User-adapted Travel Planning Algorithms for Landmark and Route Recommendation

観光スポットとルート推薦のためのユーザ適応型旅行プラン生成アルゴリズム

February 2020

Siya BAO

鮑 思雅

User-adapted Travel Planning Algorithms for Landmark and Route Recommendation

観光スポットとルート推薦のためのユーザ適応型旅行プラン生成アルゴリズム

February 2020

Waseda University

Graduate School of Fundamental Science and Engineering

Department of Computer Science and Communications Engineering,

Research on Information System Design

Siya BAO

鮑 思雅

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Due to the demands for higher standards of living, travel has become easier and more affordable, and the problem of travel planning has gained attention recently.

Travel planning is still an important and troublesome task before departure. For example, a tourist may require assistance for planning his/her trip: "How to visit the most popular landmarks if I would like to leave the hotel at 9:00 while I need to take the train at 20:00.". Based on those considerations, the service of automatical travel route planning is required.

To deal with the demand for travel route planning, it addresses three main issues:

1. Where to go: Recommend user the popular and interesting landmarks to visit;

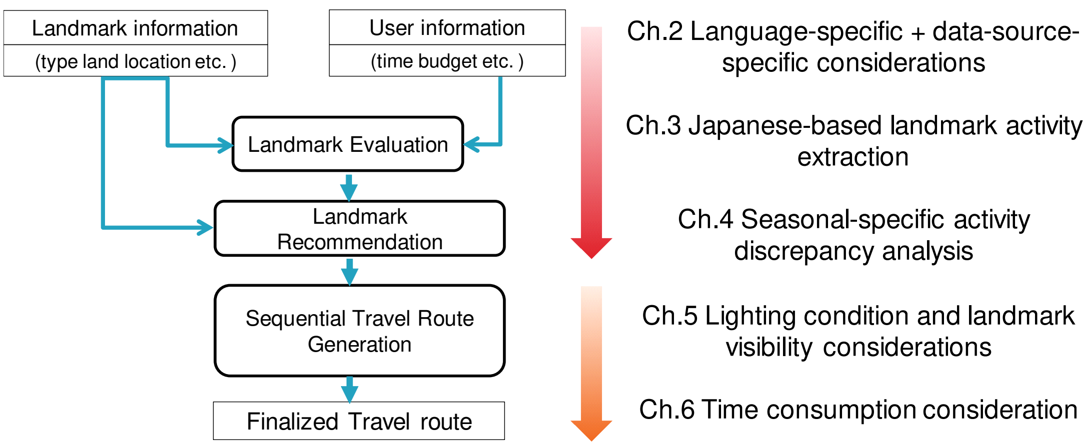


Figure 1.1: The flow of travel recommendation.

2. What to see: Recommend activities that users can experience for each landmark;

3. How to go: Recommend a travel route by integrating landmarks based on the user's multiple constraints (time, budget, etc.).

For issue (1), conventional recommendation services, such as TripAdvisor [69],concentrated on recommending the most famous landmarks in a city or a region. For example, the Golden Gate Bridge shall be recommended if we set the target destination as San Francisco. In detail, through users' historical trajectories, the popularity of each landmark can be evaluated based on factors such as the number of visitors and the overall satisfaction ratings.

However, this way of landmark recommendation may not be satisfying every user. Thus, researches considering the personalized preferences of users have attracted extensive attention in recent years [6, 20, 59].

With the rapid development of mobile devices, huge amounts of user-contributed geo-data can available on the Internet. Several location-based social network (LBSN) services such as FourSquare, Twitter and Facebook provide rich geographic and check-in information, and travel website such as TripAdvisor provide millions of travel comments recording users' unique travel experiences [36, 56]. This data benefits researchers to explore interesting landmarks for travel recommendations [30, 67, 70, 86].

Most of the existing studies mainly focus on the user's general preferences as essential factors for landmark recommendation [3, 32, 55, 81, 83].

Nevertheless, it is pointed out that users from different language groups have different satisfaction rating behaviors [16, 72]. In an instance, Japanese users intend to rate landmarks lower than Chinese and English users. Thus, we should also pay attention to how deviances in languages will influence landmark recommendation.

Moreover, due to a large amount of those data, how to efficiently co-operate data from various sources is still a challenge, landmark information including landmark coverages, satisfaction ratings, and type descriptions do largely vary across different data sources.

For issue (2), rather than provide a user with a list of individual landmarks, it is helpful to point out what exactly the attractiveness of each landmark is. Compared with common information searching, it may be too difficult for a user who is not familiar with the destination city or landmark to raise a question. For example, when visiting an art museum, people often intend to ask a general question such as *What can I see?* or *What is on the exhibition now?*. Thus, how to automatically generate meaningful activity examples with detailed descriptions is important.

In addition, instead of recommending a series of landmarks, suitable and representative seasonal activities of each landmark should also be recommended at the same time. Taking the Meguro-gawa (River) in Tokyo as an example, it is famous for cherry flower (Sakura) viewing during the period of March to April, while the cherry flower viewing is not unavailable during the other months. Therefore, the seasonal activity differences for each landmark should be included in order to match users' travel time schedule.

For issue (3), several studies focused on providing the travel route that allows users to visit each landmarks on the recommendation list [77]. But users are not able to visit all the landmarks under most conditions due to the limitations of travel time. By concerning with the user's time budget, some researchers focus on generating travel routes for a sequence of landmarks based on the route distance and landmark attractiveness [37, 77].

However, there is still an issue that a user may arrive at a landmark beyond the business time of the landmark. As mentioned in [29], the satisfaction of visiting landmarks is highly related to the arrival time. In other words, a good travel route should guarantee that a user can visit the landmark during its business time, and meanwhile, the total visiting time should not exceed the user's time limitation.

Moreover, as safety is another vital factor in traveling especially for users who are unfamiliar with the destination [23, 68], how to construct a safe route for users should be taken into account either.

In the first half of this dissertation, the personalized landmark and landmark activity recommendation algorithms are proposed in Chapter 2, Chapter 3 and Chapter 4 using online travel comments while dealing with the issues of collaborating multiple data sources and identify language impact as mentioned above. It refers to *Landmark Evaluation* and *Landmark Recommendation* in Fig. 1.1

In the last half of this dissertation (Chapter 5 and Chapter 6), time-concerning travel route recommendation algorithm, safe and comprehensive route recommendation algorithm are proposed. It refers to *Sequential Travel Route Generation* in Fig. 1.1

In this dissertation, we propose three recommendation algorithms in support of personalized landmark and landmark activity recommendations with data-source-specific and language-specific considerations to cope with the discrepancies between users from different backgrounds, followed by proposing a document-level sentiment prediction algorithm for comment quality improvement. We also propose two travel route planning algorithms in terms of safety and comprehensive route recommendation and one-day travel route recommendation.

1.2 Dissertation Overview

This dissertation is organized as follows:

Chapter 2 [A Personalized Landmark Recommendation Algorithm for Language-Specific Users by Open Data Mining] proposes a personalized landmark recommendation algorithm aiming at exploring new sights into the determinants of landmark satisfaction prediction. We gather 1,219,048 user-generated comments in Tokyo, Shanghai and New York from four travel websites. We find that users have diverse satisfaction on landmarks according to their preferred languages and travel websites. With those findings, we propose an effective algorithm for personalized landmark satisfaction prediction. Our algorithm provides the top-6 landmarks with the highest satisfaction to users for a one-day trip plan in Tokyo, Shanghai, and New York. The results show that our proposed algorithm has better performances than previous studies from the viewpoints of landmark recommendation and landmark satisfaction prediction.

Chapter 3 [A Travel Decision Support Algorithm: Landmark Activity Extraction from Japanese Travel Comments] proposes an algorithm to construct activity queries and extract meaningful examples as detailed descriptions based on the linguistic characteristics of Japanese. The proposed algorithm concentrates on analyzing the feasibility of exploring landmark activity queries and representative examples from travel comments To deal with those two questions, we utilize the advantages of travel comments posted by thousands of other travelers. The proposed algorithm includes 4 steps: (1) phrase all comments in the entire comment set; (2) obtain the keyword for each landmark; (3) construct the query with the keyword in step (2); (4) extract top-5 examples with the query. An evaluation of activity-example extraction is conducted in two case studies through 18,939 travel comments. Based on the experimental results, with a relatively small scale of travel comments, it still allows us to explore rich landmark activity information.

Chapter 4 [Landmark Seasonal Travel Distribution and Activity Prediction Based on Language-specific Analysis] proposes a seasonal activity prediction algorithm based on user comments over the period of 2012 to 2017 in different language groups. We take advantage of online user comments which provide visiting time for each landmark and detailed activity description. With the accumulation of 417,787 user comments on TripAdvisor for 300 landmarks in three big cities, we analyze the language-specific differences in travel distributions. After that, the prediction of future travel distribution for each language group is generated. Then potential peak and off seasons of each landmark are distinguished and representative seasonal activities are extracted through comment contents for peak and off-seasons, respectively. Experimental results in the three cities show that the proposed algorithm is more accurate in terms of peak season detection and seasonal activity prediction than previous studies.

Chapter 5 [A Safe and Comprehensive Route Finding Algorithm for Pedestrians Based on Lighting and Landmark Conditions] proposes a safe and comprehensive route finding algorithm for pedestrians based on lighting and landmark conditions. Safety and comprehensiveness can be predicted by the five possible indicators: (1) lighting conditions, (2) landmark visibility, (3) landmark effectiveness, (4) turning counts along a route, and (5) road widths. We first investigate the impacts of these five indicators on pedestrians' perceptions of safety and comprehensiveness during route findings. After that, a route finding algorithm is proposed for pedestrians. In the algorithm, we design the *score* based on the indicators (1), (2), (3), and (5) above and also introduce a turning count reduction strategy for the indicator (4). Thus we find out a safe and comprehensive route through them. In particular, we design daytime score and nighttime score differently and find out an appropriate route depending on the periods. Experimental simulation results demonstrate that the proposed algorithm obtains higher scores compared to several existing algorithms. We also demonstrate that the proposed algorithm can find out safe and comprehensive routes for pedestrians in real environments by questionnaire results.

Chapter 6 [A Personalized Landmark and Route Recommendation Algorithm for a One-Day Trip] proposes a personalized travel recommendation algorithm with time planning for a one-day trip. The proposed algorithm consists of three steps: (1) recommend top-6 personalized landmarks based on landmark categorization and region clustering; (2) build a travel map to generate all possible travel routes based on top-6 personalized landmarks in step (1); (3) generate a realistic travel route for a one-day visit based on evaluations on the number of landmarks to visits and travel time consumptions. Experimental results confirm the advantages of our proposed algorithm beyond previous studies from the viewpoints of landmark recommendation precision and travel time optimization.

Chapter 7 [Conclusion] summarizes this dissertation and presents future works.

Chapter 2

A Personalized Landmark Recommendation Algorithm for Language-Specific Users by Open Data Mining[1]

2.1 Introduction

Travelers usually try to perceive a general image about how travel destinations or landmarks will be like before the departure. Due to this sense, it is not surprising that users' travel decisions are strongly influenced by travel comment contents in many aspects including where to go, and what to see and do [6, 20, 59].

Several studies focus on predicting users' satisfaction on landmarks through social media comments [55], some of which try to discover the connectedness between users' travel behaviors and backgrounds [72]. Those studies take into consideration new ideas for efficiently predicting users' satisfaction on landmarks but their accuracy is often limited to the sample size (mostly, the number of samples $N \leq 600$) and the usage of a single data source. Moreover, deviance in language discrepancies is not considered though users from different backgrounds have different satisfaction rating behaviors [16].

In this chapter, we utilize large datasets using heterogeneous open data sources and divide users into three languages (Chinese, English and Japanese) groups. We firstly collect 1,219,048 user-generated comments from the travel websites of Ctrip [13], Jaran [31], 4travel [1] and TripAdvisor [69] for 194 landmarks in Tokyo, Japan, 189 landmarks in Shanghai, China and 196 landmarks in New York, USA. We analyze users' average satisfaction on landmarks and landmark coverage differences between the travel

[1] Technical contents in this chapter have been presented in the publications ⟨4⟩ and ⟨5⟩.

websites in three cities. Then we extract 1,046,395 user comments from 1,219,048 user comments and divide them into three languages groups. Users' language-specific satisfaction and favorable landmark types are examined. Finally, an algorithm is proposed to predict the users' satisfaction over each landmark according to their preferences on landmark types, languages and travel websites.

Our contributions are highlighted as follows:

- We analyze data-source-specific and language-specific landmark satisfaction differences with 1,219,048 user comments through four travel websites.
- We analyze pairwise landmark type relationships in order to correct error types in the existing travel websites.
- We propose a personalized landmark recommendation algorithm based on landmark satisfaction prediction. Our algorithm can recommend landmarks that fit the user's preferences with an accuracy over 82% and successfully predicts the user's satisfaction on landmark with the error rate lower than 7.5%, which outperforms the previous studies.

This chapter belongs to *Landmark Evaluation* and *Landmark Recommendation* in Fig. 1.1.

2.2 Related Works

Travel comments are a clear reflection of how users' travel experiences satisfy their expectation [36, 56]. For example, when their experiences exceed their expectation, there usually will be of a 5-star rating. Otherwise, when their experiences fail to meet their expectation, there usually will be of a 2-star rating or less.

With the consideration of the experience-expectation, there are two challenges that we have to face for reliable personalized travel route recommendation.

(1) Collaborating multiple data sources: Many studies have made an attempt to discover user travelogues (comments, GPS trajectory, check-in data etc.) to increase the diversity and quantity of experimental data [30,67,70,86]. In an instance, Zhou et.al [86] and Sun [67] extract users' preferences by photo tags through Flickr. However, it is found that landmark information including landmark coverages, satisfaction ratings, and type descriptions do largely vary across different data sources, but existing studies on travel recommendation using travelogues usually use the sole data source. For example, the Rainbow Bridge in Tokyo has a high rating of 4.5 stars on Ctrip[2], while the rating of it is 4.0 stars on TripAdvisor [3]. In other words, it is helpful to consider discrepancies

[2]http://you.ctrip.com/sight/tokyo294/132603.html

[3]https://www.tripadvisor.jp/Attraction$_R$eview-g14134368-d555410.html

between different data sources for more reliable landmark recommendation.

(2) Identify language impact: Most of the existing studies only focus on the user's general preferences as essential factors for landmark recommendation [3, 55, 81]. Nevertheless, it is pointed out that users from different language groups have different satisfaction rating behaviors [16]. In an instance, Japanese users intend to rate landmarks lower than Chinese and English users. Thus, we should also pay attention to how deviances in languages will influence landmark recommendation.

2.3 Problem Statement

2.3.1 Notations

Key notations used in this chapter are listed in Table 2.1. Key concepts include the followings.

- Type: Category that a landmark belongs to. We mainly concentrate on eight types of landmarks: *History, Nature, Entertainment, Art, Sport, Food and Drink, Shopping*, and *Night Life* inspired by [81].
- Type weight: User's rating for 8 landmark types ranging from *1–5* (very dislike to very like).
- Data source: Four leading websites which are frequently used including Ctrip, Jaran, 4travel and TripAdvisor.
- Language: Three main languages groups in the three websites including Chinese, English, and Japanese.
- User profile: Includes the user's type weights, frequently used websites and languages.

2.3.2 Problem definition

Our personalized travel route recommendation problem in this chapter is defined as follows: Given a set of landmarks and a user's profile, recommend top-k landmarks that fit the user profile. As it is difficult for users to directly derive explicit answers from the travel comments, we correlate and analyze travel comments by exploring differences in various website and language groups, and recommend a series of personalized landmarks.

Table 2.1: Notation description

Notation	Description
l	landmark
L	a set of landmarks
u	user
web	travel website
$lang$	language
w	weight for a landmark type
t	landmark type
LT	a set of 8 landmark types
ls	language-specific satisfaction
ds	data-source-specific satisfaction

2.4 Data-Collect Process

We collect user comments from four leading travel websites in Ctrip (China) [13], Jaran (Japan) [31], 4travel (Japan) [1], and TripAdvisor (United States) [69], with the aim to deal with the small sample size issue in the previous studies as discussed in Section 6.1. A data-collection program was developed in R, which took approximately 20 days to crawl all the data.

All user comments collected from each website was before February 1st, 2018, and Tokyo, Shanghai and New York are famous travel destinations. Particularly, the landmark which is labeled as "region", for example, "Shibuya District", is not considered as it does not have a specific type or location. Let N_L be the number of total landmarks in each website. Then we have for Tokyo, $N_L = 835$ in Ctrip, $N_L = 1140$ in Jaran, and $N_L = 1336$ in TripAdvisor; for Shanghai, $N_L = 4420$ in Ctrip, $N_L = 1497$ in 4travel, and $N_L = 1680$ in TripAdvisor, and for New York, $N_L = 631$ in Ctrip, $N_L = 227$ in 4travel, and $N_L = 4450$ in TripAdvisor. We extracted top-k ranked landmarks (we set $k = 100$), each of which has at least five user comments from each website.

Table 2.2: The statistics of user comment among Ctrip, Jaran and TripAdvisor in Tokyo, Shanghai and New York.

Tokyo	User comment	Chinese	English	Japanese	Other-language
Ctrip	25018	>99%	<1%	>1%	>1%
Jaran	134945	>1%	>1%	>99%	>1%
Tripadvisor	110276	6794 (6.16%)	43638 (39.57%)	42091 (38.17%)	17753 (16.10%)
Total	270239	31812 (11.77%)	43638 (16.15%)	177036 (65.51%)	17753 (6.57%)
Shanghai	User comment	Chinese	English	Japanese	Other-language
Ctrip	66282	>99%	<1%	>1%	>1%
4travel	2598	>1%	>1%	>99%	>1%
Tripadvisor	55052	7357 (53.25%)	29316 (39.57%)	14137 (25.68%)	17753 (32.25%)
Total	123932	73639 (59.42%)	29316 (23.65%)	16735 (13.50%)	17753 (14.32%)
New York	User comment	Chinese	English	Japanese	Other-language
Ctrip	1495	>99%	<1%	>1%	>1%
4travel	3684	>1%	>1%	>99%	>1%
Tripadvisor	806187	6140 (0.76%)	481790 (59.76%)	186289(23.11%)	131968 (16.37%)
Total	811366	7635 (0.94%)	481790 (59.38%)	189973 (23.41%)	131968 (16.26%)

Since Jaran is the website for Japanese sightseeing, we extract user comments for Tokyo using it but it does not contain landmarks in other countries. Instead, we use 4travel.jp to extract user comments including Japanese for Shanghai and New York.

For each landmark in a travel website, we collected each user's satisfaction on the landmark, comment content and the time of writing. User's satisfaction on each landmark is five-scale rated from 1 (very dislike) to 5 (very like). We also collected the rank of the landmark at each travel website Note that, how to rank landmarks at every travel website is not open. It can be just decided based on the user satisfaction and the number of user comments on each website.

As a result, 1,219,048 user comments were collected for further research (seeTable 4.1).

After data collection, we conducted language-detection through R. Three representative characters, including Chinese (de), English (is) and Japanese (no), were utilized to distinguish a specific language.

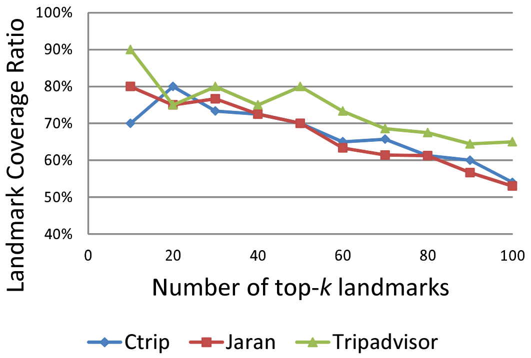
The statistics of user comments from three travel websites in three cities is shown in Table 4.1. As a result, the statistics show that Ctrip mainly includes Chinese users as the ratio of Chinese comments is always over 99%, and thus we consider Ctrip's comments all as Chinese comments. Similarly, we consider Jaran's comments all as Japanese comments. In cases of Tokyo and New York, we also consider Ctrip's comments all as Chinese comments, and Jaran's and 4travel's all as Japanese comments. In addition, other languages in TripAdvisor provides a ratio no more than 13.89% of the total comment set. For this reason, we only concentrate on Chinese, English and Japanese for language-specific analysis in Section 2.6.

2.5 Data-Source-Specific Analysis

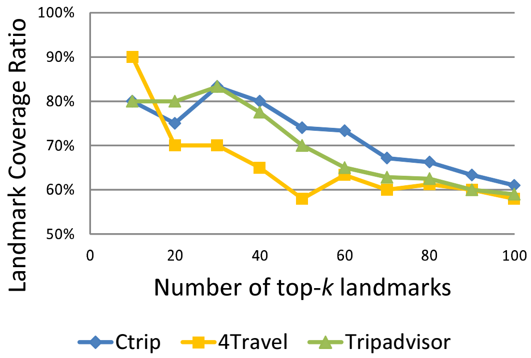
In this section, we investigate variances between different groups of users. Data-source-specific analysis are conducted. Subsequently, it is found that data-source-specific widely do exist and it suggests that this should be considered as an important factor in the personalized landmark recommendation.

2.5.1 Comparison of Landmark Coverage

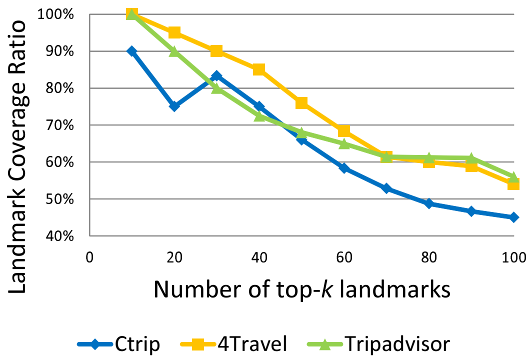
We investigate the coverage ratio of top-k ranked landmarks between the websites in three cities (see Fig. 2.3). "Coverage" is the ratio of landmarks in one website's top-k ranked landmarks that also occur in the other websites' top-k ranked landmarks. The lower the coverage ratio is, the higher the landmark uniqueness of each website is. Fig. 2.3 shows that the coverages of landmarks dramatically differ in the different websites in all city cases. For example, in the case of New York, when $k = 10$, the coverage ratios of all



(a) Tokyo.



(b) Shanghai.



(c) New York.

Figure 2.1: Data-source-specific landmark coverage ratio among Ctrip, Jaran, 4travel and TripAdvisor.

the three websites are more than 90%. When $k \geq 50$, the coverage ratios of all the three websites degrade to 76% or less.

Filtering is one of the most convenient methods in the landmark recommendation [60] but it usually assumes that the coverages of landmarks in all data sources always stay 100%. Oppositely, our results suggest that the coverages of landmarks remarkably vary by the travel websites. Therefore, our findings highlight the importance of collaborating heterogeneous data sources to resolve the problem of usage of a single data source as we discuss in Section 6.1.

2.5.2 Comparison of Average Satisfaction

The data-source-specific average satisfaction AS_k^{web} of the top-k landmarks in a travel website web is defined as follows:

$$AS_k^{web} = \frac{\sum_{i=1}^{k} ds(l_i)^{web}}{k} \tag{2.1}$$

where $ds(l_i)^{web}$ is the data-source-specific satisfaction and it is the average of all users' satisfaction for the i-th ranked landmark l in a particular travel website. Then we compare the differences between average satisfaction for the top-k ranked landmarks in the four websites (see in Fig. 2.2).

In Fig. 2.2, Jaran always has the lowest average satisfaction in the case of Tokyo, and 4travel has the lowest average satisfaction in the cases of Shanghai and New York. On the other hand, Ctrip and TripAdvisor has higher average satisfaction than Jaran or 4travel. This might be explained as the differences between various travel website user groups. In other words, Ctrip and TripAdvisor users tend to have a higher satisfaction rate on landmarks.

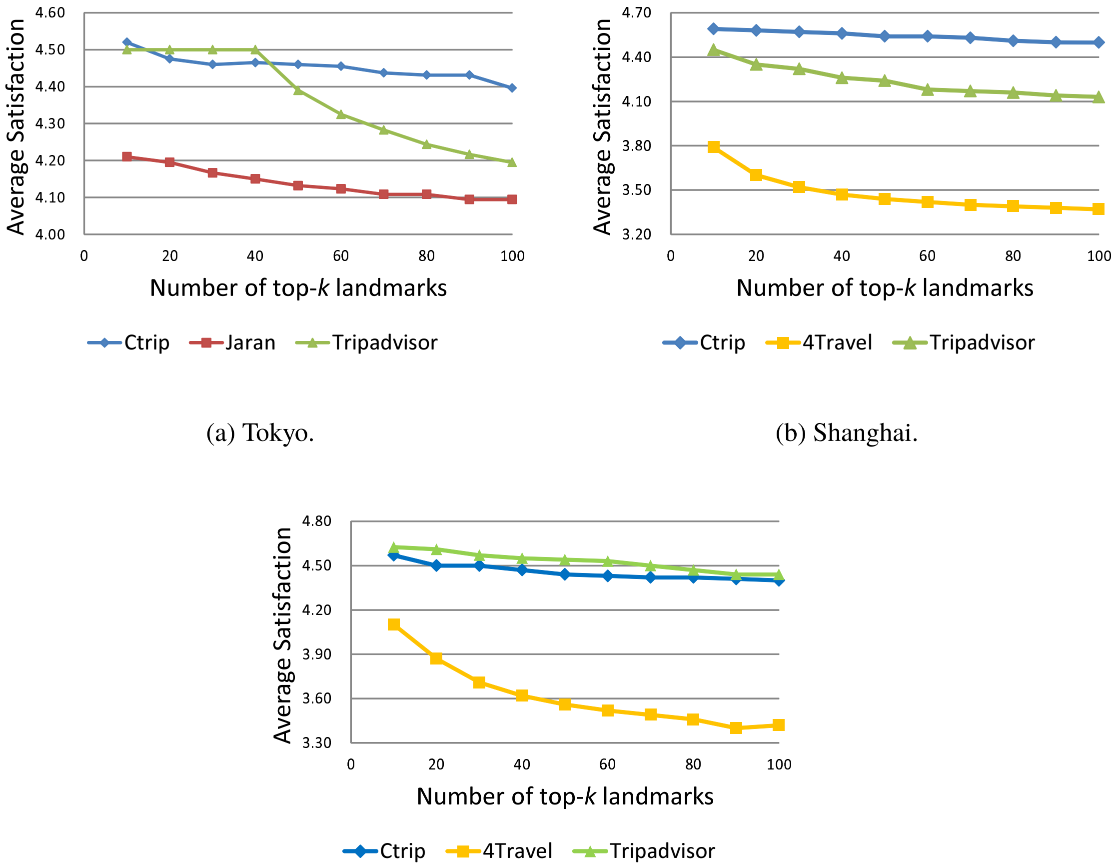
2.5.3 Comparison of Type Information

A landmark type represents the characteristics of the landmark. However, every travel website has its own way to describe types. Thus we have re-arranged all landmarks' types into eight types of *Art, Entertainment, Food and drink, History, Nature, Night life, Shopping*, and *Sport*. Let LT be a set of these eight landmark types. It is worth mentioning that, a landmark can have more than one types, for example, "Statue of Liberty" has two landmark types of *Art* and *History*.

