An Innovative Tour Recommendation System for Tourists in Japan

Quang Thai LE*, Davar PISHVA**

*Faculty of International Management, Ritsumeikan Asia Pacific University, Beppu, Japan

** Faculty of Asia Pacific Studies, Ritsumeikan Asia Pacific University, Beppu, Japan

quanle11@apu.ac.jp, dpishva@apu.ac.jp

***Abstract*[1]— The paper demonstrates prototype of system that is capable of suggesting optimal touring plans which are composed of various points of interest (POI) and take travelers' preferences and context into account. It systematically collects and analyzes information on thousands of tourists attraction areas and geographical nodes of Japan Railway (JR) train stations together with concurrent weather information, estimated travel time, associated expenses, and lists of multiple cultural events in order to demonstrate practicality as well as reliability of the system. A programmatic approach based on the heuristic greedy search is employed for transforming the obtained data into informative routes. It demonstrates the feasibility of the approach through its mobile prototype on web platform and tests it under various scenarios in eight different places in Japan which includes Tokyo, Osaka, Kyoto, Kobe, Yokohama, Nagoya, Fukuoka and Sapporo. Its result and the performance can be considered as a stepping stone towards a more localized and practical recommendation system in the field of tourism in the near future.**

Keywords— e-tourism, travel planning system, web scraping, modeling, and data mining.

I. Introduction

Olympic Games usually provide economic stimulus and since Japan won the bid to host the 2020 Summer Olympic Games in September 2014, various think-tank groups have announced their estimates for its economic impacts. The Olympic 2020 will, therefore, be imposing both challenges and opportunities to the current stream of globalization in the country [1]. According to Japan National Tourist Organization (JNTO), 12 million tourists visited Japan in the year 2014, but the number is expected to increase to 20 million by 2020 and as many as 8.5 million foreign tourists are projected to visit the country just during the Olympic [2]. Hence, the next 5 years will be a good opportunity for Japan to leverage its "world class" brand in tourism and hopefully become a tourism-oriented country in the near future [3]. Similarly, in April 2014, Japanese Ministry of Education decided to provide 37 selected universities in Japan with an annual subsidy of about $3.6 million each for the next 10 years so as to support them to become the so-called top global universities [4]. This is just one of the many efforts used for attracting international professors, researchers, students and cooperating with other prestigious institutions around the world, welcoming about 300,000 talented foreign students to visit and study in Japan until 2020 [5]. Despite its academic nature, possible tourism dimension of such huge number of foreign consumers is undeniably an open issue for the tourism industry.

Another challenge facing Japan is the sledding down of its world's economic ranking. The government believes that through artificial inflation (i.e., devaluation of its currency), not only Japanese products could become more completive in the world market, but also the country as a whole could become a more attractive tourist destination. This idea seems to be working since a 40% depreciation of Japanese currency during the past three years has increased both Japanese products export as well as the number of tourists visited the country. In fact for the first time during the past 45 years, the number tourists that visited Japan during the six months period surpassed the number Japanese tourists who visited other countries (OBS News, July 2015). Hence, the 2020 Summer Olympic is not just a precious opportunity for the country to leverage the domestic tourism to recover its economy, but also a momentum for the nation to strengthen its image of a globalizing country.

As an effort to bridge the gap between academia and empirical application in the field of tourism, our paper proposes a practical recommendation system for foreign tourists in Japan. It initially shares some background information regarding characteristics of Japan tourism and previous related works and then addresses the need of an innovative touristic recommendation system for Japan. Section 3 provides a detailed explanation on the methodologies employed in the system, including a brief introduction to the system's architecture, its data retrieval technique, the employed heuristic cluster-first-route-second approach, and demonstration of a web-based prototype application which has been developed to test our approach. Section 4 shows the results obtained on various traveling scenarios across Japan and compares them with results of other approaches. Section 5 points out current limitations and many promising future directions along this research. Section 6 subsequently summarizes and concludes the paper.

II. Literature Reviews

A. Characteristics of Tourism Activities in Japan

Japan tourism has seasonal flavor and different marketing strategies are used during each of the four seasons. For example, Japanese Railway (JR) Company issues

Manuscript received on May 8, 2015. This work is a follow-up of the invited journal to the accepted conference paper of the 17th International Conference on Advanced Communication Technology.

Thai Quang LE is a student at Ritsumeikan Asia Pacific University (APU) in the field of Business Administration focusing on Strategic Management and candidate for graduation.

