Personalized Tour Recommendation based on User Interests and Points of Interest Visit Durations

Kwan Hui Lim[*†], **Jeffrey Chan**[*], **Christopher Leckie**[*†] and **Shanika Karunasekera**[*]
[*]Department of Computing and Information Systems, The University of Melbourne, Australia
[†]Victoria Research Laboratory, National ICT Australia, Australia
{limk2@student., jeffrey.chan@, caleckie@, karus@}unimelb.edu.au

Abstract

Tour recommendation and itinerary planning are challenging tasks for tourists, due to their need to select Points of Interest (POI) to visit in unfamiliar cities, and to select POIs that align with their interest preferences and trip constraints. We propose an algorithm called PERSTOUR for recommending personalized tours using POI popularity and user interest preferences, which are automatically derived from real-life travel sequences based on geotagged photos. Our tour recommendation problem is modelled using a formulation of the Orienteering problem, and considers user trip constraints such as time limits and the need to start and end at specific POIs. In our work, we also reflect levels of user interest based on visit durations, and demonstrate how POI visit duration can be personalized using this time-based user interest. Using a Flickr dataset of four cities, our experiments show the effectiveness of PERSTOUR against various baselines, in terms of tour popularity, interest, recall, precision and F_1-score. In particular, our results show the merits of using time-based user interest and personalized POI visit durations, compared to the current practice of using frequency-based user interest and average visit durations.

1 Introduction

Tour recommendation and itinerary planning are challenging tasks due to the different interest preferences and trip constraints (e.g., time limits, start and end points) of each unique tourist[1]. While there is an abundance of information from the Internet and travel guides, many of these resources simply recommend individual Points of Interest (POI) that are deemed to be popular, but otherwise do not appeal to the interest preferences of users or adhere to their trip constraints. Furthermore, the massive volume of information makes it a challenge for tourists to narrow down to a potential set of POIs to visit in an unfamiliar city. Even after the tourist finds a suitable set of POIs to visit, it will take considerable time and effort for the tourist to plan the appropriate duration of visit at each POI and the order in which to visit the POIs.

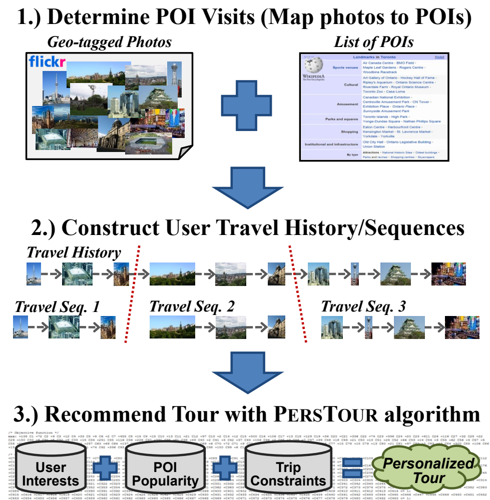

Figure 1: Tour Recommendation Framework

To address these issues, we propose the PERSTOUR algorithm for recommending personalized tours where the suggested POIs are optimized to the users' interest preferences and POI popularity. We formulate our tour recommendation problem based on the Orienteering problem [Tsiligirides, 1984], which considers a user's trip constraints such as time limitations and the need for the tour to start and end at specific POIs (e.g., POIs near the tourist's hotel). Using geo-tagged photos as a proxy for tourist visits, we are able to extract real-life user travel histories, which can then be used to automatically determine a user's interest level in various POI categories (e.g., parks, beaches, shopping) as well as the popularity of individual POIs. As tourists have different preference levels between POI popularity and POI relevance to their interests, our PERSTOUR algorithm also allows tourists to indicate their preferred level of trade-off between POI popularity and his/her interest preferences.

Our main contributions are as follows:

- We propose the PERSTOUR algorithm for recommending personalized tours with POIs and visit duration based on POI popularity, users' interest preferences and

[1]We use the terms "tourist" and "user" interchangeably.

This paper is published in *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI'15)*.
The dataset described in Section 5.1 is publicly available at https://sites.google.com/site/limkwanhui/datacode#ijcai15.

trip constraints. Our tour recommendation problem is modelled in the context of the Orienteering problem (Section 3).

- We introduce the concept of *time-based user interest*, where a user's level of interest in a POI category is based on his/her time spent at such POIs, relative to the average user. We also compare our time-based user interest to the current practice of using frequency-based user interest, and show how time-based user interest results in recommended tours that more accurately reflect real-life travel sequences (Section 3.1).
- We demonstrate the *personalization of POI visit duration* using time-based user interest. Our results show that personalized visit durations more accurately reflect the real-life POI visit durations of users, compared to the current practice of using average visit duration (Section 3.1).
- We implement a framework (Fig. 1) for extracting real-life user travel histories, which are then used for training our PERSTOUR algorithm and serve as ground truth for our subsequent evaluation (Section 4).
- We evaluate different variants of PERSTOUR against various baselines using a Flickr dataset spanning four cities. Our results show that PERSTOUR out-performs these baselines based on tour popularity, user interest, recall, precision and F_1-score (Sections 5 and 6).

For the rest of the paper: Section 2 discusses related work in tour recommendation, and Section 7 concludes our paper.