We combine the landmarks and their types of the top-100 ranked landmarks in each website and the results are shown in Table 2.3. In an instance of landmark combination, we have 100 landmarks in each website in Tokyo and obtain totally 194 different landmarks by discarding the redundant ones. In detail, "Yu Garden" in Shanghai is labeled as *Art* and *History* in Ctrip, *Nature* in TripAdvisor and *Art* and *Nature* in 4travel. Then we combine all the types in the three websites and discard the redundant parts. The refined types for "Yu Garden" are *Art*, *History* and *Nature*.

It is interesting that a travel website provides more type information for its local or domestic landmarks compared with foreign landmarks. This can be explained by the travel website companies are more familiar with the local landmarks, and may describe the landmarks with more details.

Our average type number is of a maximum improvement rate of 27.71% compared with the original four travel websites (see Table 2.3). The results indicate that it is



(a) Tokyo.

(b) Shanghai.

(c) New York.

Figure 2.2: Data-source-specific average satisfaction among Ctrip, Jaran, 4travel and TripAdvisor.

necessary to collaborate information among different data sources to enrich landmark type information.

2.6 Language-specific analysis

In this section, we investigate variances between different groups of users. Language-specific analysis are conducted. Subsequently, it is found that anguage-specific discrepancies widely do exist and it suggests that this should be considered as an important factor in the personalized landmark recommendation.

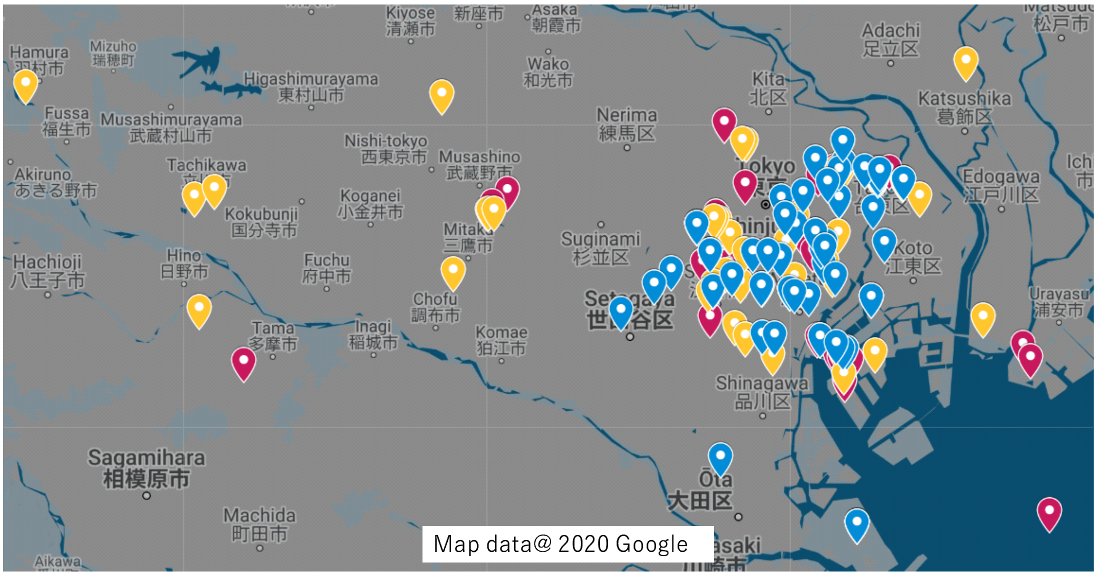


Figure 2.3: An example of data-source-specific landmark coverage in Tokyo. Red mark: Ctrip, yellow mark: Jaran, blue mark: TripAdvisor

2.6.1 Comparison of average satisfaction

We analyze the user average satisfaction based on language discrepancies in three cities using the user comments in TripAdvisor. In total, 92523 user comments are used for the case of Tokyo; 50810 user comments are used for the case of Shanghai and 674219 user comments are used for the case of New York.

The language-specific average satisfaction for Chinese is calculated as the average satisfaction of all Chinese users' satisfaction on the top-k ranked landmarks in each city as follows. The average satisfaction for Japanese and English is calculated similarly.

$$AS_k^{lang,TripAdvisor} = \frac{\sum_{i=1}^{k} ls^{lang,TripAdvisor}(l_i)}{k} \tag{2.2}$$

where $ls^{lang,TripAdvisor}(l_i)$ is the average of all $lang$ users' satisfaction for the i-th ranked landmark l_i in TripAdvisor, where $lang$ is either of Chinese, English or Japanese.

Fig. 2.4 portrays the results of language-specific average satisfaction in three cities. Chinese and English users average satisfaction is relatively similar and is always higher than Japanese users' average satisfaction. Even though Japan shares a great cultural similarity with China, Japanese users' average satisfaction is significantly deviating from the Chinese groups.

From Table 2.4, it shows that the sane conclusion that Chinese and English users' average satisfaction is relatively similar, while Japanese users' average satisfaction is

Table 2.3: Landmark type comparison with Ctrip, Jaran, 4travel and TripAdvisor.

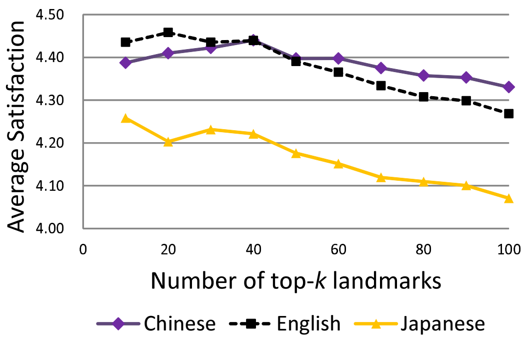
Tokyo	k	Total landmark type	Average landmark type	Improvement rate
Ctrip	100	146	1.46	+23.97%
Jaran	100	178	1.78	+1.69%
TripAdvisor	100	158	1.58	+14.56%
Ours	194	351	1.81	-
Shanghai	**k**	**Total landmark type**	**Average landmark type**	**Improvement rate**
Ctrip	100	178	1.78	+2.25%
4travel	100	161	1.61	+13.05%
TripAdvisor	100	152	1.52	+19.74%
Ours	189	344	1.82	-
New York	**k**	**Total landmark type**	**Average landmark type**	**Improvement rate**
Ctrip	100	184	1.84	+13.13%
4travel	100	163	1.63	+27.71%
TripAdvisor	100	205	2.05	+1.54%
Ours	196	408	2.08	-

Table 2.4: Language-specific average satisfaction among Ctrip, Jaran and TripAdvisor in Tokyo.

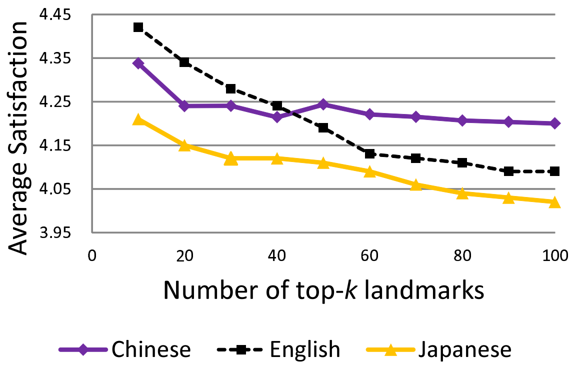
Language	Average Satisfaction
Chinese	4.35
English	4.27
Japanese	4.12

significantly deviating from the other two language groups.

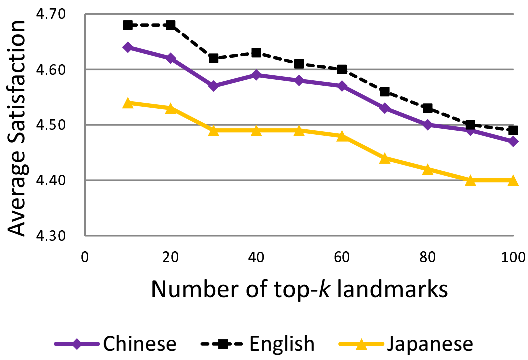
To sum up, the results say that users from different cultural backgrounds value the landmark's average satisfaction differently and the conclusion stays the same no matter which city's data we use. Therefore, it is crucial to examine the users' language backgrounds into consideration for accurate satisfaction prediction.



(a) Tokyo.



(b) Shanghai.



(c) New York.

Figure 2.4: Language-specific average satisfaction among Chinese, English and Japanese.

2.6.2 Comparison of overall type preference

We assume that all users have similar preferences on landmark types no matter which language they use and we demonstrate our assumption as follows:

We calculate frequencies and ranks of the eight landmark types for the three language groups using 1,038,087 user comments for each city. In order to visualize the similarities between the three language groups, we calculate the cosine similarity *cosim* of the type preferences in two ways:

We first define a vector

$$t_r(\vec{lang}) = (r_{Art}, r_{Entertainment}, \ldots, r_{Sport}) \tag{2.3}$$

for a language $lang$, where $\vec{r_t}$ is defined that, if the landmark type t is the i-th rank in the language $lang$, then r_t is $(1/i)$ in this language in each city. For example, in the case of $t_r(\overrightarrow{Chinese})$ in Shanghai, *Art* is the second place then $r_{Art} = 1/2 = 0.5$ and *Sport* is the last place with $r_{Sport} = 1/8 = 0.125$.

In the same way, we define the other vector as follows:

$$t_f(\overrightarrow{lang}) = (f_{Art}, f_{Entertainment}, \ldots, f_{Sport}) \tag{2.4}$$

for a language $lang$, where f_t is the frequency of the landmark type t in the language $lang$ in each city. For example, in the case of $t_f(\overrightarrow{Chinese})$ in Shanghai, f_{Art} and f_{Sport} become 0.26 and 0.00, respectively.

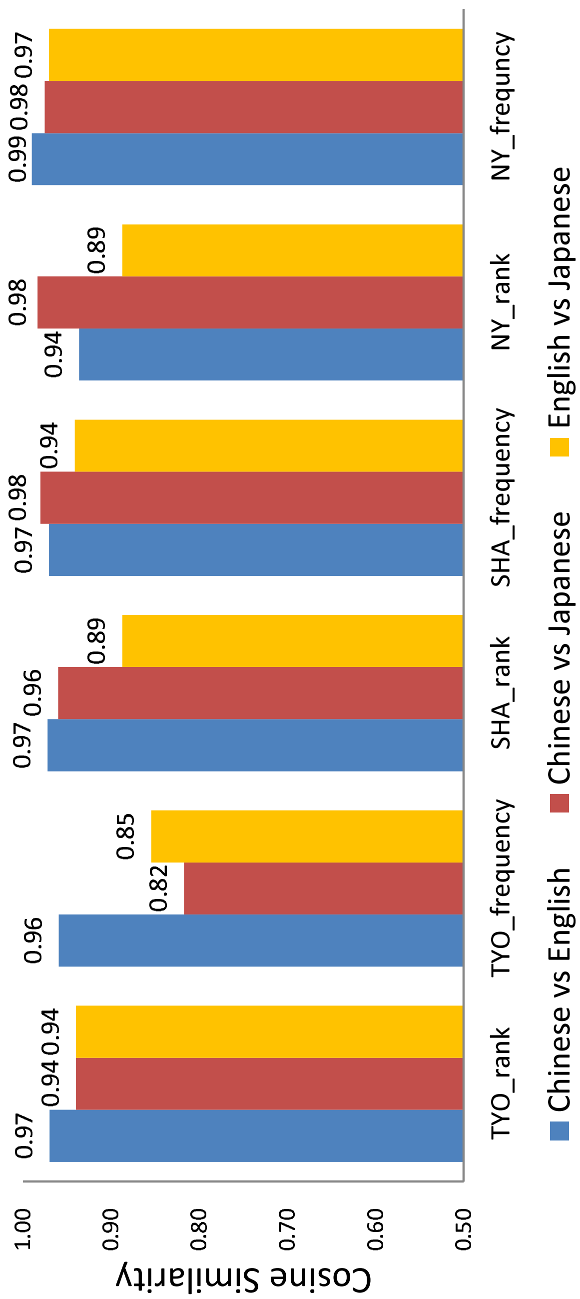


Figure 2.5: Cosine similarity between three language groups.

Then Eqn. (2.5) shows the $cosim_r$ value weighted by the type rank and Eqn. (2.6) shows the $cosim_f$ value weighted by the type frequency:

$$cosim_r(lang_i, lang_j) = \frac{< t_r(\vec{lang_i}) \cdot t_r(\vec{lang_j}) >}{\|t_r(\vec{lang_i})\| \cdot \|t_r(\vec{lang_j})\|} \tag{2.5}$$

$$cosim_f(lang_i, lang_j) = \frac{< t_f(\vec{lang_i}) \cdot t_f(\vec{lang_j}) >}{\|t_f(\vec{lang_i})\| \cdot \|t_f(\vec{lang_j})\|} \tag{2.6}$$

where $< \vec{v_1} \cdot \vec{v_2} >$ shows the inner product of two vectors $\vec{v_1}$ and $\vec{v_2}$ and $\| \cdot \|$ shows the L^2 norm.

The results of $cosim_r$ and $cosim_f$ are shown in Fig. 2.5. It can be seen that the $cosim_r$ values or the $cosim_f$ values between the three language groups are of high similarities over 0.82. The results indicate that the users' overall preferences of landmark types do not have a direct association with the language that they use. This confirms our assumption at the beginning of this subsection.

Thus, we concentrate on the users' personalized preferences on landmark types, rather than taking the users' overall preferences into account.

2.7 Personalized Landmark Recommendation algorithm

2.7.1 Landmark Database Establishment

We build a landmark database of three cities. After combining types and eliminating redundant landmarks, 194 landmarks are kept for Tokyo, 189 landmarks are kept for Shanghai and 196 landmarks are kept for New York.

Unfortunately, error types still exist in some cases. For example, in the case of Jimbocho Bookstore Area (a bookstore street), it is labeled as *Art* + *History* + *Shopping* in one website, where *Shopping* seems not proper. Thus, we analyze the relationship between type pairs. Fig. 2.6 presents the top-3 strong relations between the eight types. The size of the circle of each type shows the frequency at which it occurs in the database and the thickness of an arrow link between a pair of types shows how strong the relationship is.

Then we conduct the pairwise comparison between types for each landmark. If one pair of types is not included in the top-3 relations, then it refers that the relation between the pair is not strong enough and it will be labeled as a potential error pair. Error pairs are manually re-checked. For example, *History* does not has a strong relation with *Shopping* and we discard the type of *Shopping* for Jimbocho Bookstore Area.

Our finalized database contains 579 landmarks and 1103 types (see Table 2.3). Each landmark l in the finalized database include:

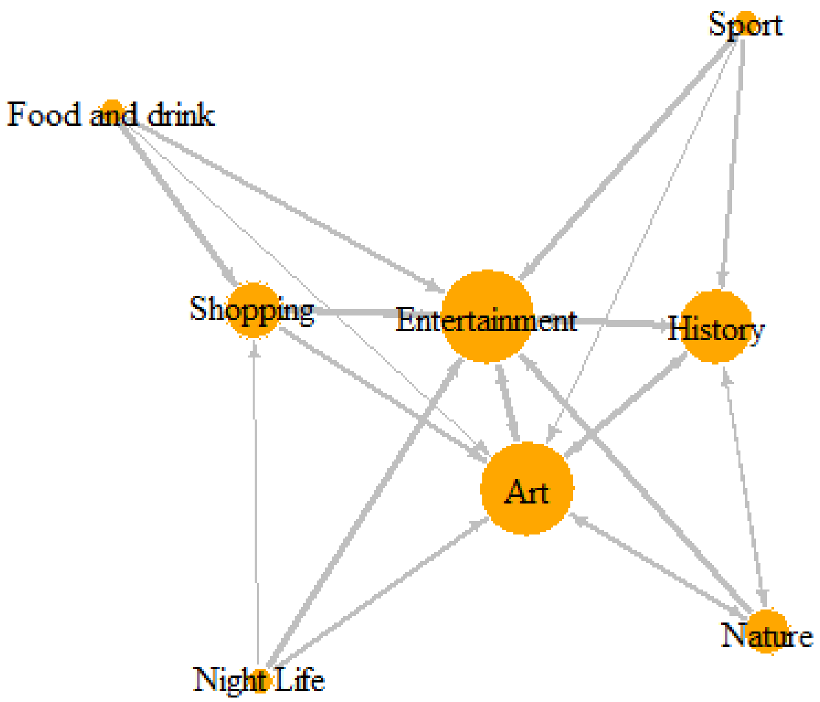


Figure 2.6: Representation of pairwise type relations.

- a language-specific satisfaction rating $ls^{lang}(l)$ for each language $lang$, which is defined by the average satisfaction over all the comments to l in a specific language.
- a data-source-specific satisfaction satisfaction $ds^{web}(l)$ for every travel website web, which is defined by the average satisfaction over all the comments to l in a specific website.

2.7.2 Satisfaction prediction

In order to predict user's satisfaction on every landmark, we propose a mathematical model to simulate the relation between user satisfaction and three variables, which are user's preferences on landmark types, user's language(s) and commonly visited travel website(s). We consider a linear relation, which is used in many related studies [15].

We introduce $S_{u,l}$ as the prediction satisfaction on a landmark l of a user u. $S_{u,l}^{type}$ is the satisfaction on l depending on u's type preferences. Likewise, $S_{u,l}^{lang}$ and $S_{u,l}^{web}$ are the the satisfaction on l depending on u's commonly used language(s) and travel website(s), respectively. $S_{u,l}$ is as follows:

$$S_{u,l} = \alpha \times S_{u,l}^{type} + \beta \times S_{u,l}^{lang} + (1 - \alpha - \beta) \times S_{u,l}^{web} + \theta \tag{2.7}$$

where α and β are the two constants weighting the significance of the three variables ($0 \leq (\alpha + \beta) \leq 1$). In an instance, if $\alpha = 1$, $S_{u,l}$ is only affected by the type preferences. We have both α and β equal to $1/3$ for the following analysis, which we assume that the three variables to be of equivalent importance. The calculation of the three variables is introduced as follows:

For $S_{u,l}^{type}$, firstly, a user is required to rate his/her preference $w(t)$ on each landmark type t with a five-scale rating: very dislike (1), dislike (2), fair (3), like (4), and very like (5). Let $LT(l)$ be a set of types that the landmark l contains and then $S_{u,l}^{type}$ is computed as follows:

$$S_{u,l}^{type} = \frac{\sum_{t \in LT(l)} w(t)}{|LT(l)|} \tag{2.8}$$

where $|LT(l)|$ is how many landmark types that $LT(l)$ has.

For $S_{u,l}^{lang}$, $ls^{lang}(l)$ is a language-specific satisfaction of a landmark l for a specific language $lang$ and $LS(l)$ is a set of language-specific satisfactions of all the languages that l has. If the user's language is $lang$ and l has the corresponding satisfaction $ls^{lang}(l)$, then $S_{u,l}^{lang}$ is equal to $ls(l)^{lang}$. Otherwise, $S_{u,l}^{lang}$ is the average of its other languages satisfaction as below:

$$S_{u,l}^{lang} = \begin{cases} ls^{lang}(l) & \text{(if } LS(l) \text{ includes } lang\text{)} \\ \frac{\sum_{\text{all languages}} ls^{lang}(l)}{|LS(l)|} & \text{(otherwise)} \end{cases} \tag{2.9}$$

where $|LS(l)|$ represents how many language-specific satisfaction that $LS(l)$ has.

For $S_{u,l}^{web}$, $ds^{web}(l)$ is a data-source-specific satisfaction in a specific travel website web for l and $DS(l)$ is a set of data-source-specific satisfactions of all the websites that l has. If the user's preferred website is web, and l has the corresponding satisfaction $ds^{web}(l)$, then $S_{u,l}^{web}$ is equal to $ds^{web}(l)$. Otherwise, $S_{u,l}^{web}$ is the average of its other travel website satisfaction as below:

$$S_{u,l}^{web} = \begin{cases} ds^{web}(l) & \text{(if } DS(l) \text{ includes } web\text{)} \\ \frac{\sum_{\text{all websites}} ds^{web}(l)}{|DS(l)|} & \text{(otherwise)} \end{cases} \tag{2.10}$$

where $|DS(l)|$ shows how many travel website satisfaction that $DS(l)$ includes.

In addition, we set a bonus constant θ as an error parameter [15]. Let $\theta = (n(l) - 1) + 0.1$, where $n(l)$ represents the number of times the landmark l included in the top-100 ranked landmarks in a city among the three travel websites .

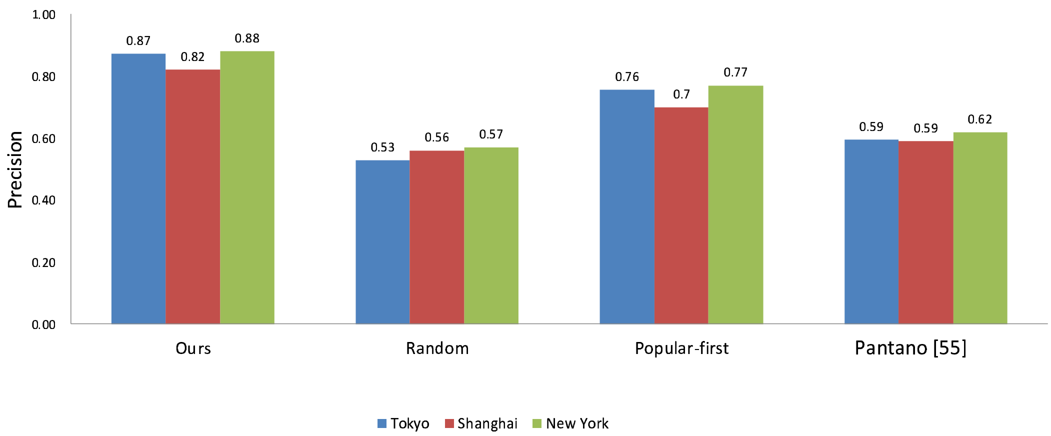


Figure 2.7: Evaluation result on landmark recommendation precision.

2.8 Experimental Results

2.8.1 Experimental settings

In this chapter, the experimental areas are set to be Tokyo, Shanghai and New York and the 194, 196 and 189 landmarks in our database are, respectively, used for landmark recommendation for each city. The algorithm is coded in R.

We test our algorithm with 12 different user profiles. We have interviewed 12 users, aging from 20–60, 4 males and 8 females. They are 6 Japanese users, 4 Chinese users and 2 English users. They are also 4 Ctrip users, 4 Jaran users, 1 Jaran and 4travel user and 3 TripAdvisor users. 12 users have visited at least three landmarks in Tokyo, Shanghai and 10 users have visited at least three landmarks in New York.

2.8.2 Landmark type recommendation precision

In this subsection, landmark type recommendation precision of our algorithm is analyzed. Users were required to fill in a questionnaire about the preferences with the five-scale rating on the landmark types, also their commonly used languages, and commonly used travel websites were recorded.

The proposed algorithm first calculates the $S_{u,l}$ values for each landmark l in the database by the user's profile. Duration time on a landmark usually takes around 1–2 hours, we assume that a user visits at most 6 landmarks in a one-day trip.

Our algorithm provides each user with the top-6 landmarks with the highest prediction satisfaction $S_{u,l}$. Three other algorithms are used for comparisons. Random algorithm randomly recommends six landmarks through the top-100 ranked landmarks in TripAdvisor, Popular-first algorithm recommends the top-6 ranked landmarks in TripAdvisor and the algorithm in [55] recommends six landmarks that match the user's type preferences. The six recommended landmarks are recorded as l_i^{rec} ($i = 1, 2, \ldots, 6$). Each trail in each city is implemented by a sub-dataset with 50 randomly selected landmarks.

Next, we compare the precision of the six recommended landmarks of our proposed algorithm with that of the other three comparison algorithms. We consider that the landmark type t fits the user's preferences if the user has rated three points or more to t. Then the number of true positives (TP) is defined by the number of the recommended landmarks' types that successfully fits the user's preferences. We consider that the landmark type t fails to fit the user's preferences if the user has rated two points or less to t. Then the number of false positives (FP) is defined by the number of the recommended landmarks' types that fails to fit the user's preferences.

For example, in the case of Tokyo, assume that the ratings of the user u_1 to the landmark types of *Art* and *Nature* are four and five points, respectively. Assume also that our algorithm recommends for the user u_1 the first landmark l_1^{rec} = *Sensoji*, of which types are labeled by *Art* and *Nature* in our database. Then TP for *Sensoji* is two and its FP is zero, since the two types *Art* and *Nature* fit the user's preferences. The computation of TP and FP for the other five recommended landmarks is the same. Then the *precision* for a user u can be obtained by:

$$Precision(u) = \frac{TP}{TP + FP} \tag{2.11}$$

Fig. 2.7 shows the average precision of 12 users by our proposed algorithm, Random algorithm, Popular-first algorithm and the algorithm in [55]. For Random algorithm, it is reasonable that it obtained the lowest precision because landmarks are randomly extracted with no concerns over landmarks types or user preferences. Popular-first algorithm recommends the top-6 ranked landmarks rated by a group of users in TripAdvisor. This algorithm improves precision as it considers users' general preferences. However, because this algorithm does not take the user's personalized preferences into consideration and hence our proposed algorithm's precision is around 10% higher than that of the Popular-first algorithm.

In [55], although it considers users' personalized preferences, it has poor performances with only around 60% precision. Because it is restricted to a small sample size of $N = 500$, as we discussed in Section 6.1 and it fails to resolve the problem of medium satisfaction (1, 2, 3 and 4 points). Pantano et al. assume that a user only has the very dislike (0 points) and the very like (5 points). But the truth is that the ratio of very dislike (0 points) cases in our landmark database rarely exists, which means 0 point cases should be considered less important compared with the medium satisfaction groups. Thus medium preferences should be considered rather than dividing users' preferences into only two extreme negative or positive groups.

2.8.3 Landmark satisfaction prediction accuracy

We evaluate if our proposed accurately predicts the user's real satisfaction. The three lists, each of which contains 30 landmarks randomly chosen from each city database, are prepared and the users were required to write their real satisfaction on the landmarks in the list which they have visited before as many as possible. A user's real satisfaction on a landmark l is denoted as ${}^{real}S_{u,l}$. Through the questionnaires, 30 valid ${}^{real}S_{u,l}$ values were used as the ground-truths for each city.

