of Transport Modes.*** Future tour itinerary recommendation research can consider multiple modes of transport (e.g., walking, bus, train, taxi, car), instead of a single type of transport, with an objective of minimizing changing and waiting time when switching between transport modes. Another research direction is incorporate constraints on transport modes to cater to different demographics group, e.g., adults (all transport modes), families with babies or the elderly (all transport modes except walking).
- ***Dynamic Tour Recommendation.*** As environment may change during the course of a tour, another possibility for future work is to develop dynamic tour recommendation algorithms that adapt to these changing contexts during a pre-planned tour, e.g., bad weather, human fatigue, traffic congestion. For example, a tourist completed half of a recommended tour itinerary before it started to rain, in which case the algorithm should modify the remaining tour itinerary to include POIs that are indoor and linked by shelter.
- ***Country-specific Tour Recommendation.*** Politics, economy and culture are also important considerations for a tourist to visit a specific city or country, e.g., tourists might prefer visa-free countries or countries speaking the same language. Thus, an interesting future direction would be to consider the home country of the tourist, before recommending tours in countries that matches his/her preferences for certain political, economical and cultural status.
- ***Explicit Feedback and Improvements.*** Another possible future enhancement is to incorporate mechanisms for obtaining explicit feedback (e.g., user ratings) or implicit signals (e.g., length of time at POIs), which can serve: (i) as an evaluation metric to measure user satisfaction; (ii) to build a more accurate interest model for returning user; and (iii) to improve future tour recommendations by avoiding POIs with negative feedback.
- ***Sentiment and Activity-based User Interests.*** Current tour recommendations typically model user interests based on visit counts to specific POIs or POI categories, but do not consider the sentiments towards these POIs or the type of activities. Future work can utilize sentiment analysis or topic modelling techniques to better model sentiments and activities at specific POIs, leading to a more fine-grained model of user interests for tour recommendations.

- **Cold-start Tour Recommendation.** For the widespread adoption of automated tour recommendation systems, there is a need to address the cold-start problem of determining the interest preferences of new users with no previous travel history. Future work can infer interest preferences of users based on other sources of information, e.g., demographic details or interests of friends.

References

- [1] S. Amer-Yahia, S. B. Roy, A. Chawlat, G. Das, and C. Yu. Group recommendation: Semantics and efficiency. In *Proc. of VLDB'09*, pages 754–765, 2009.
- [2] A. Anagnostopoulos, R. Atassi, L. Becchetti, A. Fazzone, and F. Silvestri. Tour recommendation for groups. *Data Mining and Knowledge Discovery*, pages 1–32, 2016.
- [3] L. Ardissono, A. Goy, G. Petrone, M. Segnan, and P. Torasso. Intrigue: personalized recommendation of tourist attractions for desktop and hand held devices. *Applied Artificial Intelligence*, 17(8-9):687–714, 2003.
- [4] R. Baeza-Yates and B. Ribeiro-Neto. *Modern information retrieval*. Addison-Wesley, 1999.
- [5] J. Bao, D. W. Yu Zheng, and M. Mokbel. Recommendations in location-based social networks: A survey. *GeoInformatica*, 19(3):525–565, 2015.
- [6] R. Baraglia, C. I. Muntean, F. M. Nardini, and F. Silvestri. Learnnext: learning to predict tourists movements. In *Proc. of CIKM'13*, pages 751–756, 2013.
- [7] I. Benouaret and D. Lenne. A composite recommendation system for planning tourist visits. In *Proc. of WI'16*, pages 626–631, 2016.
- [8] I. Benouaret and D. Lenne. A package recommendation framework for trip planning activities. In *Proc. of RecSys'16*, pages 203–206, 2016.
- [9] P. Bolzoni, S. Helmer, K. Wellenzohn, J. Gamper, and P. Andritsos. Efficient itinerary planning with category constraints. In *Proc. of SIGSPATIAL'14*, pages 203–212, 2014.