2 Related Work

Tour recommendation has been a well-studied field, with many developed applications [Vansteenwegen and Oudheusden, 2007; Castillo *et al.*, 2008; Brilhante *et al.*, 2014] and research ranging from recommending beautiful, quiet, and happy tours [Quercia *et al.*, 2014] to tour recommendation using random walks with restart [Lucchese *et al.*, 2012]. In this section, we focus on research related to our work, and refer readers to [Souffriau and Vansteenwegen, 2010] and [Damianos Gavalas, 2014] for an overview on the general field of tour recommendation.

[Choudhury *et al.*, 2010] was one of the earlier tour recommendation studies based on the Orienteering problem, where recommended tours start and end at specific POIs while trying to maximize an objective score. Using a modified Orienteering problem, [Gionis *et al.*, 2014] utilized POI categories such that recommended tours are constrained by a POI category visit order (e.g., museum $\rightarrow$ park $\rightarrow$ beach). Similarly, [Lim, 2015] used a modified Orienteering problem constrained by a mandatory POI category, which corresponds to the POI category a user is most interested in. Based on user-indicated interests and trip constraints (e.g., time budget, start and end locations), [Vansteenwegen *et al.*, 2011a] recommended tours comprising POI categories that best match user interests while adhering to these trip constraints.

In contrast, [Brilhante *et al.*, 2013] formulated tour recommendation as a Generalized Maximum Coverage problem [Cohen and Katzir, 2008], with the objective of finding an optimal set of POIs based on both POI popularity and user interest. Thereafter, [Brilhante *et al.*, 2015] extended upon the former by using a variation of the Travelling Salesman Problem, with the main aim of finding the shortest route among the set of optimal POIs recommended in [Brilhante *et al.*, 2013]. In addition to user interests in tour recommendation, [Chen *et al.*, 2014] also considered travelling times based on different traffic conditions, using trajectory patterns derived from taxi GPS traces. With further considerations for different transport modes, [Kurashima *et al.*, 2010; 2013] used a combined topic and Markov model to recommend tours based on both user interests and frequently travelled routes.

While these earlier works are the state-of-the-art in tour recommendation research, our proposed work differs from these earlier works in two main aspects: (i) instead of using frequency-based user interest (by POI visit frequency) or requiring users to explicitly indicate their interest preferences, we derive a relative measure of *time-based user interest* using a user's visit durations at POIs of a specific category, relative to the average visit durations of other users; (ii) we recommend a *personalized POI visit duration* to individual users based on their time-based user interests, instead of using the average POI visit duration for all users or not considering visit duration at all.

3 Background and Problem Definition

3.1 Preliminaries

If there are m POIs for a particular city, let $P = \{p_1, ..., p_m\}$ be the set of POIs in that city. Each POI p is also labelled with a category Cat_p (e.g., church, park, beach) and latitude/longitude coordinates. We denote a function $Pop(p)$ that indicates the popularity of a POI p, based on the number of times POI p has been visited. Similarly, the function $T^{Travel}(p_x, p_y)$ measures the time needed to travel from POI p_x to p_y, based on the distance between POIs p_x and p_y and the indicated travelling speed. For simplicity, we use a travelling speed of 4km/hour (i.e., a leisure walking speed).[2]

Definition 1: Travel History. Given a user u who has visited n POIs, we define his/her travel history as an ordered sequence, $S_u = ((p_1, t^a_{p_1}, t^d_{p_1}), ..., (p_n, t^a_{p_n}, t^d_{p_n}))$, with each triplet $(p_x, t^a_{p_x}, t^d_{p_x})$ comprising the visited POI p_x, and the arrival time $t^a_{p_x}$ and departure time $t^d_{p_x}$ at POI p_x. Thus, the visit duration at POI p_x can be determined by the difference between $t^a_{p_x}$ and $t^d_{p_x}$. Similarly, for a travel sequence S_u, $t^a_{p_1}$ and $t^d_{p_n}$ also indicates the start and end time of the itinerary respectively. For brevity, we simplify $S_u = ((p_1, t^a_{p_1}, t^d_{p_1}), ..., (p_n, t^a_{p_n}, t^d_{p_n}))$ as $S_u = (p_1, ..., p_n)$.

Definition 2: Travel Sequence. Based on the travel history S_u of a user u, we can further divide this travel history into multiple travel sequences (i.e., sub-sequences of S_u). We divide a travel history S_u into separate travel sequences if

[2] $T^{Travel}(p_x, p_y)$ can be easily generalized to different transport modes (e.g., taxi, bus, train) and also consider the traffic condition between POIs (e.g., longer travel times between two POIs in a congested city, compared to two equal-distanced POIs elsewhere).

$t^d_{p_x} - t^a_{p_x+1} > \tau$. That is, we separate a travel history into distinct travel sequences if the consecutive POI visits occur more than τ time units apart. Similar to other works [Choudhury *et al.*, 2010; Lim, 2015], we choose $\tau = 8\ hours$ in our experiments. These travel sequences also serve as the ground truth of real-life user trajectories, which are subsequently used for evaluating our PERSTOUR algorithm and baselines.

Definition 3: Average POI Visit Duration. Given a set of travel histories for all users U, we determine the average visit duration for a POI p as follows:

$$\bar{V}(p) = \frac{1}{n} \sum_{u \in U} \sum_{p_x \in S_u} (t^d_{p_x} - t^a_{p_x}) \delta(p_x = p), \qquad \forall\, p \in P \quad (1)$$

where n is the number of visits to POI p by all users and $\delta(p_x{=}p) = \{^{1,\ p_x=p}_{0,\ otherwise}$. $\bar{V}(p)$ is commonly used in tour recommendation as the POI visit duration for all users [Brilhante *et al.*, 2013; 2015; Chen *et al.*, 2014], while many earlier works do not factor in POI visit durations at all. In our work, we show how recommended POI visit durations can be personalized to individual users based on their interest (Definition 5), and use $\bar{V}(p)$ as a comparison baseline (i.e., the non-personalized POI visit duration).