We compare our proposed algorithm with two baseline algorithms as follows:

- MAP (moving average predictor) represents the users' average satisfaction in a certain period [21], where we set the period from 2017.1.1–2019.1.1.

- $AVMAP$ (a variation moving average predictor) represents a landmark's expectation satisfaction computed by the Dirichlet distribution [35].

Both the MAP and the AVMAP use a single data source.

The prediction precision is analyzed by (a) average *error rate*, (b) maximum error rate (-/+) and (c) standard deviation (SD). *Error rate* ER is computed by:

$$ER = \frac{|{}^{pre}S_{u,l} - {}^{real}S_{u,l}|}{{}^{real}S_{u,l}} \times 100\% \tag{2.12}$$

where ${}^{pre}S_{u,l}$ is the predicted satisfaction of a user u on a landmark l by our algorithm, the MAP, or the AVMAP, and ${}^{real}S_{u,l}$ is a user u's the real satisfaction on a landmark l.

Table 6.4 lists the comparison results. In Table 6.4, it indicates that the proposed algorithm has the lowest error rate around 7% which is better than the MAP and the AVMAP, where only a single travel website comment data is used. This indicates that our proposed algorithm is effective in predicting users' satisfaction on landmarks in any cases and it is necessary to use multiple travel website data for better prediction accuracy.

Generally speaking, our proposed algorithm is very effective in predicting users' satisfaction on landmarks.

Table 2.5: Evaluation result on satisfaction prediction accuracy.

Tokyo	Average error rate	Maximum error rate(-)	Maximum error rate(+)	SD
Ours	6.88%	-12.80%	11.90%	0.08
MAP [21][1]	10.94%	-26.00%	18.42%	0.12
MAP [21][2]	11.68%	-20.00%	18.42%	0.13
AVMAP [35][1]	11.67%	-30.00%	20.00%	0.13
AVMAP [35][2]	10.55%	-18.40%	14.75%	0.11
Shanghai	Average error rate	Maximum error rate(-)	Maximum error rate(+)	SD
Ours	7.42%	-14.75%	11.90%	0.10
MAP [21][1]	11.97%	-30.00%	20.75%	0.13
MAP [21][1]	9.72%	-20.00%	15.75%	0.11
AVMAP [35][1]	10.58%	-30.00%	20.97%	0.13
AVMAP [35][2]	11.67%	-20.00%	18.42%	0.14
New York	Average error rate	Maximum error rate(-)	Maximum error rate(+)	SD
Ours	7.47%	-10.00%	11.90%	0.08
MAP [21][1]	10.63%	-24.00%	9.67%	0.10
MAP [21][2]	11.43%	-20.00%	12.50%	0.12
AVMAP [35][1]	9.46%	-30.00%	14.75%	0.11
AVMAP [35][2]	10.05%	-20.00%	10.53%	0.10

[1] Used comments from Ctrip only.

[2] Used comments from TripAdvisor only.

2.9 Conclusion

We propose an algorithm that can effectively predict user's satisfaction on a landmark by the user's preferences on landmark type, language and travel websites. The results demonstrate that our proposed algorithm has a high degree of precision in terms of the landmark recommendation and landmark satisfaction prediction compared with previous studies.

In the future, we need to further confirm the effectiveness of merging data sources, and conduct language-specific analysis for other languages such as Korea and German.

Chapter 3

A Travel Decision Support Algorithm: Landmark Activity Extraction from Japanese Travel Comments[1]

3.1 Introduction

Travel decisions are grounded in the real experiences of users' travelogues. Due to sense, investigations on travelogues potentially leads to more reliable travel decisions than simply browsing web pages [58].

Many existing studies utilized travelogues to predict travel activities/purposes using open-domain data such as Twitter and Facebook. These studies includes activity classifications [14, 39], location extractions [27, 75] and daily activity pattern predictions [12, 26]. Cui et al. [14] predicted general 4 types of activities for each landmark using Twitter contents. Lian et al. [39] recognized users' activity types based check-in and transition histories. Hoang et al. [27] identified location-related information in social media contents. Cranshaw et al. [12] focused on mapping daily activity area patterns in urban cities and Hasan et al. [26] used geo-location data to identify activity categories and then predicted weekly activity patterns of individual users. To fully explore the value of large-scale travelogues in travel decision making, there still are issues remaining unsolved.

(1) Dissemination of information: These travelogue data were collected directly from SNS (Social Networking Service) which contains a large amount of irrelevant information including promotions and advertisements [**?**,9]. This disadvantage may result in the time-consuming data cleaning process. Thus, it is important and efficient to use reliable data such as travel comments on trustworthy travel sites such as TripAdvisor [20].

[1]Technical contents in this chapter have been presented in the publications ⟨2⟩.

(2) Limited information-searching capability: Compared with common information searching, it may be too difficult for a user who is not familiar with the destination city or landmark to raise a question. For example, when visiting an art museum, people often intend to ask a general question such as *What can I see?* or *What is on the exhibition now?* rather than a specific question such as *Is that painting painted in 1800?*. The input question query such as *Why do people like bass low frequencies on music* in [62] is not available in the travel searching case. Thus, how to automatically generate the query is important.

(3) Ambitious description: After inputting the question query, we should search for information related. Short messages such as tags or event phrases may be helpful for describing the landmarks but still could be less meaningful sometimes. For example, most of the tags for *Sensoji* on Twitter are geo-tags such as *#Tokyo* and *#ExploreTokyo* which is ambitious and does not include any details of activities. Thus, illustrating a series of specific examples, such as *The 80's painting exhibition is great!*, will better help users make travel decisions.

(4) Lack of Japanese analysis: Furthermore, most of the query searching problems are implemented in English while the differences among languages are not considered yet. Thus, to fill in the blank in Japanese information searching. We focus on the investigation of travel comments written by Japanese users.

In this chapter, to deal with those issues mentioned above, we propose an algorithm to construct activity queries and extract meaningful examples as detailed descriptions. Our contributions are highlighted as follows:

- Extract activity keywords based on three types of frequency factors (Section 3.4.2).
- Construct activity queries with considerations on Japanese linguistic features (Section 3.4.3).
- Develop the activity score to select representative activity examples (Section 3.4.3).
- Two case studies applying the algorithm with 18,939 comments on Jaran (a Japanese travel website) [31] (Section 3.5).

This chapter belongs to *Landmark Evaluation* and *Landmark Recommendation* in Fig. 1.1.

3.2 Data-Collect Process

Table 3.1: The statistic of travel comments.

Landmark (Japanese)	三鷹の森 ジブリ美術館	浅草寺	上野動物園	国営昭和記念公園	明治神宮
Landmark (English)	Ghibli Museum	Senso Temple	Ueno Zoo	Showakinen Park	Meiji Jingu
Category	Museum	Temple	Zoo	Park	Shrine
Number of comments	1936	7101	4170	2531	3201
Max length of comments[1]	334	391	517	382	391
Median length of comments[1]	41	40	42	41	41
Min length of comments[1]	5	4	4	3	3

[1] Counted in Japanese characters.

The goal of our proposed algorithm is to extract interesting activities for a landmark based on a large corpus of travelogues. In other words, we ensure our algorithm so as to match the domain nature of travelogues and problem domain as much as possible. Instead of using travelogues such as tweets of which contents are not limited to travel experiences but also other reports such as food, sports, and health, we set our data domain as a corpus of travel comments obtained from the professional travel website. The advantages of using travel comments are as follows:

1. Using travel comments is efficient and convenient as the traditional data-cleaning process conducted by crowdsources is no longer needed. This is because non-experiential information such as advertisements in Tweets rarely exists in travel comments on big travel websites [20].

2. A travel comment usually is longer than a tweet which has the limitation of 140 characters. The median length of comments that we collected is over 40 Japanese characters (see Table 3.1). It means that a travel comment usually contains around 2-3 completed sentences in Japanese. This advantage ensures the informativeness of each comment.

To sum up, due to these two advantages, using travel comments is convenient and could guarantee the quality of our experimental datasets.

In this chapter, we set Tokyo as the target city, and Jaran [31], which is domestic travel website, as the target website in Japan. We focused on 5 famous public landmarks in Tokyo on Jaran, which are

- 三鷹の森ジブリ美術館 (Ghibli Museum)
- 浅草寺 (Senso Temple)
- 上野動物園 (Ueno Zoo)
- 国営昭和記念公園 (Showakinen Koen)
- 明治神宮 (Meiji Jingu)

For each landmark, we collected all travel comments of it through its full timeline. Table 3.1 summarizes the data statistic of travel comments collected. All travel comments were written before November, 2018[2] and were collected through a data-collection program in R. Comments written in English (less than 1%) and comments only contained photos were not included. As a result, 18,939 travel comments were accumulated in Japanese.

[2]The date is users' visiting date, not the written date.

3.3 Problem definition

Given a landmark, the goal is to explore potential activity descriptions that may benefit users in travel decision making. The inputs contain a landmark name and a large corpus of travel comments in Japanese. The outputs are the most likely activity query and top-5 examples in the form of an activity-example table. In summary, our activity-example extraction algorithm for a landmark is described as follows:

Step A1: Phrase all comments in the entire comment set.

Step A2: Obtain the keyword using the algorithm described in Section 3.4.2.

Step A3: Construct the query with the keyword using the algorithm described in Section 3.4.3.

Step A4: Extract top-5 examples with the query using the algorithm described in Section 3.4.4.

3.4 Activity-Example Extraction algorithm

The goal in this section is to provide user with the most important activity information, thus only the most representative noun is extracted and only examples related to that are extracted. Note that, if information of other activities are further required, those information can be extracted by searching the second, third important noun repeating the steps in section 3.4.2-3.4.4.

3.4.1 Japanese Linguistic feature

Japanese is quite different from English. For the sake of clear understanding, we introduce three basic structures in Japanese. Unlike English, there is no space between words in Japanese. For this reason, in order to better understand the structures of Japanese, we separate each important elements for each sentence in the following three examples.

1. 金さん は　男性です。 (Kin is a male.)
2. 金さん が 走る。 (Kin runs.)
3. 金さん は　優しいです。 (Kin is thoughtful.)

Those structures can be generally considered as *S (Subject)+ ST (Statement)* as Japanese is a topic-prominent language [57]. In example (1), *ST* is *male* which is a *N*

Table 3.2: Example of element phrasing.

Example 1	ジブリの作品が好き (Like the works by Ghibli)				
Element	ジブリ	の	作品	が	好き
POS	Noun	AUX	Noun	AUX	Adjective verb
Example 2	ジブリの世界観に浸れる (Immersed into the world of Ghibli)				
Element	ジブリ	の	世界観	に	浸れる
POS	Noun	AUX	Noun	AUX	Verb

noun. In example (2), *ST* is run which is a V verb. In example (3), *ST* is $thoughtful$ which is an A adjective.

Moreover, there are usually AUX auxiliary words between segments in Japanese. In example (1), は is the AUX, and it connects the parts 金さん *(Kin)* and 男性 *(male)* and is similar to the meaning of is in English. In example(2), が is the AUX, and it connects the parts 金さん *(Kin)* and 走る *(run)* which emphasizes the subject 金さん *(Kin)*.

Thus, our goal is to construct a query with a structure of *S (Subject)*+ *ST (Statement)* and we consider *S* is a noun.

For each comment, we extract the dependency relationships and part of speech tags (POS) through RMeCab [61]. Table 3.2 shows two examples of phrased travel comments in Japanese. Note that, we do not consider the differences of tenses in the phrasing process. Also, POS including prefix and postfix are not considered either.

3.4.2 Keyword finding

Based on the discussion in Section 3.4.1 above, to construct a query, we first identify the noun N for the subject S. N should be representative, in other words, N is the keyword for the landmark.

The comment set I_l contains all comments for the landmark l, and L is a set of all landmarks. In order to extract the most representative keyword, we uses the concept of $TF - IDF$ and evaluate each noun $n \in I_l$ via TF and IDF. In details, TF is the words which are frequently occurs in comments of a landmark (TF), while IDF is the words rarely appears in comments of other landmarks.

We mainly focus on three types of frequencies as in [75]. The first frequency is noun frequency (nf), the second on is comment frequency (cf), and the last one is landmark

Table 3.3: Example of the keyword ranking for Ueno Zoo.

Noun	Rank
パンダ (Panda)	1st
動物園 (Zoo)	2nd
子供 (Child)	3nd

frequency (lf). Then, for a noun n and a landmark l, we compute $nf(I_l, n)$ which is the number of appearances that have n in I_l. In order to avoid the overestimation of nf caused by the repetition of the same noun in the same comment, we introduce $cf(I_l, n)$ which means the number of comments that include n in I_l.

Moreover, words such as 人（*people*） and 私（*I*） should be excluded as those words may be frequently appear but do not contribute to activity descriptions much. To deal with the problem, we compute $lf(I_L, n)$ which is the number of comments having n in the entire comment set I_L.

The keyword noun score KNS of a noun n is computed as follows using these three types of frequencies:

$$KNS(n) \propto \frac{nf(I_l, n) \times cf(I_l, n)}{lf(I_L, n)} \tag{3.1}$$

A noun n is labeled as the keyword key if $KNS(n)$ is the largest for all $n \in I_l$. Table 3.3 shows an example list of keyword ranking for Ueno Zoo.

In addition, geo-specific nouns which represent a landmark's spatial locations such as *Shanghai* and *China* may frequently appear, but those words are noisy terms as it has fewer associations with activities. In an instance, *Shanghai* often comes with a comment *It is garden/park/building in Shanghai*. To quantify the keyword extraction, if a keyword key is a geo-specific noun, then we use the noun with the second high $KNS(n)$ value.

In summary, our keyword find algorithm for a landmark l is described as follows:

Step K1: For a landmark l, extract the keyword key of it as follows:

(1-1) Calculate nf, cf and lf values for a noun n.

(1-2) Calculate $KNS(n)$.

Step K2: Select the non-geo-specific noun with the highest $KNS(n)$ value as the keyword key.

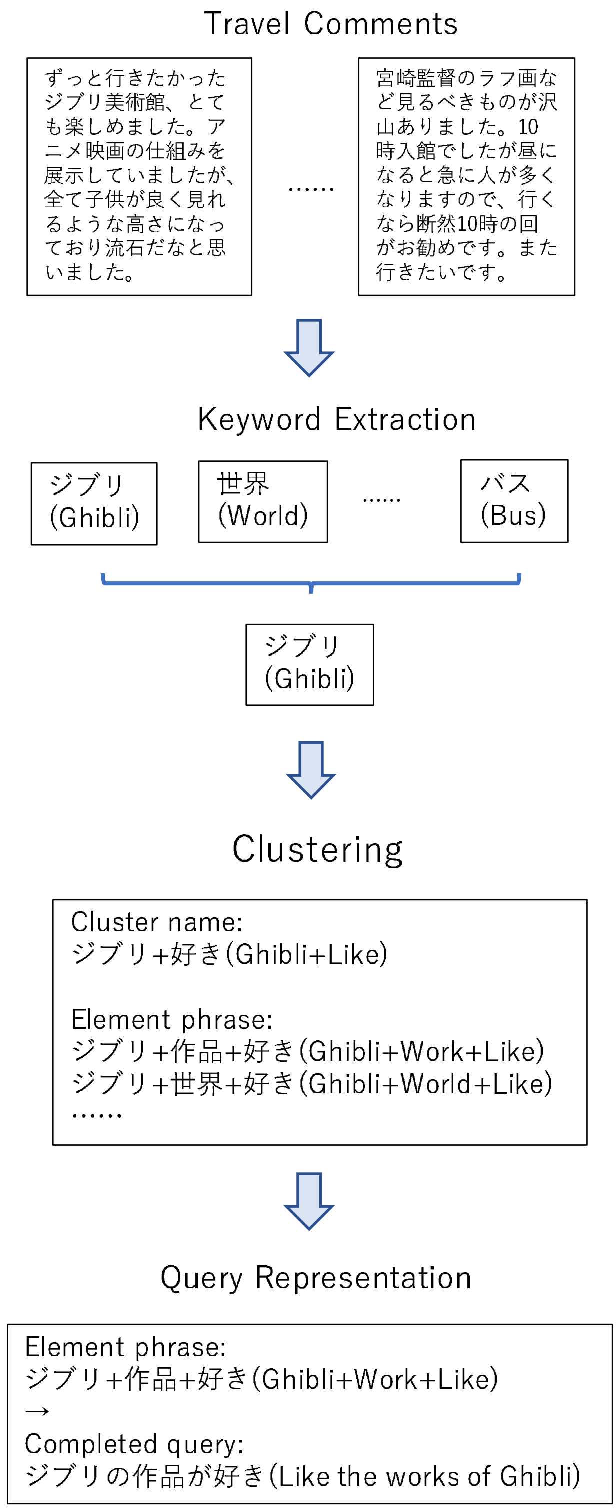


Figure 3.1: Query construction.

3.4.3 Query construction

After extracting the keyword key, we complete the query with the part of statement *ST* through:

1. Clustering: We search for 3-grams phrases using key obtained in Section 3.4.2 among *Noun, Verb, Adjective, Adjective verb, Adverb*. Users may express their personal experiences in different ways when writing the comments, and this results in identifying queries with substantially the same meanings. To handle this problem, we cluster similar 3-grams phrases based on the cosine similarity. If the cosine similarity between two 3-grams phrases is

   over the threshold d_{th} (we set $d_{th} = 0.67$ in the experiment), then we merge the phrases. The cluster name is the most repeated elements in the cluster.

2. Query Representation: We extract the most hit element phrase from the cluster. Then, we predict the AUX between any two elements in the element phrase. Given two elements e_1, 作品 *(Wokr, Noun)* and e_2, 好き *(Like, Adjective verb)* (see query representation in Fig 3.1), let A be a set of all aux's in the I_L, and for $aux \in A$, we define two probability factors:

   - The first probability factor is generic probability $gp(I_L, aux, e_1, e_2)$, it is the probability of aux connects $Noun$ and $Adjectiveverb$ in the entire comment set I_L.
   - The second probability factor is specific probability $sp(I_l, aux, e_1, e_2)$, it is the probability that aux connects 作品 *(Wokr)* and 好き *(Like)* in the comment set of the landmark I_l.

With these two types of probability factors, we define a connection probability $cp(aux, e_1, e_2)$ predicting the relationship between aux and two elements e_1, e_2 as follows:

$$cp(aux, e_1, e_2) \propto gp(I_L, aux, e_1, e_2) \times sp(I_l, aux, e_1, e_2) \tag{3.2}$$

We consider that the aux with the highest $cp(aux, e_1, e_2)$ value connects e_1 and e_2. By repeating connecting all elements in the element phrase, we finally obtain a complete query q.

3.4.4 Example searching

The next step is to find representative comments match the query q. We define a scoring mechanism called activity score *AS* to select example comments. We select comments containing or partly containing the query q and score them by *AS*.

To score each comment, we use three factors:

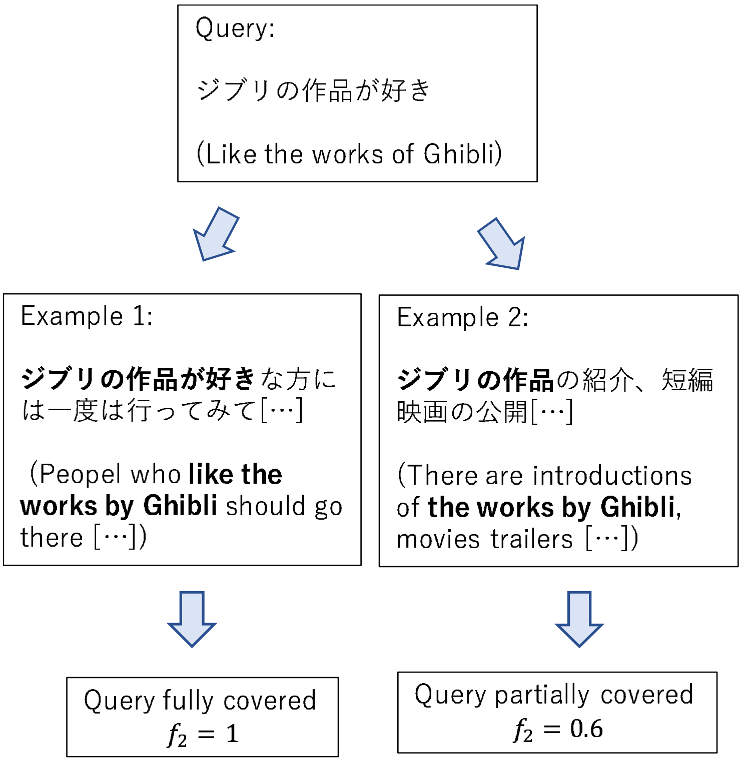


Figure 3.2: Coverage factor.

1. *Comparison factor*: It is widely used in decision making [34]. This factor is computed as a binary variable. It is 1 when the query q appears in the comment and otherwise it is 0.

2. *Coverage factor*: It is defined as the percentage of a query q contained in a comment and it is normalized between 0 and 1 (see Fig. 3.2).

3. *Length factor*: A comment with a longer length is more informative. The factor is also normalized between 0 to 1 compared with all comments in I_l. The longest comment is 1 and the shortest one is 0.

We define the AS through a weighed linear relationship:

$$AS = w_1 f_1 + w_2 f_2 + w_3 f_3 \tag{3.3}$$

when w_i is the weight for $i - th$ factor. f_1, f_2, f_3 represents comparison, coverage and length factors, respectively. At last, top-5 comments are picked as the examples for detailed descriptions of the landmark activity.

3.5 Experimental Result

Our goal is to process a full dataset of travel comments, but we first implement the small-scale test to confirm the effectiveness of our proposed algorithm. In this section, we evaluate two case studies by deriving meaningful activity queries and example comments.

3.5.1 Experimental settings

We evaluate the quality of experimental results in the following criteria:

- Whether the constructed queries are meaningful
- Whether the extracted examples are proper.

We use the 18,939 collected in Section 3.2 as the entire comment set I_L.

For the first case study, we have a subset of all Japanese travel comments for the Ghibli Museum with 1936 comments during the period of May 2007 to October 2018 [3]. For the second case study, we have a subset of all Japanese travel comments for the Ueno Zoo with 4170 comments during the period of April 2006 to October 2018 [4].

We invited 5 volunteers as evaluators. 3 of them are in computer science and engineering fields. 2 of them are English speakers, and the other 3 volunteers are Japanese and English speakers. All of them are not authors of this chapter.

For each study case, we derived one query and corresponding examples with top-5 AS values. The query and examples were originally in Japanese and translated into English lately. Volunteers were asked to first rate to what extent they understood the queries with a scale from 1 (do not understand)–5 (fully understand), and then, they were asked to rate the relevance of the top-5 examples, from 1 (least relevant)–5 (most relevant). Based on the ground-truth, we calculated normalize Discounted Cumulative Gain ($nDCG$) to evaluate the quality of the obtained examples.

We invited 5 volunteers as evaluators. 3 of them are in computer science and engineering fields. 2 of them are English speakers, and the other 3 volunteers are Japanese and English speakers. All of them are not authors of this chapter.

For each study case, we derived one query and corresponding examples with top-5 AS values. The query and examples were originally in Japanese and translated into English lately. Volunteers were asked to first rate to what extent they understood the queries with a scale from 1 (do not understand)–5 (fully understand), and then, they were asked to rate the relevance of the top-5 examples, from 1 (least relevant)–5 (most relevant). Based on the ground-truth, we calculated normalize Discounted Cumulative Gain ($nDCG$) to evaluate the quality of the obtained examples.

[3]https://www.jalan.net/kankou/spt_13204cc3302011245/

[4]https://www.jalan.net/kankou/spt_13106cc3310040182/

Table 3.4: $nDCG_p$ of our proposed algorithms in two case studies.

	Case 1: Ghibli Museum	Case 2: Ueno Zoo
Avg. $nDCG_1$	1.000	1.000
Avg. $nDCG_2$	0.900	1.000
Avg. $nDCG_3$	0.920	0.952
Avg. $nDCG_4$	0.966	0.991
Avg. $nDCG_5$	0.969	0.992

Evaluations on query construction

For the evaluation of query quality, we compared our proposed algorithm with two baseline algorithms with disabling parts of our algorithm as follows:

- Keyword query: It has the most hit noun as a query based on steps from Section 3.4.2.
- Incomplete query: It constructs a query based on steps from Section 3.4.2 to Section 3.4.3 without predicting the AUX's.

Table 3.5 shows the results of query evaluation. Generally speaking, it is observed that our algorithm outperforms the other two algorithms in both case studies. Our algorithm is at least 10% higher than the other two baseline algorithms. It is interesting to mention that in case 2, there is no AUX between 一番 *(most)* and 人気 *(popular)* by our proposed algorithm due to the common usage in Japanese.

- Ours vs Keyword query: Since the word *Panda* is more comprehensive compared to the word *Ghibli*, which is a Japanese word. Thus, Keyword query obtained a higher score in case 2 than that in case 1. On the other hand, it indicates that our algorithm is useful to help users have a more comprehensive image of the landmark activity dealing with descriptions in the foreign languages that they are not familiar with.
- Ours vs Incomplete query: Although Incomplete query provides more information, our algorithm has a completed query rather than a bag of words. Thus, it is reasonable that our algorithm is easier to understand and obtains higher scores in both case studies.

Table 3.5: Results on query construction.

Case 1: Ghibli Museum		
Algorithm	Avg. score	Query
Ours	4.2	ジブリの作品が好き) (Like the works by Ghibli)
Keyword query	2.6	ジブリ (Ghibli)
Incomplete query	3.2	ジブリ+作品+好き (Ghibli+Work+Like)
Case 2: Ueno Zoo		
Algorithm	Avg. score	Query
Ours	4.4	パンダが一番人気 (Panda is the most popular)
Keyword query	3.4	パンダ (Panda)
Incomplete query	4.0	パンダ+一番+人気 (Panda+Most+Popular)

Evaluations on nDCG

To extract the top-5 examples, we evaluated each comment using AS. We set the following weights for the three factors in AS with $w_1 = 0.25$, $w_2 = 0.5$ and $w_3 = 0.25$, which achieved the best performance according to the prior experiments.

Then, DCG is computed firstly as follows:

$$DCG_p = r_1 + \sum_{i=2}^{p} \frac{r_i}{log(i)} \tag{3.4}$$

where r_i is the score of $i-th$ example, and DCG_p is the accumulated score at i_{th} example.

