Sentiment-Aware and Personalized Tour Recommendation

Prarthana Padia*, Kwan Hui Lim†, Jeffrey Chan‡, and Aaron Harwood*
*School of Computing and Information Systems, The University of Melbourne, Australia
†Information Systems Technology and Design Pillar, Singapore University of Technology and Design, Singapore
‡School of Science, RMIT University, Australia
Email: ppadia@student.unimelb.edu.au, kwanhui_lim@sutd.edu.sg, jeffrey.chan@rmit.edu.au, aharwood@unimelb.edu.au

Abstract—**Itinerary planning is one of the most important tasks in tourism. A well-planned itinerary enhances the tourist experience and their visit satisfaction in new cities. However, the task of planning personalized tour itineraries is complicated by tourists with different interest preferences. Furthermore, there is an added complexity of recommending an itinerary with discrete budget, time and cost. Due to an increase in web-technologies and online geo-location services, there is emerging research targeting itinerary recommendation based on each tourist's interest, preferences and trip constraints. While several research works consider tourist interest, they adopt a simple measure based on the number of times a tourist has visited a place or the number of photos taken by the tourist at a place. Our research proposes an improved sentiment-aware personalized tour planner that considers each tourist's interests based on his/her sentiments on specific categories relative to his/her overall preferences. Unlike the previous approaches that do not consider the actual opinion based preferences, our proposed approach determines user interests based on their sentiments associated with their written text about a place of their visit. This interest measure is based on the intuition that users are more likely to post favorable comments about places they like. Using a dataset from Twitter, we compare our proposed algorithm against the baseline and experimental results show that our algorithm obtained superior performance in terms of tour precision, recall, F1-score and overall popularity.**

Index Terms—**Tour Recommendations; Trip Planning; Recommendation Systems; Personalization; Sentiment Analysis**

I. INTRODUCTION

Tour planning is an important but difficult task for people visiting unfamiliar cities. Personalized tour recommendation is a popular but challenging problem as every traveller has different interest preferences. Visitors are faced with the challenge of identifying popular places aligned with their personal interests. In addition, there is an added complexity due to the need to schedule visits to all recommended places while considering the available tour budget, time and cost.

During such tour visits, people frequently upload pictures and posts relating to the places they have visited in their travel itinerary. Due to the growing popularity of social media, it has become a significant data source for analyzing user preferences and planning personalized itinerary strategies [1], [2]. Despite the availability of information on Internet about travel guides and famous places, visitor's personal interests and their trip constraints are not considered.

With the advancement in web technologies and geo-location services, there is an increase in the amount of online geo-tagged photos that facilitate the modelling of user interest, preferences and trip constraints while strategizing itinerary planning. The time-stamp along with the geo-tag helps to identify the amount of time people spend at a particular place at different hours of the day and different days of the week. While earlier works consider user interest, they adopt a simple measure based on the number of times a user has visited a place.

Recent works solve this tour planning problem by generating itineraries based on photo frequency of the user at different Points of Interests (POIs) [3], user visit duration at POIs [4], weather conditions and seasons [5]–[7], traffic conditions [8], [9], queuing time awareness at POIs [10], aspatial semantics involving user preferences, other users' visiting sequences [11], etc. Along with spatial-temporal information, the tags and text of the geo-tagged post provide crucial information about the personal user experience during the visit. The text of the post can be used to analyze the sentiments of each user, associated with his/her POI visit. Unlike conventional approaches that do not consider the actual opinion based preferences of the users, our sentiment-aware approach considers user preferences based on their sentiments associated with their written text at a place of their visit. We propose an enhanced tour recommendation system that considers user sentiment based approach along with the interest preferences and other trip constraints to suggest better personalized tour itineraries.

Main Contributions

Unlike earlier works which adopt a simplistic definition of user interest based on visit counts or photo frequency, we propose a tour recommendation system that utilizes a novel user-sentiment relative measure of interest preference build upon the Orienteering problem.

We propose two approaches of user-specific interest preferences based on large collection of geo-tagged tweets available

online. The first approach is used as a baseline approach that aims to recommend an itinerary by considering the visit frequency of each POI by all users. It utilizes the visit frequency to determine the popularity of each POI and recommends the itinerary based on the most popular POIs. The second proposed approach uses sentiment-awareness on textual content obtained from location based social network posts for recommending personalized itineraries, by combining sentiment-analysis techniques with personalization and path planning algorithm. This approach focuses on user opinions and views derived from textual content that a user posted about a specific visited POI, to cater to specific category of interest to the user. For instance, if a user has shown positive sentiment towards museum and multiplexes but not outdoors and beaches, then the recommendation system considers that while building the personalized itinerary.

II. RELATED WORK

Traditionally, tour recommendation systems are based on greedy heuristics, calculating the shortest path in terms of distance or time [12], [13]. Recent research include more personalized and detailed approaches for automatic generation of itineraries based on parameters like user preferences, location history, trip constraints etc [3], [4], [12]–[25]. The data source used for these experiments is frequently obtained from location based social networks (LBSNs) or similar social networks with geo-tagged content.

Recommending an optimal trip that maximizes user's experience and considers the budget constraints, is an NP hard problem. There are several approaches of modelling similar tour recommendation problem and a large number of them are based on the Orienteering problem and its variants [3], [4], [20], [26]. In the following sections, we briefly review the Orienteering problem followed by some key researches in the field of tour recommendation.

A. Orienteering problem overview

The Orienteering problem aims to determine routing among multiple nodes where each node has a score associated with it. The objective is to maximize the total score which is the result of visiting several nodes within the provided time constraint and budget [27]. [28], [29] provide an in-depth review of Orienteering problem. [28] introduces the competition/sport of Orienteering as an event where the event is scored based on an optimization problem of maximizing the total obtainable score given a limited competition time. On the other hand, [29] reviews the Orienteering problem, its implications by comparing and discussing published approaches and established heuristics for solving the Orienteering problem.