Definition 4: Time-based User Interest. As described earlier, the category of a POI p is denoted Cat_p. Given that C represents the set of all POI categories, we determine the interest of a user u in POI category c as follows:

$$Int^{Time}_u(c) = \sum_{p_x \in S_u} \frac{(t^d_{p_x} - t^a_{p_x})}{\bar{V}(p_x)} \delta(Cat_{p_x}{=}c),\ \forall\, c \in C \quad (2)$$

where $\delta(Cat_{p_x}{=}c) = \{^{1,\ Cat_{p_x}=c}_{0,\ otherwise}$. In short, Eqn. 2 determines the interest of a user u in a particular POI category c, based on his/her time spent at each POI of category c, relative to the average visit duration (of all users) at the same POI. The rationale is that a user is likely to spend more time at a POI that he/she is interested in. Thus, by calculating how much more (or less) time a user is spending at POIs of a certain category compared to the average user, we can determine the interest level of this user in POIs of this category.

Definition 5: Personalized POI Visit Duration. Based on our definition of time-based user interest (Eqn. 2), we are able to personalize the recommended visit duration at each POI based on each user's interest level. We determine the personalized visit duration at a POI p for a user u as follows:

$$T^{Visit}_u(p) = Int^{Time}_u(Cat_p) * \bar{V}(p) \quad (3)$$

That is, we are recommending a personalized POI visit duration based on user u's relative interest level in category Cat_p multiplied by the average time spent at POI p. Thus, if a user is more (less) interested in category Cat_p, he/she will spend more (less) time at POI p than the average user.

Definition 6: Frequency-based User Interest. We also define a simplified version of user interest, denoted $Int^{Freq}_u(c)$, which is based on the number of times a user visits POIs of a certain category c (i.e., the more times a user visits POIs of a specific category, the more interested this user is in that category). As using $Int^{Freq}_u(c)$ is the current practice in tour recommendation research [Brilhante *et al.*, 2013; Lim, 2015; Brilhante *et al.*, 2015], we include it for a more complete study and as a comparison baseline to our proposed $Int^{Time}_u(c)$.

3.2 Problem Definition

We now define our tour recommendation problem in the context of the Orienteering problem and its integer problem formulation [Tsiligirides, 1984; Vansteenwegen *et al.*, 2011b; Lim, 2015]. Given the set of POIs P, a budget B, starting POI p_1 and destination POI p_N, our goal is to recommend an itinerary $I = (p_1, ..., p_N)$ that maximizes a certain score S while adhering to the budget B. In this case, the score S is represented by the popularity and user interest of the recommended POIs using the functions $Pop(p)$ and $Int(Cat_p)$, respectively. The budget B is calculated using the function $Cost(p_x, p_y) = T^{Travel}(p_x, p_y) + T^{Visit}_u(p_y)$, i.e., using both travelling time and personalized visit duration at the POI. One main difference between our work and earlier work is that we personalize the visit duration at each recommended POI based on user interest (Definition 5), instead of using the average visit duration for all users or not considering visit duration at all. Formally, we want to find an itinerary $I = (p_1, ..., p_N)$ that:

$$Max \sum_{i=2}^{N-1} \sum_{j=2}^{N} x_{i,j} \Big(\eta Int(Cat_i) + (1-\eta) Pop(i) \Big) \quad (4)$$

where $x_{i,j} = 1$ if both POI i and j are visited in sequence (i.e., we travel directly from POI i to j), and $x_{i,j} = 0$ otherwise. We attempt to solve for Eqn. 4, such that:

$$\sum_{j=2}^{N} x_{1,j} = \sum_{i=1}^{N-1} x_{i,N} = 1 \quad (5)$$

$$\sum_{i=1}^{N-1} x_{i,k} = \sum_{j=2}^{N} x_{k,j} \leq 1, \qquad \forall\, k = 2, ..., N-1 \quad (6)$$

$$\sum_{i=1}^{N-1} \sum_{j=2}^{N} Cost(i,j) x_{i,j} \leq B \quad (7)$$

$$2 \leq p_i \leq N, \qquad \forall\, i = 2, ..., N \quad (8)$$

$$p_i - p_j + 1 \leq (N-1)(1 - x_{i,j}), \qquad \forall\, i, j = 2, ..., N \quad (9)$$

Eqn. 4 is a multi-objective function that maximizes the popularity and interest of all visited POIs in the itinerary, where η is the weighting given to the popularity and interest components. Eqn. 4 is also subject to constraints 5 to 9. Constraint 5 ensures that the itinerary starts at POI 1 and ends at POI N, while constraint 6 ensures that the itinerary is connected and no POIs are visited more than once. Constraint 7 ensures that the time taken for the itinerary is within the budget B, based on the function $Cost(p_x, p_y)$ that considers both travelling time and personalized POI visit duration. Given that p_x is the position of POI x in itinerary I, constraints 8 and 9 ensure that there are no sub-tours in the proposed solution, adapted from the sub-tour elimination used in the Travelling Salesman Problem [Miller *et al.*, 1960].