As DCG depends on the size of the extracted examples, the more examples are extracted, the larger DCG becomes. In other words, it is not explicitly to figure out whether the extracted examples are qualified or not. Thus, to further normalize the results we compute $nDCG$ either as follows.

Table 3.6: Example comments for Ghibli-museum.

Query	Examples	*AS*
ジブリの作品が好き (Like the works by Ghibli)	**ジブリの作品が好き**な方には一度は行ってみて [...] トトロだけではなく他作品も含めて [...]	0.782
	昔から**ジブリの作品が好き**でよく見ています。ここはジブリの世界に入り込んだ [...]	0.774
	ジブリの作品が好きなので、リピートしています。何度行っても、童心に帰れて [...]	0.774
	[...] **ジブリの作品が好き**で暇さえあれば見ています。ジブリ美術館に行くとジブリの世界 [...]	0.772
	[...] **ジブリの作品**の紹介、短編映画の公開、ネコバス、天空の城ラピュタの再現 [...]	0.717

$$nDCG_p = \frac{DCG_p}{IDCG_p} \tag{3.5}$$

where $IDCG_p$ is the ideal rank.

The results are listed in Table 3.4. Our algorithm achieves high *nDCG* values in both case studies at any ranks. It indicates that our extracted examples are quite relevant to the landmark activity descriptions and can help users better understand what is interesting to do and see when visiting the landmark. Generally speaking, according to the experimental results, our algorithm could benefit users in the better travel decision-making process.

Extracted examples

The top-5 examples extracted for two case studies are listed in Table 3.6 and Table 3.7. The queries contained in the examples are bold.

In case 1, 1st example to 4th example contains the complete query (ジブリの作品が好き *(Like the works by Ghibli)*), while 5th example contains a part of the query (ジブリの作品 *(The works by Ghibli)*). For better understanding, English translations of first to fifth examples of Ghibli Museum are listed as follows:

- *People who* ***like the works by Ghibli*** *should go there. Reservation is needed. It does not only include Totoro but also includes other works of Ghibli[...].*

Table 3.7: Example comments for Ueno Zoo.

Query	Examples	*AS*
パンダが一番人気 (Panda is the most popular)	[...] **パンダが一番人気**であることが分かります。食事中のパンダは [...]	0.810
	[...] **パンダが一番人気**でとても可愛いくて癒されます。アリクイなど他にも [...]	0.779
	[...] **パンダが一番人気**だけど、珍しい変テコな動物もいっぱいいてビックリします [...]	0.774
	[...] **パンダが一番人気**だと思います。入り口入って右側にパンダのエリアがある [...]	0.774
	[...] **パンダが一番人気**で多くの人が行列をして見ていました。パンダがえさを食べてる姿 [...]	0.771

- *Pretty* ***like the works by Ghibli*** *a long time ago and I often watch the movies. It feels like getting in the world of Ghibli [...].*
- *[...] I like* ***the works by Ghibli****, I have been there for several times. Whenever you go to the museum, It feels like going back to your childhood time [...] .*
- *[...] I like to watch to* ***the works by Ghibli*** *in my spare time, movies trailers. If you go there, it feels like getting in the world of Ghibli [...] .*
- *[...] There are introductions of* ***the works by Ghibli****, movies trailers, models of the cat bus, Castle Laputa in the Sky [...] .*

In case 2, all examples contain the complete query (パンダが一番人気 *(Panda is the most popular)*). For better understanding, English translations of first to fifth examples of Ghibli Museum are listed as follows:

- *[...]* ***Pandas are the most popular****. It is great to have a chance to see panda slowly eating[...].*
- *[...]* ***Pandas are the most popular*** *and are so cute that heal the soul. There are other animals such as anteater [...].*
- *[...]* ***Pandas are the most popular****, also it is so surprising that there so many other rare animals [...].*

- *[...] I think **Pandas are the most popular**. The panda area is on the right side of the zoo entrance [...].*
- *[...] Because **Pandas are the most popular**, I saw them after waiting a lot in lines. Pandas were eating their food [...].*

3.6 Conclusion

In this chapter, we concentrate on answering the question: Can we accurately and directly tell users what to do and see during their visits via previous travel comments?

Based on the experimental results, with a relatively small scale of travel comments, it still allows us to explore rich landmark activity information.

For future works, we are continuing to deepen our works and intend our scope to other language-specific travel comments such as English and Chinese. Meanwhile, the workloads of the applications will be evaluated.

Chapter 4

Landmark Seasonal Travel Distribution and Activity Prediction Based on Language-specific Analysis[1]

4.1 Introduction

With the flourish of social media, it brings new insights into intelligent travel recommendation. Travelers usually try to perceive a general image about how the travel destination will be like before the departure and, due to this sense, it is not surprising that users will check the comments made by former users on the trustworthy travel websites such as TripAdvisor [20]. These comments represent users' personal travel experiences and users' tourism decisions are strongly influenced by comment contents in many aspects [2]. Websites such as TripAdvisor and Yelp have influenced the way that users decide where to go, when to go and what to see and do during holiday. [59].

In this chapter, we propose an algorithm that can effectively predict seasonal activity for a landmark for language-specific users. Specifically, we describe the following points:

- How users' travel distribution varies through language backgrounds and how the travel distributions in each language group changes between years.
- Current studies often recommend popular activities through user comments in 12 months. Can we target peak seasons ahead and only check the user comments in the peak seasons in order to reduce computation time on loading and reading data?
- *TF-IDF* are recently used in textual data retrieval (See Section 2.3 in detail). Can we improve its performance on the activity extraction that alleviates the demands of extraction for recently appeared activities?

[1]Technical contents in this chapter have been presented in the publications ⟨3⟩.

This chapter belongs to *Landmark Evaluation* and *Landmark Recommendation* in Fig. 1.1 and our contributions are highlighted as follows:

1. We collect 417,787 user-generated comments on TripAdvisor for each of the top-100 ranked landmarks in Tokyo, Shanghai, and New York City. Those comments are divided into three languages groups, including Chinese, Japanese and English.

2. We analyze the differences in travel distribution for each landmark among three language groups and differences in yearly travel distribution for each language group.

3. The future travel distribution for each landmark in each language is predicted by seasonal ARIMA, *SARIMA*. All the travel distributions are divided into two clusters based on *seasonality*. Peak-off seasons in each landmark can be detected based on its *seasonality*.

4. We introduce an exponential *TF-IDF* (*ETF-IDF*) to extract the most suitable seasonal activity in the recent years. Based on experimental results in three cities, our algorithm has a better performance in terms of language-specific peak-off season detection and language-specific seasonal activity recommendation which outperforms the previous studies [5, 33].

4.2 Related work

Many travel recommendation systems have been developed for travel decision support. Existing works on popular landmarks recommendation includes the four types of social media data, GPS trajectory [41, 73, 85] , check-in logs [39, 75], geo-tags [43, 52], and travel logs [25, 56].

Instead of recommending a series of landmarks, suitable and representative seasonal activities of each landmark should also be recommended at the same time. Taking the Meguro-gawa (River) in Tokyo as an example, it is famous for cherry flower (Sakura) viewing during the period of March to April, while the cherry flower viewing is not unavailable during the other months. Therefore, the seasonal activity differences for each landmark should be included in order to match users' travel time schedule.

For seasonal analysis , Liang et al. [40] and Yamasaki et al. [78] analyze the differences of travel distributions in different seasons without high-lightening any seasonal activity information. Fang et al. [17] extracting seasonal activities based on word frequencies in each month. However, words such as "cherry flower" are highly appeared in comments of all months, although flower viewing is only available during March-April. It indicates that word extraction only based on word frequencies is invalid in some cases.

Table 4.1: The statistics of user comments from TripAdvisor in Tokyo, Shanghai and New York.

City	User	Chinese	English	Japanese
Tokyo	88,175	6,546	42,733	39,252
Shanghai	38,850	5,524	19,119	14,207
New York City	290,762	7,050	173,878	110,834
Total	417,787	19,120	235,374	164,293

In addition, some social media data corpus can be unreliable. For example, geo-tags are often given with non-geographical posts [11], such as “explorejapan” and textual contents of how they feel that the landmarks are often unstructured or ambiguous [42], as the language used in social media such as Twitter and Facebook deviates from normal language usage. For example, users tend to use emoji, abbreviation or acronyms such as “www” (many laughs in Japanese) and the detailed activities, what they have done during the visit, are usually missing.

Thus, those issues in current studies make the data collection and filtering process time-consuming and degrade the accuracy of activity estimation.

To overcome the challenges that we discussed above, we leverage the advantages of TripAdvisor with a large amount of trustworthy travel comments. Also, we extract landmarks’ seasonal activities through peak seasons. Thus, for a word such as "cherry flower", although it frequently appears in comments of all months in landmarks such as Meguro-gawa, we only consider this word as the representative words for peak-seasons such March and April.

4.3 Data source and data-collect process

We collect user comments from TripAdvisor (United States) [69], which is one of the biggest leading travel websites.

Collecting data from TripAdvisor provides the two advantages: Firstly, TripAdvisor offers large quantities of free user-generated comments which have high availability. Secondly, we can group users comments based on the languages accordingly. TripAdvisor supports multiple languages interfaces on the user comment, and hence language division is possible. Therefore, the two issues, lack of language deviance consideration and unstructured and ambiguous textual data, in the previous studies as discussed in Section 6.1 can be resolved.

A data-collection program was developed using R and it collected user comments of the 100-ranked landmarks in three popular travel destinations, Tokyo (TYO), Shanghai (SHA), and New York City (NYC). All the user comments collected from TripAdvisor are during the period of 1st January 2012 to 31st December 2017. Particularly, the landmark which is labeled as "region", for example "Shibuya District", is not included as it does not have a specific type or location. Finally, it took about 10 days to crawl the user comment data.

For each landmark, we collected each user's comment content and the time of writing. We also collected the rank of each landmark, where the rank of each landmark is decided by the user satisfaction and the number of user comments on the travel website.[2]

Language detection was developed through R, where the representative characters in the comment content, such as English (is), were used to automatically identify a language. Table 4.1 shows the statistics of user comments in TripAdvisor. As user comments in Chinese, English, and Japanese are available in all the landmarks. On the other hand, user comments besides the three languages are excluded in some landmarks. For example, in the case of Shinagawa Aquarium in Tokyo, user comments in French is excluded. In this sense, for language-specific analysis in Section 4.4.1, we focus on totally 417,787 user comments consisting of Chinese, Japanese and English.

4.4 Observation of travel distribution

4.4.1 Comparison of language-specific travel distribution

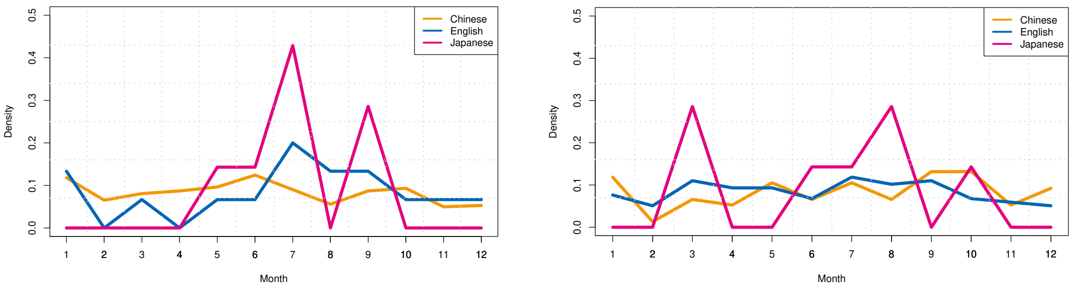
We analyze the travel distribution from the years 2014–2017 of the top-100 ranked landmarks in the collected dataset for each language group in the three cities. "Travel distribution" of each year for a landmark l in a particular language group $lang$ is described by:

$$TD_l^{lang} = \{des_{l,1}^{lang}, des_{l,2}^{lang}, ..., des_{l,11}^{lang}, des_{l,12}^{lang}\} \tag{4.1}$$

where $des_{l,i}^{lang}$ is the density of comments appeared in i-th month for a particular language group $lang$ in the entire comment set for l. $des_{l,i}^{lang}$ is defined by:

$$des_{l,i}^{lang} = \frac{N_{l,i}^{lang}}{\sum_{k=1}^{12} N_{l,k}^{lang}} \tag{4.2}$$

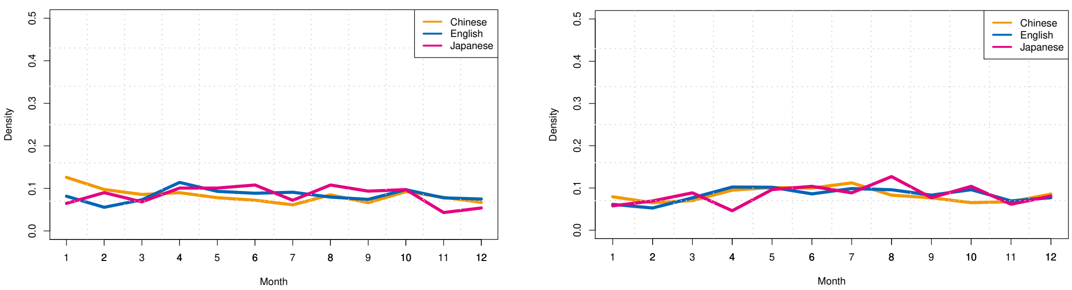
[2]Note that, how to rank landmarks on every travel website is not open. It is just decided based on user satisfaction and the number of user comments.



(a) Tokyo Camii (Church), Tokyo.

(b) Shanghai former provisional government site of the Republic of Korea (Historic spot), Shanghai.

Figure 4.1: Examples of landmark with language-specific travel distributions.



(a) Meiji Jingu (Temple), Tokyo.

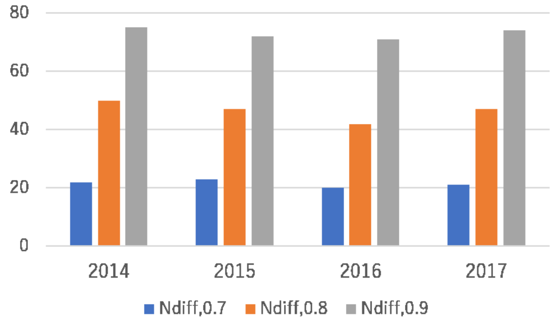
(b) Central Park (Park), New York City.

Figure 4.2: Examples of landmark without language-specific travel distributions.

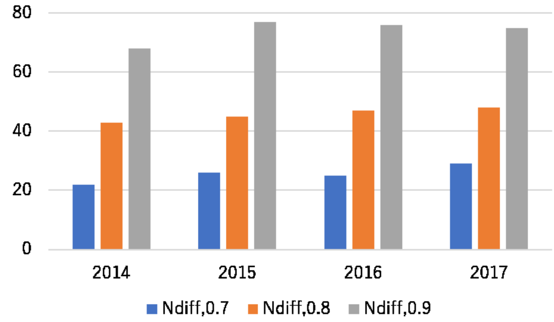
where $N_{l,i}^{lang}$ is the number of comments for l in the i-th month for a particular language group $lang$.

Figs. 4.1 (a) and (b) show the examples of landmarks that is with obvious differences in its language-specific travel distributions in Shanghai and Tokyo. Figs. 4.2 (a) and (b) show the examples of landmarks that is without obvious differences in its language-specific travel distributions in Tokyo and New York City.

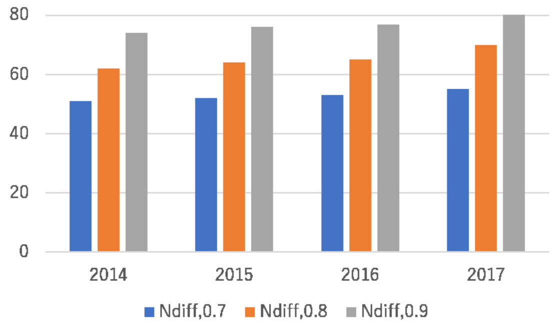
To find the differences of travel distribution for different language groups, we calculate the cosine similarity $cosim_l^{lang_{i,j}}$ between $TD_l^{lang_i}$ and $TD_l^{lang_j}$.



(a) Tokyo.



(b) Shanghai.



(c) New York City.

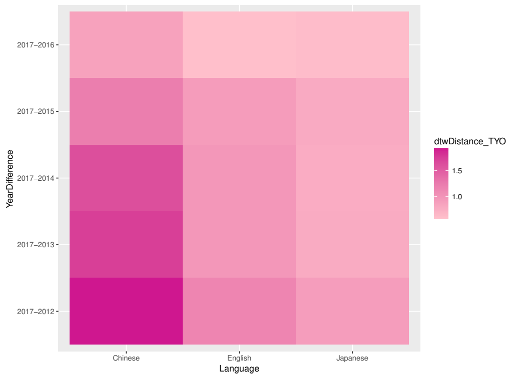
Figure 4.3: The number of $N_{diff,\theta}$ in Tokyo, Shanghai and New York.

$$cosim_l^{lang_{i,j}} = \frac{< \boldsymbol{TD}_l^{lang_i} \cdot \boldsymbol{TD}_l^{lang_j} >}{\|\boldsymbol{TD}_l^{lang_i}\| \cdot \|\boldsymbol{TD}_l^{lang_j}\|} \tag{4.3}$$

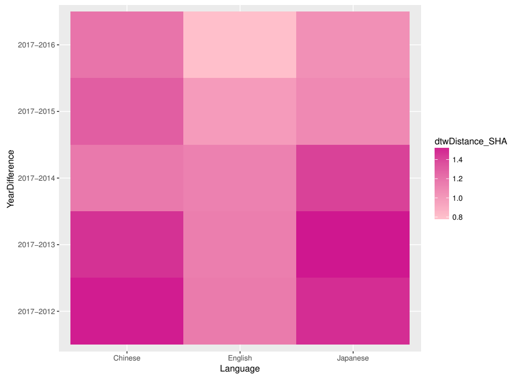
where $< \boldsymbol{v_1} \cdot \boldsymbol{v_2} >$ shows the inner product of two vectors $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ and $\| \cdot \|$ shows the L^2 norm.

If a landmark has $cosim_l^{lang_{i,j}} \leq \theta$ in any two of language groups, then we consider it as a landmark that there are obvious differences in its language-specific travel distribution. $N_{diff,\theta}$ is defined by the number of landmarks that have obvious differences in its language-specific travel distribution in the top-100 ranked landmarks at the level of θ.

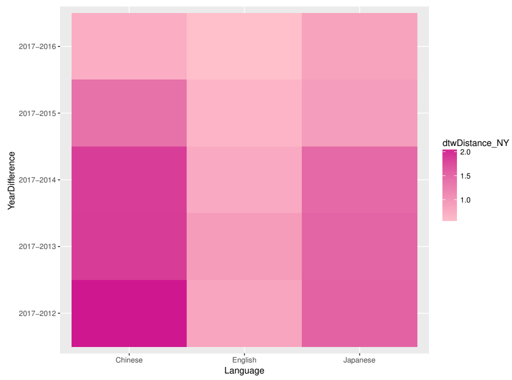
We set $\theta = 0.7, 0.8, 0.9$ and the results are shown in Fig. 4.3. The N_{diff} value increases as θ increases. Even though θ is set at a low level of 0.7, we still have a high N_{diff} value in any city cases. Thus, it is necessary to consider language-specific



(a) Tokyo



(b) Shanghai



(c) New York City

Figure 4.4: DTW distance comparison of language-specific yearly travel distribution in Tokyo, Shanghai and New York City, 2014–2017.

differences in the landmark activity recommendation.

4.4.2 Comparison of yearly travel distribution

Most of the current studies simply use the average travel distribution as the future ones. Instead, it is better for us to consider the changes between years because the popular landmark activity may differ from year to year. To confirm our assumption, we use a dynamic time warping algorithm (DTW algorithm) to analyze the dissimilarity between years quickly [38].

The DTW algorithm is an algorithm that compares the similarities between two series. It aims at finding the optimum match between them under certain restrictions [48]. It is widely used in the field of data mining to compensate for some demerits of Euclidean distance.

For each language, let TD_l^y and TD_l^{y+k} be two travel distribution vectors with the length of 12 defined by:

$$TD_l^y = \{des_{l,1}^y, des_{l,2}^y, \ldots, des_{l,12}^y\} \tag{4.4}$$

$$TD_l^{y+k} = \{des_{l,1}^{y+k}, des_{l,2}^{y+k}, \ldots, des_{l,12}^{y+k}\} \tag{4.5}$$

where $des_{l,i}^y$ is the density of i-th month in the year k for landmark l for a particular launguage. For $1 \le i \le 12$, let n_i and m_i be the integers of $1 \le n_i \le 12$ and $1 \le m_i \le 12$ The goal of the DTW algorithm is to find a mapping path $p = \{(n_1, m_1), \ldots, (n_{12}, m_{12})\}$ such that the total $DTW(TD_l^y, TD_l^{y+k})$ distance on this path is minimized [48], where

$$DTW(TD_l^y, TD_l^{y+k}) = \sum_{i=1}^{12} |des_{l,n_i}^y - des_{l,m_i}^{y+k}|. \tag{4.6}$$

Then, the average numbers of $DTW(TD_l^y, TD_l^{y+k})$ for the top-100 ranked landmarks in Tokyo, Shanghai and New York are calculated for the three language groups for $y = 2012$ and $k = 1, 2, 3, 4, 5$. Fig. 4.4 shows the results in the three city cases. It is obvious that the DTW distance between two travel distributions increases when k increases in all the language groups. It means that the travel distribution varies each year and can be more closely mirrored by that in a year close to it. Thus, the results confirm that we cannot simply use the average value of travel distribution to predict the future travel distribution.

4.5 Travel distribution algorithm

4.5.1 Travel distribution prediction by seasonal ARIMA (SARIMA)

Future travel distribution prediction

The goal of this step is to predict the accurate future travel distribution over seasonality and yearly changes. Therefore, we use the SARIMA to realize that [63].

In our proposed seasonal ARIMA (SARIMA) model, the seasonal AR (autoregressive model) term and MA (moving-average model) term generate the prediction of TD_{2018} with data obtained from 2014–2017. The shorthand notation of our proposed model is as follows:

$$SARIMA : (p, d, q)(P, D, Q)S \tag{4.7}$$

where p, d and q represents the three non-seasonal terms including non-seasonal AR order, non-seasonal differencing, and non-seasonal MA order. Similarly, P, D and Q

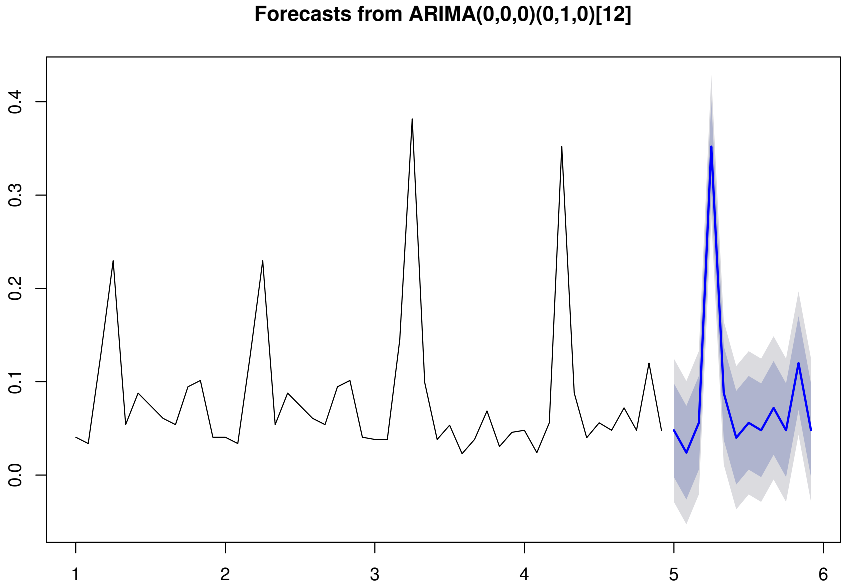


Figure 4.5: Prediction of travel distribution in 2018 for Chidorigafuchi of Japanese group by SARIMA. The travel distribution in 2018 is from period 5–6 and parameters used are (0,1,1)(0,1,0)[12].

represents the three seasonal terms including seasonal AR order, seasonal differencing, and seasonal MA order. S is the time span of regular travel distributions repeats, and we set $S = 12$ (months per year).

For the parameters set, the parameters used to predict the distribution of 2018 are the ones that best fits the distribution of 2017 based on the data of 2014–2016. Also, we always set $q = 1$ to create the exponential decay of importance of the past travel distribution with the consideration of yearly differences in Section 4.4.2. Fig. 4.5 shows an example of the Chidorigafuchi (a large Park in Tokyo). In the x-axis, 1 to 4 correspond to the years of 2014 to 2017 and 5 corresponds to the year 2018.

Seasonality clustering

In addition, SARIMA can automatically detect the seasonality in those travel distributions. Seasonality means a regular pattern of changes in a time sequence recurring every S time intervals. The landmarks with the obvious seasonality are clustered into C_1 whereas the landmarks that do not have the obvious seasonality are clustered into C_2. As shown in the examples in Fig. 4.1, it is clear that Fig. 4.1 (a) and (b) has the obvious

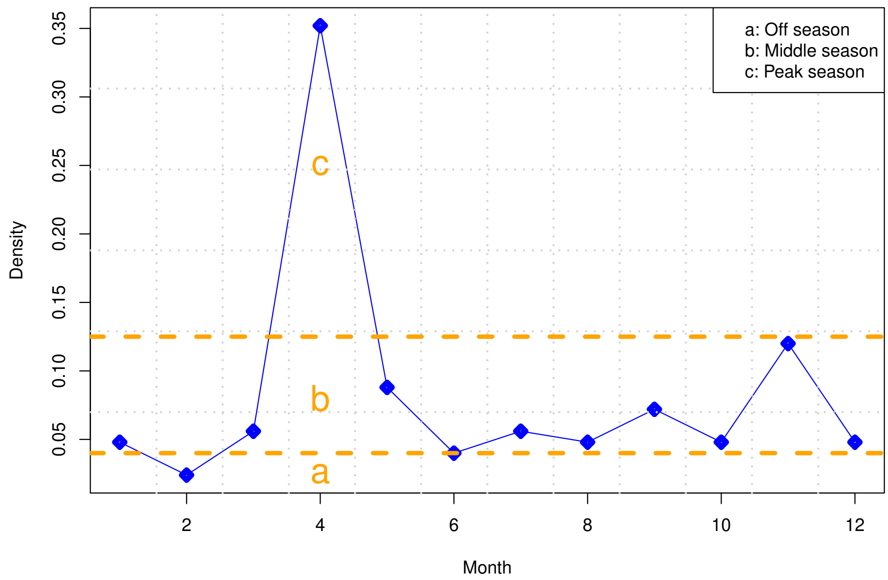


Figure 4.6: Peak-off season detection for Chidorigafuchi for Japanese group in Tokyo.

seasonality (peak seasons) in July and September for the Japanese group. Oppositely, Fig. 4.2 (a) and (b) has no seasonality in any language groups.