- [10] L. Boratto and S. Carta. State-of-the-art in group recommendation and new approaches for automatic identification of groups. In *Information Retrieval and Mining in Distributed Environments*, pages 1–20. Springer Berlin Heidelberg, 2011.
- [11] J. Borràs, A. Moreno, and A. Valls. Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, 41(16):7370–7389, 2014.
- [12] A. Botea and S. Braghin. Contingent versus deterministic plans in multi-modal journey planning. In *Proc. of ICAPS'15*, pages 268–272, 2015.
- [13] A. Botea, E. Nikolova, and M. Berlingerio. Multi-modal journey planning in the presence of uncertainty. In *Proc. of ICAPS'13*, pages 20–28, 2013.
- [14] O. Braysy and M. Gendreau. Vehicle routing problem with time windows. *Transportation Science*, 39(1):104–118, 2005.
- [15] I. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso. Where shall we go today? Planning touristic tours with TripBuilder. In *Proc. of CIKM'13*, pages 757–762, 2013.
- [16] I. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso. TripBuilder: A tool for recommending sightseeing tours. In *Proc. of ECIR'14*, pages 771–774, 2014.
- [17] I. R. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso. On planning sightseeing tours with TripBuilder. *Information Processing & Management*, 51(2):1–15, 2015.
- [18] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton. A survey of monte carlo tree search methods. *IEEE Tran. on Computational Intelligence and AI in Games*, 4(1):1–43, 2012.
- [19] R. Burke. Hybrid web recommender systems. In *The Adaptive Web*, pages 377–408. Springer Berlin/Heidelberg, 2007.
- [20] G. Cai, K. Lee, and I. Lee. Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos. *Expert Systems with Applications*, 94:32–40, 2018.
- [21] L. Castillo, E. Armengol, E. Onaindía, L. Sebastián, J. González-Boticario, A. Rodríguez, S. Fernández, J. D. Arias, and D. Borrajo. SAMAP: An user-oriented adaptive system for planning tourist visits. *Expert Systems with Applications*, 34(2):1318–1332, 2008.
- [22] C. Chekuri and M. Pal. A recursive greedy algorithm for walks in directed graphs. In *Proc. of FOCS'05*, pages 245–253, 2005.
- [23] C. Chen, X. Chen, Z. Wang, Y. Wang, and D. Zhang. Scenicplanner: planning scenic travel

- routes leveraging heterogeneous user-generated digital footprints. *Frontiers of Computer Science*, 11(1):61–74, 2017.
- [24] C. Chen, D. Zhang, B. Guo, X. Ma, G. Pan, and Z. Wu. TripPlanner: Personalized trip planning leveraging heterogeneous crowdsourced digital footprints. *IEEE Tran. on Intelligent Transportation Systems*, 16(3):1259–1273, 2015.
 - [25] D. Chen, C. S. Ong, and L. Xie. Learning points and routes to recommend trajectories. In *Proc. of CIKM'16*, pages 2227–2232, 2016.
 - [26] Y.-Y. Chen, A.-J. Cheng, and W. H. Hsu. Travel recommendation by mining people attributes and travel group types from community-contributed photos. *IEEE Tran. on Multimedia*, 15(6):1283–1295, 2013.
 - [27] Z. Chen, H. T. Shen, and X. Zhou. Discovering popular routes from trajectories. In *Proc. of ICDE'11*, pages 900–911, 2011.
 - [28] A.-J. Cheng, Y.-Y. Chen, Y.-T. Huang, W. H. Hsu, and H.-Y. M. Liao. Personalized travel recommendation by mining people attributes from community-contributed photos. In *Proc. of MM'11*, pages 83–92, 2011.
 - [29] M. D. Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, and C. Yu. Automatic construction of travel itineraries using social breadcrumbs. In *Proc. of HT'10*, pages 35–44, 2010.
