Practical Application of Tour Plan Mining System

Tour Miner

- Web Application and Case Study -

Pattara Leelaprute, Pittayut Sothanakul
Department of Computer Engineering, Faculty of Engineering,
Kasetsart University
Bangkok 10900, Thailand
Email: pattara.l@ku.ac.th, pittayut.sot@ku.th

Shingo Yamaguchi
Graduate School of Sciences and Technology for Innovation,
Yamaguchi University
2-16-1 Tokiwadai, Ube 755-8611, Japan
Email: shingo@yamaguchi-u.ac.jp

Abstract—We have been researching and developing the system called "Tour Miner" which creates a tour plan suitable for individual preferences or interests from the information accumulated in SNS. In our previous research, we proposed a method to extract travel records from check-in information shared on the SNS, "Foursquare", and constructed a database of travel records. In addition, we also proposed a method to select travel records matching individual preferences or interests from the database. In this paper, we developed "Tour Miner" as a web application based on the assumption of actual use by applying the methods and results proposed by our research team in the previous research. The web application can select travel records which match individual preferences or interests. We then released our web application to the public at the special exhibition at Yamaguchi Prefectural Museum for 1 month. According to the feedback from 425 users who used our Web Application, the selected travel records matched with their interests, and those travel records were useful in their travel plans preparation.

Keywords—SNS, mining, Foursquare, check-in, travel, tour, interest, Web Application

I. INTRODUCTION

The mainstream of travel has changed from group travel to individual travel based on individual preferences, interests and interests. It is also common for travelers to post travel records on SNS about which places they visited and what they experienced. The travel record based on the actual tourist experiences will be an important source of information for tourists.

Hasegawa et al. [1] proposed a method of extracting tweets related to a specific travel experience by considering the spatio-temporal continuity of Twitter. This research targets only their own tweets and does not include tweets of the others. Nakajima et al. [2] proposed a method of recommending tour routes by searching tweets with location information tweeted around sightseeing spots. Although this method is more promising, however it is only 5% of the total records that come with the location information [3], and it can be said that the information that can be extracted is limited.

We are researching and developing the system called "Tour Miner" which creates a tour plan suitable for individual preferences or interests from the information accumulated in SNS [4]. In our previous research [5], we proposed a method to extract travel records from check-in information accumulated in SNS sharing location information "Foursquare" and constructed a database of travel records. In addition, we proposed a method to select travel records matching individual preferences or interests from the database. In this paper, we developed "Tour Miner" as web application based on the assumption of actual use by applying the methods and results proposed by our research team in the previous research. The web application can select travel records which match individual preferences or interests. We then released our web application to the public at the Special Exhibition at Yamaguchi Prefectural Museum [6] from August 1st to September 1st, 2019. According to the feedback from 425 users who used our Web Application, the selected travel records matched with their interests, and those travel records were useful in preparing for their travel plans.

II. FOURSQUARE AND TOUR MINER

A. Foursquare

Foursquare is an SNS that shares check-in information to places such as stores and other facilities. The check-in information includes user's ID, location ID, date, latitude/longitude, category of the place, and country information. Each place is categorized into more than 400 categories, and such information is being utilized. For example, "Kyoto Imperial Palace" belongs to the category of "Historic Site". Examples of other categories related to travel are Museum, Theme Park, Temple, Shrine, Restaurant, etc. Note that privacy is maintained for Foursquare, like other SNS such as Twitter. This is because it is not the user’s real name, but is the user's ID that is published. Thus, we are not invading the user’s privacy.
B. Tour Miner

The system of Tour Miner comprises 4 steps; (1) build a database of travel records from the information accumulated in SNS, (2) select a travel record that matches the user's preferences and interests from that database, (3) newly synthesize tour plans, and (4) customize the tour plan according to the conditions of the user, thereby generating a practical tour plan. In Ref. [7] we proposed a method of creating a travel route including sightseeing spots with high visiting frequency by combining multiple travel records.

The web application we developed in this paper covers the functions explained in step (1) and step (2).

III. EXTRACTION METHOD OF TRAVEL RECORDS FROM CHECK-IN INFORMATION

A. Travel Record Database

We proposed a method to extract travel records from Foursquare and built a database of travel records. The sequence of check-in information of Foursquare does not necessarily represent a travel record. It is because people may also check-in to places they visit in their daily lives. Therefore, among check-in information, it is necessary to extract only the travel records occurred during their travel.

The authors focused on the category "Home (private)" which represents home and calculated the "home country" for each user. And we propose a method to consider the check-in information which has been placed outside their home country as traveling abroad and extract it as a travel record. Furthermore, because the extracted travel records were the travel records of more than 30,000 people, we built a database of travel records on the cloud using Google's BigQuery. As an index for each travel record, we attached the visiting country and the number of visits of each category, to make it possible to search relevant travel records quickly.

In this research, we used the Global-scale Check-in Dataset [8] that Yang obtained from Foursquare. This data set includes 33,278,683 check-in information by 266,909 users, during the 18 month period from April 2012 to September 2013.

B. Extraction Method of Travel Records from the Database

We propose a method to select a travel record suitable for individual preferences or interests (hereinafter referred to as "Interest") from the constructed database. When inputting the country name, individual preference or interests (Interest), the proposed method outputs a set of travel records that matched the interest.

Example: If you are interested in temples and shrines, the expression of that Interest will be "Temple, Shrine".

The similarity between the interest and each travel record is calculated and the travel record that the value exceeds the optimum threshold value is selected.

As an example, let's think about a user who is interested in temples and shrines in Japan. Figure 2 shows the result (partially), when Japan was designated as the destination country of travel and "Temple, Shrine" was specified for the Interest.