Unlike the previous studies which take great costs to analyze all comments, in our algorithm, only the seasonal activities in the peak seasons of the landmark in C_1 will be extracted and the seasonal activity of a random month of the landmark in C_2 will be extracted.

4.5.2 Seasonal activity extraction

Peak-off season detection by SAX

After dividing all landmark travel distributions into two clusters, we first identify the exact peak-off seasons for travel distributions in C_1. We use the Symbolic Aggregate approximation (SAX) which is fast symbolic approximation of time series and could be applied to any time series analysis [66]. Each travel distribution is transformed into a sequence of three levels, which includes a (off season), b (middle season) and c (peak season). Note that, before processing with SAX, the values of time series should be normalized. For more details, see [66]. Fig. 4.6 shows an example of peak-off season distribution in Chidorigafuchi, Tokyo.

Seasonal activity extraction by exponential *TF-IDF* **(***ETF-IDF***)**

We have pre-processed each user comment d for every landmark l in the language $lang$ in the format as follows:

$$[u][l][lang][y][m][w_1][occu_1]...[w_n][occu_n] \tag{4.8}$$

where $[u]$ is the user id, y and m are the writing year and the month of the comment, w_i is a noun appeared in the comment and $occu_i$ is the how many times w_i appeared in the comment.

We improve the basic model of TF-IDF [45]. In TF-IDF, for a noun w_i in a comment d_j, the weight W_{w_i,d_j} is given by:

$$W_{w_i,d_j} = TF_{w_i,d_j} \times \log \frac{N}{DF_{w_i}} \tag{4.9}$$

where TF_{w_i,d_j} is the number of occurrences of w_i in a comment d_j (which is equal to $[occu_i]$), DF_{w_i} is the number of comments including the term w_i, and N is the total number of used comments.

With considerations of deviance in years in Section 4.4.2, to obtain the weight $W'_{w_i,d_j}(y+1,m)$ of w_i in a particular month m in the future year $y+1$, we define the exponential TF-IDF (ETF-IDF) as follows:

$$\begin{aligned} &W'_{w_i,d_j}(y+1,m) = \\ &\alpha \times (W_{w_i,d_j}(y,m)) + (1-\alpha) \times W'_{w_i,d_j}(y-1,m) \end{aligned} \tag{4.10}$$

α is the smoothing or decay factor defined by $\alpha = 2/(1+Y)$, where Y is the total number of years [51].

We extract the most weighted noun in peak seasons for a landmark in C_1 and extract the most weighted noun in a random month for a landmark in C_2 by ETF-IDF. The extracted noun is regarded as the recommended seasonal activity.

We have conducted other textual information retrieval methods including frequency counts [5] and average word embedding [8], but our preliminary experiments turn out that the exponential TF-IDF has the best performance.

4.6 Performance evaluation

We evaluate the proposed algorithm on the datasets collected in Section 4.3 with 417,789 user comment from TripAdvisor. We divided the performance evaluation into two parts.

- The first part focuses on the evaluation for seasonality clustering.
- The second part focuses on the comparison with previous studies on seasonal activity recommendation.

Table 4.2: Clustering result summary

City	Chinese		English		Japanese	
	C_1	C_2	C_1	C_2	C_1	C_2
TYO	65%	35%	60%	40%	50%	50%
SHA	40%	60%	47%	53%	54%	56%
NY	51%	49%	52%	48%	68%	32%

Table 4.3: The average number of months used in seasonal activity recommendation

City	Chinese		English		Japanese	
	C_1	C_2	C_1	C_2	C_1	C_2
TYO	≤ 3	N/A	≤ 3	N/A	≤ 3	N/A
SHA	≤ 3	N/A	≤ 3	N/A	≤ 3	N/A
NY	≤ 2	N/A	≤ 3	N/A	≤ 3	N/A

4.6.1 Seasonality clustering

The clustering results are summarized in Table 4.2 through *SARIMA* analysis in Section 4.5.1. Table 4.2 lists the results of seasonality clustering of the three language groups in the three cities. Table 4.2 indicates that, for each language group, landmarks with the obvious seasonality do exist and the ratios of those landmarks are no less than 40%. For C_1, only user comments in peak seasons are investigated and, for C_2, only user comments in a randomly chosen month are investigated and it is shown as "N/A".

Unlike the previous studies that investigate user comments in all months, the proposed algorithm can only search through user comments in at most 3 months according to the results in Table 4.3. In other words, our algorithm is definitely more effective in extracting activity noun with around 70% improvement compared with the previous studies.

4.6.2 Comparison with previous studies on activity recommendation

Evaluation criteria

We consider that seasonal activity recommendation is valid if it fits the two conditions (labeled as ◯ in Table 4.6) :

- Condition 1: It is an available activity (activity is correct).

Table 4.4: An example of the most weighted noun extraction for Japanese group in Tokyo

Landmark	Cluster	Season-activity description
Chidorigafuchi	C_1	Sakura (Apr)
Meguro-gawa	C_1	Sakura (Apr), Illumination (Nov–Dec)
Kasai Sea Life Park	C_2	Tuna (all months)
Sky Tree	C_2	Observation (all months)
Tokyo Tower	C_2	Observation (all months)

- Condition 2: The activity is available in the recommended months (the months are correct).

Then the number of true positives (TP) is defined by the number of the recommended seasonal activities that successfully fit the two conditions (1) and (2). We consider that the recommended seasonal activity is invalid if it does not meet both of the two conditions (1) and (2). The number of false positives (FP) is defined by the number of the recommended activities that fail to meet either one of the two conditions. We check the official website for activity information confirmation.

Then the *precision* can be obtained by:

$$Precision = \frac{TP}{TP + FP} \tag{4.11}$$

Table 4.4 presents five valid examples of recommended seasonal activities in Tokyo for Japanese groups. For example, in the case of Meguro-gawa (River), our algorithm recommends two activities for Japanese groups, Sakura (April) and Illumination (November to December). Both of the two activities are available according to the information on official websites. TP, in this case, becomes 2 and FP is 0.

Procedures

We compare the proposed algorithm (Ours) to Ours+*TF-IDF*, [33] and [5]. For our proposed algorithm, we use the comments from the years 2014–2017, where $Y = 4$ and $\alpha = 0.4$ for *ETF-IDF*. Ours+*TF-IDF* considers no exponential decay in noun extraction. Jiang et. al [33] divides 12 months into four seasons, with spring, summer, autumn, and winter and it extracts topic tags for the most popular season by using the frequency count. Frequency count [5] is a baseline algorithm which extracts the most hit noun as recommended seasonal activities for all months.

Table 4.5: Comparison on average precision on activity recommendation.

Algorithm	Avg. precision
Ours (Chinese)	63.3%
Ours (English)	63.3%
Ours (Japanese)	66.7%
Ours+*TF-IDF*[1]	60.0%
[33][1]	56.7%
[5][1]	50.0%

[1] Used comments in Japanese.

We randomly select 10 landmarks from each city. For Ours, we recommend seasonal activities for each landmark in three languages groups. In total, 90 trials were carried out for Ours for each city. For Ours+*TF-IDF*, [33] and [5], the seasonal activity for each landmark is derived without language consideration. For these three algorithms, 30 trials were conducted for each city.

Results analysis

Table 4.5 presents the comparison of average precision on seasonal activity recommendation. It is observed that our algorithm has the best performance with at least 4% improvement.

Ours (Chinese) vs Ours (English) vs Ours (Japanese): All of them achieve relatively high average precision in Table 4.5, which confirms the effectiveness of our proposed algorithm. According to the examples in Table 4.6, our proposed algorithm in English groups did not extract the peak season in November–December, while the other two groups share the same results. It indicates that it is possible that users in different groups may have different travel distribution and the language-specific analysis is necessary. This suggests that, the recommended seasonal activities for another language group can be used as references in future research.

Ours (Japanese) vs Ours+*TF-IDF*: According to the example in Table 4.6, both of the algorithms can successfully detect the two peak seasons, but Ours+*TD*IDF* does not take the year deviances which cannot correctly extract the second seasonal activity. Using the example in Table 4.6, the noun “Illumination” is recently most weighted in November to December in the year 2016–2017, while the noun of “Sakura” is weighted the most in November–December in the year 2014–2017.

Thus, the Ours+*TF-IDF* take the "Sakaru" as the most important noun. On the other hand, Ours takes the year deviances into account, and hence the recently appeared noun “Illumination” can be successfully extracted. This may be the reason why our algorithm outperforms Ours+*TF-IDF*.

Ours vs [33]: As the season interval is fixed in [33], it has a possibility that the recommended seasonal activity cannot be available in the entire interval. For example, in Table 4.6, Ref. [33] extracts “Sakura” in spring, but the “Sakura” only lasts from March to April but not to March, and hence this is not considered to be valid. Moreover, Ref. [33] fails to detect the second peak season (November–December).

Ours vs [5]: As Ref. [5] just simply extracts the most appeared noun as the recommendation for all months, it should not be applied to the landmark travel distributions with seasonality. For example, in Table 4.6, “Sakura” is not available for all months. Also, it fails to detect the second peak season (November–December).

Table 4.6: An example of the most weighted noun extraction for Meguro-gawa in Tokyo

Algorithm	Meguro-gawa season-activity description			
Ours (Chinese)	Sakura (Apr)	○[2]	Illumination (Nov–Dec)	○
Ours (English)	Sakura (Apr)	○	—	-
Ours (Japanese)	Sakura (Apr)	○	Illumination (Nov–Dec)	○
Ours+*TF-IDF*[1]	Sakura (Apr)	○	Sakura (Nov–Dec)	× (Condition 2 fails)
[33][1]	Sakura (March–May)	×[3] (Condition 2 fails)	—	-
[5][1]	Sakura (all months)	× (Condition 2 fails)	—	-

[1] Used comments in Japanese.

[2] Seasonal activity recommendation is valid.

[3] Seasonal activity recommendation is invalid.

Table 4.7: Statistics of error activity recommendation in three cities.

Error type	Number	Ratio
Type-related	46	47.9%
Location-related	40	41.7%
Others	10	10.4%

Generally speaking, our proposed algorithm is effective in detecting peak-off seasons and recommended seasonal landmark activities.

4.6.3 Error analysis

Based on the results on Table 4.5, all of our experiments do not exceed 67% precision level. To further understand what types of errors the proposed algorithm makes, we check each error and it is found that extractions of non-activity related nouns lead to most of the errors. Table 4.7 lists three categories of errors. It is found that 89.6% errors belong to the *Type*-related and *Location*-related errors. Thus, we focus on analyzing these two categories of errors. Two examples of landmarks are shown in Table 4.8 and Table 4.9. For easy understanding, example comments in Chinese and Japanese are translated into English. Discussions on these two error categories are as follows:

1. *Type*-related: Taking the example of "Yoyogi Park" in Table 4.8, unlike "Meguro-gawa" in Table 4.6 which has a very attractive activity "Sakura", a park itself does not have attractive activities to write or record (activity such as "walking" obviously seems not attractive enough). Thus, users prefer to just describe how they feel about the landmark itself, such as "Beautiful park for walking [...]", rather than to write what they have done in the comments.

   So far, we have not considered the writing style differences among different types of landmarks. How to identify the writing styles in the different type of landmarks should be investigated in the future.

2. *Location*-related: For this error category, we assume that description of *location* is considered more important than *activity*. It cannot be ignored that users often declare the location in textual data. Taking the example of "Yu Yuan" (Garden) in Table 4.9, for foreign users in English and Japanese groups, they prefer to use "China" (country) to declare the location. Oppositely, for domestic users in Chinese group, "China" (country) is not frequently used but they use the more detailed noun, "Shanghai" (city), to declare the location.

Table 4.8: Examples of *Type*-related error activity recommendation.

Example landmark: Yoyogi Park (Park)				
Case	City	Language	Error description	Example Comment
1	TYO	Chinese	“Park”	It is a beautiful park! [...]
2	TYO	English	“Park”	A quiet park [...]
3	TYO	Japanese	“Park”	Beautiful park for walking [...]

Table 4.9: Examples of *Location*-related error activity recommendation

Example landmark: Yu Yuan (Garden)				
Case	City	Language	Error description	Example Comment
1	SHA	Chinese	“Shanghai”	A very famous spot in Shanghai [...]
2	SHA	English	“China”	Beautiful ancient China architecture [...]
3	SHA	Japanese	“China”	China garden [...]

4.7 Conclusion

We have proposed an algorithm that can effectively predict language-specific seasonal activity for a landmark, with the experiments through 418,788 user comments from TripAdvisor. Our proposed algorithm outperforms the other previous studies in terms of activity recommendation precision.

Moreover, based on additional error analysis, it suggests that future work will concentrate on analyzing the writing styles in the different type of landmarks and improving the quality of user comment data.

As the limited time, we use the user comments on TripAdvisor in this chapter. Therefore, we will carry out further experiments using the user comments in other two travel websites including a Chinese travel website and a Japanese travel website.

Chapter 5

A Safe and Comprehensive Route Finding Algorithm for Pedestrians Based on Lighting and Landmark Conditions[1]

5.1 Introduction

Walking as one of the eco and active transport modes contributes to reducing negative environmental influences, relieving traffic congestion and benefiting personal health [74]. Safe and comprehensive walking environment can increase the utility of walking [23]. Under a dark environment, especially in the nighttime, pedestrians feel anxiety and face a potential risk of crimes and violence [18]. Moreover being in a complicated and unfamiliar environment, pedestrians are easy to get lost [68].

We can have two main issues to consider in the walking environments for pedestrians:

1. **Route safety:** Potential dangers due to the lack of visual accessibility under the dark environments causes the route unsafety.

2. **Route comprehensiveness:** Poor cognition of surrounding environment due to the inferior ability in route learning makes the route finding more difficult.

To solve the first problem, the common way is to enhance the lighting condition in public exterior pedestrian areas [24]. In [53], Pain et al. show that improving road lighting condition could effectively help people reduce the fear of crime, because good lighting condition can facilitate clear obstacle detection for uneven patches, visual orientation

[1]Technical contents in this chapter have been presented in the publications ⟨1⟩ and ⟨8⟩.

to observe surrounding road and buildings, facial recognition to promote a sense of security and general comfort for pedestrians. Also, in [10], Bullough et al. point out that improving observation and recognition of large objects referred to by *landmarks*, is very effective in reducing people's fear of crimes. In summary, (1) lighting conditions and (2) landmark visibility are quite significant for increasing pedestrian's feelings of safety. Especially, their influence may differ in daytimes and nighttimes and we have to take it into consideration.

In terms of the second issue concerning the ability of route learning, Spiers et al. consider the route learning as an internal wayfinding behavior [65]. Sigels et al. points out the three phases in route-learning process: [64]: (1) the first phase requires the ability to recognize visualizations of landmarks; (2) the second phase requires the ability to recall orders of the turns and landmarks associated with a route; (3) the last phase requires the ability to reconstruct the route based on phases (1) and (2).

Based on this three-phase strategy, many researches demonstrate that proper selection and representation of landmarks and turning points are effective during wayfinding, especially for a new environment [19]. To sum up, (2) landmark visibility, (3) how much every landmark contributes to the route finding, i.e., landmark effectiveness, and (4) turning counts along a route are quite important in comprehensive route findings. Clearly, (2) landmark visibility is dependent on lighting, especially in the dark environment, i.e., (1) lighting conditions are also important to comprehensive route findings because we are not able to see well the route if it is unlighted in the nighttimes.

Moreover, (5) road widths have significant impacts both on the safety and comprehensiveness of a route. Wide sidewalk roads offer pedestrians enough spaces to walk at their chosen pace, stand, or merely observe their surroundings. Wider sidewalks also offer more spaces for landscaping and amenities, making the environment more attractive and also acting as a buffer between traffic and pedestrians. Furthermore, in [79], they describe that there are continuing interactions between pedestrians' walking behaviors and physical environments, which implies that evaluations should be conducted not only through objective criteria but also pedestrians' viewpoint-dependent judgments.

Based on the discussions above, we propose a safe and comprehensive route finding algorithm for pedestrians based on the lighting and landmark conditions in this chapter. Safety and comprehensiveness can be predicted by the following five indicators (1) lighting conditions, (2) landmark visibility, (3) landmark effectiveness, (4) turning counts along a route, and (5) road widths. Note that these indicators affect walking environments differently during the daytimes and nighttimes. With those indicators, we propose a safe and comprehensive algorithm. In particular, we design daytime score and nighttime score differently and find out an appropriate route depending on the time periods. Experimental simulation results demonstrate that the proposed algorithm obtains higher scores compared to several existing algorithms. We demonstrate that our proposed algorithm generates safe and comprehensive routes for pedestrians in the real environments through pedestrians' viewpoints.

Our main contributions in this chapter are summarized as:

1. We propose five indicators to affect safe and comprehensive route findings for pedestrians under different time periods.

2. We also evaluate the impacts on the five indicators through questionnaires and clarify how they affect safety and comprehensiveness.

3. We design the *score* to reflect the indicators (1), (2), (3) and (5). Then with a turning count reduction strategy based on the indicator (4), we propose a safe and comprehensive route finding algorithm. In particular, we design daytime score and nighttime score differently and find out an appropriate route depending on the time periods.

4. Experiments through pedestrians' viewpoints based on several real outdoor environments confirm the effectiveness and efficiency of the proposed algorithm.

This chapter belongs to *Sequential Travel Route Generation* in Fig. 1.1.

5.2 Related Works

There are several previous works with relevance to safe and comprehensive route findings for pedestrians. In [44, 50, 82], they consider safety on block crossings or intersections. On the other hand, in [7,19,49], they use landmarks to improve pedestrians' route learning abilities by placing landmarks at junctions or turnings in order to help pedestrians recall paths better. However, pedestrians require safety and comprehensiveness not only at intersections but along their entire routes. Some studies have evaluated safety along the entire route by indicators such as accident risks, lighting conditions, and road lanes. In [4,47,76,80], although these studies consider some safe indicators, there is a possibility that the route obtained are complicated due to the lack of eye-catching landmarks and the difference between the daytimes and nighttimes are unclear and not confirmed. Therefore the routes suggested by these systems are not always optimal for pedestrian navigation.

5.3 Five Indicators for Safe and Comprehensive Pedestrians Route Finding

As we mentioned in the previous section, the safety and comprehensiveness of pedestrian routes must be dependent on the five indicators below:

(1) lighting conditions,

(2) landmark visibility,

(3) landmark effectiveness,

(4) turning counts along a route and

(5) road widths.

For lighting conditions, landmark visibility, those data can not be directly obtained from ZENRIN's map database. For road widths, roads in ZENRIN's map databased are labeled as level 1,2,3, etc. which represents different type of roads.

In order to investigate how much these indicators affect pedestrians' safety and comprehensiveness, we have conducted the preliminary questionnaires to 40 testers composed of 20 males and 20 females with ages ranging from 20 to 52. We have used a five-point scale that one (1) shows "strongly disagreement (not important)" and five (5) shows "strongly agreement (very important)."

Based on these questionnaires feedbacks, we assign appropriate scores to every road during pedestrian route finding.

5.3.1 (1) Lighting Condition

Lighting conditions must contribute to both pedestrians' feelings of safety and comprehensiveness during route findings. Since lighting conditions must differ under different time circumstances, we have to evaluate them both in the daytimes and nighttimes. For example, consider an unlighted road intersection in a rural area. We can see the intersection in the daytime but we may not notice it well in the nighttime due to the poor lighting conditions.

To confirm the impacts of lighting conditions on route safety and comprehensiveness between daytimes and nighttimes, 40 pictures of roads with various lighting conditions under 15 p.m. (daytime) and 19 p.m. (nighttime) were presented to testers (see Fig. 5.1). Then the testers filled out questionnaires about their feelings on two time periods.

The questionnaire results are summarized in Tables 5.1 and 5.2, where Mean and SD show the mean points and the standard deviation over 40 testers. "Lighting" has more than four points in both tables, which indicates that it is highly helpful in safe and comprehensive route findings in both time periods. Pedestrian's unsafe feeling and nervousness can be relieved as good lighting condition provides better views and helps pedestrians find their locations more easily and accurately.

Table 5.1: The impact of road lighting conditions on the road safety.

Factor \ Time	Daytime		Nighttime	
	Mean	SD	Mean	SD
Lighting	4.25	0.73	4.53	0.50
Unlighting	2.58	0.70	1.95	0.80

Table 5.2: The impact of road lighting conditions on the road comprehensiveness.

Factor \ Time	Daytime		Nighttime	
	Mean	SD	Mean	SD
Lighting	4.35	0.69	4.55	0.55
Unlighting	2.93	0.57	1.88	0.64

Moreover, as higher scores were assigned to the nighttimes both in safety and comprehensiveness compared to that in the daytimes, it confirms our assumption that the time-of-day will affect people's perceptions.

In summary, we have to design the road score taking into account lighting conditions as well as time periods, daytimes or nighttimes.

5.3.2 (2) Landmark Visibility

Landmarks are objects on land that is easy to see and recognize and must be main lighting sources. For example, landmarks such as train stations, restaurants, and convenience stores have bright metal or LED light board outside which is easy and clear to observe. If a road has many lighting landmarks nearby, it must have a good lighting condition even in the nighttimes and contribute to both route safety and comprehensiveness according to the discussions in Section 5.3.1.

Landmark lighting can be evaluated by to what extent we can see the landmark. Now we evaluate the visibilities of every landmark type through the questionnaires by testers. Especially we evaluate the discrepancies of landmark visibilities between daytimes and nighttimes. Testers have completed the questionnaires on how much every landmark type affects route findings in the daytimes and nighttimes.

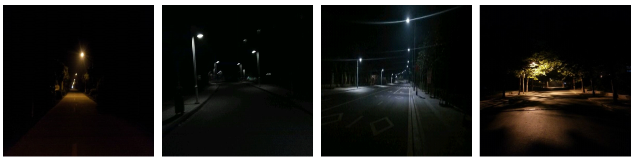
Table 5.3 summarizes the landmark types in alphabetical order used in the questionnaire.[2] In Table 5.3, the landmark types of which average points are larger than three in

[2]The landmark types listed in Table 5.3 are also used in the experiments in Sections 5.5 and 5.6. Note that, we can use another landmark type, such as shrines and hotels/inns in Kyoto area, if needed. In this

(a) Lighted condition in the daytime.

(b) Unlighted condition in the daytime.

(c) Lighted condition in the nighttime.



(d) Unlighted condition in the nighttime.

Figure 5.1: Example pictures used for visual confirmations on lighting conditions in the daytime and nighttime.

the daytimes are bank, bus station, convenience store, gas station, hospital, library, park, post office, shopping mall, sport stadium, temple, and train station. Most of them are easy to be observed due to their large sizes or special shapes. For a convenience store, it is common to see in the daily life and thus it must be a very useful guidance although it is not so large. The landmark types of which average points are larger than three points in the nighttimes are the convenience store, gas station, hospital, library, restaurant,

sense, Table 5.3 shows an example set of landmark types.

Table 5.3: The impact of landmark types on the visibility.

Landmark type t	Daytime		Nighttime	
	Mean	SD	Mean	SD
Bank	3.60	0.88	2.25	0.99
Bus station	4.00	0.53	1.50	0.49
Convenience store	4.00	0.83	4.20	0.88
Gas station	3.90	0.70	3.50	0.70
Hospital	3.70	0.73	3.13	0.64
Library	4.50	0.45	3.10	0.76
Park	3.70	1.16	1.70	1.05
Post office	4.00	0.83	2.90	0.64
Public infrastructure	2.30	1.73	1.25	1.24
Restaurant	2.70	0.73	4.10	0.53
School	1.80	0.99	0.70	0.49
Shopping mall	3.60	1.03	4.10	0.70
Sport stadium	4.00	0.76	3.80	0.76
Temple	3.30	0.49	1.70	0.83
Train station	5.00	0.00	4.90	0.35

shopping mall, sport stadium, and train station. All of them have shiny signboards which make them easy to catch even at nighttimes.

The average scores in Table 5.3, i.e., "Mean" values in Table 5.3, directly give the landmark visibilities of a type t in the daytimes and the nighttimes, which are defined by $V(t, day)$ and $V(t, night)$, respectively. For example, $V(t, day) = 1.80$ and $V(t, night) = 0.70$ if t = School from Table 5.3. The pedestrian route with many visible landmarks has good lighting conditions and thus it must lead to pedestrians' safe and comprehensive feelings.

5.3.3 (3) Landmark Effectiveness

Landmarks can also be helpful for pedestrians locate themselves and find out correct routes. If pedestrians can observe the appropriate landmarks along their routes, they can always select the right way quickly and correctly by confirming these landmarks without anxieties. However, their features differ from each other and thus we design a weight $W(t)$ for every landmark type t.

As pedestrians have a tendency to assign a high value to a landmark that is *scarce* such as a train station or a large hospital since it gives a feeling of uniqueness or distinctiveness. Then we can consider $W(t)$ to weigh each landmark type t as follows:

Let $L(t)$ be a set of landmarks in a given area of which type is t. For example, if there are two parks, Park_1 and Park_2, and three restaurants, Restaurant_1, Restaurant_2, and Restaurant_3, in a given area, then $L(\text{Park}) = \{\text{Park}_1, \text{Park}_2\}$ and $L(\text{Restaurant}) = \{\text{Restaurant}_1, \text{Restaurant}_2, \text{Restaurant}_3\}$ in this area. First we compute the frequency, $f(t)$, for a landmark type t defined by:

$$f(t) = \frac{|L(t)|}{A} \tag{5.1}$$

where $|L(t)|$ is the number of landmarks of the type t in the given area. For the example above, $|L(Park)| = 2$ and $|L(Restaurant)| = 3$ in this area. A is the size of the given area with the unit of km^2. We assign more weights to landmarks with relatively less quantity. The weight of a landmark type t is given by:

$$W(t) = \frac{1}{f(t)} \tag{5.2}$$

Table 5.4 shows an example of frequency and weight calculation results, which lists up 15 types of landmarks in Takadanobaba area with the size of 1km^2.

5.3.4 (4) Turning Counts

We assume that the typical distance of a pedestrian's walking route is around 1km which is justified for a dominant trip or neighborhood walking [46] and ask testers how many times they can have tolerances on turning counts from their start to goal. Table 5.5 summarizes the results.

In either the day or night, the tolerance of turning counts are around four. As too many turnings may cause pedestrians to pass the right way or understand when to turn, the results suggest that turning counts in both the daytimes and nighttimes should not be over four.