Based on this problem definition, we can then proceed to solve our tour recommendation problem as an integer programming problem. For solving this integer programming problem, we used the lpsolve linear programming package [Berkelaar *et al.*, 2004]. We denote our proposed algorithm for personalized tour recommendation as PERSTOUR, and shall describe our overall framework and the different PERSTOUR variants in the following section.

4 Tour Recommendation Framework

Fig. 1 outlines our overall tour recommendation framework. This framework requires a list of POIs (with lat/long coordinates and POI categories) and a set of geo-tagged photos (with lat/long coordinates and time taken), which can be easily obtained from Wikipedia and Flickr, respectively. Thereafter, the main steps in our framework are:

Step 1: Determine POI visits (Map photos to POIs). We first determine the POI visits in each city by mapping the set of geo-tagged photos to the list of POIs. In particular, we map a photo to a POI if their coordinates differ by <200m based on the Haversine formula [Sinnott, 1984], which is used for calculating spherical (earth) distances.

Step 2: Construct Travel History/Sequences. Based on the POI visits from Step 1, we can construct the travel history of each user by sorting their POI visits in ascending temporal-order (Definition 1). Using each user's travel history, we then proceed to group consecutive POI visits as an individual travel sequence, if the consecutive POI visits differ by <8 hours (Definition 2). Thus, we are also able to determine the POI visit duration based on the time difference of the first and last photo taken at each POI.

Step 3: Recommend Tours using PERSTOUR. As described in Section 3.2, there can be different variants of PERSTOUR, based on the value of η and the type of interest function chosen. The value of η indicates the weight given to either POI popularity or user interest, while the interest function can be either time-based interest (Int_u^{Time}) or frequency-based interest (Int_u^{Freq}). We experiment with the following variants:

- **PERSTOUR using** η=0 **(PT-0)**. PERSTOUR with full emphasis on POI popularity, ignoring user interest (i.e., no need to choose between Int_u^{Time} or Int_u^{Freq}).
- **PERSTOUR using** Int_u^{Time} **and** η=0.5 **(PT-.5T)**. PERSTOUR with balanced emphasis on optimizing both POI popularity and *time-based* user interest.
- **PERSTOUR using** Int_u^{Freq} **and** η=0.5 **(PT-.5F)**. PERSTOUR with balanced emphasis on optimizing both POI popularity and *frequency-based* user interest.
- **PERSTOUR using** Int_u^{Time} **and** η=1 **(PT-1T)**. PERSTOUR with full emphasis on optimizing *time-based* user interest, ignoring POI popularity.
- **PERSTOUR using** Int_u^{Freq} **and** η=1 **(PT-1F)**. PERSTOUR with full emphasis on optimizing *frequency-based* user interest, ignoring POI popularity.

These variants allow us to best evaluate the effects of different η values, and compare between time-based interest and frequency-based interest. As PT-0 does not consider user interest, there is no need to choose between time-based or frequency-based user interest.

Table 1: Dataset description

City	No. of Photos	No. of Users	# POI Visits	# Travel Sequences
Toronto	157,505	1,395	39,419	6,057
Osaka	392,420	450	7,747	1,115
Glasgow	29,019	601	11,434	2,227
Edinburgh	82,060	1,454	33,944	5,028

5 Experimental Methodology

5.1 Dataset

For our experiments, we use the Yahoo! Flickr Creative Commons 100M (YFCC100M) dataset [Thomee *et al.*, 2015], which consists of 100M Flickr photos and videos. This dataset also comprises the meta information regarding the photos, such as the date/time taken, geo-location coordinates and accuracy of these geo-location coordinates. The geo-location accuracy range from world level (least accurate) to street level (most accurate).

Using the YFCC100M dataset, we extracted geo-tagged photos that were taken in four different cities, namely: Toronto, Osaka, Glasgow and Edinburgh. To ensure the best accuracy and generalizability of our results, we only chose photos with the highest geo-location accuracy and experimented on four touristic cities around the world. A more detailed description of our dataset is shown in Table 1.

5.2 Baseline Algorithms

Similar to our PERSTOUR approach, these baseline algorithms commence from a starting POI p_1 and iteratively choose a next POI to visit until either: (i) the budget B is exhausted; or (ii) the destination POI p_N is reached. The sequence of selected POIs thus forms the recommended itinerary, and the three baselines are:

- **Greedy Nearest (GNEAR)**. Chooses the next POI to visit by randomly selecting from the three *nearest*, unvisited POIs.
- **Greedy Most Popular (GPOP)**. Chooses the next POI to visit by randomly selecting from the three *most popular*, unvisited POIs.
- **Random Selection (RAND)**. Chooses the next POI to visit by *randomly selecting* from all unvisited POIs.

GNEAR and GPOP are meant to reflect tourists behavior by visiting nearby and popular POIs respectively, while RAND shows the performance of a random-based approach.