IV. PRACTICAL APPLICATION

We developed a practical application based on the method described in chapter III. This Web Application was released to the public at a special exhibition of the Yamaguchi Prefectural Museum from August 1st to September 1st, 2019.

A. Design Policy

The following requirements were set for practical use;

Requirement 1: The web application has to sufficiently accept various users' interests.
Requirement 2: The web application has to accurately extract travel records.
Requirement 3: The web application has to be available for multiple users at the same time.

For Requirement 1, we examined the expected users to select the appropriate place category. Since we assumed that users at Yamaguchi Prefectural Museum are young family with children, then we chose the place categories that are easily understandable for them. The category names we selected are; Temple, Shrine, Castle, Historic Site, Museum, Theme Park, Mall, Japanese Restaurant, Baseball and Garden.

For Requirement 2, we analyzed the travel records of the tourists in the database to examine their travelling trends. As a result, it was found that out of 931 foreign tourists visiting Japan, 733 (78.7%) traveled only in one region (see Graph 1).
If a travel record is extracted without designating a specific region, a location (check-in location) in the region that the user is not interested in or does not plan to go to may be selected as an output. Therefore, we thought that it would be possible to extract the travel records that the user wants more accurately by letting the user specify “region”.

For Requirement 3, we chose to implement our application in the form of a web application. This enables each individual to use the application with his/her own smartphone.

Please note that our web application at present is specifically designed for Japan.

For the second step, as shown in Fig. 4, users select place categories. In this example, Temple and Japanese Restaurant is selected. For the convenience of screen display, “Restaurant” is displayed, but the actual place category is “Japanese Restaurant”.

For the third step, as shown in Fig.5, the user selects which region the place category selected in the previous step belongs to. In this example, “Kanto” (the eastern region of Japan's main island) and “Kansai” (the southern-central region of Japan's main island) are selected.

### B. Implementation & Application Example

As the overall application flow, after inputting user information (gender and age group) and his/her interest (place category and region of interest), the application calculates the similarity, and generates output of five travel records with the highest similarity score. As for details of each travel record, apart from displaying each check-in information namely the specific place name and check-in date/time, those places are also displayed on Google Map.

Next, using actual examples, we explain usage examples of the web application step-by-step.

For the first step, as shown in Fig.3, the user selects the gender and age group so we know who has come to use our Web Application.

For the second step, as shown in Fig. 4, users select place categories. In this example, Temple and Japanese Restaurant is selected. For the convenience of screen display, “Restaurant” is displayed, but the actual place category is “Japanese Restaurant”.

For the third step, as shown in Fig.5, the user selects which region the place category selected in the previous step belongs to. In this example, “Kanto” (the eastern region of Japan's main island) and “Kansai” (the southern-central region of Japan's main island) are selected.
Then the application calculates similarity based on the place category (Temple, Japanese Restaurant) and the region name (Kanto, Kansai) entered, and then the application generates outputs of five travel records with the highest similarity score. From Fig. 6 to Fig. 9, only the first two travel records outputs are shown as examples. Fig. 6 and Fig. 7 show the travel record of the traveler with user ID: 122107 from Korea. Fig. 8 and Fig. 9 show the travel record of the traveler with User ID: 42299 from Thailand. Fig. 6 and Fig. 8 show a wide area map, while Fig. 7 and Fig. 9 show a specific area zoomed up. Please note that each travel record contains a considerable amount of check-in information but all cannot be displayed due to limitations of the paper in terms of space. For reference, we show the output log of the actual travel records of travelers from Korea and Thailand in Fig. 10.

C. Analysis of the generated result

As a result of verifying the output results explained in paragraph B., every travel record was reflected in the entered interest of the user. In other words, both of inputted place categories which are “Temple” and “Japanese Restaurant”, and the region names which are “Kanto” and “Kansai” are factored in the output travel records which properly included the actual Japanese restaurant and temple check-in locations in those regions. Therefore, it can be said that “Tour Miner” provides useful information for planning a trip according to the region and place category in which the user is interested. In addition, these check-in locations are displayed on Google Maps, so you can see the distance between each location and that can help you make plans for transportation which is indispensable in planning your trip.
D. Analysis of user's actual application usage results

There were 425 users while the app was being released to public at the special exhibition of the Yamaguchi Prefectural Museum from August 1st to September 1st, 2019.

We verify whether the 10 place categories previously selected as options for the user to select were appropriate or not. Analysis of user's actual application usage results showed that, as shown in Graph 2, although “Theme Park” and “Japanese Restaurant” were most selected, all of 10 place categories were selected evenly.

Besides, statistical data shows that users of age group under 20 years old accounts for 56.70%, between 20-29 years represents 7.29%, between 30-39 years is 9.41%, between 40-49 years is 16.47%, between 50-59 years is 4%, and over 60 years is 6.11%.

According to feedback from many users, the output travel records are relevant to their interests and are useful when planning trips.

### Fig.10. Log of actual travel records of travelers from Korea (User ID: 122107) and Thailand (User ID: 42299)

![Graph 2. Actual numbers of place category selected by web application users.](image)

V. CONCLUSION

In this paper, we developed a web application based on actual use in order to be a part of a system called “Tour Miner” which creates a travel plan tailored to individual preferences or interests from the information accumulated in SNS. The developed web application applies the methods and results proposed in our previous research to select travel records that match the user’s personal preferences or interests from check-in information stored in Foursquare.

We then released our web application to public at the special exhibition at Yamaguchi Prefectural Museum, and according to the feedback from 425 users, the selected travel records matched with their interests, and those travel records were useful in preparing their travel plans.

Our future work will be to utilize the past history of users to designate interests, and to complete developing all steps of the “Tour Miner” to arrive the practical tour plan.
REFERENCES