Davar Pishva is a professor in ICT at Ritsumeikan Asia Pacific University (APU) Japan (corresponding author phone: +81-977-78-1261, fax: +81-977-78- 1261, e-mail:dpishva@apu.ac.jp).

promotional free train tickets during a certain period in spring, summer or winter (e.g., Seishun 18 ticket) [6]. The country is also famous for its numerous annual cultural events. No wonder Japan was chosen as one of the World's top tourist destinations in 2011 by the CNN [7], and as the world's ninth most tourist-friendly country [8]. Its transportation infrastructure consists of modern buses and highly efficient train system equipped with rapid-transit railway network that link thousands of stations located throughout the country [6]. Tokyo is by far the most popular destination in Japan, accounting for 57.4% of foreign visitors [9].

Even though touring Japan seems to connote economic affluence, the recent depreciation of the Japanese Yen has made it possible for people from emerging countries to come and experience the Japanese life [9]. Despite continuous efforts to cope with the increasing flow of foreign tourists, there are still some major problems from the foreign tourists' perspectives in Japan. Specifically, in 2011, it is reported that public transport related issues (route details, usage and fees), lack of public LAN services, difficulties in communication are the top problems, accounting for about 45%, 37%, and around 25% respectively among all of the foreign correspondents [10]. Furthermore, the task of selecting a tourism destination in Japan is very important [11], and it requires "comprehensive information" so as to be "comfortable" [12]. This is because; there is still a lack of communication means in order to effectively bridge the tourists with the domestic tourism assets. Hence, the exploration of a collective delivery of tourism related knowledge to the right target audience under a particular context is an important practical research topic. Therefore, not only infrastructure metrics such as international airport capacity, wireless network or, foreign language signs availability, etc., but also intangible services such as multilingual promotion websites, tour-guide profession, etc. have to be further enhanced [13].

B. *Previous Works on Tourism Recommendation System and Research Objectives*

1) *A Brief Literature Review*

Previous approaches in implementing digital recommender system (RS) for the tourism industry mainly focused on providing more specialized products to customers. For example, in 2014, Damianos et al. did an intensive review on this topic with 19 different touristic recommendation systems, discovering that more and more researches take into account tourists' contextual information in the recommendation algorithm such as "attracts already visited", "user mobility pattern", "transportation mode", etc. [14]. The review also suggested consideration of weather conditions, more practical route and flexible usage of transportation as prospect future research directions [14]. Moreover, as clearly stated in the paper, most of the recommendation in tourism provides functions mainly either for suggesting tourism attractions or tourism services (e.g., restaurants, hotels, etc.), and there is a lack of systems that can combine these two [14]. In the same year, Ricardo, Lino, Ana, and Paulo introduced a system called PSiS. PSiS Mobile is a sightseeing tours recommendation system also takes the concurrent weather conditions as a real-time constraint [15]. The system also recognizes an interesting implementation problem when certain places are filtered out according to their operation time schedule, even though sometimes the architectonic outer appearance is sometimes the selling point [15]. Moreover, the system also emphasized device context (e.g., battery, connectivity) and dynamic context (e.g., real-time adaptation) [15]. In another research, Chiang and Huang (2015) made use of the user interface to continuously receive feedback from the users, as well as facilitate them to adjust any unsatisfied recommended results [16].

Regarding recommendation algorithms, most of the existing studies in the field has employed similarity or collaborative filtering (CF) to achieve the task. This probably results from the ability to match the user's preferences from previous cases, with the assumption that users who are similar are likely to have the same POIs [17]. To illustrate, Gulcin and Buse (2011) introduced a generic intelligent system applying similarity algorithm to find and match customers' preferences to previous cases that shared common attributes the most [18]. The paper also summarized a list of factors that are influential to decision-making (e.g., travel budget, local knowledge, hobbies, gender, age, etc.), of which "origins of customers" plays an important role [18]. Even though the CF algorithm was proven to draw good results on the POIs ranking tasks for new users, the fact that it requires historical data of actual cases with travelers' profile in details, disprove its practicality in the early stage (e.g., "cold start" problem). Kai, Huagang, Peng, Nenghai proposed the use of geo-tagged textual information of photos of various attractions retrieved from Panoramio with CF to recommend POIs to new users. Even though the research claims that its approach can generally tackle the "cold start" problem of CF, it does not consider factors such as users' sex, nationality, gender, etc., which are claimed to be very important in the recommendation process [18]. More importantly, since CF algorithm usually ends up only with a list of places that are highly attractive to a certain user's profile, the nature of the algorithm does not consider the geographical influence, or a whole route planning with logistics consideration, which is claimed to be very essential for a realistic system [17]. Hence, some previous researches with solely CF implementation failed to demonstrate their system's practicality.

Data sources are indispensable in any recommendation research. Even though there have been a number of previous researches investigating recommendation systems in tourism, there is yet any commonplace regarding metadata (e.g., geographical coordinates), sources, and type of input data. This might result from not only whether a list of POIs or whole itinerary is being achieved by the system, but also which algorithms are being employed. Particularly, some researches make use of collaborative filtering technique on previous tourism cases collected from a local tourism agency of a specific city to suggest a touristic planning for the user, while others make use of available data sources such as public photos from Flicker [17], etc., to heuristically picture different cases of sightseeing trips at some famous destinations around the world, which are then can also be fed into the CF algorithm. As for travelers' preferences, they are usually extracted from social network platforms such as foursquare or OpenSocial API [14].