5.3 Evaluation

We evaluate PERSTOUR and the baselines using leave-one-out cross-validation [Kohavi, 1995] (i.e., when evaluating a specific travel sequence of a user, we use this user's other travel sequences for training our algorithms). Specifically,

Table 2: Comparison between Time-based User Interest (PT-.5T and PT-1T) and Frequency-based User Interest (PT-.5F and PT-1F), in terms of Recall (T_R), Precision (T_P) and F_1-score (T_{F_1}).[4]

Toronto			
Algo.	*Recall*	*Precision*	*F_1-score*
PT-.5F	.760±.009	.679±.013	.708±.012
PT-.5T	**.779±.010**	**.706±.013**	**.732±.012**
PT-1F	.737±.010	.682±.013	.698±.012
PT-1T	**.744±.011**	**.710±.013**	**.718±.012**
GNEAR	.501±.010	.512±.015	.487±.011
GPOP	.440±.009	.623±.015	.504±.011
RAND	.333±.007	.495±.011	.391±.009

Osaka			
Algo.	*Recall*	*Precision*	*F_1-score*
PT-.5F	.757±.025	.645±.037	.687±.032
PT-.5T	**.759±.026**	**.662±.037**	**.699±.033**
PT-1F	.679±.023	.582±.032	.616±.027
PT-1T	**.683±.025**	**.622±.032**	**.641±.029**
GNEAR	.478±.026	.433±.038	.441±.030
GPOP	.439±.034	.649±.038	.517±.035
RAND	.354±.021	.488±.032	.406±.024

Glasgow			
Algo.	*Recall*	*Precision*	*F_1-score*
PT-.5F	.819±.017	.758±.024	.780±.021
PT-.5T	**.826±.017**	**.782±.022**	**.798±.020**
PT-1F	**.748±.017**	.728±.022	.726±.019
PT-1T	.739±.018	**.736±.021**	**.728±.019**
GNEAR	.498±.020	.519±.028	.490±.022
GPOP	.418±.015	.592±.024	.480±.017
RAND	.340±.012	.462±.017	.386±.013

Edinburgh			
Algo.	*Recall*	*Precision*	*F_1-score*
PT-.5F	.740±.006	.607±.010	.654±.009
PT-.5T	**.740±.007**	**.633±.010**	**.671±.008**
PT-1F	**.678±.007**	.572±.009	.605±.008
PT-1T	.668±.007	**.601±.009**	**.618±.008**
GNEAR	.471±.007	.429±.010	.427±.008
GPOP	.486±.008	.640±.010	.539±.008
RAND	.336±.006	.479±.009	.384±.006

Table 3: Comparison between Personalized and Non-personalized Visit Durations, in terms of RMSE (T_{RMSE}).

Toronto		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**147.57±10.85**
	Non-personalized	152.44±9.84
PT-.5F	Personalized	**146.33±10.85**
	Non-personalized	152.61±10.09
PT-.5T	Personalized	**143.56±10.89**
	Non-personalized	150.65±10.09
PT-1F	Personalized	**137.07±11.40**
	Non-personalized	145.54±10.78
PT-1T	Personalized	**145.20±11.79**
	Non-personalized	148.18±11.29

Osaka		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**51.35±11.41**
	Non-personalized	54.91±11.91
PT-.5F	Personalized	**56.71±12.43**
	Non-personalized	60.06±13.09
PT-.5T	Personalized	57.09±12.39
	Non-personalized	**55.84±13.18**
PT-1F	Personalized	**56.62±13.21**
	Non-personalized	62.24±14.60
PT-1T	Personalized	**53.44±13.05**
	Non-personalized	58.88±14.63

Glasgow		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**75.98±11.53**
	Non-personalized	85.76±12.07
PT-.5F	Personalized	**88.20±13.03**
	Non-personalized	92.71±12.92
PT-.5T	Personalized	**76.40±11.34**
	Non-personalized	90.33±12.35
PT-1F	Personalized	**79.67±12.27**
	Non-personalized	86.24±12.85
PT-1T	Personalized	**73.29±11.94**
	Non-personalized	91.06±13.45

Edinburgh		
Algo.	*Visit Duration*	*RMSE*
PT-0	Personalized	**91.08±4.85**
	Non-personalized	113.15±5.21
PT-.5F	Personalized	**84.56±4.96**
	Non-personalized	99.54±5.14
PT-.5T	Personalized	**89.76±5.85**
	Non-personalized	100.15±5.27
PT-1F	Personalized	**69.61±5.04**
	Non-personalized	78.89±5.31
PT-1T	Personalized	**72.11±6.09**
	Non-personalized	74.48±5.29

we consider all real-life travel sequences with ≥3 POI visits and evaluate the algorithms using the starting POIs and destination POIs of these travel sequences. Thereafter, we evaluate the performance of each algorithm based on the recommended tour itinerary I using the following metrics:

1. **Tour Recall: $T_R(I)$.** The proportion of POIs in a user's real-life travel sequence that were also recommended in itinerary I. Let P_r be the set of POIs recommended in itinerary I and P_v be the set of POIs visited in the real-life travel sequence, tour recall is defined as: $T_R(I) = \frac{|P_r \cap P_v|}{|P_v|}$.

2. **Tour Precision: $T_P(I)$.** The proportion of POIs recommended in itinerary I that were also in a user's real-life travel sequence. Let P_r be the set of POIs recommended in itinerary I and P_v be the set of POIs visited in the real-life travel sequence, tour precision is defined as: $T_P(I) = \frac{|P_r \cap P_v|}{|P_r|}$.

3. **Tour F_1-score: $T_{F_1}(I)$.** The harmonic mean of both the recall and precision of a recommended tour itinerary I, defined as: $T_{F_1}(I) = \frac{2 \times T_P(I) \times T_R(I)}{T_P(I) + T_R(I)}$.