B. Variants of Tour Recommendation

Tour recommendation is a well-studied field that includes various research works with different approaches. The approaches typically focus on maximizing user preferences within the given trip constraints [3], [16], [18]–[20], [30]. [14] suggests an approach to recommend itineraries to people travelling to a new city with no previous history based on offline modelling and online recommendation of places learnt from social opinions of local experts. [12] proposed a trip builder where user tour itineraries were matched to touristic POIs from Wikipedia and modelled as an instance of Generalized Maximum coverage problem. As an extension of the work of [12], [13] proposed a solution based on an instance of the Travelling Salesman Problem. [16] proposed a unique method of using POI and route information as features to a machine learning algorithm to recommend tour routes. There are also several researches that modelled the tour recommendation problem based on the Orienteering problem [3], [31], [32] considering several POI visit sequences and trip constraints. Several research works also considered real life constraints like POI availability, travelling time uncertainty and POI diversity to recommend personalized tours [24], [25]. Parameters modelling real time constraints like queuing time awareness [10], visit duration [33], visit recency [4], photo frequency [3], route attractiveness [34]–[36] and data on pedestrian crowdedness [23] significantly aid the itinerary recommendations. Furthermore, there has been development of various mobile and online applications for the purpose of tour recommendation [21], [37]–[40].

C. Differences with earlier works

Our proposed approach is based on sentiment analysis and differs from the conventional tour recommendation works in various aspects. We derive user-based interest from user's post (tweet) by analyzing the sentiment associated with the text of the post about POIs of specific categories. While there are interesting approaches that uses sentiment for recommendation purposes [41]–[43], many of these works focus on individual item recommendation, whereas the tour recommendation tasks includes the additional consideration of various trip constraints and multiple items. The current state-of-art tour recommendation approaches use either time-based user interest (based on POI visit duration), visit frequency based user interest, photo-frequency based user interest or explicitly mentioned user interest preferences for itinerary recommendation. These approaches do not consider the actual user opinion about the place of visit and provides an approximation of user interest, where interest could have positive or negative views associated with it. In contrast, we propose an enhanced tour recommendation system that considers user sentiment based approach along with the interest preferences and other trip constraints to suggest better personalized tour itineraries.

III. PROBLEM FORMULATION AND ALGORITHMS

In this section, we first present our problem formulation along with the constraints, followed by the definitions of the baselines and proposed algorithms. Following which, we discuss the dataset used for the experiments.

A. Problem Formulation

Similar to many earlier works [3], we model our recommendation problem based on the Orienteering problem [27]–[29].

In this tour recommendation problem, our main objective is to recommend a tour itinerary $I = (p_1, ..., p_N)$ that maximizes the total profit from visiting the list of POIs p_1 to P_N, while ensuring that the tour itinerary can be completed within a specific time budget B. Given a set of POIs P, we are trying to optimize for:

$$Max \sum_{p_i \in P} \sum_{p_j \in P} Path_{p_i,p_j} \left(\eta Int_u(p_i) + (1 - \eta) Pop(p_i) \right) \quad (1)$$

where $Path_{p_i,p_j} = 1$ if a path between POI p_i and p_j is selected as part of the itinerary, and $Path_{p_i,p_j} = 0$ otherwise. $Int_u(p_i)$ represents a user-specific interest score of how interesting POI p_i is to user u, while $Pop(p_i)$ indicates the general popularity of POI p_i.

In addition, Equation 1 is subjected to the following constraints:

(i) starting and ending at specific POIs

(ii) connectivity of POIs in the itinerary

(iii) completing the itinerary within a specific time or distance budget B.

$$\sum_{p_i \in I} Path_{p_s,p_i} = \sum_{p_j \in I} Path_{p_j,p_d} = 1 \quad (2)$$

Constraint (i) (Equation 2) ensures that the recommended itinerary starts at a specific POI p_s, and ends at another specific POI p_d. In real-life, this starting and destination POIs would correspond to POIs near the hotel that a tourist is staying at.

$$\sum_{p_i,p_k \in I} Path_{p_i,p_k} = \sum_{p_j,p_k \in I} Path_{p_k,p_j} \leq 1 \quad (3)$$

Constraint (ii) (Equation 3) ensures that the recommended itinerary fulfills two conditions, namely: (i) all selected paths are connected as a full itinerary; and (ii) no POIs are visited more than once.

$$\sum_{p_i \in I} \sum_{p_j \in I} Cost(p_i, p_j) Path_{p_i,p_j} \leq B \quad (4)$$

Constraint (iii) (Equation 4) ensures that the recommended itinerary can be completed within a specific time or distance budget B.

B. Algorithms and Baselines

The two approaches applied for recommending itinerary are as follows:

(i) Baseline approach (PA_V): This approach is based on the visit frequency of POIs. Given a set of travel history of all users U, the popularity of the POI is determined using average visit frequency at each POI.

(ii) Proposed sentiment-aware approach (PA_S): This approach is based on the sentiment scores of users associated with each POI. Given a set of travel history of all users U, the interest relevance of the POI is determined using the average sentiment score at each POI.

Multiple categories of places can comprise multiple POIs in a city. Consider m POIs for a particular city. Let $P = \{p_1, .., p_m\}$ be the set of POIs in that city. Each POI p has a category Cat_p (e.g., retail, education centre, community use) and latitude/longitude coordinates associated with it. Following which, we present the key notations and definitions used in this paper.

Definition 1: Travel History.

1) PA_V: Given a user u who has visited n POIs, the travel history is determined as an ordered sequence, $S_V = ((p_1, v_{p1}), (p_2, v_{p2})..., (p_n, v_{p_n}))$, where each duplet (p_x, v_{px}) comprises the visited POI p_x, and the number of visits at POI p_x.

2) PA_S: Given a user u who has visited n POIs, the travel history is determined as an ordered sequence, $S_S = ((p_1, s_{p1}), (p_2, s_{p2})...(p_n, s_{p_n}))$, where each duplet (p_x, s_{p_x}) comprises the visited POI p_x, and the sentiment score at POI p_x.

Definition 2: Average POI popularity.

1) PA_V: Given a set of travel histories of all users U, the system determines the popularity of the POI using average visit frequency at each POI.

$$\overline{PV}(p) = \frac{1}{n} \sum_{u \in U} \sum_{p_x \in S_V} (v_{px}) \delta(P_x = P) \forall p \in P \quad (5)$$

where n is the number of visits at POI p by all users U and $\delta(p_x = p) = 1$, if $px = p$ and 0, otherwise.