System architecture of previous system mainly employs formal mobile application system, comprising of user preferences or queries input module, core algorithmic recommendation engine module, customized recommendations module, and a presentation module (visualized maps, GIS, etc.) [14]. There are just a few pieces of research which emphasize detail technical concerns regarding the performance of the implementation platform such as network connectivity (e.g., Wi-Fi network), or battery consumption. Real-time feature, further enhancing the recommendation system with the flexibility to alter the suggested tour trip according to changes in user's context or the environment (e.g., user's current location, unexpected breaks, weather changes, attraction closure, etc.), seems to be quite unique and realistic module.

2) A Need of Localization of Tourism Recommendation System

Recommendation system in tourism is very different from other fields such as movies, or music recommendations. Besides the fact that user's preferences in the case of tourism recommendation are much harder to capture because of a big gap in the activity frequency, context is also another perspective that distinguish the two cases. Particularly, recommended objects such as films, music, news, etc. are very generic to any user's situation or preferences; hence the recommendation process can work well with solely user's preferences and object's contents. In contrast, the practicality of recommendation system depends much on the context, or the environment such as the country destination, regional regularity, public transportation, or seasonal factors. The variety of input data spaces of previous researches can further support this statement. Many researchers, even though tend to build a generic recommendation system, exploit data sources that seem to be only available from certain resources (e.g., historical cases from local tourism agency) or for a certain regions (e.g., public sightseeing photos are only sufficient to famous places). The disjoint in the data input or data sources is possibly one of the reasons for the current lack of evaluation tests for this kind of system, which usually requires formal field studies [14]. Nevertheless, most of the previous system concentrates more on the theoretical approach of the recommending engine, making the applicable scope become too general that unexpectedly reduce the practicality of actual implementation. Thus, we call for a demand of touristic recommendation systems that are more localized, or tailored to a specific regional area such as city, or country. Only by shifting the focus to this direction, we can ensure that future research on recommendation system would be able to not only share and reuse different resources (user's preferences, data input space, test cases, etc.) but also better bridge tourists with regional tourism promotional activities and policies.

3) An Innovative Recommendation System for Foreign Tourists in Japan

Our research investigates four out of seven new prospects in tourist touristic planning service, which are clearly stated in a thorough survey research in the field [14].

a) Most of the related work proposes more of a generic program rather than systems that could tackle specific situations in a region or country. However, countries such as Japan impose unique difficulties, such as language barriers, complex train network, and unique culture to the tourists. Moreover, by targeting a specific tourism region in Japan one can focus on leveraging local advantages such as well-developed subway networks, bus, a variety of cultural events, etc. in building a better recommendation system. Since to the best of our knowledge, there are not many tours planning recommendation system designed specifically for Japan targeting foreign tourists, our research may be considered a pioneer along this direction.

b) Our system integrates logistic planning during the recommendation process. Particularly, aside from user's context (e.g., current location, weather condition), we consider train public transportation as a foundation for generating recommendation result. Such perspective is emphasized as a necessity in this line of research [19], which is different from public transportation advisory services (e.g., [20],[21]) that can be requested by the user after a list of attractions or services suggestion. Furthermore, because our suggested attractions are clustered and selected simultaneously with geographical influence, our approach ensures the practicality of generated solution while improving user's satisfaction [17].

c) We examine a "unified attractions/tourist services recommendation" [14] with time and budget constraints. Because of the availability and diversity of input data sources, we are able to suggest not only attractions or point of interests, but also services such as restaurants, cultural events, foreign tourist support spots, etc. Since services cost much more than sightseeing scenarios, which usually do not put much financial concerns on the users, the involvement of services recommendation would further raise a concern about budget constraints. This matter is also addressed and resolved by our system, which generates results that satisfy user's constraints including budget and time window.

d) Since we consider the practicality of the system as our priority in the research process, we thrive to make the solution as realistic as possible. As clearly stated by Damianos at el.: a realistic tour had better provide travelers with breaks, either for a meal or resting in a nearby parks [14]. This function is also equipped in the prototype of the system in a way that the user can interrupt the touring anytime for a break, and the system would immediately suggest parks, coffee shops, restrooms, etc. that are adjacent to the user's current location.

III. Methodology

This research proposes an empirical and practical approach to the described context above so as to support individual tourists, tourism policy makers, as well as tourism managers and provide more fruitful experiences to foreign tourists. The rest of the paper is as follows.