4. **Root-Mean-Square Error (RMSE) of POI Visit Duration: $T_{RMSE}(I)$.** The level of error between our recommended POI visit durations in itinerary I compared to the real-life POI visit durations taken by the users. Let $I^s \in I$ be the recommended POIs that were visited in real-life[5], and D_r and D_v be the recommended and real-life POI visit durations respectively, RMSE is defined as: $T_{RMSE}(I) = \sqrt{\frac{\sum_{p \in I^s}(D_r - D_v)^2}{|I^s|}}$.

5. **Tour Popularity: $T_{Pop}(I)$.** The overall popularity of all POIs in the recommended itinerary I, defined as: $T_{Pop}(I) = \sum_{p \in I} Pop(p)$.

6. **Tour Interest: $T^u_{Int}(I)$.** The overall interest of all POIs in the recommended itinerary I to a user u, defined as: $T^u_{Int}(I) = \sum_{p \in I} Int_u(Cat_p)$.

7. **Popularity and Interest Rank: T^a_{Rk}.** The average rank of an algorithm a based on its T_{Pop} and T_{Int} scores ranked against other algorithms (1=best, 8=worst).

We selected these metrics to better evaluate the following: (i) time-based vs frequency-based user interest, using Metrics 1-3; (ii) personalized vs non-personalized POI visit durations, using Metric 4; and (iii) PERSTOUR vs baselines, using Metrics 5-7. As personalized POI visit durations only apply to PERSTOUR and not the baselines, we only report T_{RMSE} scores for the PT-0, PT-.5F, PT-.5T, PT-1F and PT-1T algorithms. Our baseline for comparing T_{RMSE} are variants of PERSTOUR that use non-personalized POI visit durations, i.e., average POI visit durations.

6 Results and Discussion

6.1 Comparison between Time-based and Frequency-based User Interest

We first study the performance difference between using time-based user interest and frequency-based user interest, as shown in Table 2. Comparing the T_{F_1} scores between PT-.5T and PT-.5F, and between PT-1T and PT-1F, the results show

[4]PT-.5T out-performs PT-.5F in terms of T_R (.7402 vs .7398) for Edinburgh, although both values are rounded to .740 in Table 2.

[5]We can only compare POI visit durations for POIs in itinerary I that were "correctly" recommended (i.e., visited in real-life).

Table 4: Comparison of PERSTOUR (PT) against baselines, in terms of Popularity (T_{Pop}), Interest (T_{Int}) and Rank (T_{Rk}). Number within brackets indicate the rank based on Popularity and Interest scores, where 1=best and 8=worst.

Toronto				Osaka			
Algo.	*Popularity*	*Interest*	*Rk*	*Algo.*	*Popularity*	*Interest*	*Rk*
PT-0	2.204±.069 (1)	0.904±.048 (5)	3	PT-0	1.263±.094 (1)	0.791±.166 (6)	3.5
PT-.5F	2.053±.063 (2)	1.088±.060 (4)	3	PT-.5F	1.126±.095 (3)	1.151±.213 (3)	3
PT-.5T	1.960±.064 (3)	1.223±.061 (2)	**2.5**	PT-.5T	1.144±.093 (2)	1.171±.206 (2)	**2**
PT-1F	1.583±.048 (4)	1.137±.061 (3)	3.5	PT-1F	0.809±.075 (5)	1.137±.211 (4)	4.5
PT-1T	1.419±.044 (7)	1.351±.069 (1)	4	PT-1T	0.737±.067 (6)	1.205±.211 (1)	3.5
GNEAR	1.424±.049 (6)	0.773±.054 (6)	6	GNEAR	0.500±.059 (7)	0.853±.183 (5)	6
GPOP	1.566±.050 (5)	0.443±.029 (8)	6.5	GPOP	0.837±.062 (4)	0.223±.066 (8)	6
RAND	0.581±.032 (8)	0.467±.037 (7)	7.5	RAND	0.433±.055 (8)	0.305±.089 (7)	7.5

Glasgow				Edinburgh			
Algo.	*Popularity*	*Interest*	*Rk*	*Algo.*	*Popularity*	*Interest*	*Rk*
PT-0	1.701±.101 (1)	0.459±.069 (5)	3	PT-0	2.269±.046 (1)	1.047±.053 (5)	3
PT-.5F	1.562±.089 (3)	0.563±.091 (3)	3	PT-.5F	2.016±.042 (2)	1.383±.068 (4)	3
PT-.5T	1.601±.089 (2)	0.625±.084 (2)	**2**	PT-.5T	2.012±.043 (3)	1.579±.069 (2)	**2.5**
PT-1F	1.128±.069 (5)	0.562±.090 (4)	4.5	PT-1F	1.541±.038 (5)	1.430±.070 (3)	4
PT-1T	1.001±.052 (6)	0.676±.096 (1)	3.5	PT-1T	1.336±.034 (6)	1.722±.076 (1)	3.5
GNEAR	0.874±.064 (7)	0.339±.070 (6)	6.5	GNEAR	1.269±.033 (7)	0.939±.054 (6)	6.5
GPOP	1.399±.075 (4)	0.217±.049 (8)	6	GPOP	1.775±.039 (4)	0.577±.033 (7)	5.5
RAND	0.483±.048 (8)	0.229±.041 (7)	7.5	RAND	0.656±.025 (8)	0.526±.033 (8)	8

that PERSTOUR using time-based user interest (PT-.5T and PT-1T) consistently out-performs its counterpart that uses frequency-based user interest (PT-.5F and PT-1F). This observation highlights the effectiveness of time-based user interest in recommending tours that more accurately reflect real-life tours of users, compared to using frequency-based user interest. While PT-1T under-performs PT-1F in terms of T_R for Edinburgh and Osaka, we focus more on the T_{F_1} scores as it provides a balanced representation of both T_R and T_P. Moreover, PT-.5T and PT-1T (time-based user interest) consistently results in higher T_P scores, compared to its PT-.5F and PT-1F counterparts (frequency-based user interest). Another observation is that all PERSTOUR variants also consistently out-perform the three baselines, in terms of T_{F_1} scores.