 - [30] M. D. Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, and C. Yu. Constructing travel itineraries from tagged geo-temporal breadcrumbs. In *Proc. of WWW'10*, pages 1083–1084, 2010.
 - [31] R. Cohen and L. Katzir. The generalized maximum coverage problem. *Information Processing Letters*, 108(1):15–22, 2008.
 - [32] V. Conitzer and T. Sandholm. Complexity of mechanism design. In *Proc. of UAI'02*, pages 103–110, 2002.
 - [33] R. Coulom. Efficient selectivity and backup operators in monte-carlo tree search. In *Proc. of CCG'06*, pages 72–83, 2006.
 - [34] M. Dorigo, M. Birattari, and T. Stützle. Ant colony optimization. *IEEE Computational Intelligence Magazine*, 1(4):28–39, 2006.
 - [35] D. Eppstein. Finding the k shortest paths. *SIAM J. on Computing*, 28(2):652–673, 1998.
 - [36] J. E. Flórez, A. T. A. de Reyna, C. L. L. Javier García, A. G. Olaya, and D. Borrajo. Planning multi-modal transportation problems. In *Proc. of ICAPS'11*, pages 66–73, 2011.
 - [37] Z. Friggstad, S. Gollapudi, K. Kollias, T. Sarlos, C. Swamy, and A. Tomkins. Orienteering Algorithms for Generating Travel Itineraries? In *Proc. of WSDM'18*, pages 180–188, 2018.
 - [38] C.-Y. Fu, M.-C. Hu, J.-H. Lai, H. Wang, and J.-L. Wu. Travelbuddy: Interactive travel route recommendation with a visual scene interface. In *Proc. of MMM'14*, pages 219–230, 2014.
 - [39] E. Galbrun, K. Pelechris, and E. Terzi. Safe navigation in urban environments. In *Proc. of UrbComp'14*, 2014.
 - [40] E. Galbrun, K. Pelechris, and E. Terzi. Urban navigation beyond shortest route: The case of safe paths. *Information Systems*, 57:160–171, 2016.
 - [41] H. Gao, J. Tang, and H. Liu. Exploring social-historical ties on location-based social networks. In *Proc. of ICWSM'12*, pages 114–121, 2012.
 - [42] I. Garcia, L. Sebastia, and E. Onaindia. On the design of individual and group recommender systems for tourism. *Expert Systems with Applications*, 38(6):7683–7692, 2011.
 - [43] I. Garcia, L. Sebastia, E. Onaindia, and C. Guzman. A group recommender system for tourist activities. In *Proc. of EC-WEB'09*, pages 26–37, 2009.
 - [44] D. Gavalas, V. Kasapakis, C. Konstantopoulos, G. Pantziou, N. Vathis, and C. Zaroliagis. The ecompass multimodal tourist tour planner. *Expert Systems with Applications*, 42(21):7303–7316, 2015.
 - [45] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou. A survey on algorithmic approaches for solving tourist trip design problems. *J. of Heuristics*, 20(3):291–328, 2014.
 - [46] D. Gildea and T. Hofmann. Topic-based language models using em. In *Proc. of EuroSpeech'99*, 1999.
 - [47] A. Gionis, T. Lappas, K. Pelechris, and E. Terzi. Customized tour recommendations in urban areas. In *Proc. of WSDM'14*, pages 313–322, 2014.
 - [48] A. Gunawan, H. C. Lau, and P. Vansteenwegen. Orienteering problem: A survey of recent variants, solution approaches and applications. *Euro. J. of Operational Research*, 255(2):315–332, 2016.
 - [49] J. Han, J. Pei, B. Mortazavi-Asl, H. Pinto, Q. Chen, U. Dayal, and M. C. Hsu. Prefixspan:

- Mining sequential patterns efficiently by prefix-projected pattern growth. In *Proc. of ICDE'01*, pages 215–224, 2001.
- [50] U. Hanani, B. Shapira, and P. Shoval. Information filtering: Overview of issues, research and systems. *User Modeling and User-Adapted Interaction*, 11(3):203–259, 2001.