2) PA_S: Given a set of travel history of all users U, the system determines the interest relevance of the POI using average sentiment score of each POI.

$$\overline{PS}(p) = \frac{1}{n} \sum_{u \in U} \sum_{p_x \in S_S} (s_{px}) \delta(P_x = P) \forall p \in P \quad (6)$$

where n is the total sentiment scores of POI p by all users U and $\delta(p_x = p) = 1$, if $px = p$ and 0, otherwise.

Definition 3: User Interest functions.

As mentioned earlier, the category of a POI p is represented as Cat_p. Given that C denotes the set of all POI categories, the interest of a user u in POI category c is denoted as follows:

1) PA_V:

$$Int_u^V(c) = \sum_{p_x \in S_V} \frac{(v_{px})}{\overline{PV}(p_x)} \delta(Cat_{px} = C) \; \forall c \in C \quad (7)$$

$$\text{where } \delta(Cat_{px} = C) = \begin{cases} 1, & \text{if } Cat_{px} = C \\ 0, & \text{otherwise} \end{cases}$$

2) PA_S:

$$Int_u^S(c) = \sum_{p_x \in S_S} \frac{(s_{px})}{\overline{PS}(p_x)} \delta(Cat_{px} = C) \; \forall c \in C \quad (8)$$

$$\text{where } \delta(Cat_{px} = C) = \begin{cases} 1, & \text{if } Cat_{px} = C \\ 0, & \text{otherwise} \end{cases}$$

In short, the above equations for our proposed sentiment-aware approach are used to model the interest of a user in a particular POI category c based on sentiment score of each POI of category c, relative to the average sentiment score of all users at the same POI. Our reasoning for this formulation is that a user is likely to write more positive tweet text about the POI that he/she is interested in. Thus, by calculating the sentiment score (positive and negative) of a user, the interest level of that user in POIs of this category is determined. We use the popular open-source *vader sentiment-analysis* module [44] for calculating the sentiment scores of the tweet text.

Based on the above problem definitions, we then proceed to solve the tour recommendation problem as a variant of Integer Programming (IP) problem. The *lpsolve* linear programming module [45] is used to solve the IP problem.

IV. Dataset

For our experiment and analysis, we utilized a Twitter dataset collected for Melbourne, Australia similar to [46], which comprises 2.2 million geo-tagged tweets that are not mapped to POIs yet. Twitter is a suitable data source for this task and has been frequently used for similar location-based works such as event detection [47]–[50]. These geo-tagged tweets were then mapped to a list of POIs based on their respective entries on the City of Melbourne's Open Data Platform [51], i.e., proximity of geo-tagged tweets to the entries of POIs based on their latitude/longitude coordinates. In a similar manner, the categories of POIs are based on their respective entries listed on the City of Melbourne's Open Data Platform. The dataset furthermore comprises information like the date and timestamp of tweets, geo-location coordinates of tweets, user ID, tweet text, hashtags etc. The descriptive statistics of the dataset is shown in Table I.

V. Data Analysis Results

In this section, we discuss the analysis of our sentiment-based approach, performed over geo-tagged posts from Twitter (as per Table I). The following sections describe several data graphs and charts involving comparison of sentiment scores of

TABLE I
DATASET DESCRIPTION

| POI theme	No. of POIs	No. of Users	No. of Tweets
Community Use	19	12,296	73,280
Education Centre	13	1,190	3,385
Health Services	11	776	2,527
Leisure/Recreation	62	9,556	41,735
Mixed Use	11	5,776	22,035
Office	11	2,021	3,946
Place Of Assembly	40	8,670	42,197
Place of Worship	31	3,717	15,535
Purpose Built	4	2,253	5,114
Residential Accommodation	2	89	164
Retail	3	3,044	8,324
Specialist Residential Accom.	1	72	143
Transport	18	5,193	20,114
Vacant Land	3	189	433
Total: 14	**229**	**54,842**	**238,942**
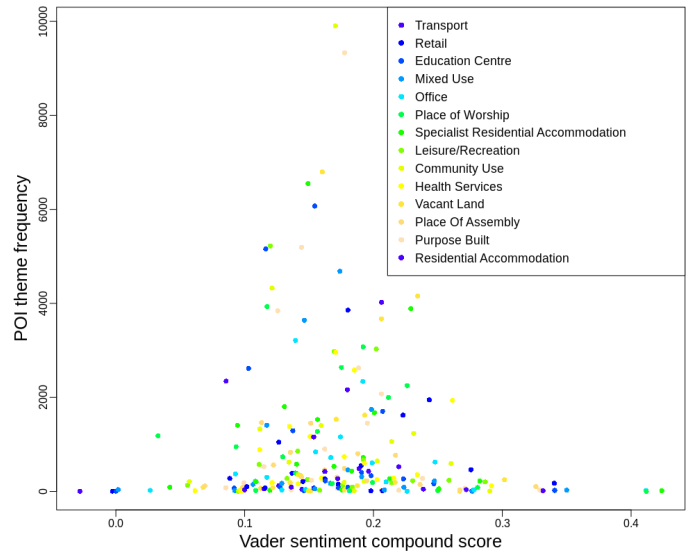
Fig. 1. Scatter plot of Vader compound sentiment score against frequency of visit for each POI with filled coloured circles and group legend.

posts with several other features like visit frequency, time of post and hashtags. The aim is to infer the relationships between post sentiments and other mentioned features of posts, which will show the usefulness of sentiment scores as a user interest function for recommending itineraries.