A. Data Collection

1) Data Source and Description:

A collection of appropriate data is crucial for any successful projects, and this is especially a conundrum in the field of tourism where relevant data is scattered across wide networks rather than being aggregated at a fixed location. Not to mention that tourism activities usually involves other domains such as public infrastructures, policy and promotional activities, weather forecasting, etc. which picture tourism related data as much prevalent. Particularly, the result shown in this paper is based on the data that were obtained from four relevant websites and services; namely, Navitime (http://www.navitime.co.jp), TripAdvisor (http:// tripadvisor.com), and Jalan (http://www.jalan.net). The Navitime*.com* website contains a list of relevant information on all of the JR train stations in Japan (e.g., their names, addresses, geographical coordinates), which have already been categorized into 47 different prefectures and operation companies. The website has also an extensive list of tourism guidance offices and regional specialty shops located throughout Japan [22]. The TripAdvisor.com is a worldwide famous entity for its rich database on tourism products. Its portal provides an extensive list of the most famous attractions in Japan, which are ranked by actual travelers, and are neatly classified into various categories such as “sight & landmarks”, “natural & parks”, or “museums” [23]. The Jalan.net is a domestic tourism portal that provides insight information on regularly updated cultural events and intangible attributes such as average visiting time, restaurants’ price and rating, best time to visit, etc., that have been collected from travelers [24]. Lastly, Yahoo API is used as a reliable concurrent weather forecasting.

2) Collecting Data Using Web Scraping Technique

To equip the system with the capability of suggesting a unified combination of both tangible and intangible tourism products, and incorporating public transportation during the process, the work of collecting relevant data from the three stated online sources is indispensable. However, the fact that the three sites do not have a unified structure and their relevant data records (e.g., details regarding a specific place) are displayed on a separate web page; makes the task quite cumbersome. Fortunately, this problem can be resolved using web-scraping technique, a programmatic approach that enables generating numerous virtual web agents that are capable of interacting websites and extracting their data in a systematic manner. As such, this research employs Ruby programming language to build a program that utilizes “Watir” and “Nokogiri” library to facilitate the task. Moreover, “Parallel” library is also used to shorten the execution time by processing multiple URLs simultaneously. All of the retrieved data is subsequently stored in a local MySQL database, which acts as an intermediary data library for the recommendation system.

B. Overall System Architecture

This section provides an illustrative example of situations in which our system could be beneficial.
Instead of focusing on a total tour recommendation package, which usually suggest a fix accommodation facility (e.g., hotel, hostel, etc.) for the whole trip and lasts for couple of days, our approach centers on a practical scenario that involves two separate module. While the first module gets executed once, the second module operates dynamically and continuously adjust the recommendation to real-time needs:

a) The first module involves generation of short-term touristic trip (e.g., within a day) which passes by various physical attractions, places, sceneries, or cultural events in accordance with public transportation starting from an origin (e.g., user’s current location), and satisfying other preferences (e.g., favorite attractions categories, budget and time constraints, concurrent weather condition, etc.).

b) Second process more or less operates in real-time and recommends the so-called “on-demand services” such as restaurants (e.g., lunch, dinner, etc.), resting places (e.g., nearby coffee shops, parks, public restrooms, etc.), and attend to immediate inquiries and demands of the users (e.g., passing by tourism office in the vicinity). These are called “on-demand services” since their occurrences are not fixed and depend on different contexts. This is because, different people might have different desires for meal-time, break-time or a nearby vicinity maybe unpredictable, etc. Furthermore, since during the first process, the system also incorporates user’s preferences that may dynamically change the trip (e.g., weather condition, etc.), the real-time adjustment capability of the second module could become quite handy.

Figure 1 shows the system flow and the next subsections explains the methodology used in its implementation.

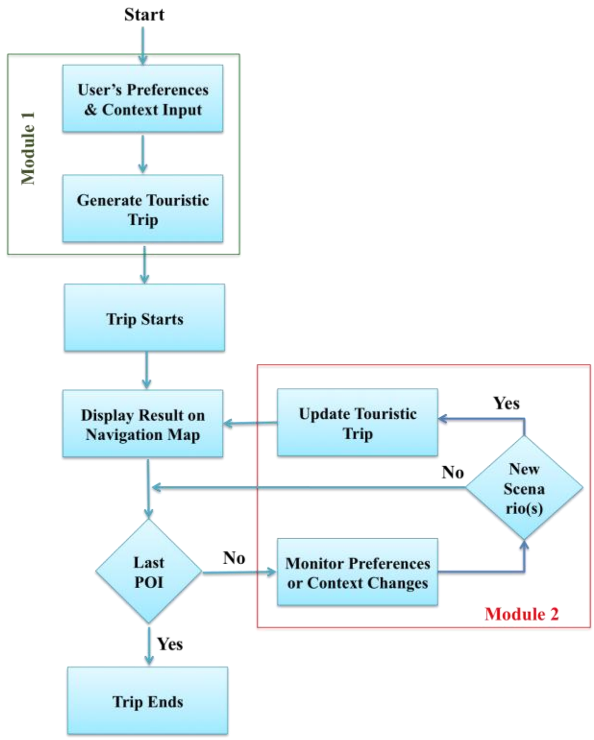

Fig. 1. System flow and its two modules (shown in green and red color).