 - [51] T. Hofmann. Probabilistic latent semantic indexing. In *Proc. of SIGIR'99*, pages 50–57, 1999.
 - [52] L. Hu, J. Cao, G. Xu, L. Cao, Z. Gu, and W. Cao. Deep modeling of group preferences for group-based recommendation. In *Proc. of AAAI'14*, pages 1861–1867, 2014.
 - [53] A. Jameson, S. Baldes, and T. Kleinbauer. Enhancing mutual awareness in group recommender systems. In *Proc. of ITWP'03*, 2003.
 - [54] S. Jiang, X. Qian, T. Mei, and Y. Fu. Personalized travel sequence recommendation on multi-source big social media. *IEEE Tran. on Big Data*, 2(1):43–56, 2016.
 - [55] A. Keler and J. D. Mazimpaka. Safety-aware routing for motorised tourists based on open data and vgi. *J. of Location Based Services*, 10(1):64–77, 2016.
 - [56] J. Kiseleva, M. J. Mueller, L. Bernardi, C. Davis, I. Kovacek, M. S. Einarsen, J. Kamps, A. Tuzhilin, and D. Hiemstra. Where to go on your next trip?: Optimizing travel destinations based on user preferences. In *Proc. of SIGIR'15*, pages 1097–1100, 2015.
 - [57] S. Kisilevich, F. Mansmann, and D. Keim. P-dbscan: a density based clustering algorithm for exploration and analysis of attractive areas using collections of geo-tagged photos. In *Proc. of COM.Geo'10*, page 38, 2010.
 - [58] R. Kohavi, A. Deng, B. Frasca, T. Walker, Y. Xu, and N. Pohlmann. Online controlled experiments at large scale. In *Proc. of KDD'13*, pages 1168–1176, Aug 2013.
 - [59] S. Kotiloglu, T. Lappas, K. Pelechris, and P. Repoussis. Personalized multi-period tour recommendations. *Tourism Management*, 62:76–88, 2017.
 - [60] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura. Travel route recommendation using geotags in photo sharing sites. In *Proc. of CIKM'10*, pages 579–588, 2010.
 - [61] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura. Travel route recommendation using geotagged photos. *Knowledge and Information Systems*, 37(1):37–60, 2013.
 - [62] K. W.-T. Leung, D. L. Lee, and W.-C. Lee. Clr: a collaborative location recommendation framework based on co-clustering. In *Proc. of SIGIR'11*, pages 305–314, 2011.
 - [63] R. A. Lewis, J. M. Rao, and D. H. Reiley. Here, there, and everywhere: Correlated online behaviors can lead to overestimates of the effects of advertising. In *Proc. of WWW'11*, pages 157–166, Mar 2011.
 - [64] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma. Mining user similarity based on location history. In *Proc. of SIGSPATIAL'08*, page 34, 2008.
 - [65] X. Li, G. Cong, X.-L. Li, T.-A. N. Pham, and S. Krishnaswamy. Rank-geofm: a ranking based geographical factorization method for point of interest recommendation. In *Proc. of SIGIR'15*, pages 433–442, 2015.
 - [66] D. Lian, V. W. Zheng, and X. Xie. Collaborative filtering meets next check-in location prediction. In *Proc. of WWW'13*, pages 231–232, 2013.
 - [67] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik. Predictive trip planning-smart routing in smart cities. In *Proc. of MUD'14*, pages 331–338, 2014.
 - [68] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik. Dynamic route planning with real-time traffic predictions. *Information Systems*, 64:258–265, 2017.
 - [69] K. H. Lim. Recommending tours and places-of-interest based on user interests from geo-tagged photos. In *Proc. of SIGMOD'15 PhD Symposium*, pages 33–38, 2015.
 - [70] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie. Personalized itinerary recommendation with queuing time awareness. In *Proc. of SIGIR'17*, pages 325–334, 2017.