A. Correlation between POI visit frequency and user sentiment scores corresponding to POIs

The correlation plot in Figure 1 contains the analysis performed over the tweet data, described in Table I. The x-axis of the plot corresponds to the average compound (overall)
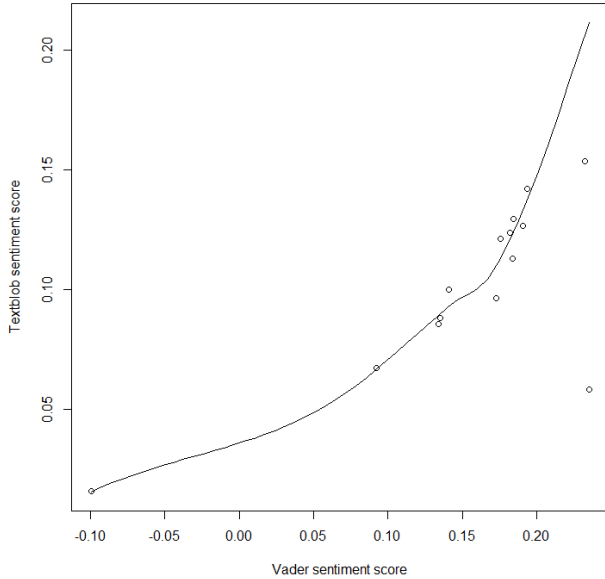
Fig. 2. Comparison of average sentiment scores of each POI category using Vader sentiment analyzer and textblob sentiment analyzer (based on python 'nltk' and 'pattern') modules.

sentiment score of each POI ranging from -1 to 1. The compound sentiment score comprises the weighted positive, negative and neutral sentiment scores of the text. The y-axis corresponds to the average visit frequency of each POI. The legend shows the respective categories that each of the 229 POIs belong to.

The plot shows that POIs with lower sentiment score and higher sentiment score have less visit frequency compared to POIs with moderate sentiment score. This trend indicates that places like residential accommodation, place of assembly which are occasionally visited gets extreme (either low or high) sentiment scores from people. Due to the calculation of sentiment scores based on fewer tweets and visit frequency, the overall score for such POIs end up being extreme. As the visit frequency increases and the place becomes frequently visited, a higher number of people start posting their views about that place. Thus places belonging to the categories of "education centre" and "community use" results in generally moderate sentiment score due to increased number of positive/negative sentiment tweets that average to a moderate scale.

B. Comparison of sentiment analysis modules

The scatter plot in Figure 2 shows the POI category-wise correlation between the two open source sentiment analysis modules – *Vader* [44] and *Textblob* [52] – analyzed over a sample set comprising of 5% (13,300+ tweets) of the whole dataset. The aim of this comparison is to find out the level of similarity between the scores of the two modules, in order to check the validity of the scores used for modelling the sentiment-based interest approach.

The plot includes the best fit increasing curve, deviating only slightly from the ideal positive correlation line. Except for a few outliers, not in a far range, the score data adheres to the increasing curve. This result indicates that both sentiment analysis modules give out quite similar scores for analysis, thus showing that our approach is able to work properly, independent of the different choices of sentiment analysis technique.

C. POI categories and sentiments across hours and days

The plots in Figure 3 show the average compound sentiment scores (along with the standard deviation bounds of the distribution) for different hours of the day and different days of the week, corresponding to five POI categories: (i) community user; (ii) education centre; (iii) leisure/recreation; (iv) retail; and (v) transport.

Sentiments across hours: POI categories like education centre, community use and retail show a peak in sentiment score during morning hours. This correlates with the usual visit duration of these category of places, i.e. morning. The standard error shows higher deviation in some categories during 4 - 5am time. This could be due to higher variability in the scores or lesser number of tweets due to those (contextually) odd hours.

Sentiments across days: POI categories like retail, transport, education centre display noticeable rise in sentiment score during the weekdays. As these places are related to the daily routine of people, they are expected to have higher scores in weekdays. Categories like leisure/recreation and community use show a slightly higher sentiment towards the weekend. People are expected to visit these places more during the weekends compared to the weekdays. The standard error limits are comparatively lower and constant compared to the hour wise plots. Note that Day 1 corresponds to Monday, Day 2 Tuesday and so on.

Overall, the plots show the expected pattern of user sentiments with respect to various times of the post. For instance, generally users are more likely to post favourable comments during the day time, at the above mentioned categories, than at nights. The figures describe the anticipated user pattern with respect to user sentiments, validating the use of sentiment scores as a user interest function to model the sentiment-based interest approach.

D. Word clouds of hashtags for POI categories

The word clouds in Figure 4 highlight the prominent hashtags from tweets regarding specific POI categories like (i) transport; (ii) retail; (iii) leisure/recreation; (iv) education centre; and (v) transport.

Hashtags like *unimelb*, *rmit*, *artschool*, *university*, *endomondo*, *endorphins* are prominent in the category of "Education centre". These indicate the education centres and topics popular in Melbourne. Whereas, *queensvictoriamarket*, *nightmarket*,
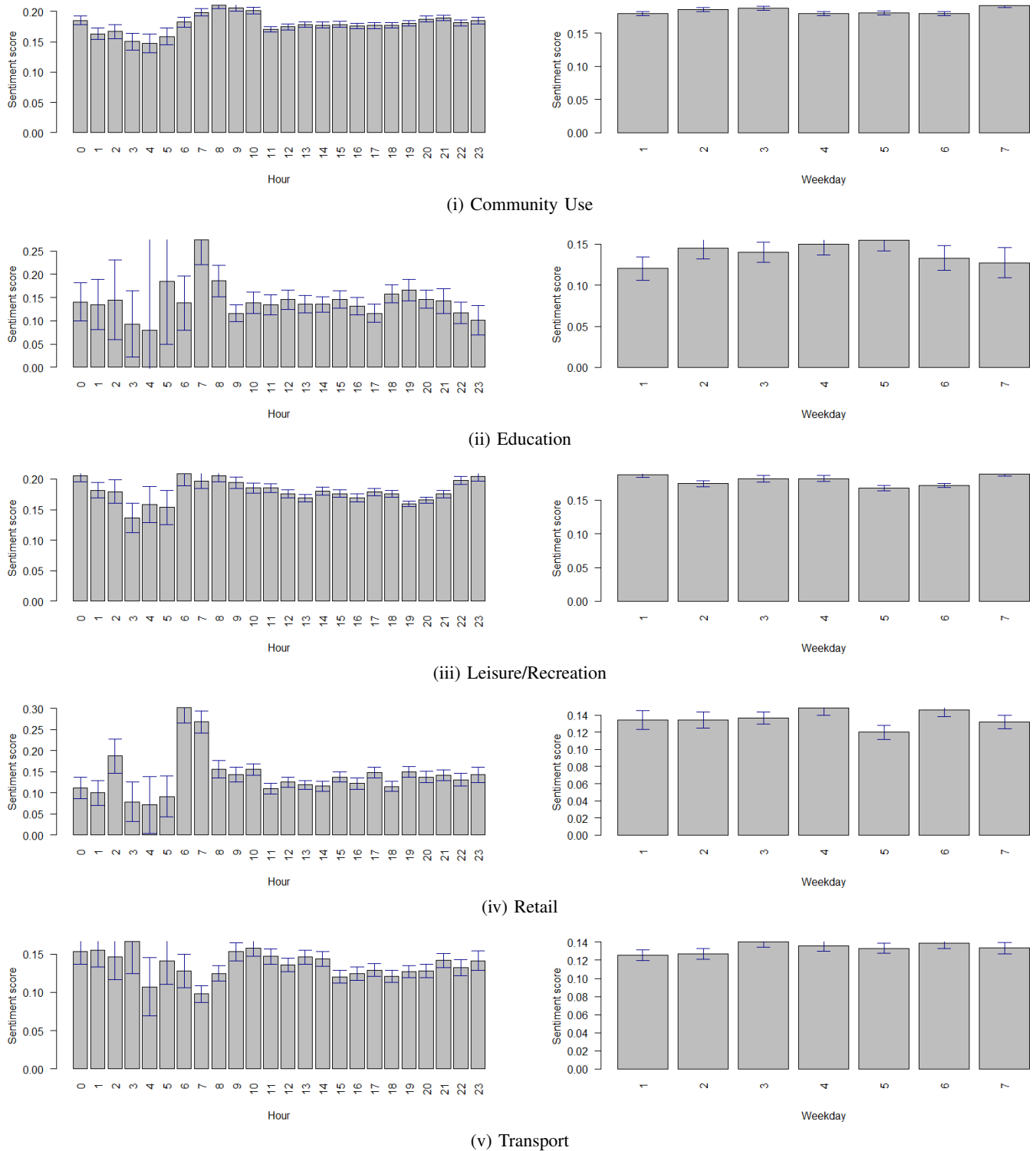
Fig. 3. Tweet hour and weekday comparison against their respective average Vader sentiment compound score for the following POI categories: (i) Community Use; (ii) Education; (iii) Leisure/Recreation; (iv) Retail; (v) Transport