C. Module 1: Touristic Trip Recommendation Using Cluter-first-route-second Heuristics.

Our system employs cluster-first-route-second which is inspired by the nature of touristic trip in Japan. Public transportations such as trains, subways, and buses are highly developed and utterly integrated into daily lives of local people. This has consequently shaped many aspects of the tourism industry, notably those in the touristic planning process. Particularly, as an effort to leverage the well-established network of more than 10,000 train stations in Japan with the 1-day free train pass ticket to reduce huge amount of transportation cost, a typical touring trip usually involves commuting through train stations, stopping at some of them and walking to surrounding areas, nearby sceneries and attraction sites. In fact, this observation aligns with many of researches on power-law distribution which predicts the user’s likely favor nearby attractions rather than those that far away [25][26]. This is also further supported by the argument that travelers, after touring one attraction, have a tendency to move to an adjacent point of interest [17]. Therefore, in the linear programming context, looking for a sequence of point of interests out of a selection pool of thousands together with an optimal travelling route among train stations can simply be interpreted as searching for a route that passes by a sequence of train stations which satisfies user’s touristic preferences and context. The following procedures explain details of this process.

1) Category and Weather Based POIs Filtering

From a pool of all of available attractions, we first filter them based on the user’s selected categories and current weather information (if requested) so as to get a list those

POIs that meet the user's preferences. Depending on the weather conditions (e.g., raining, snowing), outdoor POIs such as natural parks or bridges may not get included in the result list. When this category option is unselected, the algorithm outputs a list of the most popular places in the examining area.

2) Heuristic POIs Clustering Based on Nearest JR Train Station and Walking Preferences:

In this step, our approach subsequently clusters all of the filters POIs into different geographical groups associated with a nearby JR train station, hence the distance among the station and attractions located inside each of the clusters become the walking distance travelled by the user. Walking preference is one of the interesting factors that the paper takes into consideration. Different people have different walking preferences when traveling which can be traced to differences in demographic and cultural preferences [27]. So, the most likely problem that has to be addressed is how to search for POIs that are located within a favorable walking distance from a nearby JR station. Therefore, the measurement of walking distances among thousands of the examining stations and attractions is indipensable, which is yet very expensive in terms of computation. Instead of using available public API for finding the nearest station to/from a place which would require long processing time to go through thousands of places, this paper suggests the Euclidean measurement to heuristically facilitate the categorization. Undoubtedly the Euclidean measurement is not a perfect approach for calculating the walking distance among places; the trade-off between its results and performance is acceptable, however.

As shown in Figure 2, on a 2-dimensional geographical map, from each of the JR station, a circle, having the relative walking distance as its radius, can be drawn to indicate which POIs can be reached from the station. As a result, one can easily compare relative distances when traveling from places to places quite efficiently and obtain lists of nearby stations for each of the POIs for later access. However, since walking distance preferences vary among users, an addition requirement to the user's input is necessary to retrieve the suitable walking distances used in the clustering process. So, we categorize users' walking distance preferences into three levels: "Not really", "Fairly", and "Definitely". Because there are trade-off between preferable walking distance and number of POIs can be covered during the clustering process, walking distances associating with each of three levels are set depending on different touristic areas to ensure that an appropriate portion of POIs can be covered. To illustrate,

Table 1 describes the result of POIs coverage in percentage in accordance with seven different radius distance used during the clustering process within eight different major cities and downtown areas as recommended by the JNTO. The finding shows that three different levels of walking preferences (shown in bold) are selected.

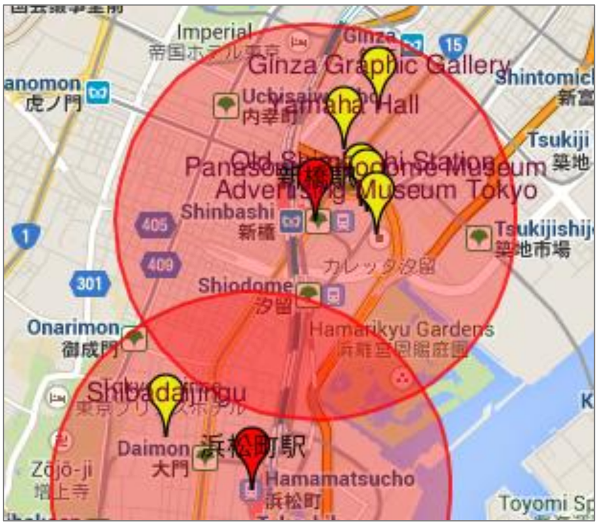

Fig. 2. Red circles with 800-meter radius are drawn to estimate the Euclidean distance from each of station (red markers) to nearby POIs (yellow markers).