 - [71] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera. Personalized tour recommendation based on user interests and points of interest visit durations. In *Proc. of IJCAI'15*, pages 1778–1784, 2015.
 - [72] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera. Towards next generation touring: Personalized group tours. In *Proc. of ICAPS'16*, pages 412–420, 2016.
 - [73] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera. Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. *Knowledge and Information Systems*, 54(2):375–406, Feb 2018.
 - [74] K. H. Lim, E.-P. Lim, B. Jiang, and P. Achananuparp. Using Online Controlled Experiments to Examine Authority Effects on User Behavior in Email Campaigns. In *Proc. of HT'16*, pages 255–260, Jul 2016.
 - [75] K. H. Lim, X. Wang, J. Chan, S. Karunasekera, C. Leckie, Y. Chen, C. L. Tan, F. Q. Gao,

- and T. K. Wee. PersTour: A personalized tour recommendation and planning system. In *Proc. of HT'16*, 2016.
- [76] Q. Liu, S. Wu, L. Wang, and T. Tan. Predicting the next location: A recurrent model with spatial and temporal contexts. In *Proc. of AAAI'16*, pages 194–200, 2016.
- [77] X. Lu, C. Wang, J.-M. Yang, Y. Pang, and L. Zhang. Photo2trip: generating travel routes from geo-tagged photos for trip planning. In *Proc. of MM'10*, pages 143–152, 2010.
- [78] C. Lucchese, R. Perego, F. Silvestri, H. Vahabi, and R. Venturini. How random walks can help tourism. In *Proc. of ECIR'12*, pages 195–206, 2012.
- [79] A. Majid, L. Chen, H. T. Mirza, I. Hussain, and G. Chen. A system for mining interesting tourist locations and travel sequences from public geo-tagged photos. *Data & Knowledge Engineering*, 95:66–86, 2015.
- [80] I. Memon, L. Chen, A. Majid, M. Lv, I. Hussain, and G. Chen. Travel recommendation using geo-tagged photos in social media for tourist. *Wireless Personal Communications*, 80(4):1347–1362, 2015.
- [81] J. Pennebaker. *The Secret Life of Pronouns: What Our Words Say About Us*. Bloomsbury, 2013.
- [82] T.-A. N. Pham, X. Li, and G. Cong. A general model for out-of-town region recommendation. In *Proc. of WWW'17*, pages 401–410, 2017.
- [83] D. Quercia, R. Schifanella, and L. M. Aiello. The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city. In *Proc. of HT'14*, pages 116–125, 2014.
- [84] I. Refanidis and A. Alexiadis. Deployment and evaluation of selfplanner, an automated individual task management system. *Computational Intelligence*, 27(1):41–59, 2011.
- [85] I. Refanidis, C. Emmanouilidis, I. Sakellariou, A. Alexiadis, R.-A. Koutsiamanis, K. Agnantis, A. Tasidou, F. Kokkoras, and P. S. Efraimidis. myVisitPlanner GR: Personalized itinerary planning system for tourism. In *Proc. of SETN'14*, pages 615–629, 2014.
- [86] I. Refanidis and N. Yorke-Smith. A constraint-based approach to scheduling an individual's activities. *ACM Tran. on Intelligent Systems and Technology*, 1(2):12, 2010.
- [87] S. B. Roy, L. V. Lakshmanan, and R. Liu. From group recommendations to group formation. In *Proc. of SIGMOD'15*, pages 1603–1616, 2015.
- [88] W. Souffriau and P. Vansteenwegen. Tourist trip planning functionalities: State-of-the-art and future. In *Proc. of ICWE'10*, pages 474–485, 2010.
- [89] W. Souffriau, P. Vansteenwegen, J. Vertommen, G. V. Berghe, and D. V. Oudheusden. A personalized tourist trip design algorithm for mobile tourist guides. *Applied Artificial Intelligence*, 22(10):964–985, 2008.