vicmarket, *market*, *bourkestreet* are prominent in "Retail" category, indicating the market areas (retail) of Melbourne. *flinderstreetstation*, *melbournecentralstation*, *flinderstreet*, *docklands* are notable in "Transport" category, which are popular stations and transport destinations of Melbourne. Lastly, *ausopen*, *australiaopen*, *tennis*, *mcg*, *afl* are outstanding in the category of "Leisure and Recreation". These hashtags are related to popular sports being played in Australia. Though these hash-

Fig. 4. Word cloud of Twitter hashtags prominent for the following POI categories (clockwise from top left): Community Use; Education; Leisure/Recreation; Transport; and Retail.

tags can be dynamic and are not constant events for attracting tourists, these are the famous sport events that play a role in attracting tourists to Australia. The above word clouds comprise of the most frequently used hashtags from the posts, corresponding to their respective POI categories.

VI. EXPERIMENT METHODOLOGY

Evaluation and Metrics.

We evaluated our algorithm and the baselines using leave-one-out cross-validation, which involves evaluating a specific travel sequence of a user, while using his/her other travel sequences as training data. The following evaluation metrics were used:

1) **Tour Precision:** $T_p(I)$. This metric is defined as the proportion of POIs recommended in itinerary I, that matched 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 actual travel sequence then the tour precision is defined as: $T_p(I) = \frac{|P_r \cap P_v|}{|P_r|}$.

2) **Tour Recall:** $T_r(I)$. This metric is defined as the proportion of POIs in a user's actual 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 then the tour recall is defined as: $T_p(I) = \frac{|P_r \cap P_v|}{|P_v|}$.

3) **Tour F_1−score:** $T_{F_1}(I)$. This metric is the harmonic mean of both the recall and precision of a recommended tour itinerary I. Mathematically, it is defined as: $T_{F_1}(I) = \frac{2 \times T_p(I) \times T_r(I)}{T_p(I) + T_r(I)}$.

4) **Tour Popularity:** $T_{pop}(I)$. This metric is the overall popularity of all POIs in the recommended itinerary I. Also defined as $T_{pop}(I) = \sum_{p \in I} Pop_p$.

5) **Tour Interest:** $T_{Int}^u(I)$. This metric is the overall interest of all POIs in the recommended itinerary I to a user u. Also defined as: $T_{Int}^u(I) = \sum_{p \in I} Int_u(Cat_p)$.

VII. EXPERIMENT RESULTS AND DISCUSSION

The comparison between Sentiment based User Interest($PA_{S.25}$, $PA_{S.5}$, PA_{S1}), Visit frequency based User Interest($PA_{V.25}$, $PA_{V.5}$, PA_{V1}) and a combination of both Sentiment and Visit based User Interests($PA_{VS.25}$, $PA_{VS.5}$, PA_{VS1}) with respect to interest weights 0.5 and 1 has been carried out with a dataset comprising of 31000+ user sequences, which was 40% of the total dataset. The interest weight indicates the balance between place popularity and user interest level while modeling the objective function. 0.5 interest weight indicates an equal balance between popularity

TABLE II

COMPARISON OF SENTIMENT-BASED USER INTEREST($PA_{S.25}$, $PA_{S.5}$, PA_{S1}), VISIT FREQUENCY BASED USER INTEREST($PA_{V.25}$, $PA_{V.5}$, PA_{V1}) AND A COMBINATION OF BOTH SENTIMENT AND VISIT FREQUENCY BASED USER INTERESTS($PA_{VS.25}$, $PA_{VS.5}$, PA_{VS1}) WITH RESPECT TO INTEREST WEIGHTS 0.5 AND 1, IN TERMS OF PRECISION (T_P), RECALL (T_R) AND F1-SCORE (T_{F1}), POPULARITY (T_{pop}) AND INTEREST (T_{int}).