TABLE I.
CLUSTERING POIS BY NEARBY JR STATION

Distance (meter)		**300**	**500**	**800**	**1000**	**1500**	**2000**	**3000**
Tokyo	Cover	18	32	**53**	61	**78**	**88**	95
	Top 20	11	21	**35**	41	**53**	**60**	64
	Top 50	14	25	**42**	49	**63**	**72**	77
	Top 100	15	28	**46**	53	**69**	**77**	83
Yokohama	Cover	6	20	**41**	50	**57**	**63**	88
	Top 20	4	11	**25**	30	**36**	**41**	63
	Top 50	5	16	**33**	39	**46**	**51**	73
	Top 100	5	18	**37**	44	**50**	**56**	79
Kyoto	Cover	1	4	**11**	**16**	**24**	63	81
	Top 20	1	3	**7**	**11**	**15**	49	60
	Top 50	1	3	**9**	**13**	**18**	54	68
	Top 100	1	3	**10**	**14**	**20**	58	73
Osaka	Cover	19	33	**63**	**72**	**94**	96	99
	Top 20	13	22	**42**	**51**	**69**	71	72
	Top 50	16	27	**51**	**60**	**79**	81	84
	Top 100	17	29	**55**	**64**	**85**	87	90
Kobe	Cover	11	**42**	**64**	**70**	75	77	83
	Top 20	8	**32**	**47**	**50**	55	57	60
	Top 50	9	**36**	**54**	**58**	64	65	70
	Top 100	10	**38**	**58**	**63**	69	70	76
Fukuoka	Cover	5	9	**17**	22	**33**	**44**	83
	Top 20	1	4	**10**	12	**19**	**26**	55
	Top 50	3	6	**12**	17	**25**	**36**	70
	Top 100	4	7	**14**	18	**28**	**38**	75
Nagoya	Cover	5	9	16	**26**	**46**	**72**	87
	Top 20	3	6	11	**18**	**34**	**53**	64
	Top 50	3	6	13	**21**	**38**	**60**	74
	Top 100	4	8	14	**22**	**41**	**65**	79
Sapporo	Cover	7	26	34	38	50	60	72
	Top 20	3	19	21	23	31	39	48
	Top 50	6	22	26	28	39	48	59
	Top 100	7	25	30	33	43	52	63

Furthermore, by clustering POIs, each of the clusters is identified not only by the central station, but also by a ranking score. As implied by its name, the ranking score represents the relative popularity of the cluster and can be calculated by summing up the rankings of all the included POIs. However, the lower the score, the better its cluster's ranking. Even though our approach is not focusing on the most well-known attractions, ranking is yet an assistive metric in the following searching process.

3) Heuristics Based Clustering and POIs Searching

By using the greedy algorithm, we iteratively search for adjacent stations surrounded by POIs that posses the best ranking positions. As illustrated in Figure 3, this searching process can be examined by modeling it on a directed graph. Each vertex of the graph describes location of a train station, which can either be the one nearest to the user's current position, or to the desired destination, or the ones to pass through on the trip. Each of those nodes will have a weighted ranking score. The lower the score, the more satisfying it is. Edges connecting vertexes represent the travel duration between them. The overall goal in solving this problem is to accumulatively select the best available cluster and included POIs that are located as close to the current position as possible, and satisfy all of other input constraints such as available budget and time window.

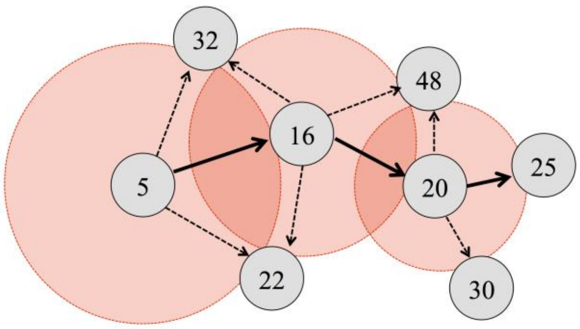

Fig. 3. Multiple circles with gradually increasing radius is drawn from each of vertex (small gray circles) to recognize its nearest next best station.

a) Searching for next best cluster: Through the greedy algorithm, we try to obtain a good tour by searching locally from one vertex to another in order to choose the next best vertex to visit. The next best vertex (or station), must be as close to the current vertex (or station) as possible, and posses the most satisfied weighted ranking score (e.g., Figure 2). In order to find nearby vertex from a specific location, a similar approach being used during the POIs clustering process that employs Euclidean measurement can be done. This method to replace extensive distance calculation between all of the vertexes hence reduce the computational burden on the system.

b) Searching for the next best POI: At each of the vertex, greedy algorithm with local search technique is subsequently used to select the best POIs located within a specified preferred walking distance from central station. So as to pick the POIs that both has high rankings, and satisfy the constraints variables, all of the POIs is ascendingly sorted by their rankings, and the selection process is done from the top in accordance with the remaining budget and time. Following each of the POI selection, those constraints variables are updated. The search would stop once either all of the constraints are met or a balanced number of POIs from each of the desired categories is obtained. A sample result is illustrated in Figure 4 as follows.