- [90] E. Spyrou and P. Mylonas. A survey on flickr multimedia research challenges. *Engineering Applications of Artificial Intelligence*, 51:71–91, 2016.
- [91] Y. Su, X. Li, W. Tang, J. Xiang, and Y. He. Next Check-in Location Prediction via Footprints and Friendship on Location-Based Social Networks. In *Proc. of MDM'18*, pages 251–256, 2018.
- [92] D. Tang, A. Agarwal, D. O'Brien, and M. Meyer. Overlapping experiment infrastructure: more, better, faster experimentation. In *Proc. of KDD'10*, pages 17–26, Jul 2010.
- [93] K. Taylor, K. H. Lim, and J. Chan. Travel Itinerary Recommendations with Must-see Points-of-Interest. In *Proc. of WWW'18*, pages 1198–1205, 2018.
- [94] H. Tong and C. Faloutsos. Center-piece subgraphs: problem definition and fast solutions. In *Proc. of KDD'06*, pages 404–413, 2006.
- [95] M. Toyoshima, M. Hirota, D. Kato, T. Araki, and H. Ishikawa. Where Is the Memorable Travel Destinations? In *Proc. of SocInfo'18*, pages 291–298, 2018.
- [96] T. Tsiligrirides. Heuristic methods applied to orienteering. *J. of the Operational Research Society*, 35(9):797–809, 1984.
- [97] UNWTO. United Nations World Tourism Organization (UNWTO) annual report 2015, 2016. <http://www2.unwto.org/annual-reports>.
- [98] P. Vansteenwegen and D. V. Oudheusden. The mobile tourist guide: An OR opportunity. *OR Insight*, 20(3):21–27, 2007.
- [99] P. Vansteenwegen, W. Souffriau, G. V. Berghe, and D. V. Oudheusden. Iterated local search for the team orienteering problem with time windows. *Computers & Operations Research*, 36(12):3281–3290, 2009.
- [100] P. Vansteenwegen, W. Souffriau, G. V. Berghe, and D. V. Oudheusden. The city trip planner: An expert system for tourists. *Expert Systems with Applications*, 38(6):6540–6546, 2011.

- [101] P. Vansteenwegen, W. Souffriau, and D. V. Oudheusden. The orienteering problem: A survey. *Euro. J. of Operational Research*, 209(1):1–10, 2011.
- [102] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. of CVPR'01*, pages 511–518, 2001.
- [103] J. Wang, Y. Feng, E. Naghizade, L. Rashidi, K. H. Lim, and K. E. Lee. Happiness is a Choice: Sentiment and Activity-Aware Location Recommendation. In *Proc. of WWW'18*, pages 1401–1405, 2018.
- [104] X. Wang, C. Leckie, J. Chan, K. H. Lim, and T. Vaithianathan. Improving personalized trip recommendation by avoiding crowds. In *Proc. of CIKM'16*, pages 25–34, 2016.
- [105] World Travel and Tourism Council. 2016 economic impact annual update summary, 2016. <https://www.wttc.org/research/economic-research/economic-impact-analysis/>.
- [106] A. Yahi, A. Chassang, L. Raynaud, H. Duthil, and D. H. P. Chau. Aurigo: An interactive tour planner for personalized itineraries. In *Proc. of IUI'15*, pages 275–285, 2015.
- [107] L. Yao, Q. Z. Sheng, Y. Qin, X. Wang, A. Shemshadi, and Q. He. Context-aware point-of-interest recommendation using tensor factorization with social regularization. In *Proc. of SIGIR'15*, pages 1007–1010, 2015.
- [108] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *Proc. of SIGIR'11*, pages 325–334, 2011.
- [109] P. Yin, M. Ye, W.-C. Lee, and Z. Li. Mining gps data for trajectory recommendation. In *Proc. of PAKDD'14*, pages 50–61, 2014.
- [110] Z. Yu, H. Xu, Z. Yang, and B. Guo. Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints. *IEEE Tran. on Human-Machine Systems*, 46(1):151–158, 2016.