Algo	Precision	Recall	F1-Score	Interest	Popularity
$PA_{V.25}$	0.729 ± 0.026	0.827 ± 0.015	0.760 ± 0.022	1.109 ± 0.113	0.674 ± 0.045
$PA_{S.25}$	0.729 ± 0.026	0.827 ± 0.015	0.760 ± 0.022	0.806 ± 0.102	0.674 ± 0.045
$PA_{VS.25}$	$\mathbf{0.734 \pm 0.027}$	$\mathbf{0.833 \pm 0.015}$	$\mathbf{0.764 \pm 0.023}$	$\mathbf{1.120 \pm 0.114}$	$\mathbf{0.706 \pm 0.045}$
$PA_{V.5}$	0.670 ± 0.023	0.796 ± 0.013	0.710 ± 0.020	$\mathbf{1.340 \pm 0.106}$	0.628 ± 0.041
$PA_{S.5}$	$\mathbf{0.701 \pm 0.023}$	$\mathbf{0.818 \pm 0.013}$	$\mathbf{0.738 \pm 0.020}$	1.019 ± 0.115	0.647 ± 0.040
$PA_{VS.5}$	0.654 ± 0.024	0.799 ± 0.013	0.698 ± 0.020	1.259 ± 0.104	$\mathbf{0.660 \pm 0.040}$
PA_{V1}	$\mathbf{0.695 \pm 0.031}$	$\mathbf{0.798 \pm 0.019}$	$\mathbf{0.729 \pm 0.027}$	1.083 ± 0.127	0.416 ± 0.049
PA_{S1}	$\mathbf{0.695 \pm 0.031}$	$\mathbf{0.798 \pm 0.019}$	$\mathbf{0.729 \pm 0.027}$	0.998 ± 0.130	0.416 ± 0.049
PA_{VS1}	0.679 ± 0.031	0.797 ± 0.018	0.717 ± 0.027	$\mathbf{1.224 \pm 0.142}$	$\mathbf{0.423 \pm 0.050}$

and user interest whereas 1 considers only user interest level. With similar tour budget and cost, the experiment results show that sentiment based algorithm outperforms the baseline in most terms.

The evaluation between Sentiment based User Interest($PA_{S.25}$, $PA_{S.5}$, PA_{S1}), Visit frequency based User Interest($PA_{V.25}$, $PA_{V.5}$, PA_{V1}) and a combination of both Sentiment and Visit frequency based User Interests($PA_{VS.25}$, $PA_{VS.5}$, PA_{VS1}) has been carried out with tweets across the city of Melbourne, Australia. Three interest weights of 0.25, 0.5 and 1 have been used for the evaluation. The experimental dataset comprises of 31k user sequences, randomly sampled from 40% of the total dataset. With similar tour budget and cost, the experiment results show that sentiment based algorithm outperforms the baseline in most cases, which we will discuss further. The experiment results are shown in Table II.

A. Comparison in terms of Popularity and Interest metrics.

Table II shows that in terms of popularity (T_{pop}) metric, the proposed sentiment-aware interest algorithms with interest weights of 0.25 and 0.5 ($PA_{S.25}$ and $PA_{S.5}$) outperform the baseline algorithms. Whereas, the combined approach methods ($PA_{VS.25}$, $PA_{VS.5}$ and PA_{VS1}) perform better than the two stand-alone approaches in terms of overall tour popularity. Overall, the algorithms with interest weight 0.25 outperform the rest in terms of tour popularity, implying visit recommendations to more popular and famous places.

Furthermore, in terms of interest (T_{int}) metric, the proposed sentiment-aware approaches have slightly lower performance compared to the baseline and the combined algorithms. The interest level of the combined approach is mostly higher or comparable to the stand-alone baseline approach. Though

the interest metric is used to determine the most optimal recommendation solution, it may not be necessarily realistic in terms of recommendation of places.

B. Comparison in terms of Precision, Recall and F1-Score.

From Table II, we can observe that with regards to precision (T_p), recall (T_r) and F1-score (T_{F_1}) the proposed sentiment-aware algorithm with interest weight of 0.25 ($PA_{S.25}$) offers a similar performance as the baseline ($PA_{V.25}$). Although the combination of both sentiment-based approach and visit frequency approach ($PA_{VS.25}$) outperforms the previous two methods in terms of all the three metrics. The combination approach also outperforms the rest of the proposed and baseline methods of interest weights 0.5 and 1, standing out to be the best. Considering interest weights 0.5 and 1, the proposed approach with interest weight of 0.5 ($PA_{S.5}$) outperforms the visit frequency baseline algorithms ($PA_{V.5}$ and PA_{V1}) and the combined interest algorithms ($PA_{VS.5}$ and PA_{VS1}) in the experiments by several percentages, standing out to be superior. Though PA_{S1} performs at par with the baseline PA_{V1} when interest weight is considered to be 1. The baseline algorithms $PA_{V.5}$ and PA_{V1} outperforms their respective combined algorithms having the same interest weights.

Overall, the algorithms with interest weight of 0.25 ($PA_{V.25}$, $PA_{S.25}$ and $PA_{VS.25}$) have better performance compared to the other approaches. This indicates that the algorithms that give a smaller emphasis on user interest and larger emphasis on popularity provide the best quality of prediction for tour recommendation. Furthermore, the results suggest that the combined approach (for interest weight of 0.25) offers the best performance, followed by the sentiment-aware approach (for interest weight of 0.5), which performs at par with the baseline approach. Hence, the approaches considering both user interest

and place popularity give better performance for predicting where a user is more likely to travel in reality, compared to the stand-alone visit frequency baseline approach. The combined approach is the best followed by the proposed sentiment-aware approach, which is better or comparable to the visit frequency baseline, in terms of relevance of the recommended itinerary.

The overall results suggest that the algorithms using the combined approach performs better than the algorithms using stand-alone approaches in terms of tour popularity and user interest levels.

VIII. CONCLUSION

We modelled the tour recommendation problem as an instance of Orienteering problem and proposed a sentiment-aware approach for recommending personalized tours. We used geo-tagged tweets to determine sentiments of the users associated with specific POIs and automatically derived user interest and popularity of POI to train the algorithm. Our work improves upon the earlier research in several ways: (i) we introduce sentiment-aware user interest derived from user posts at a POI of a specific category, unlike earlier works based on visit or photo frequency based user interest; and (ii) we improve upon the personalization level for each user by recommending itineraries based on user's personal views and opinions and identifying the positive or negative connotation associated with the interest level.