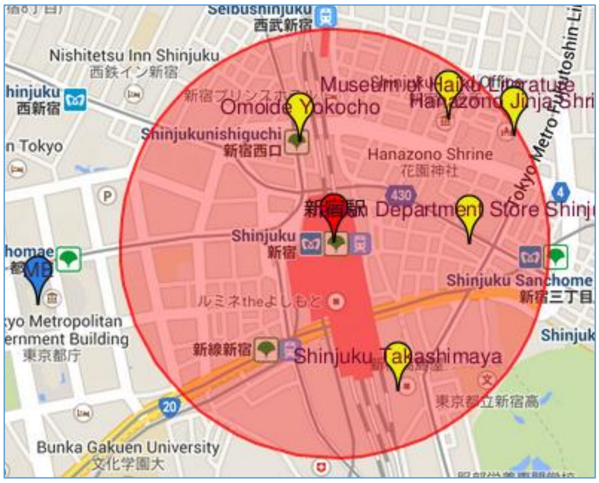

Fig. 4. Search result of a suggested trip of about 4 hours in cloudy weather condition, starting from the Sofitel Hotel Tokyo (blue marker), visit a nearby Shinjuku eki (red marker), and subsequently walks through each POIs (yellow marker) which include "Isetan department store", "Omoide Yokocho", "Hanazono Jinja", "Museum of Haiku", and "Shinjuku Takashimaya", elements of specified categories "Landmarks", "Specialty Museums", "Religious Sites", and "Shopping".

4) Route Construction & Optimization

After obtaining a list of suggested POIs and its associated intermediary stations, an optimal sequence of traveling route that passes through all the places is investigated. Basically, the overall route will comprise of the traveling order through stations, and the walking order within a group of POIs located at a single station. Despite the fact that many genetic algorithms such as tabu search, etc., [28][29] facilitate the route optimization procedure, we used Google Maps API because of its proven success and the practical nature of our research approach. Apart from the optimal route passing by multiple stations and POIs that can be easily calculated with the support of the API, in order to present the most appropriate touristic planning to the user, the system undergoes two main routing and optimization procedures as follows.

a) Taking traveling time among stations, and POIs into account: Because this factor was not considered during the local search to enhance the performance, a more thorough time duration needed to finish the whole trip is calculated. Since travelling time between two station might vary during the day, depending on the type of train that the user gets on, it is can be requested by sending parameters including estimated departure time, the origin and as well as the destination station to Google API service. Similarly, the walking time inside each of the clusters is also obtainable. By doing this, we can reconsider the initial solution by eliminating any redundant POIs, if necessary, backward from the trip to better match with the original time constraint.

b) Re-matching POIs: Through observations, after the searching process, there might be cases where POIs are more accessible from nearby POIs or a station belonging to another geographical cluster (e.g., Figure 5). In such cases, an optimization process is designed to iteratively examine each of the selected POIs and move them to a more appropriate central station, and remove any of the redundant stations from the initial route. Moreover, the optimization process also considers POIs that initially eliminated in the clustering process due to further distance compared with the user's preferences as possible replacements.

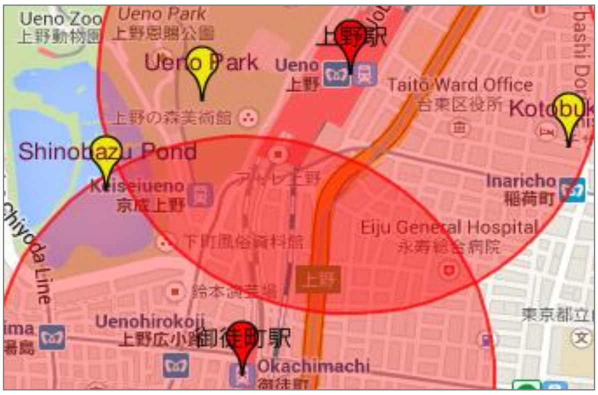

Fig. 5. A better solution can be achieved by matching the "Shinobazu Pond" POI (yellow markers) to the upper station to reduce the commuting time between the two stations (red markers).

The illustrated POIs clustering and filtering, heuristic local search, post-optimization steps of the first module are further elaborated by pseudo-codes in Figures 6, 7 and 8 respectively.