The reason for the more accurate recommendations of time-based user interest compared to frequency-based user interest is due to its use of POI visit durations instead of POI visit frequency. Consider user A who only visited two parks but spent three or more hours at each of them and user B who visited five parks but spent less than 15 minutes at each of them. Frequency-based interest incorrectly classifies user B as having more interest in the parks category due to his/her five visits, compared to user A's two visits. On the other hand, time-based interest more accurately determines that user A has a higher interest in the parks category due to his/her long visit duration, despite user A only visiting two parks. Furthermore, time-based interest can more accurately capture a user's level of interest based on how much longer this user spends at a POI compared to the average user (e.g., a user is more interested if he/she spends 3 hours at a POI when the average time spent is only 30 minutes). With the availability of user interest levels, we can better personalize POI visit duration for each unique user, which we evaluate next.

6.2 Comparison between Personalized and Non-personalized Visit Durations

The T_{RMSE} scores in Table 3 show that our recommendation of a personalized POI visit duration (Definition 5) out-performs the non-personalized version in 19 out of 20 cases[6], based on a smaller error in the recommended POI visit durations. This result shows that personalizing POI visit duration using time-based user interests more accurately reflects the real-life POI visit duration of users, compared to the current standard of simply using average POI visit duration. Apart from recommending accurate POIs (T_{F_1} scores), recommending the appropriate amount of time to spend at each POI is another important consideration in tour recommendation, which has not been explored in earlier works.

Previously, we have observed how time-based interest results in more accurate POI recommendations based on the T_{F_1} scores. Our personalized POI visit duration builds upon this success by customizing the POI visit duration to each unique user based on his/her relative interest level (i.e., spend more time in a POI that interests the user, and less time in a POI that the user is less interested in). Accurate POI visit durations have another important implication in tour recommendation, where spending less time at un-interesting POIs frees up the time budget for more visits to POIs that are more interesting to the user. Similarly, a user might prefer to spend more time visiting a few POIs of great interest, compared to visiting many POIs of less interest to the user.

6.3 Comparison of Popularity and Interest

Based on the T_{Rk} scores in Table 4, we observe that PT-.5T (time-based user interest) is consistently the best performer, out-performing all baselines as well as its PT-.5F counterpart that uses frequency-based user interest. In addition, we also observe that PT-1T (time-based user interest) out-performs its PT-1F counterpart (frequency-based user interest) for three out of four cities. These results show the effectiveness of time-based user interest over frequency-based user interest, based on the T_{Rk} scores.

The effects of the η parameter can be observed in the T_{Pop} and T_{Int} scores. A value of $\eta = 0$ (PT-0) results in the best performance in T_{Pop} and worst performance in T_{Int}, while a value of $\eta = 1$ (PT-1F and PT-1T) results in the opposite. While we include the T_{Pop} and T_{Int} scores for completeness, we are more interested in T_{Rk} as it gives a balanced measurement of both T_{Pop} and T_{Int}.

7 Conclusion

We modelled our tour recommendation problem based on the Orienteering problem and proposed the PERSTOUR algorithm for recommending personalized tours. Our PERSTOUR algorithm considers both POI popularity and user interest preferences to recommend suitable POIs to visit and the amount of time to spend at each POI. In addition, we implemented a framework where geo-tagged photos can be used to automatically detect real-life travel sequences, and determine POI popularity and user interest, which can then be used to

[6]Except for PT-.5T on the Osaka dataset.

train our PERSTOUR algorithm. Our work improves upon earlier tour recommendation research in two main ways: (i) we introduce *time-based user interest* derived from a user's visit durations at specific POIs relative to other users, instead of using a frequency-based user interest based on POI visit frequency; and (ii) we *personalize POI visit duration* based on the relative interest levels of individual users, instead of using the average POI visit duration for all users or not considering POI visit duration at all.

Using a Flickr dataset across four cities, we evaluate the effectiveness of our PERSTOUR algorithm against various baselines in terms of tour popularity, interest, precision, recall, F_1-score, and RMSE of visit duration. In particular, our experimental results show that: (i) using time-based user interest results in tours that more accurately reflect the real-life travel sequences of users, compared to using frequency-based user interest, based on precision and F_1-score; (ii) our personalized POI visit duration more accurately reflects the time users spend at POIs in real-life, compared to the current standard of using average visit duration, based on the RMSE of visit duration; and (iii) PERSTOUR and its variants generally outperform all baselines in most cases, based on tour popularity, interest, precision, recall and F_1-score.

Acknowledgments. National ICT Australia (NICTA) is funded by the Australian Government through the Department of Communications and the Australian Research Council through the ICT Centre of Excellence Program.

References

[Berkelaar *et al.*, 2004] Michel Berkelaar, Kjell Eikland, and Peter Notebaert. lpsolve: Open source (mixed-integer) linear programming system, 2004. http://lpsolve.sourceforge.net/.