Using a Twitter dataset across several POI categories, we present a data analysis on how different POI categories correlate with different sentiment scores at varying time scales. Our proposed algorithms are evaluated against the frequency-based user interest baseline and the combination approach (includes both sentiment-based and frequency-based user interests) in terms of precision, recall, F1-score, tour popularity and interest. The experiment results show that: (i) the combined approach outperform both individual user interest approaches, followed by sentiment-based user interest approach in terms of precision, recall, F1-score and overall tour popularity; and (ii) sentiment-based interest approach gives a better prediction of where a user is more likely to travel compared to the baseline approach by considering the reviews and opinions of the users and using the positive or negative connotation as a proxy for their interest levels.

IX. FUTURE WORK

Some potential directions for future work are: (i) using improved and sophisticated sentiment analysis techniques that allows us to generalize across sentiments in different types of language in addition to English; (ii) using sentiment awareness approach to recommend tour itineraries for groups of users sharing similar interests and preferences; and (iii) recommending personalized tour itineraries considering the public transport arrival and departure time to minimize the waiting time and facilitate realistic tour planning by modelling real time uncertainty of public transport.

REFERENCES

[1] H. Q. Vu, G. Li, R. Law, and B. H. Ye, "Exploring the travel behaviors of inbound tourists to hong kong using geotagged photos," *Tourism Management*, vol. 46, pp. 222–232, 2015.

[2] B. Kadar, "Measuring tourist activities in cities using geotagged photography," *Tourism Geographies*, vol. 16, no. 1, pp. 88–104, 2014.

[3] P. Padia, B. Singhal, and K. H. Lim, "User-relative personalized tour recommendation," in *Joint Proceedings of the ACM Intelligent User Interfaces 2019 Workshops*, 2019.

[4] K. H. Lim, J. Chan, S. Karunasekera, and C. Leckie, "Tour Recommendation and Trip Planning using Location-based Social Media: A Survey," *Knowledge and Information Systems*, vol. 60, no. 3, p. 1247–1275, 2019.

[5] 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*, vol. 95, 2015.

[6] T. Yamasaki, A. Gallagher, and T. Chen, "Personalized intra-and inter-city travel recommendation using large-scale geotags," in *Proceedings of the 2nd ACM International Workshop on Geotagging and Its Applications in Multimedia (GeoMM'13)*, 2013, pp. 25–30.

[7] S. Jiang, X. Qian, T. Mei, and Y. Fu, "Personalized travel sequence recommendation on multi-source big social media," *IEEE Transactions on Big Data*, vol. 2, no. 1, pp. 43–56, 2016.

[8] T. Liebig, N. Piatkowski, C. Bockermann, and K. Morik, "Dynamic route planning with real-time traffic predictions," *Information Systems*, vol. 64, pp. 258–265, 2017.

[9] ——, "Predictive trip planning-smart routing in smart cities," in *Proceedings of EDBT/ICDT Workshop on Mining Urban Data (MUD'14)*, 2014, pp. 331–338.

[10] A. Gunawan, Z. Yuan, and H. C. Lau, "A mathematical model and metaheuristics for time dependent orienteering problem," in *Proceedings of the 10th International Conference of the Practice and Theory of Automated Timetabling (PATAT'14)*, 2014, pp. 202–217.

[11] G. Cai, K. Lee, and I. Lee, "Itinerary recommender system with semantic trajectory pattern mining from geo-tagged photos," *Expert Systems with Applications*, vol. 94, pp. 32–40, 2018.

[12] I. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso, "Where shall we go today? Planning touristic tours with TripBuilder," in *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM'13)*, 2013, pp. 757–762.

[13] I. R. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso, "On planning sightseeing tours with TripBuilder," *Information Processing & Management*, vol. 51, no. 2, pp. 1–15, 2015.

[14] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in *Proceedings of the 20th International Conference on Advances in Geographic Information Systems (SIGSPATIAL'12)*, 2012, pp. 199–208.

[15] P. Bolzoni, S. Helmer, K. Wellenzohn, J. Gamper, and P. Andritsos, "Efficient itinerary planning with category constraints," in *Proceedings of the 22nd ACM SIGSPATIAL international conference on advances in geographic information systems*. ACM, 2014, pp. 203–212.

[16] C. Chen, D. Zhang, B. Guo, X. Ma, G. Pan, and Z. Wu, "TripPlanner: Personalized trip planning leveraging heterogeneous crowdsourced digital footprints," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1259–1273, 2015.

[17] D. Chen, C. S. Ong, and L. Xie, "Learning points and routes to recommend trajectories," in *Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM'16)*, 2016, pp. 2227–2232.

[18] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, "Travel route recommendation using geotags in photo sharing sites," in *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM'10)*, 2010, pp. 579–588.

[19] ——, "Travel route recommendation using geotagged photos," *Knowledge and Information Systems*, vol. 37, no. 1, pp. 37–60, 2013.

[20] K. H. Lim, J. Chan, C. Leckie, and S. Karunasekera, "Towards Next Generation Touring: Personalized Group Tours," in *Proceedings of the 26th International Conference on Automated Planning and Scheduling (ICAPS'16)*, Jun 2016, pp. 412–420.

[21] 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 *Hellenic Conference on Artificial Intelligence*, 2014, pp. 615–629.

[22] K. Taylor, K. H. Lim, and J. Chan, "Travel itinerary recommendations with must-see points-of-interest," in *Companion of the The Web Conference 2018 on The Web Conference 2018*, 2018, pp. 1198–1205.

[23] X. Wang, C. Leckie, J. Chan, K. H. Lim, and T. Vaithianathan, "Improving Personalized Trip Recommendation to Avoid Crowds Using Pedestrian Sensor Data," in *Proceedings of the 25th ACM International Conference on Information and Knowledge Management (CIKM'16)*, Oct 2016, pp. 25–34.

[24] C. Zhang, H. Liang, K. Wang, and J. Sun, "Personalized trip recommendation with poi availability and uncertain traveling time," in *Proceedings of the 24th ACM International Conference on Information and Knowledge Management (CIKM'15)*, 2015, pp. 911–920.

[25] C. Zhang, H. Liang, and K. Wang, "Trip recommendation meets realworld constraints: Poi availability, diversity, and traveling time uncertainty," *ACM Transactions on Information Systems*, vol. 35, no. 1, 2016.