5.3.5 (5) Road Widths

We classify a set of roads into two types: the roads with the width of more than 4m are the *main roads*, which are recorded as main general roads in the map database; oppositely, the narrow roads with the width of up to 4m are the *branch roads*, which are recorded as narrow roads in the map database.[3] Note that highways are excluded in our pedestrians

[3] We use a commercial pedestrian road database given by ZENRIN [84], where the roads with the width of more than 4m are classified into main roads and those with the width of up to 4m are classified into branch roads.

Table 5.4: The frequency and weight of landmarks.

Landmark type t	Frequency $f(t)$	Weight $W(t)$
Bank	1.33	0.75
Bus station	10.2	0.10
Convenience store	9.78	0.10
Gas station	0.89	1.13
Hospital	5.33	0.19
Library	2.22	0.45
Park	0.89	1.13
Post office	1.33	0.75
Public infrastructure	1.33	0.75
Restaurant	4.00	0.25
School	8.44	0.12
Shopping mall	0.44	2.25
Sport stadium	1.33	0.75
Temple	2.67	0.38
Train station	0.44	2.25

Table 5.5: The tolerance of turning counts on the route finding.

Time	Mean	SD
Daytime	4.10	0.79
Nighttime	3.70	0.96

network. We evaluate the impacts of road widths on road safety and comprehensiveness through the questionnaires by testers.

The results obtained from tester's feedbacks are summarized in Table. 5.6. Since the main roads have a higher score than the branch roads, it suggests that road widths affect pedestrians' feelings both on safety and comprehensiveness. This is because a wider road offers clearer sight views whereas a narrow road often come up with crimes, assault, and intimidating, and it is easy to have pedestrians' nervous and get lost. In the nighttimes, both the main roads and the branch roads have lower points. Especially, the branch roads in the nighttimes have the lowest point.

Table 5.6: The impact of road widths on the road safety and comprehensiveness.

Factor \ Time	Daytime		Nighttime	
	Mean	SD	Mean	SD
Main Road	4.30	0.44	3.90	0.51
Branch Road	3.90	0.79	3.00	0.74

5.4 Safe and Comprehensive Pedestrians Route Finding Algorithm

The evaluations in Section 5.3 clearly demonstrate the impacts of the five indicators on pedestrians' feelings on safety and comprehensiveness on the surrounding environments. Based on these evaluations, we first design the *score* to effectively represent the indicators (1), (2), (3) and (5) discussed in the previous section. Then with considering turning count reduction based on the indicator (4), we propose a safe and comprehensive route finding algorithm for pedestrians based on lighting and landmark conditions.

Based on the evaluations in Section 5.3, we first determine how many landmarks may be observed from every road or edge as lighting sources and guidance while walking along the road (Section 3.1). Then we give an edge score taking into account the five indicators to evaluate time-depending route safety and comprehensiveness scores (Section 3.2). After that, we generate a pedestrian route which can reduce the turning counts (Section 3.3). Section 3.4 summarizes our proposed algorithm.

5.4.1 Problem Definition

A pedestrian network is given by a graph $G = (N, E)$, where a node $n_i \in N$ represents a road crossing point and an edge $e_{ij} = (n_i, n_j) \in E$ corresponds to a road between nodes n_i and n_j. The road width for every edge $e \in E$ and a set L of landmarks with their landmark type t are also given beforehand. Let LT be a set of landmark types. Note that, since we use the 15 landmark types as listed in Table 5.4 in our experimental areas in Sections 5.5 and 5.6, LT is composed of these 15 landmark types in our experiments, i.e., $LT = \{\text{Bank}, \text{Bus station}, \ldots, \text{Train station}\}$. Time periods (daytime and nighttime) are also given as inputs. Then the *route finding problem* here is to find a safe and comprehensive route from a starting node $s \in N$ to a destination node $z \in N$. Fig. 5.2 shows an example of a pedestrian network associated with road widths and landmarks. In this figure, the blue number shows the node index i of each node $n_i \in N$ in the pedestrian network $G = (N, E)$.

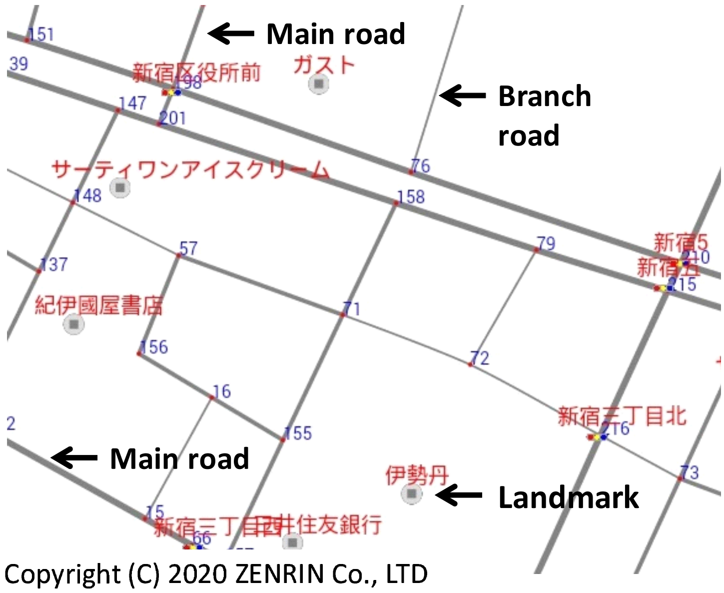


Figure 5.2: An example of a pedestrian network in Shinjuku.

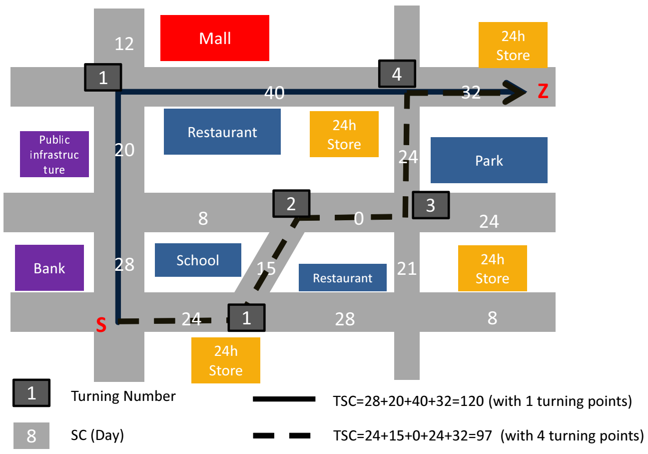
5.4.2 Edge Landmark

We first define an *edge landmark*. Let $e \in E$ be an edge or pedestrian road in a pedestrian network of $G = (N, E)$. Then for every landmark $l \in L$, we compute the shortest distance $d(l, e) = \left[(x_p - x_l)^2 + (y_p - y_l)^2)\right]^{1/2}$ between the landmark l and the middle point of e, where (x_p, y_p) are coordinates of the middle point p of the edge e and (x_l, y_l) are coordinates of the landmark l. [4]

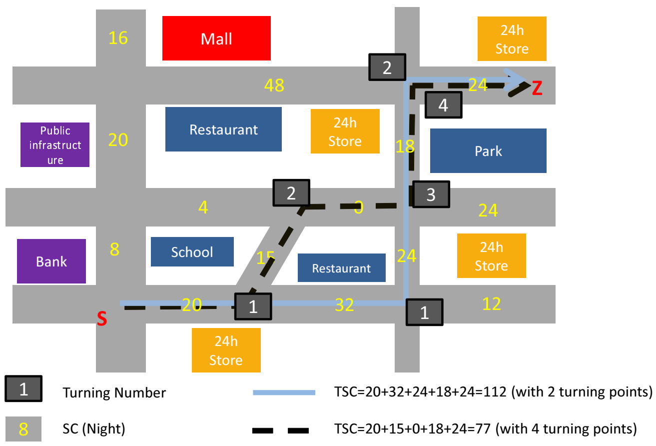
To judge whether the landmark l is close enough to the road for observation, we set K_{th} to be the threshold value. In Japan, the main road (general road) width is usually 6m–8m and the branch road width is under 4m. In this chapter, we set K_{th} to be the half of road width, i.e., 4m for the main roads and 2m for the branch road, to ensure pedestrians can see the landmarks within these distances [87]. Then, if $d(l, e) \leq K_{th}$, then the landmark l belongs to the edge e and this landmark is called *an edge landmark*. After extraction of all edge landmarks, a set of edge landmarks associated with the edge e is denoted by $L(e)$.

Let $L_t(e) \subseteq L(e)$ be a set of edge landmarks with landmark type $t \in LT$ and then $L(e)$ can be written by $L(e) = \bigcup_{t \in LT} L_t(e)$.

[4]The coordinate (x_l, y_l) of the landmark l corresponds to the representative point registered on the commercial pedestrian road database that we use.



(a) SC score in the daytime.



(b) SC score in the nighttime.

Figure 5.3: Safe and comprehensive routes in the daytime and nighttime.

5.4.3 Edge Score Design

Edge landmarks definitely contribute to lighting conditions and comprehensiveness of the routes. If an edge has many edge landmarks with higher $V(t, day)$ or $V(t, night)$ scores as well as $W(t)$, this edge must provide safe feelings and comprehensiveness of route learning.

Then we can first design the score on landmark effectiveness LE as follows: Let $LE(e, day)$ and $LE(e, night)$ be scores on landmark effectiveness on every edge $e \in E$

in the daytimes and nighttimes. They can be calculated by:

$$LE(e, day) = \sum_{t \in LT} |L_t(e)| \times V(t, day) \times W(t) \tag{5.3}$$

$$LE(e, night) = \sum_{t \in LT} |L_t(e)| \times V(t, night) \times W(t). \tag{5.4}$$

where $V(t, day)$ and $V(t, night)$ show the visibilities of the landmark type t during the daytimes and nighttimes defined by Section 2.2 and $W(t)$ shows the weight of a landmark type t defined by Section 2.3.

The LE score represents the landmark effectiveness of the edge e, which means how much edge landmarks impacts on route findings in the daytimes and nighttimes. As the LE score is defined by the sum of scores of edge landmarks, thus it is obvious to know that an edge e has a higher $LE(e, day)$ or $LE(e, night)$ value if e has more edge landmarks.

In order to further take into account the road width factor, safety and comprehensiveness degrees of an edge $e \in E$ in the daytimes and nighttimes are calculated by:

$$SC(e, day) = LE(e, day) \times Wid(e, day) \tag{5.5}$$

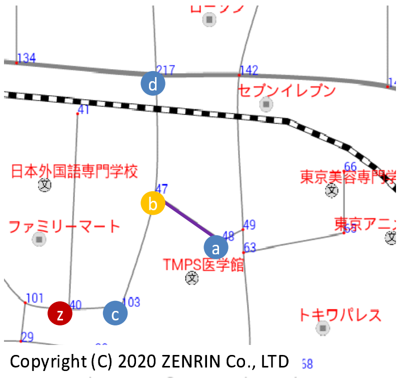
$$SC(e, night) = LE(e, night) \times Wid(e, night) \tag{5.6}$$

where $Wid(e, day)$ and $Wid(e, night)$ show road width factor in the daytimes and nighttimes, respectively, and we set these values as listed in Table 5.6, i.e., Wid for a main road e is given by $Wid(e, day) = 4.30$ and $Wid(e, night) = 3.90$, Wid for a branch road e is $Wid(e, day) = 3.90$ and $Wid(e, night) = 3.00$.

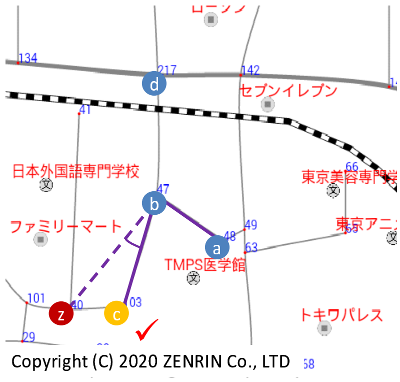
Finally, given a threshold value SC_{th}, we mainly consider a set of edges satisfying $SC(e, day) \geq SC_{th}$ or $SC(e, night) \geq SC_{th}$ in route findings in Section 5.4.4.

By introducing SC_{th} above, the algorithm can mainly search a set of roads with high priority whose SC score is equal to or higher than SC_{th}. We expect that the final route is composed of safe and comprehensive roads and thus the entire route also must be safe and comprehensive.

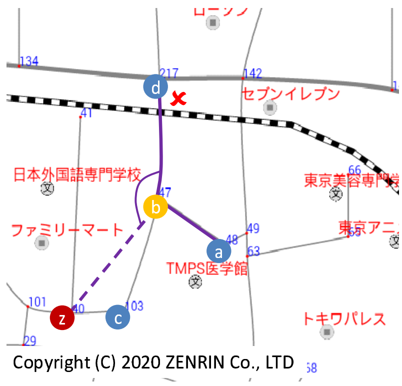
Our preliminary experiments in various areas demonstrated that the average product sums of landmark visibilities and weights given by Eqs. (5.3) and (5.4) in daytimes and nighttimes were around 2.44 and 2.31, respectively. The average road width factors $Wid(e, day)$ and $Wid(e, night)$ were 4.10 and 3.45, respectively. Hence we set $SC_{th} = 2.42 \times 4.10 \approx 10.00$ in daytimes and $SC_{th} = 2.31 \times 3.45 \approx 8.00$ in nighttimes, which require around one bright landmark. These edges must contribute to both safety and comprehensiveness by taking into account lighting conditions, landmark visibilities and road widths.



(a) The sub-route (a, b) is searched. We assume that all the sub-roads have high SC scores.



(b) Since $\angle zbc \le \pi/3$, the node c is selected as a next node.



(c) Since $\angle zbd > \pi/3$, the node d is not selected as a next node.

Figure 5.4: Turing reduction procedure.

5.4.4 Turning Reduction for Route Planning

Reduction of turning counts can ensure the route going towards the destination as accurate and fast as possible [65]. We realize it by the following strategy:

Assume that the edge (a, b) is just searched as a sub-route (see Fig. 5.4(a)), i.e., the current edge is (a, b) and the current node is b. The next nodes will be c and d and assume that the edges (b, c) and (b, d) have SC scores larger than or equal to SC_{th}. Let z be the destination node.

In this case, since $\angle zbc \le \pi/3$ (see Fig. 5.4(b)), then we select the edge (b, c) as a next sub-route. Likewise, as $\angle zbd > \pi/3$ (see Fig. 5.4(c)), the edge (b, d) deviates from the direction towards the destination node z and will not be selected as a next sub-route.

We set the threshold to be $\pi/3$ here since the field of view of single normal human eye covers around 60 degrees horizontally [28] and pedestrians feel that they go to the opposite side against the destination if the threshold is greater than $\pi/3$ according to our preliminary experiments.

Let i be a current node and z be the destination node. As in this example, among the non-visited neighbor nodes with the score of SC_{th} or lager, we select the next node j with the highest score satisfying $\angle zij \leq \pi/3$. If there are no such neighbor nodes, we just select the non-visited neighbor node with the highest SC score among the neighbors. Note that, if we have no non-visited neighbor nodes here, we backtrack to the previous node and perform the neighbor node search again.

Our strategy above does not directly reduce the turning counts for route searching and thus it may not always reduce the turning counts of a generated route. However, we search a next sub-route within the threshold angle against the destination node as depicted in Fig. 5.4 and thus the sub-route selected here must have no significant detours against the destination node. This means that the sub-route directly goes toward the destination node and we resultantly expect that the turning counts can be reduced indirectly. In fact, the experiments in Section 5.5 confirm that our strategy above reduces the turning counts compared to the existing ones.

Note that we will discuss the effectiveness of the threshold $\pi/3$ above in Section 5.6.

5.4.5 The Algorithm

Algorithm 1 summarizes our proposed route finding algorithm. After extracting edge landmarks and setting SC scores for all the edges in a given pedestrian network, we start the route finding from the start node s. The neighbor node search is done according to the discussion in Section 5.4.4. We repeat this process until we reach the destination node z. The obtained route can go towards the destination directly and hence the turning counts can be reduced. The sum of the SC scores along the obtained route can be large enough. Overall we believe that our algorithm gives a safe and comprehensive walking route.

Example 1. *Fig. 5.3 portrays an example of our route finding algorithm.*

In Fig. 5.3(a), the SC score is given to each road in the daytime. If we search the shortest route from the start node s to the destination node z, we can have a dotted route. The number of turnings is four and the sum of all the SC scores along this route becomes 97. If we use our algorithm, we can find out the bold route. The number of turnings decreased to one and the sum of all the SC scores along this route increases to 116.

Algorithm 1: Safe and comprehensive route finding algorithm.

Input : Pedestrian network $G = (N, E)$, road widths, a set L of landmarks, start node s, destination node z, and time periods (daytime or nighttime)

Output : Safe and comprehensive route in G

begin
- Extract edge landmarks for each edge $e \in E$ from L according to Section 5.4.2;
- Calculate SC score for each edge $e \in E$ according to Section 5.4.3;
- the edge cost $c(e)$ for each edge $e \in E$;
- Let i to be a current node; Initialize $i \leftarrow s$;
- **repeat**
  - Find a neighbor node of i according to Section 5.4.4 and update the current node i.
- **until** *the current node i becomes* z;

end

Since the route obtained by our algorithm has many visible landmarks and uses the wide roads, it must be comprehensive to pedestrians.

In Fig. 5.3(b), the SC score is given to every road in the nighttime. In the same way, the dotted route shows the shortest route whereas our algorithm finds out the bold route. The route obtained by our algorithm passes through many lighting roads, which contributes to safe and comprehensive route to pedestrians, even in the nighttime.

5.5 Simulation Evaluation

Our proposed route finding algorithm is implemented in Java running on Nexus 7 with Android 6.0 and compared it to three conventional route search algorithms including very recent ones [4, 49, 80].

5.5.1 Setup

We select three typical test fields in Japan including the urban areas of Shinjuku and Takadanobaba and the rural area of Nishitokyo. Urban areas usually have more landmarks than rural areas. The total number of landmarks in the test fields of Shinjuku, Takadanobaba, and Nishitokyo are 472, 237, and 137, respectively.

We assume just one walking distance and thus we set the distance of each route to be around 1km from the starting node to the destination node. For every test field, we perform random 50 trails with different starting nodes and destination nodes and compare

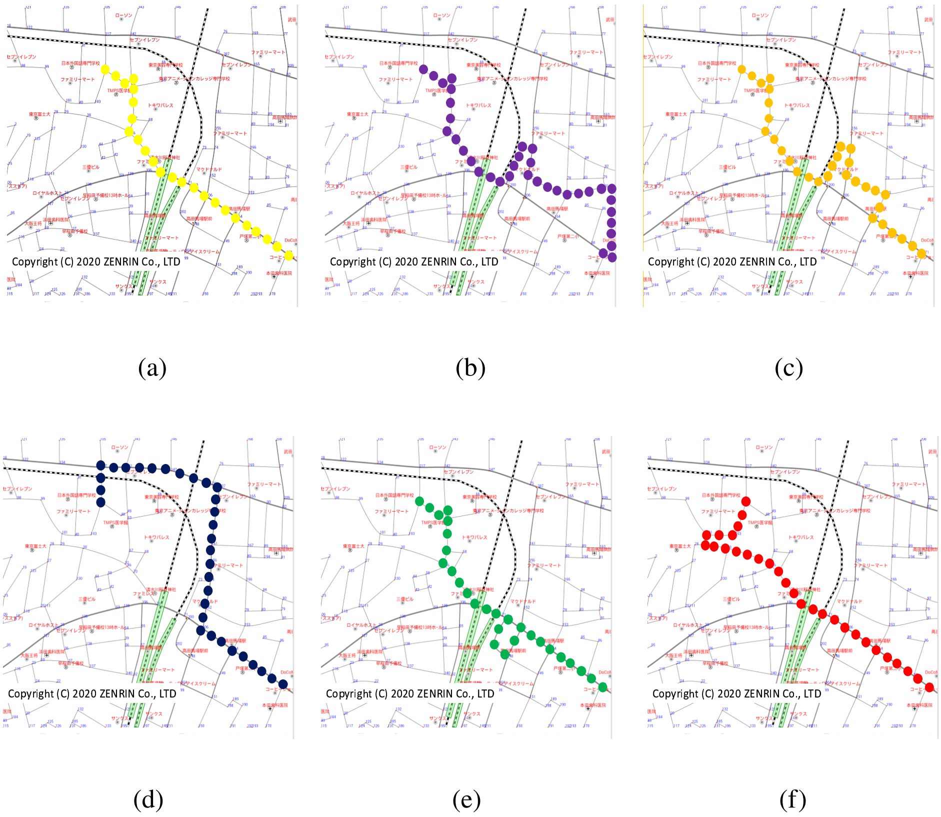


Figure 5.5: Generated routes in Takadanobaba (urban area). (a) Shortest route. (b) Route generated by [80]. (c) Route generated by [4]. (d) Route generated by our daymode algorithm. (e) Route generated by [49]. (f) Route generated by our nightmode algorithm.

the following cases.

Case 1 (Daytime): The four daytime route finding algorithms are compared:

1. **Shortest route:** We obtain the shortest route from a starting node to a destination node.
2. **The algorithm in [80]:** We obtain a pedestrian route based on the risk (accident rate).
3. **The algorithm in [4]:** We obtain a pedestrian route based on various safety index for urban areas.
4. **Ours (daytime):** Assuming the daytime, we obtain a pedestrian route using our proposed algorithm.

Case 2 (Nighttime): The two nighttime route finding algorithms are compared:

1. **The algorithm in [49]:** Assuming the nighttime, we obtain a pedestrian route based on the landmark visibility.
2. **Ours (nighttime):** Assuming the nighttime, we obtain a pedestrian route using our proposed algorithm.

5.5.2 Comparison results

To illustrate the advantage of our proposed algorithm, three main evaluation criteria are used to analyze. First, the route distance is the physical distance between the starting node and the destination node for each route. Then TSC is the sum of all SC scores along the obtained routes. Lastly, TC is the turning counts along with an entire route.

The results are shown in Tables 5.7–5.9, where the results obtained by the shortest route algorithm is normalized to 1.00. Fig. 5.5 shows the route examples in the Takadanobaba area. In these figures, the yellow line represents the outcome of the shortest route. The purple line represents the outcome of [80]. The orange line shows the outcome of [4]. The green line represents the outcome of [49]. The blue and red lines show the outcomes of our daytime result and nighttime result, respectively.

According to Table 5.7, the distances obtained by [4, 49, 80] and our algorithm both in the daytimes and nighttimes are longer than the shortest routes in every test field. This demonstrates that by taking safety and comprehensiveness factors into consideration, the walking distances will be increased definitely.

As listed in Table 5.8, all of the scores obtained by our algorithm are higher than those obtained by the shortest route algorithm and the other three conventional route finding algorithms. It indicates that our algorithm must be a positive impact on improving the safe and comprehensive environment. Note that, as there are many landmarks in Shinjuku and Takadanobaba areas, the differences in TSC values between daytimes and nighttimes are not significant. On the other hand, in Nishitokyo area, where not so many landmarks exist compared to Shinjuku or Takadanobaba, the TSC of values Nishitokyo area become lower and the differences are relatively larger.

From Table 5.9, all of the turning counts obtained by the other three conventional route finding algorithm are higher than those obtained by our algorithm. The results suggest that our turning reduction algorithm in Section 5.4.4 performs better than the route finding algorithms without considering the turning issue. That is to say, our algorithm can help pedestrians have the comprehensive understanding of the route.

Table 5.7: Normalized distance evaluation results of generated routes in three test fields.

Mode	Algorithm	Shinjuku	Takadanobaba	Nishitokyo
Daytime	Shortest route	1.00	1.00	1.00
	[80]	1.55	1.57	1.31
	[4]	1.30	1.44	1.18
	Ours (daytime)	1.21	1.25	1.15
Nighttime	[49]	1.22	1.22	1.08
	Ours (nighttime)	1.22	1.24	1.16

Table 5.8: Normalized TSC evaluation results of generated routes in three test fields.

Mode	Algorithm	Shinjuku	Takadanobaba	Nishitokyo
Daytime	Shortest route	1.00	1.00	1.00
	[80]	1.34	1.24	1.06
	[4]	1.10	1.12	1.02
	Ours (daytime)	1.44	1.39	1.26
Nighttime	[49]	1.35	1.24	1.10
	Ours (nighttime)	1.46	1.37	1.16

5.5.3 Discussion on turning count threshold

In order to further explain the effectiveness of the threshold $\pi/3$ introduced in Section 5.4.4, we have conducted the additional experiments to observe the performances of various thresholds. In the same way as in Section 5.5.2, we have compared the turning counts generated by each algorithm in 20 routes in the three areas and we set the distance of each route to be around 1km from the starting node to the destination node.

Table 5.10 summarizes the average number of turning counts over 20 routes. As described in Section 5.4.4, the human has an eye range from 0 to $\pi/3$ for each eye [28] and thus the turning threshold should not be larger than $\pi/3$. Hence we test the three thresholds of $\pi/3$, $\pi/4$, and $\pi/6$. "$*$" in Table 5.10 represents the cases that have the same number of turning counts compared with the shortest route and "$**$" represents the cases that have the decreased number of turning counts compared with the shortest route. Note that the results are slightly different from the ones in Table 5.9, since we have randomly reselected the starting node and the destination node in the test areas for the additional experiments.

From the results in Table 5.10, the turning counts of the three conventional algorithms

Table 5.9: Normalized TC evaluation results of generated routes in three test fields.

Mode	Algorithm	Shinjuku	Takadanobaba	Nishitokyo
Daytime	Shortest route	1.00	1.00	1.00
	[80]	1.91	1.52	1.16
	[4]	1.88	1.56	1.13
	Ours (daytime)	0.94	0.97	0.94
Nighttime	[49]	1.75	1.54	1.18
	Ours (nighttime)	0.90	0.93	0.97

Table 5.10: Turning reduction evaluation results of generated routes in three test fields (the threshold θ is changed in our algorithm).

Mode	Algorithm	Shinjuku	Takadanobaba	Nishitokyo
Daytime	Shortest route	2.25	3.15	2.05
	[80]	3.50	5.90	2.30
	[4]	3.40	6.00	2.30
	Ours ($\theta = \pi/3$, daytime)	2.20**	3.10**	2.05*
	Ours ($\theta = \pi/4$, daytime)	2.45	3.20	2.10
	Ours ($\theta = \pi/6$, daytime)	2.15**	3.20	2.05*
Nighttime	[49]	3.45	5.50	2.50
	Ours ($\theta = \pi/3$, nighttime)	2.25*	3.15*	2.05*
	Ours ($\theta = \pi/4$, nighttime)	2.35	3.20	2.15
	Ours ($\theta = \pi/6$, nighttime)	2.20**	3.25	2.05*

[4, 49, 80] are quite increased compared to the shortest route. It suggests that, the more factors are taken into considerations for route-searching, the more complex the route will be and thus it will generate redundant turnings.