- [111] Q. Yuan, G. Cong, and C.-Y. Lin. COM: a generative model for group recommendation. In *Proc. of KDD'14*, pages 163–172, 2014.
- [112] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann. Time-aware point-of-interest recommendation. In *Proc. of SIGIR'13*, pages 363–372, 2013.
- [113] Q. Yuan, G. Cong, and A. Sun. Graph-based point-of-interest recommendation with geographical and temporal influences. In *Proc. of CIKM'14*, pages 659–668, 2014.
- [114] C. Zhang, H. Liang, and K. Wang. Trip recommendation meets real-world constraints: Poi availability, diversity, and traveling time uncertainty. *ACM Tran. on Information Systems*, 35(1):5, 2016.
- [115] C. Zhang, H. Liang, K. Wang, and J. Sun. Personalized trip recommendation with poi availability and uncertain traveling time. In *Proc. of CIKM'15*, pages 911–920, 2015.
- [116] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with gps history data. In *Proc. of WWW'10*, pages 1029–1038, 2010.
- [117] Y. Zheng and X. Xie. Learning travel recommendations from user-generated gps traces. *ACM Tran. on Intelligent Systems and Technology*, 2(1):2, 2011.
- [118] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W.-Y. Ma. Recommending friends and locations based on individual location history. *ACM Tran. on the Web*, 5(1):5, 2011.
- [119] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma. Mining interesting locations and travel sequences from gps trajectories. In *Proc. of WWW'09*, pages 791–800, 2009.

Author Biographies



Kwan Hui Lim is an Assistant Professor at the Information Systems Technology and Design Pillar, Singapore University of Technology and Design. Previously, he was a Research Fellow at the School of Computing and Information Systems, University of Melbourne, Research Engineer at the Living Analytics Research Centre, Singapore Management University, and Research Intern at IBM Research - Australia. He received his PhD from the University of Melbourne, and MSc (Research) and BCompSci (1st Class Honours) from the University of Western Australia. He is a recipient of the 2016 Google PhD Fellowship in Machine Learning. His research interests are in Data Mining, Machine Learning, Artificial Intelligence, Social Network Analysis, and Social Computing.



Jeffrey Chan is a Senior Lecturer at RMIT University, Melbourne, Australia. He has a BEng, BSci and PhD, all from the University of Melbourne, and have interned at the National Institute of Informatics in Japan. He was previously a senior post-doctoral fellow at the Digital Enterprise Research Institute in Ireland and a research fellow at the University of Melbourne. He has over 70 publications in graph mining, social network analysis and data mining and his research interests are in data analytics, analysing graphs and social networks and learning about new and exciting research.



Christopher Leckie received the B.Sc. degree in 1985, the B.E. degree in electrical and computer systems engineering (with first-class Honors) in 1987, and the Ph.D. degree in computer science in 1992, all from Monash University, Clayton, Vic., Australia. He joined Telstra Research Laboratories in 1988, where he conducted research and development into artificial intelligence techniques for various telecommunication applications. In 2000, he joined the University of Melbourne, Parkville, Vic., where he is currently a Professor in the Department of Computing and Information Systems. His research interests include scalable data mining, network intrusion detection, bioinformatics, and wireless sensor networks.



Shanika Karunasekera received the B.Sc. degree in electronics and telecommunications engineering from the University of Moratuwa, Sri Lanka, in 1990 and the Ph.D. degree in electrical engineering from the University of Cambridge, Cambridge, U.K., in 1995. From 1995 to 2002, she was a Software Engineer and a Distinguished Member of Technical Staff with Lucent Technologies, Bell Labs Innovations, USA. In January 2003, she joined the University of Melbourne, Victoria, Australia, and currently she is a Professor in the Department of Computing and Information Systems. Her current research interests include distributed system engineering, distributed data-mining and social media analytics.