[26] A. Gionis, T. Lappas, K. Pelechrinis, and E. Terzi, "Customized tour recommendations in urban areas," in *Proceedings of the 7th ACM International Conference on Web Search and Data Mining (WSDM'14)*, 2014, pp. 313–322.

[27] A. Gunawan, H. C. Lau, and P. Vansteenwegen, "Orienteering problem: A survey of recent variants, solution approaches and applications," *Euro. Journal of Operational Research*, vol. 255, no. 2, pp. 315–332, 2016.

[28] T. Tsiligirides, "Heuristic methods applied to orienteering," *Journal of the Operational Research Society*, vol. 35, no. 9, pp. 797–809, 1984.

[29] P. Vansteenwegen, W. Souffriau, and D. V. Oudheusden, "The orienteering problem: A survey," *European Journal of Operational Research*, vol. 209, no. 1, pp. 1–10, 2011.

[30] C. Lucchese, R. Perego, F. Silvestri, H. Vahabi, and R. Venturini, "How random walks can help tourism," in *European Conference on Information Retrieval*, 2012, pp. 195–206.

[31] M. De Choudhury, M. Feldman, S. Amer-Yahia, N. Golbandi, R. Lempel, and C. Yu, "Automatic construction of travel itineraries using social breadcrumbs," in *Proceedings of the 21st ACM conference on Hypertext and hypermedia*. ACM, 2010, pp. 35–44.

[32] ——, "Constructing travel itineraries from tagged geo-temporal breadcrumbs," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1083–1084.

[33] L. W. Dietz and W. Worndl, "How long to stay where? on the amount of item consumption in travel recommendation," in *Proceedings of the 13th ACM Conference on Recommender Systems (RecSys'19)*, 2019.

[34] Y. Zhang, P. Siriaraya, Y. Wang, S. Wakamiya, Y. Kawai, and A. Jatowt, "Walking down a different path: route recommendation based on visual and facility based diversity," in *Proceedings of The Web Conference 2018 Companion (WWW'18)*, 2018, pp. 171–174.

[35] D. Herzog, S. Sikander, and W. Worndl, "Integrating route attractiveness attributes into tourist trip recommendations," in *Proceedings of The 2019 World Wide Web Conference Companion (WWW'19)*, 2019, pp. 96–101.

[36] S. Wakamiya, P. Siriaraya, Y. Zhang, Y. Kawai, E. Aramaki, and A. Jatowt, "Pleasant route suggestion based on color and object rates," in *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining (WSDM'19)*, 2019, pp. 786–789.

[37] D. Chen, D. Kim, L. Xie, M. Shin, A. K. Menon, C. S. Ong, I. Avazpour, and J. Grundy, "Pathrec: Visual analysis of travel route recommendations," in *Proceedings of the Eleventh ACM Conference on Recommender Systems*. ACM, 2017, pp. 364–365.

[38] I. Brilhante, J. A. Macedo, F. M. Nardini, R. Perego, and C. Renso, "TripBuilder: A tool for recommending sightseeing tours," in *Proceedings of ECIR'14*, 2014, pp. 771–774.

[39] P. Vansteenwegen, W. Souffriau, G. V. Berghe, and D. Van Oudheusden, "The city trip planner: an expert system for tourists," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6540–6546, 2011.

[40] W. Wörndl and A. Hefele, "Generating paths through discovered places-of-interests for city trip planning," in *Information and Communication Technologies in Tourism*, 2016, pp. 441–453.

[41] X. Yu, Y. Liu, X. Huang, and A. An, "Mining online reviews for predicting sales performance: A case study in the movie domain," *IEEE Transactions on Knowledge and Data engineering*, vol. 24, no. 4, pp. 720–734, 2010.

[42] Y. Liu, X. Huang, A. An, and X. Yu, "Arsa: a sentiment-aware model for predicting sales performance using blogs," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval (SIGIR'07)*, 2007, pp. 607–614.

[43] J. Sun, G. Wang, X. Cheng, and Y. Fu, "Mining affective text to improve social media item recommendation," *Information Processing & Management*, vol. 51, no. 4, pp. 444–457, 2015.

[44] C. J. Hutto and E. Gilbert, "Vader: A parsimonious rule-based model for sentiment analysis of social media text," in *Proceedings of The Eighth international AAAI conference on weblogs and social media (ICWSM'14)*, 2014, pp. 216–225.

[45] M. Berkelaar, K. Eikland, and P. Notebaert, "lpsolve: Open source (mixed-integer) linear programming system," 2004, http://lpsolve.sourceforge.net/.

[46] K. H. Lim, K. E. Lee, D. Kendal, L. Rashidi, E. Naghizade, Y. Feng, and J. Wang, "Understanding Sentiments and Activities in Green Spaces using a Social Data-driven Approach," in *Smart Cities: Issues and Challenges*. Elsevier, Jun 2019, pp. 77–107.

[47] Y. Huang, Y. Li, and J. Shan, "Spatial-temporal event detection from geo-tagged tweets," *ISPRS International Journal of Geo-Information*, vol. 7, no. 4, p. 150, 2018.

[48] C. Zhang, L. Liu, D. Lei, Q. Yuan, H. Zhuang, T. Hanratty, , and J. Han, "Triovecevent: Embedding-based online local event detection in geo-tagged tweet streams," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2017, pp. 595–604.

[49] Y. George, S. Karunasekera, A. Harwood, and K. H. Lim, "Spatio-temporal Event Detection using Poisson Model and Quad-tree on Geo-tagged Social Media," in *Proceedings of the 2019 IEEE International Conference on Big Data (BigData'19)*, Dec 2019.

[50] C. Zhang, D. Lei, Q. Yuan, H. Zhuang, L. Kaplan, S. Wang, and J. Han, "Geoburst+: effective and real-time local event detection in geo-tagged tweet streams," *ACM Transactions on Intelligent Systems and Technology*, vol. 9, no. 3, p. 34, 2018.

[51] City of Melbourne, "City of melbourne's open data platform," 2018, https://data.melbourne.vic.gov.au/.

[52] S. Loria, P. Keen, M. Honnibal, R. Yankovsky, D. Karesh, and E. Dempsey, "Textblob: simplified text processing," 2018, https://textblob.readthedocs.io/en/dev/.