For our proposed algorithm, when the threshold θ is set to be $\pi/3$, $\pi/4$, and $\pi/6$, we cannot always guarantee that the turning counts could be decreased in all the cases compared to the shortest route, but the average values of turning counts do not exceed 10% of the shortest routes.

For the differences between the three thresholds, the results indicate that, even though the differences between the turning reduction among those three thresholds is not large, $\theta = \pi/3$ has the most stable performance as it is the only one that the turning counts are not increased in all the three areas both in daytime and nighttime.

Thus we can say that, when the threshold is set to be $\pi/3$, it outperforms the other two thresholds and it is effective in eliminating redundant turning counts compared to the conventional algorithms.

However, turning count reduction may depend on the area selection, and thus how to efficiently and accurately reduce the turning counts in all circumstances must be one of the future works.

5.6 Evaluation by Pedestrians

In this section, we evaluate real outdoor pedestrian routes generated by our algorithm.

5.6.1 Setup

In order to confirm the efficiency and effectiveness of our proposed algorithm, we have conducted real outdoor evaluation experiments in Takadanobaba area using the routes shown in Fig. 5.5. 20 testers participated in the experiment. The testers' mean age was around 25 (10 males and 10 females).

The experiments of the shortest route, [80], [4] and our daytime routes were carried out at 15 pm while [49] and our nighttime routes were carried out at 19 pm. All participants walked themselves followed by an assistant behind recording their walking times. For each tester, he/she walked along all the six routes separately and then evaluated the routes with five criteria below:

1. Distance (D)
2. Unit time cost (UTC)
3. Turning count (TC)
4. Safety environment (SE)
5. Comprehensive environment (CE)

D is the physical distance of a route; UTC represents the walking speed [m/s] [54]; TC shows the total number of turning counts of a route; SE represents how safe the walking environment is with a score from 1 (not safe) ~ 5 (very safe) and CE represents the understanding of the spatial relationship between roads with a score from 1 (very hard to find the way) ~ 5 (very easy to find the way). Once walking was completed, the testers were required to answer the self-report on SE and CE.

Table 5.11: Experimental results in real environments.

Mode	Algorithm	Distance [m]	UTC [m/s]	TC	SE	CE
Daytime	Shortest route	550m	1.14	4	4.00	4.25
	[80]	980m	0.83	10	3.80	3.35
	[4]	740m	0.95	10	3.90	3.00
	Ours (daymode)	730m	1.12	3	4.20	4.40
Nighttime	[49]	660m	1.16	7	4.00	4.10
	Ours (nightmode)	610m	1.22	3	4.30	4.65

5.6.2 Comparison results

Table 5.11 summarizes the results. The routes generated by our algorithm have longer distances compared to the shortest route because our algorithms take lightings and landmarks into considerations.

In the daytime, the UTC value of our algorithm is only 2% slower than that of the shortest route, even though the total distance of the proposed method was around 33% longer. This is because more main roads with multiple visible landmarks can guide the participants find their positions and directions more easily and unnecessary turning points are detruncated in the proposed algorithm. Also, the UTC value of our algorithm is higher than [80] and [4]. The algorithms in [4, 80] only take safety factors into considerations while ignoring the usefulness of landmarks as route guidances. Too many turns confuse and slow down testers during route findings.

In the nighttime, the UTC value of our algorithm is even larger than that of the shortest route and [49]. This relies on the clear and lighting landmark guidances along the roads that not only light up the sight view but also help pedestrians find the direction to the goal more confidently. In conclusion, the results refer that long distances obtained by our algorithm have a smaller negative effect on testers if enough and clear navigation information such as salient landmarks are available.

From the viewpoint of TC, the shortest route, [80], [4] and [49] have more than four turnings. Both the daytime and nighttime routes by our proposed algorithm have only three turnings.

The testers' average scores of SE and CE of our proposed algorithms are more than four points which are higher than those of the shortest route and [49]. This is definitely because our algorithm takes into account the five indicators.

Overall, our proposed algorithm has outperformed the previously proposed algorithms in terms of improving safe and comprehensive walking environments.

Questionnaire about landmark recognition

Name: _________ Gender: ____ Age:__ Date:

Q1 How quickly can you recognize the landmarks?
☐ 1. Can not recognize at all
☐ 2. Can recognize but slowly
☐ 3. Need time to recognize
☐ 4. Can recognize quickly
☐ 5. Can recognize immediately

Q2 Are there discrepancies between the landmarks' real locations and the landmarks' locations on the map?
☐ 1. Quite a lot
☐ 2. Sort of
☐ 3. Maybe sort of
☐ 4. A little
☐ 5. Not at all

Q3 To what extent do the discrepancies influence you on recognizing the landmarks?
☐ 1. Very confusing
☐ 2. Confusing
☐ 3. Sort of confusing
☐ 4. Slightly confusing
☐ 5. Not confusing at all

Thank you for your cooperation very much!

Figure 5.6: Questionnaire sheet on landmark recognition.

5.6.3 Discussion for Landmark Recognition

There is a possibility that we have a difference between the location of a landmark registered on the map and its location visually recognized by the users in the real world. In this section, we confirm to what extent users can recognize the edge landmarks introduced in Section 5.4.2.

We have randomly selected ten edge landmarks generated by our algorithm in Takadanobaba area. Then we have asked 20 testers to what extent they can connect the landmarks in the real world with the landmarks on the map. The testers' mean age was around 25 (10 males and 10 females). The testers were required to fill a questionnaire with Q1–Q3 in Fig. 5.6 and the results are listed in Table 5.12. In Table 5.12, "Q1"–"Q3" show the answers that the a tester fills in. "False recognition" shows the counts that a tester failed to recognize landmarks' locations in the real world. "Successful recognition" shows the counts that a tester successfully recognized landmarks' locations in the real world. ("False recognition" + "Successful recognition") in Table 5.12 should always be 10 since we totally used ten edge landmarks.

From Table 5.12, we obtain an average score of Q1 with 4.55 points. The results of Q1 suggest that testers could recognize the landmarks immediately. From the results of Q2, there exist some coordinate differences between the location of a landmark registered on the map and its location visually recognized in the real world. However, the results

Table 5.12: The results of landmark recognition.

User ID	Q1	Q2	Q3	False recognition	Successful recognition
User1	5	4	4	0	10
User2	5	4	5	1	9
User3	4	3	4	2	8
User4	4	4	4	1	9
User5	5	3	4	0	10
User6	5	4	4	0	10
User7	5	4	4	1	9
User8	5	4	5	1	9
User9	5	4	4	0	10
User10	5	4	5	0	10
User11	4	4	5	0	10
User12	5	5	5	0	10
User13	4	4	4	1	9
User14	4	4	4	0	10
User15	4	4	5	1	9
User16	4	4	4	2	8
User17	4	4	4	1	9
User18	4	4	4	0	10
User19	5	4	5	0	10
User20	5	4	5	1	9
Mean	4.55	3.95	4.4	0.6	9.4
SD	0.50	0.38	0.49	0.66	0.66

of Q3 indicate that the differences do not have a large negative impact on the landmark recognition for testers.

Also, although there do exist gaps between the registered coordinates and actual coordinates, as long as those landmarks are salient enough to catch, the issue can be solved in the most cases. This is the main reason why we take the “visibility” into consideration. Generally speaking, our proposed algorithm is robust but how to eliminate the coordinate discrepancies will be one of our significant future works.

5.7 Conclusion

This chapter proposes a route finding algorithm using landmark and lighting indicators to provide route guidance that enhances the safety and comprehensiveness levels during walking. Simulation and real outdoor experiments indicate that our proposed algorithm is effective enough in improving safety and comprehensiveness.

In the future, our proposed algorithm can be further improved using other indicators, such as gender, weather, and temperature.

Chapter 6

A Personalized Landmark and Route Recommendation Algorithm for One-Day Trip [1]

6.1 Introduction

One-day travel has become one of the most important ways of entertainment and several travel recommendation systems have been developed [32, 37, 83].

In this chapter, we propose a new personalized travel route recommendation algorithm. Firstly, we evaluate a given area to select the *top*-6 interesting regions based on personalized preferences. Second, we build a travel map based on geographical and time relations among landmarks. Finally, a travel route planning algorithm is proposed to recommend the best travel route. Experimental results show that our proposed algorithm outperforms previous algorithms in precision in landmark recommendation and travel time planning.

This chapter belongs to *Sequential Travel Route Generation* in Fig. 1.1

6.2 Related works

Many existing studies recommend interesting landmarks throughout social media [67, 77, 86], they provide a user with a list of popular landmarks based on users' preferences. However, it is significant to provide a user with a completed travel route rather than a list of individual landmarks. Xu et al. [77] focus on generating travel routes for a sequence of landmarks based on the route distance and landmark attractiveness. But there is still an

[1]Technical contents in this chapter have been presented in the publications ⟨7⟩ and ⟨10⟩.

issue that a user may arrive at a landmark beyond the business time. As the satisfaction of visiting landmarks is highly related to the arrival time [29]. In other words, a good travel route should guarantee that a user can visit the landmark during its business time, and meanwhile, the total visiting time should not exceed the user's time limitation.

Other studies take time limitations into considerations [22, 71]. Gionis et al. [22] focuses on the order of types of landmarks visited, and distances between any two landmarks, while the opening time is ignored. Vansteenwegen et al. [71] provides personalized travel routes with considerations of users' preferences on landmark types, opening time and break time. However, only art-related landmark types are considered which are not suitable for landmarks such as stadiums or cinemas. Also, transpiration ways (train, taxi and bus etc.) between landmarks are not considered which is not suitable if two landmarks are far way from each other.

Therefore, in this paper, we define 8 landmark types to ensure that every landmark can be described properly. Also, the shortest transportation time and the most convenient ways between any two landmarks are provided in order to help user move smoothly.

6.3 Proposed Route Recommendation Algorithm

6.3.1 Landmark Categorization and User Profile

We set user's current position S as starting point and visiting area is set to be a square around S with the width of 10 km and height of 10 km. Then we establish a landmark database with all landmarks in the area. Let L be a set of landmarks in this area. LT is a set of eight landmark types, which are *History, Nature, Entertainment, Art, Sport, Food and Drink, Shopping*, and *Night Life*. Table 6.1 lists 8 example landmarks' type. Every landmark $\ell \in L$ is characterized by its type(s) t, its name label and its coordinates. A landmark can have more than one types. For example, MeijiJingu shrine has two landmark types of *History* and *Nature*.

After collecting nearby landmark information, then we assign a weight to a landmark type. A user creates a user profile with his/her interest weights $w(t)$ for every landmark type $t \in LT$ by the five-point scale, with an explicit definition of each possible weight: very dislike (1), dislike (2), fair (3), like (4), and very like (5).

6.3.2 Region Evaluation

We divide the given 10 km×10 km area into 100 small regions with the width of 1 km and height of 1 km (see Fig. 6.1(a)). We assume that landmarks are distributed evenly in the target area and every small region r has one region landmark $p(r)$ in most of the

Table 6.1: Landmark type in Shinjuku area.

Landmark type	Landmarks
History	Meiji Jingu, Shinjuku Gyoen, Imperial Palace, State Guest House
Nature	Meiji Jingu, Shinjuku Gyoen, Imperial Palace
Entertainment	Cinema I, Tokyo Tower, Tokyo Metropolitan City Hall, Yoyogi Stadium
Art	State Guest House, Museum of Modern Japanese Literature
Sport	Yoyogi Stadium
Food+Drink	Department Store I-III
Shopping	Department Store I-III, Clothing I-II
Night Life	Cinema I

Table 6.2: Region landmark weight.

Region r_i	r_1	r_2	r_3	r_4	r_5	r_6
Weight $w(t_{p(r)})$	4	3	4	5	5	4

cases. Let $t_{p(r)}$ be the landmark type of $p(r)$. If $p(r)$ has more than one landmark types, we select the one with the largest weight as $t_{p(r)}$.

With the user's preferences, the algorithm matches the personal preferences in each landmark. Then we can evaluate each region with a possible visit score. Let R be a set of such small regions. For a region $r \in R$, we assign it with a visit score $visit(r)$:

$$visit(r) = w(t_{p(r)}) \times \alpha \tag{6.1}$$

$$\alpha = \begin{cases} 1 & \text{if } w(t_{p(r)}) = 3, 4, 5 \\ 0 & \text{otherwise} \end{cases} \tag{6.2}$$

where $w(t_{p(r)})$ is a weight of the landmark type given by the user and α is a constant given above. If r does not include any landmark inside, $visit(r) = 0$.

For a one-day trip, as duration time over each landmark is around 1–2 hours, we select 6 regions with the 1st to 6th highest visit scores given by Eqn. (6.1) as *candidate regions*.

6.3.3 Travel Map

As described in Section 1, it is important to schedule the visiting time so that users can visit each region at the proper time. Thus, to further consider the open-closing time and

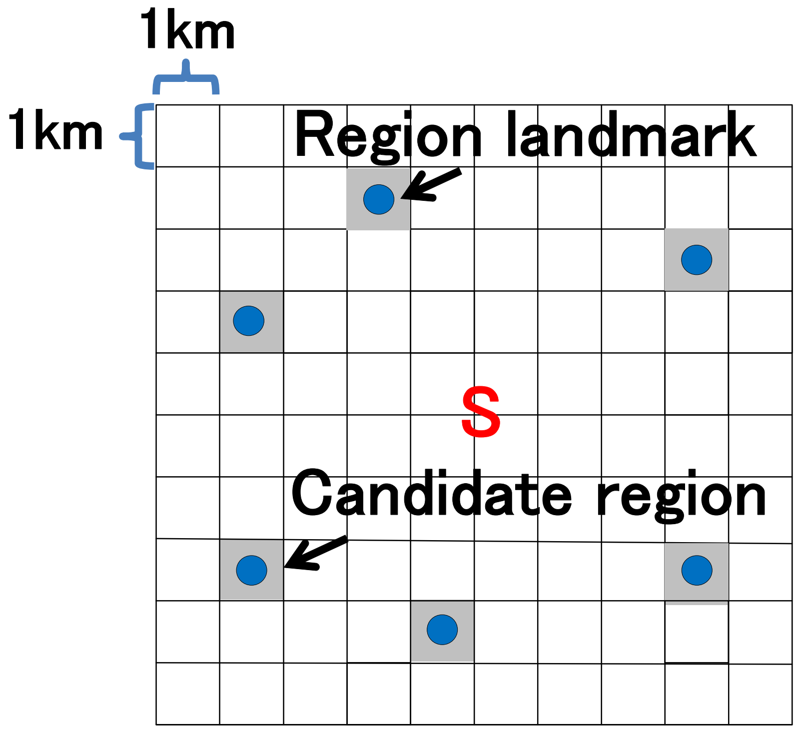


Figure 6.1: An example of the region evaluation.

transition relationships among candidate regions, a travel map among candidate regions is built as in Fig. 6.3, where, for each region r, $OT(r)$ is open time, $CT(r)$ is closing time and $DUT(r)$ is duration time of its region landmark $p(r)$.

We define that a transition between any two candidate regions r_i and r_j is valid if and only if $CT(r_i) \geq CT(r_j)$. We record them as a pair of $(head, tail) = (r_i, r_j)$. For example, for the regions r_2 and r_3, as r_2 has early closing time than r_3, r_2 should be visited first and then we consider $(head, tail) = (r_2, r_3)$ and set a directed edge from r_2 to r_3 in the travel map.

Fig. 6.3 shows an example of a travel map with user's current location S and 6 candidate regions. The transition distance $dis(r_i, r_j)$ between any two candidate regions r_i and r_j is the shortest distance between them. If the $dis(r_i, r_j)$ is longer than 1km, then the moving time $move(r_i, r_j)$ is the time cost by bus or trains. Otherwise, the moving time $move(r_i, r_j)$ is the time cost on foot. In instance, for the regions r_1 and r_2, the moving time $move(r_1, r_2)$ is 0.4 hours. Every direct edge in the travel map has its associated moving time as in Fig. 6.3.

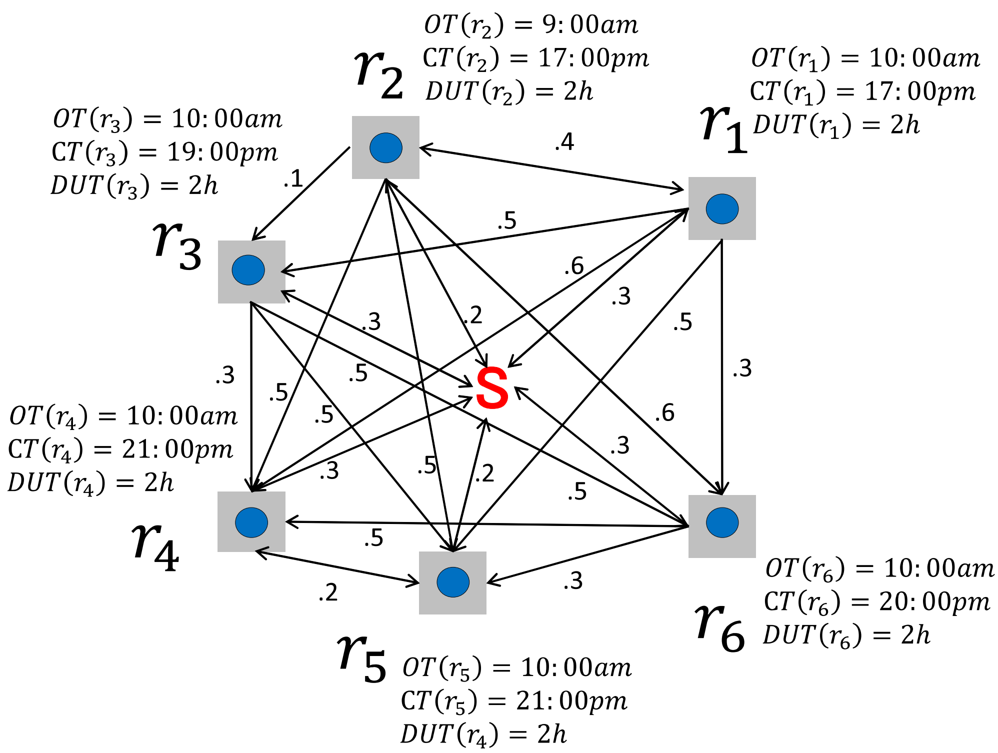


Figure 6.2: An example of travel map.

6.3.4 Travel map construction

As the challenge described in Section 6.1, it is important to schedule the visiting time so that users can visit each landmark at the proper time. Thus, we need to decide visiting orders among the candidate landmarks obtained by the algorithm in Section 6.3.2. A route searching algorithm is proposed for travel route generation.

Firstly, we build a travel map as shown in Fig. 6.3. We denote a travel map as a directed graph $G = (V, E)$, where V is the set of the staring point s and 6 candidate landmarks, i.e., $V = \{s, l_1, ..., l_6\}$. E is the set of all valid trajectories among V. Every landmark $l \in V$ has its open time $ot(l)$, closing time $ct(l)$ and duration time $dut(l)$.

Pre-processing: Every edge $e = (x, y) \in E$ has a direction. If the closing time of the landmark x is earlier than that of y, i.e., $ct(x) \leq ct(y)$, we set a directed edge (x, y) in G. This means that we have to visit the landmark x before the landmark y so that we can visit as many landmarks as possible within the limited time.

Transportation determination: The distance $dis(x, y)$ associated with every edge $e = (x, y) \in E$ is the shortest distance between two vertices x and y. If $dis(x, y)$ is longer than 1km, then the moving time $T_{move(x,y)}$ is the time cost by bus or train. Otherwise, the moving time $T_{move(x,y)}$ is the time cost on foot. In instance, for l_1 and l_2 in Fig. 6.3, the moving time $T_{move(l_1,l_2)}$ is 0.4 hours. Every edge $e = (x, y) \in E$ is associated with the moving time $T_{move(x,y)}$ between its two vertices.

6.3.5 Travel route recommendation

We regard the travel route recommendation procedure as a problem of finding all feasible routes in G that fits the user's travel query. The user u's travel query is $query(u) = \{s, T_s, T_e, q\}$, where s is the starting point, T_s is the starting time, T_e is the ending time and q is the least number of landmarks that the user wants to visit in a day.

Travel route generation: We use a route planning algorithm based on depth-first search and enumerate all the possible routes passing through all or part of the landmarks in G. We consider s as the starting point, and then, we explore as far as possible before backtracking. Meanwhile, the number of landmarks in a route should be equal to or larger than q. After that, we add s as the ending point of each route to construct a completed travel route. Let TR' be a temporal set of all such generated routes.

Travel route ranking: Then, we check if the enumerated route $tr \in TR'$ is feasible or not. When the user can visit all the landmarks in tr at their opening time in the order of tr staying every landmark during their duration time and come back to the starting point, the route tr is considered to be feasible. Let TR be a set of such feasible routes.

Finally, we introduce the Fin score to evaluate the attractiveness of a route tr by Equ. (6.3).

$$Fin(tr) = \sum_{l \in tr} S_{u,l} \tag{6.3}$$

$Fin(tr)$ shows the sum of satisfaction values $S_{u,l}$'s of all landmarks l's included in a travel route tr. Then the route $tr \in TR$ with the highest Fin score is recommended.

In a summary, our travel route recommendation algorithm for a user u is described as follows:

Step R1: Obtain the top-6 landmarks using the algorithm described in Section 2.2.

Step R2: Construct the travel map as a directed graph described in Section 2.3 and Section 2.4.

Step R3: Enumerate all the feasible routes and store them into TR.

Step R4: Evaluate the Fin score in every route in TR and recommend the highest one to the user u.

Table 6.3: Evaluation result on landmark recommendation precision..

Algorithm	Ours	Random	Popular-first
Precision	100%	10%	70%

6.4 Experimental Results

6.4.1 Experimental settings

The algorithm is coded in Java and we have examined one-day-trip cases with 15 different user profiles. We have used the landmark data in [84]. We set the experimental area to be Shinjuku area with the height of 10 km and the width of 10 km, where the starting point S is at the Shinjuku railway station. Our goal is to offer a user a travel route with visiting the *top-k* landmarks. In addition, every candidate region includes only one interesting landmark (its region landmark $p(r)$ in this experiment and these region landmarks are recommended landmarks that a user will visit.

6.4.2 Landmark type recommendation performance

In this subsection, we first recommend top-6 landmarks for 15 users in each city respectively based on their user profiles. Then, we evaluate if recommended landmarks fit users' type weights or not.

The number of true positives (TP) is defined by how many types of recommended top-6 landmarks that successfully fit the user's type weights ($w(t) \geq 3$). The number of false positives (FP) is defined by how many types of the recommended top-6 landmarks that fail to fit the user's type weights ($w(t) < 3$). *Precision* can be obtained by:

$$\text{Precision} = (\text{TP})/(\text{TP} + \text{FP}). \tag{6.4}$$

Table. 6.3 shows the precision of our proposed algorithm, the Random algorithm, and the Popular-first algorithm.

- Random algorithm randomly recommends 6 landmarks in each city.
- Popular-first algorithm always recommends *top*-6 ranked landmarks in each city on TripAdvisor.

According to Table. 6.3, it is shown that Random algorithm obtained the worst performances. It can be explained that Random algorithm has no concerns about user preferences rather than randomly choose a set of 6 landmarks. For Popular-first algorithm,

Table 6.4: Comparison result on average *Fin* score and time planning.

Algorithm	Ours	Xu [77]
Avg. *Fin* score	1.44	1.41
Avg. STPR	0.87	0.67

it recommends the *top*-6 popular landmarks rated by a group of users on TripAdvisor, and the results show that this algorithm improves the precision around 10% compared with the random algorithm. However, the recommended landmarks by the popular-first algorithm are only based on the users' general preferences. In other words, the popular-first algorithm cannot deal with personalized travel preferences. This explains why our proposed algorithm achieves 20% improvement on precision compared with the popular-first algorithm.

6.4.3 Route travel time optimization

Next, we evaluated whether the route recommendation meets users' all requirements with 15 user profiles. We have compared our proposed algorithm with the algorithm in [77], which does not take the visiting time order into consideration.

Two goodness functions are used: (1) *Fin* score based on Eqn. (6.3) and (2) success ratio that user can visit all the landmarks within business hours. (2) is the ratio of the cases that a user successfully visits all the landmarks over all the 15 cases. Table 6.4 shows the results of comparisons, where average successful time planning ratio is denoted as STPR.

In the second line in Table 6.4, for average *Fin* score, it suggests that our proposed algorithm has an even or better performance than [77], while [77] has a poor performance on scheduling visiting time as it does not consider occasions about exceeding business hours.

Also, according to the third line in Table 6.4, although we did not ensure a perfect travel time schedule every time, still our proposed algorithm is better than [77] with 20% improvement. Generally speaking, our proposed algorithm is effective in providing personalized travel recommendations with a proper time schedule.

Generally speaking, our proposed algorithm is effective in the service of personalized travel recommendation.

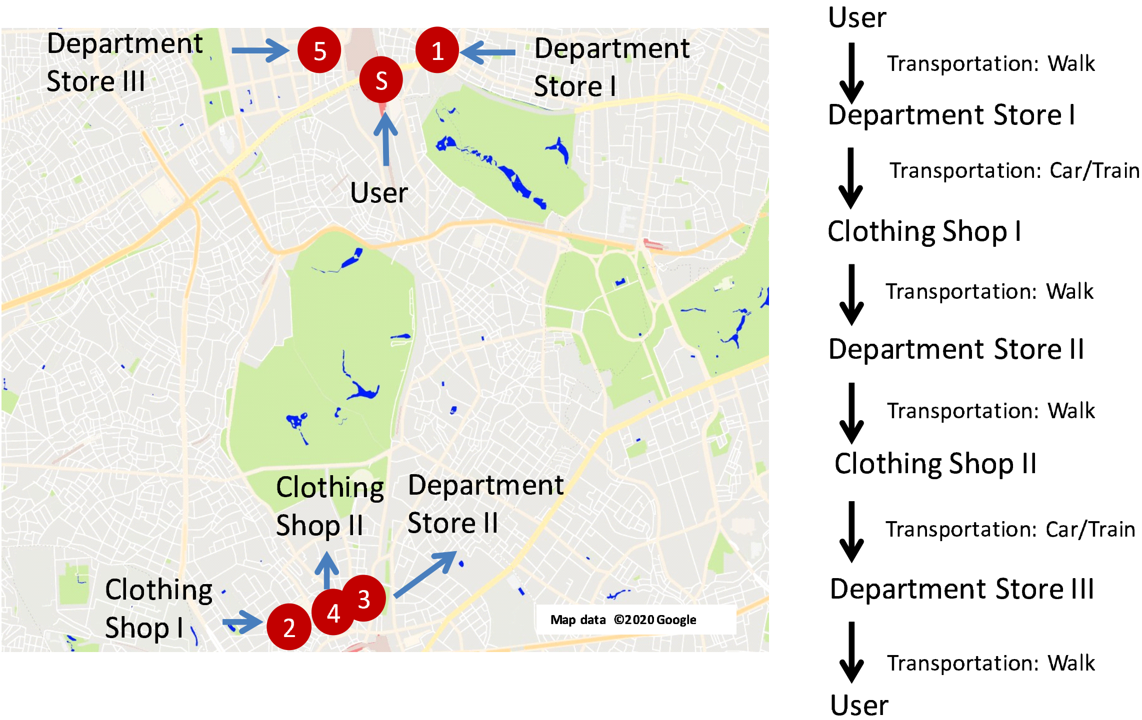


Figure 6.3: An example constructed travel map.