[Brilhante *et al.*, 2013] Igo Brilhante, Jose Antonio Macedo, Franco Maria Nardini, Raffaele Perego, and Chiara Renso. Where shall we go today? Planning touristic tours with TripBuilder. In *Proceedings of CIKM'13*, pages 757–762, 2013.

[Brilhante *et al.*, 2014] Igo Brilhante, Jose Antonio Macedo, Franco Maria Nardini, Raffaele Perego, and Chiara Renso. TripBuilder: A tool for recommending sightseeing tours. In *Proceedings of ECIR'14*, pages 771–774, 2014.

[Brilhante *et al.*, 2015] Igo Brilhante, Jose Antonio Macedo, Franco Maria Nardini, Raffaele Perego, and Chiara Renso. On planning sightseeing tours with TripBuilder. *Information Processing & Management*, 51(2):1–15, 2015.

[Castillo *et al.*, 2008] Luis Castillo, Eva Armengol, Eva Onaindía, Laura Sebastiá, Jesús González-Boticario, Antonio Rodríguez, Susana Fernández, Juan D. Arias, and Daniel Borrajo. SAMAP: An user-oriented adaptive system for planning tourist visits. *Expert Systems with Applications*, 34(2):1318–1332, 2008.

[Chen *et al.*, 2014] Chao Chen, Daqing Zhang, Bin Guo, Xiaojuan Ma, Gang Pan, and Zhaohui Wu. TripPlanner: Personalized trip planning leveraging heterogeneous crowdsourced digital footprints. *IEEE Trans. on Intelligent Transportation Systems*, 2014.

[Choudhury *et al.*, 2010] Munmun De Choudhury, Moran Feldman, Sihem Amer-Yahia, Nadav Golbandi, Ronny Lempel, and Cong Yu. Automatic construction of travel itineraries using social breadcrumbs. In *Proceedings of HT'10*, pages 35–44, 2010.

[Cohen and Katzir, 2008] Reuven Cohen and Liran Katzir. The generalized maximum coverage problem. *Information Processing Letters*, 108(1):15–22, 2008.

[Damianos Gavalas, 2014] Konstantinos Mastakas Grammati Pantziou Damianos Gavalas, Charalampos Konstantopoulos. A survey on algorithmic approaches for solving tourist trip design problems. *Journal of Heuristics*, 20(3):291–328, 2014.

[Gionis *et al.*, 2014] Aristides Gionis, Theodoros Lappas, Konstantinos Pelechrinis, and Evimaria Terzi. Customized tour recommendations in urban areas. In *Proceedings of WSDM'14*, pages 313–322, 2014.

[Kohavi, 1995] Ron Kohavi. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of IJCAI'95*, pages 1137–1145, 1995.

[Kurashima *et al.*, 2010] Takeshi Kurashima, Tomoharu Iwata, Go Irie, and Ko Fujimura. Travel route recommendation using geotags in photo sharing sites. In *Proceedings of CIKM'10*, pages 579–588, 2010.

[Kurashima *et al.*, 2013] Takeshi Kurashima, Tomoharu Iwata, Go Irie, and Ko Fujimura. Travel route recommendation using geotagged photos. *Knowledge and Information Systems*, 37(1):37–60, 2013.

[Lim, 2015] Kwan Hui Lim. Recommending tours and places-of-interest based on user interests from geo-tagged photos. In *Proceedings of the 2015 SIGMOD PhD Symposium*, 2015.

[Lucchese *et al.*, 2012] Claudio Lucchese, Raffaele Perego, Fabrizio Silvestri, Hossein Vahabi, and Rossano Venturini. How random walks can help tourism. In *Proceedings of ECIR'12*, pages 195–206, 2012.

[Miller *et al.*, 1960] Clair E. Miller, Albert W. Tucker, and Richard A. Zemlin. Integer programming formulation of traveling salesman problems. *Journ. of the ACM*, 7(4):326–329, 1960.

[Quercia *et al.*, 2014] Daniele Quercia, Rossano Schifanella, and Luca Maria Aiello. The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city. In *Proceedings of HT'14*, pages 116–125, 2014.

[Sinnott, 1984] Roger W. Sinnott. Virtues of the Haversine. *Sky and Telescope*, 68(158), 1984.

[Souffriau and Vansteenwegen, 2010] Wouter Souffriau and Pieter Vansteenwegen. Tourist trip planning functionalities: State-of-the-art and future. In *Proc. of ICWE'10*, pages 474–485, 2010.

[Thomee *et al.*, 2015] Bart Thomee, David A. Shamma, Benjamin Elizalde Gerald Friedland, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. The new data and new challenges in multimedia research. In *arXiv:1503.01817*, 2015.

[Tsiligirides, 1984] Theodore Tsiligirides. Heuristic methods applied to Orienteering. *Journal of the Operational Research Society*, 35(9):797809, 1984.

[Vansteenwegen and Oudheusden, 2007] Pieter Vansteenwegen and Dirk Van Oudheusden. The mobile tourist guide: An OR opportunity. *OR Insight*, 20(3):21–27, 2007.

[Vansteenwegen *et al.*, 2011a] Pieter Vansteenwegen, Wouter Souffriau, Greet Vanden Berghe, and Dirk Van Oudheusden. The city trip planner: An expert system for tourists. *Expert Systems with Applications*, 38(6):6540–6546, 2011.

[Vansteenwegen *et al.*, 2011b] Pieter Vansteenwegen, Wouter Souffriau, and Dirk Van Oudheusden. The Orienteering problem: A survey. *European Journal of Operational Research*, 209(1):1–10, 2011.