6.4.4 Example of Travel Route Recommendation in Shinjuku Area

Fig. 6.3 shows a recommendation example when a user specifies a one day trip in Shinjuku area, requiring a route with visiting more than 3 landmarks, starting the trip at 10:00 am and finishing the trip before 21:00 pm. We provide the user with a route which visits 5 landmarks, starting at 10:00 am and finishing at 20:20 pm.

The recommended travel route is shown as follows:

S(User) (10:00) →
Department Store I (10:10–12:10) →
Clothing Shop I (12:30–13:30) →
Department Store II (13:40–15:40) →
Clothing Shop II (15:50–17:50) →
Department Store III (18:10–20:10) →
S(User) (20:20)

6.5 Conclusion

In this chapter, we have proposed an algorithm recommending personalized travel route. Experimental results indicate that our proposed algorithm retrieves interesting travel routes for users and outperforms conventional algorithms effectively.

In the future, we will perform more field tests and focus on incorporating additional constraints such as weather and healthy conditions.

Chapter 7

Conclusion

This dissertation aimed to provide users with not only personalized landmark recommendations but also realistic travel routes under different constraints.

In **Chapter 2**, **Chapter 3** and **Chapter 4**, we focus on landmark recommendation with personalized preferences using a great deal of online travel comments. In **Chapter 2**, a personalized landmark recommendation algorithm is proposed which is aiming at exploring new sights into the determinants of personalized landmark satisfaction prediction. The proposed algorithm considers features of landmark type preference, data-source, and language, and selects the most suitable based on satisfaction prediction through the three features. Experimental results show that our proposed algorithm has better performances than previous studies from the viewpoints of landmark recommendation and landmark satisfaction prediction. In **Chapter 3**, an activity-related comment extraction algorithm is proposed based on the linguistic characteristics of Japanese. The proposed algorithm can assistant users to have a better understanding of what they can experience and enjoy while visiting landmarks selected by the algorithm in **Chapter 2**. According to user feedbacks upon two case studies, top-5 comments extracted do provide rich landmark activity information. In **Chapter 4**, a seasonal landmark recommendation algorithm is proposed with consideration of language-specific factors. The algorithm concludes 418,788 user comments from TripAdvisor and analyzes the differences in travel distribution under the impacts of year and language factors. Based on that analysis, we predict future travel distribution for each language group. Then potential peak and off seasons of each landmark are identified and representative seasonal activities are extracted through comment contents for peak and off-seasons, respectively. According to the experimental results in the three cities, it suggests that the proposed algorithm is more accurate in terms of peak season detection and seasonal activity prediction than previous studies.

In **Chapter 5** and **Chapter 6**, we focus on the generation of travel routes. In **Chapter 5**, a safe and comprehensive travel route generation algorithm is proposed. The algorithm recommend routes based on five factors (1) lighting conditions, (2) landmark visibility, (3) landmark effectiveness, (4) turning counts along a route, and (5) road widths. We conduct experiments both in rural and urban areas in the day and night periods respectively, and simulated results and user feedbacks both confirm that obtains better performances compared to several existing algorithms in terms of better safety and comprehensiveness with higher scores. In **Chapter 6**, a one-travel route recommendation algorithm is proposed. The algorithm recommends top-6 landmarks based on users' type preferences, and generate a realistic travel route for a one-day visit under the constraints of the number of landmarks to visits and travel time consumptions. Experimental results confirm the advantages of our proposed algorithm beyond previous studies from the viewpoints of landmark recommendation precision and travel time optimization.

In the future, we will be dedicated to the development of personalized travel route recommendation system with considerations of other indicators, such as gender, weather, and temperature. Also, we will be continuing to deepen our works and intend our scope to other language-specific travel comments such as Korean and Spanish.

Acknowledgment

Throughout the writing of this dissertation I have received a great deal of support and assistance. I would like to to firstly give my deepest and most heartfelt thanks to Prof. Nozomu Togawa (戸川望教授) at the Department of Computer Science and Communications Engineering of Waseda University for his all supports on my research work for almost seven years. I could not have imagined having a better advisor and mentor for my research.

I'm genuinely grateful to Prof. Masao Yanagisawa (柳澤政生教授) at the Department of Electronic and Physical Systems of Waseda University, Prof. Hayato Yamana (山名早人教授) at the Department of Computer Science and Communications Engineering of Waseda University, for their strong support and advice.

I would also like to thank Prof. Youhua Shi (史又華教授) at the Department of Electronic and Physical Systems of Waseda University for his valuable guidance on my research work. I have also received technical support from Dr. Kazushi Kawamura (川村一志講師) at Waseda University, Dr. Masashi Tawada (多和田雅師講師) at Waseda University to conduct this research.

My sincere thanks also goes to Ms. Shuko Watanabe (渡部周子氏) and all the students in the Togawa Laboratory, the Yanagisawa Laboratory and the Shi Laboratory for their kindness.

I have had a wonderful time in the last six years. with my colleagues especially Mr. Masaru Oya (大屋優氏)at Waseda University, Mr. Kento Hasegawa (長谷川健人氏)at Waseda University, Mr. Oku Daisuke (於久太祐氏) at Waseda University, Mr. Ishikawa Ryota (石川遼太氏) at Waseda University. In particular, I am grateful to Mr. Terutaka Orihara (折原照崇氏) at Hitachi, Ltd. and Ms. Huiqian Jiang (蒋慧倩氏) at NEC Corporation for enlightening me the first glance of research.

In addition, I would like to thank family and friends for their continuous support and love.

References

[1] 4travel, http://4travel.jp.

[2] D. Ahn, H. Park, and B. Yoo, "Which group do you want to travel with? A study of rating differences among groups in online travel reviews," *Electronic Commerce Research and Applications*, vol. 25, pp. 105–114, 2017.

[3] M. Amoretti, L. Belli, and F. Zanichelli, "UTravel: Smart mobility with a novel user profiling and recommendation approach," *Pervasive and Mobile Computing*, vol. 38, pp. 474–489, 2017.

[4] Z. Asadi-Shekari, M. Moeinaddini, and M. Shah, "Pedestrian safety index for evaluating street facilities in urban areas," *Safety Science*, vol. 74, pp. 1–14, 2015.

[5] R. Baayen, *Word frequency distributions*. Kluwer Academic Pulishers, 2001.

[6] S. Banerjee and A. Chua, "In search of patterns among travellers' hotel ratings in TripAdvisor," *Tourism Management*, vol. 53, pp. 125–131, 2016.

[7] S. Bao, T. Nitta, D. Shindou, M. Yanagisawa, and N. Togawa, "A landmark-based route recommendation method for pedestrian walking strategies," in *Proc. of the 2015 IEEE Global Conference on Consumer Electronics*, pp. 672–673, 2015.

[8] C. Boom, S. Canneyt, T. Demeester, and B. Dhoedt, "Representation learning for very short texts using weighted word embedding aggregation," *Pattern Recognition Letters*, vol. 80, pp. 150–156, 2016.

[9] D. Boyd, S. Golder, and G. Lotan, "Tweet, tweet, retweet: Conversational aspects of retweeting on twitter," in *Proc. of the 2010 IEEE International Conference on Hawaii International Conference on System Sciences*, pp. 1–10, 2010.

[10] J. Bullough, M. Rea, and Y. Zhou, "Analysis of visual performance benefits from roadway lighting," *Report prepared for the Transportation Research Board on NCHRP project 5-19*, 2009. http://onlinepubs.trb.org/onlinepubs/nchrp/docs/nchrp05-19_visibilitybenefits.pdf

[11] Z. Cheng, J. Caverlee, and K. Lee, "You are where you tweet: A content-based approach to geo-locating twitter users," in *Proc. of the 2010 ACM International Conference on Information and Knowledge Management*, pp. 759–768, 2010.

[12] J. Cranshaw, R. Schwartz, J. I. Hong, and N. Sadeh, "The livehoods project: utilizing social media to understand the dynamics of a city," in *Proc. of the 2012 AAAI International Conference on Weblogs and Social Media*, 2012.

[13] Ctrip, http://www.ctrip.com.

[14] Y. Cui, C. Meng, Q. He, and J. Gao, "Forecasting current and next trip purpose with social media data and Google Places," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 159–174, 2018.

[15] R. Darlington and A. Hayes, *Regression Analysis and Linear Models Concepts, Applications, and Implementation*. GUILFORD PRESS, 2016.

[16] S. Dolnicar and B. Grun, "Assessing analytical robustness in cross-cultural comparison," *International Journal of Culture, Tourism and Hospitality Research*, vol. 1, pp. 140–160, 2009.

[17] G. Fang, S. Kamer, and S. Fujita, "A Japanese Tourism Recommender System with Automatic Generation of Seasonal Feature Vectors," *International Journal of Advanced Computer Science and Applications*, vol. 8, pp. 347–354, 2017.

[18] S. Farrall, J. Bannister, J. Ditton, and E. Gilchrist, "Social psychology and the fear of crime," *British Journal of Criminology*, vol. 40, pp. 399–414, 2000.

[19] E. Farran, Y. Courbois, J. Herwegen, and M. Blades, "How useful are landmarks when learning a route in a virtual environment?" *Journal of Experimental Child Psychology*, vol. 111, pp. 571–586, 2012.

[20] R. Filieri, S. Alguezaui, and F. McLeay, "Why do travelers trust TripAdvisor? Antecedents of trust towards consumer-generated media and its influence on recommendation adoption and word of mouth," *Tourism Management*, vol. 51, pp. 174–185, 2015.

[21] E. Frees, R. Derrig, and G. Meyers, *Predictive Modeling Applications in Actuarial Science: Volume 1, Predictive Modeling Techniques*. Cambridge University Press, 2014.

[22] A. Gionis, T. Lappas, K. Pelechrinis, and E. Terzi, "Customized tour recommendations in urban areas," in *Proc. of the 2014 IEEE International Conference on Web Search and Data Mining*, pp. 313–322, 2014.

[23] Z. Guo, "Does the pedestrian environment affect the utility of walking? a case of path choice in downtown boston," *Transportation Research Part D: Transport and Environment*, vol. 14, pp. 343–352, 2009.

[24] A. Haans and Y. A. Kort, "Light distribution in dynamic street lighting: Two experimental studies on its effects on perceived safety, prospect, concealment, and escape," *Journal of Environmental Psychology*, vol. 32, pp. 342–352, 2012.

[25] Q. Hao, R. Cai, C. Wang, R. Xiao, J. Yang, Y. Pang, and L. Zhang, "Equip tourists with knowledge mined from travelogues," in *Proc. of the 2010 ACM International Conference on World Wide Web*, pp. 401–410, 2010.

[26] S. Hasan and S. Ukkusuri, "Urban activity pattern classification using topic models from online geo-location data," *Transportation Research Part C: Emerging Technologie*, vol. 44, pp. 363–38, 2014.

[27] T. Hoang and J. Mothea, "Location extraction from tweets," *Information Processing and Management*, vol. 54, pp. 129–144, 2018.

[28] I. Howard and B. Rogers, *The visual fields*. New York: Oxford University Press, 1995.

[29] H. Hsieh and C. Li, "Mining and planning time-aware routes from check-in data," in *Proc.of the 2014 IEEE International Conference on Information and Knowledge Management*, pp. 481–490, 2014.

[30] W. Hsu, Y. Wen, L. Wei, and W. Pen, "Skyline Travel Routes: Exploring Skyline for Trip Planning," in *Proc. of the 2014 IEEE International Conference on Mobile Data Management*, pp. 31–36, 2014.

[31] Jaran, https://www.jaran.jp.

[32] K. Jiang, H. Yin, P. Wang, and N. Yu, "Learning from contextual information of geo-tagged web photos to rank personalized tourism attractions," *Neurocomputing*, vol. 119, no. 7, pp. 17–25, 2013.

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

[34] N. Jindal and B. Liu, "Mining comparative sentences and relations," in *Proc. of the 2006 AAAI National Conference on Artificial Intelligence*, pp. 1331–1336, 2006.

[35] S. Khopkar and A. Nikolaev, “Predicting long-term product ratings based on few early ratings and user base analysis,” *Electronic Commerce Research and Applications*, vol. 21, pp. 38–49, 2017.

[36] K. Kim, O. Park, S. Yun, and H. Yun, “What makes tourists feel negatively about tourism destinations? Application of hybrid text mining methodology to smart destination management,” *Technological Forecasting and Social Change*, vol. 123, pp. 362–369, 2017.

[37] Q. Le and D. Pishva, “An innovative tour recommendation system for tourists in japan.” in *Proc. of the 2015 IEEE International Conference on Advanced Communications Technology*, pp. 489–494, 2015.

[38] G. Li, Y. Wang, M. Li, and Z. Wu, “Similarity match in time series streams under dDynamic time warping distance,” in *Proc. of the 2008 IEEE International Conference on Computer Science and Software Engineering*, pp. 399–402, 2008.

[39] D. Lian and X. Xie, “Collaborative activity recognition via check-in history,” in *Proc. of the 2011 ACM SIGSPATIAL International Workshop on Location-Based Social Networks*, 2011.

[40] C. Liang and W. Bi, “Seasonal variation analysis and SVR forecast of tourist flows during the year: A case study of Huangshan mountain,” in *Proc. of the 2017 IEEE International Conference on Big Data Analysis*, pp. 921–927, 2017.

[41] H. Liu, L. Wei, Y. Zheng, M. Schneider, and W. Peng, “Route discovery from mining uncertain trajectories,” in *Proc. of the 2011 IEEE International Conference on Data Mining Workshops*, pp. 1239–1242, 2011.

[42] I. Lourentzou, A. Morales, and C. Zhai, “Text-based geolocation prediction of social media users with neural networks,” in *Proc. of the 2017 IEEE International Conference on Big Data*, pp. 696–705, 2010.

[43] X. Lu, C. Wang, J. Yang, y. Pang, and L. Zhang, “Photo2trip: generating travel routes from geotagged photos for trip planning,” in *Proc. of the 2011 ACM International Conference on Multimedia*, pp. 143–152, 2011.

[44] W. Ma, D. Liao, Y. Liu, and H. Lo, “Optimization of pedestrian phase patterns and signal timings for isolated intersection,” *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 502–514, 2015.

[45] C. Manning, P. Raghavan, and H. Schutze, *Introduction to Information Retrieval*. Cambridge University Press, 2008.

[46] H. Millward, S. Jamie, and D. Scott, "Active-transport walking behavior: destinations, durations, distances," *Journal of Transport Geography*, vol. 28, pp. 101–110, 2013.

[47] H. Miura, S. Takeshima, N. Matsuda, and H. Taki, "A study on navigation system for pedestrians based on street lightings," *Human-Computer Interaction*, vol. 6883, pp. 49–55, 2011.

[48] M. M̈uller, *Information retrieval for music and motion*. Springer, 2007.

[49] K. Nakazawa, N. Kita, K. Takagi, T. Inoue, H. Shigeno, and K. Okada, "A dynamic map based on landmark's visibility," *IPSJ Journal*, vol. 49, no. 1, pp. 233–241, 2008.

[50] Y. Ni, M. Wan, J. Sun, and K. Li, "Evaluation of pedestrian safety at intersections: A theoretical framework based on pedestrian-vehicle interaction patterns," *Accident Analysis and Prevention*, vol. 96, pp. 118–129, 2016.

[51] NIST, *NIST/SEMATECH e-Handbook of Statistical Methods: Single Exponential Smoothing*. U.S. Department of Commerce, 1957.

[52] I. Önder, "Classifying multi-destination trips in Austria with big data," *Tourism Management Perspectives*, vol. 21, pp. 54–58, 2017.

[53] R. Pain, R. MacFarlane, K. Turner, and S. . Gill, "When, where, if and but. qualifying gis and the effect of street lighting on crime and fear," *Environment and Planning A*, vol. 38, pp. 2055–2074, 2006.

[54] M. Panou, K.Touliou, E. Bekiari, and E. Gaitanidou, "Pedestrian and multimodal route guidance adaptation for elderly citizens," in *Proc. of the 2010 IEEE IST-Africa*, pp. 1–7, 2010.

[55] E. Pantano, C. Priporas, and N. Stylos, "You will like it ! using open data to predict tourists ' response to a tourist attraction," *Tourism Management*, vol. 60, pp. 430–438, 2017.

[56] S. Park and D. Kim, "Assessing language discrepancies between travelers and online travel recommendation systems: Application of the Jaccard distance score to web data mining," *Technological Forecasting and Social Change*, vol. 123, pp. 381–388, 2017.

[57] G. Ramchand, *The Oxford Handbook of Linguistic Interfaces (Oxford handbooks in Linguistics)*. Oxford University Press, 2007.

[58] T. H. Rashidi, A. Abbasi, M. Maghrebi, S. Hasan, and T. S. Waller, "Exploring the capacity of social media data for modelling travel behaviour: Opportunities and challenges," *Transportation Research Part C: Emerging Technologies*, vol. 75, pp. 197–211, 2017.

[59] R.Filieri, " What makes an online consumer review trustworthy?" *Annals of Tourism Research*, vol. 58, pp. 46–64, 2016.

[60] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, *Recommender Systems Handbook*. Springer-Verlag, 2011.

[61] RMeCab, https://code.google.com/p/mecab/downloads/list.

[62] H. Shen, G. Liu, H. Wang, and N. Vithlani, "SocialQ&A: An online social network based question and answer system," *IEEE Transactions on Big Data*, vol. 3, pp. 91–106, 2017.

[63] R. Shumway and D. Stoffer, *Time series analysis and its applications: With R examples*. Springer, 2006.

[64] A. Siegel and S. White, "The development of spatial representations of large-scale environments," *Advances in Child Development and Behavior*, vol. 10, pp. 9–55, 1975.

[65] H. Spiers and E. Maguire, "The dynamic nature of cognition during wayfinding," *Journal of Environmental Psychology*, vol. 28, no. 3, pp. 232–249, 2008.

[66] V. Sugumaran, *Recent Advances in Intelligent Technologies and Information Systems*. IGI GLOBAL, 2015.

[67] C. Sun, "Tour recommendations by mining photo sharing social media," *Decision Support Systems*, vol. 101, pp. 28–39, 2017.

[68] T. Tatenami, D. Matsushita, and J. Munemo, "Transition of level of anxiety in wayfinding with pedestrian navigation system on mobile telephone," *Journal of Architecture and Planning*, vol. 608, pp. 59–64, 2006.

[69] TripAdvisor, https://www.tripadvisor.jp.

[70] P. Vansteenwegan, W. Souffriau, G. Berghe, and D. Oudheusden, "The City Trip Planner: An expert system for tourists," *Expert Systems with Applications*, vol. 38, pp. 6540–6546, 2011.

[71] P. Vansteenwegen, W. Souffriau, G. V. Berghe, and D. VanOudheusden, “The City Trip Planner: An expert system for tourists,” *Expert Systems with Applications*, vol. 38, no. 6, pp. 6540–6546, 2011.

[72] D. Weaver, A. Kwek, and Y. Wang, “Cultural connectedness and visitor segmentation in diaspora Chinese tourism,” *Tourism Management*, vol. 63, pp. 302–314, 2017.

[73] L. Wei, Y. Zheng, and W. Peng, “Constructing popular routes from uncertain trajectories,” in *Proc. of the 2012 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 195–203, 2012.

[74] L. Wen, C. Rissel, and H. Fu, “The effect of active transport, transport systems, and urban design on population health,” *Journal of Environmental and Public Health*, vol. 13, pp. 1–13, 2013.

[75] Y. Wen, K. Cho, W. Peng, J. Yeo, and S. Hwang, “KSTR: Keyword-aware skyline travel route recommendation,” in *Proc. of the 2015 IEEE International Conference on Data Mining*, pp. 449–458, 2015.

[76] X. Xu, W. Sasaki, C. Yu, and Y. Takama, “Proposal of collecting lighting situation of roads at night for recommendation of safety walking route using smartphone,” in *Proc. of the 2014 IEEE International Symposium on System Integration*, pp. 414–418, 2014.

[77] Y. Xu, T. Hu, and Y. Li, “A travel route recommendation algorithm with personal preference,” in *Proc. of the 2016 IEEE International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*, pp. 390–396, 2016.

[78] T. Yamasaki, A. Gallagher, and T. Chen, “Geotag-based travel route recommendation featuring seasonal and temporal popularity,” in *Proc. of the 2013 IEEE International Conference on Information, Communications and Signal Processing*, pp. 1–4, 2013.

[79] Y. Yang, “Interactions between psychological and environmental characteristics and their impacts on walking,” *Journal of Transport and Health*, vol. 2, pp. 195–198, 2015.

[80] K. Yew, T. H, and S. Paua, “Safe journey: A pedestrian map using safety annotation for route determination,” in *Proc. of the 2010 IEEE International Symposium on Information Technology*, pp. 15–17, 2010.

[81] Y. Ying, L. Chen, and G. Chen, "A temporal-aware POI recommendation system using context-aware tensor decomposition and weighted HITS," *Neurocomputing*, vol. 242, pp. 474–489, 2017.

[82] C. Yu, W. Ma, K. Han, and X. Yanga, "Optimization of vehicle and pedestrian signals at isolated intersections," *Transportation Research Part B: Methodological*, vol. 98, pp. 135–153, 2017.

[83] Z. Yu, H. Xu, Z. Yang, and B. Guo, "Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints," *IEEE Transactions on Human-Machine Systems*, vol. 46, pp. 151–159, 2016.

[84] ZENRIN, http://www.zenrin.co.jp/.

[85] V. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and 2010 ACM International Conference on World Wide Web, pp. 1029–1038, 2010.

[86] X. Zhou, M. Wang, and D. Li, "From stay to play-A travel planning tool based on crowdsourcing user-generated contents," *Applied Geography*, vol. 78, pp. 1–11, 2017.

[87] 総務省, "Building standards act. 1950," https://elaws.e-gov.go.jp/search/elawsSearch/elaws_search/lsg0100/.

List of Publications

論文（学術誌原著論文）

⟨1⟩ ○ **S. Bao**, T. Nitta, M. Yanagisawa, and N. Togawa, “A Safe and Comprehensive Route Finding Algorithm for Pedestrians Based on Lighting and Landmark Conditions,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E100-A, no. 11, pp. 2439–2450, Nov. 2017.

国際会議

⟨2⟩ ○ **S. Bao**, M. Yanagisawa, and N. Togawa, “A Travel Decision Support Algorithm: Landmark Activity Extraction from Japanese Travel Comments,” Studies in Computational Intelligence (Springer), Computer and Information Science, *2019 IEEE/ACIS International Conference on Computer and Information Science*, vol. 849, pp 109–123, Beijing, China, Jun. 2019.

⟨3⟩ ○ **S. Bao**, M. Yanagisawa, and N. Togawa, “Landmark Seasonal Travel Distribution and Activity Prediction Based on Language-specific Analysis,” in *Proc. of the 2018 IEEE International Conference on Big Data*, pp. 3628–3637, Seattle, USA, Dec. 2018.

⟨4⟩ ○ **S. Bao**, M. Yanagisawa, and N. Togawa, “Personalized Landmark Recommendation Algorithm Based on Language-specific Satisfaction Prediction Using Heterogeneous Open Data Sources,” in *Proc. of the 2018 IEEE International Conference on Computational Intelligence and Communication Networks,*, pp. 70–76, Esbjerg, Denmark Aug. 2018.

⟨5⟩ ○ **S. Bao**, M. Yanagisawa, and N. Togawa, “Personalized Landmark Recommendation for Language-specific Users by Open Data Mining,” Studies in Computational Intelligence (Springer), Computer and Information Science, *2018 IEEE/ACIS International Conference on Computer and Information Science*, vol. 791, pp.107–121, Singapore, Singapore Jun. 2018.

〈6〉 ○ **S. Bao**, M. Yanagisawa, and N. Togawa, “Road-illuminance level inference across road networks based on Bayesian analysis,” in *Proc. of the 2018 IEEE International Conference on Consumer Electronics*, pp. 1–6, Las Vegas, USA, Jan. 2018.

〈7〉 ○ **S. Bao**, M. Yanagisawa, and N. Togawa, “Personalized one-day travel with multi-nearby-landmark recommendation,” in *Proc. of the 2017 IEEE International Conference on Consumer Electronics Berlin*, pp. 258–261, Berlin, Germany, Sept. 2017.

〈8〉 ○ **S. Bao**,T. Nitta, K. Ishikawa, M. Yanagisawa, and N. Togawa, “A safe and comprehensive route finding method for pedestrian based on lighting and landmark,” in *Proc. of the 2016 IEEE Global Conference on Consumer Electronics*, pp. 589–593, Kyoto, Japan, Oct. 2016.

〈9〉 ○ **S. Bao**,T. Nitta, D. Shindou, M. Yanagisawa, and N. Togawa, “A landmark-based route recommendation method for pedestrian walking strategies,” in *Proc. of the 2015 IEEE Global Conference on Consumer Electronics*, pp. 672–673, Osaka, Japan, Oct. 2015.

国内学会

〈10〉 鮑思雅, 柳澤政生, 戸川望, “One-day Trip Recommendation for Nearby Spots Based on Users’ Locations and Preferences,” マルチメディア，分散，協調とモバイル（DICOMO）シンポジウム講演論文集, 札幌市, Jun. 2017.

〈11〉 鮑思雅, 柳澤政生, 戸川望, “An Evaluation Method of Road Illuminance Levels Using Road Lights and Landmarks,” 電子情報通信学会 2017 年ソサイエティ大会講演論文集, 名古屋市, Mar. 2017.

研究費・助成金

〈12〉 公益財団法人電気通信普及財団　平成 29 年度 8 月期海外渡航旅費援助, 総額 29 万円.