D. Module 2: Real-time Preferences, Context Monitor and Recommendation

Regardless of how good an initially generated result may be, a good recommendation system has to consider the possibility of incorporating dynamic changes in both user's preferences and the environmental impacts. This is especially applicable in the field of tourism. For instance, the user might want to add a new POI during the trip, take a rest, have a meal at nearby restaurants, look for restrooms; or suddenly it might start raining. The fact that our problem deals not with a single user but travelers from different countries with varied cultural preferences and habits, there is definitely a need for a certain level of flexibility in the system, which can make it more compatible to such capriciousness.

We categorized all of the possible internal (e.g., from user's perspective) and external changes as follows:

1) POIs related changes: These happen when the user wants to change the type of attractions during the trip, add, remove, or change an attraction on the list. This category also considers effects of sudden weather changes (e.g., raining, sunny, etc.).

2) Additional services demands: On an on-going trip, the user might request the system to recommend surrounding services including culinary services (e.g., restaurants, coffee shops, etc.), resting places (e.g., parks, rest room), tourism information office, free wifi-spots and so on.

Such data can be retrieved by web scraping technique from some of well-known local websites such as Navitime.net (for public rest rooms, tourism information

```
Result: Array filteredPOIs
foreach POI in POIs do
    if POI match User.category then
        if POI match currentWeather then
            ADD POI to filteredPOIs
        end
    end
end
```

Algorithm 1: POIs Filtering on Category and Weather

Fig. 7. Pseudo-code for step 1: Filtering POIs basing on user's preferences on choices of category and concurrent weather condition.

```
Result: Array finalPOIs, finalStations
nextStation = User.currentLocation;
while checkTripSatisfied(finalPOIs) == false do
    POIS = findPOIsNearStation(filteredPOIs, nextStation);
    POIS = sortByRankAscending(POIS);
    foreach POI in POIS do
        if checkTripSatisfied(finalPOIs) == false then
            if POI.timeNeeded < User.timeConstraint AND
            POI.budgetNeeded < User.budgetConstraint then
                User.timeConstraint -= POI.timeNeeded;
                User.budgetConstraint -= POI.budgetNeeded;
                POI.nearByStation = nextStation;
                ADD POI to finalPOIs;
                if nextStation not in finalStations then
                    ADD nextStation to finalStations;
                end
            end
        end
        else
            Exit For;
        end
    end
    nearStations = findNearStationsFromLocation(nextStation);
    if nearStations is not null then
        nextStation = findStationHighestScore(nearStations);
    end
    else
        Exit While;
    end
end
```

Algorithm 2: POIs and Station Searching

Fig. 6. Pseudo-code for step 2: Local searching for POIs and Station with user's constraints using the Greedy algorithm.

office), FreeSpot.com (for public free Wi-Fi spots), Jalan.net (for culinary services). For its achievement, a background module is set-up in the system to continuously capture and response such changes at a specific interval of 5 seconds. By re-using part of the 1st module to generate touristic result basing on the newly captured input, this module equips the system with more flexibility and responsiveness. The details of practical implementation architecture, and as well as some technologies can be employed to resolve some of recognized connectivity and energy consumption of portable devices bill be addressed in the next section of the paper.

E. Prototype Implementation

A prototype of the proposed system on web-based mobile platform is built and tested so as to further explain some empirical perspectives on its practical implementation, namely system architecture, GUI, connectivity and energy consumption.

1) System Architecture

```
Result: Array finalPOIs, finalStations
foreach POI in finalPOIS do
    bestNearByStation = findBestNearbyStation(POI, finalStations);
    if POI.nearByStation != bestNearByStation then
        POI.nearByStation = bestNearByStation;
    end
end
foreach POI in allPOIsInArea do
    if POI not in finalPOIs then
        replacablePOI = findReplacablePOI(POI, finalPOIs);
        if replacablePOI then
            replacePOI(POI, replacablePOI, finalPOIs);
        end
    end
end
removeRedundantStations(finalPOIs, finalStations);
```

Algorithm 3: Post Optimization

Fig. 8. Pseudo-code for step 3: Optimizing the list of stations and POIs by re-matching each of POIs to their most optimal station, searching for better POI in the area, and removing any redundant stations.

Overall, the system architecture comprises of a central server with an installed database such as MySQL, Mongo DB, etc., and several clients connected. The proposed recommendation algorithm can be programmed to compile under native mobile systems (e.g., iOS, Android, Windows phone) and be easily loaded into tourist portable devices (e.g., smart phones, tablet, etc.), together with its language preference (e.g., English, Vietnamese, Chinese, Korean, etc.) and other resource files. This software acts as a client, connects to the server and downloads or updates data packages on user's request, or on a regular basis. Different touristic zones and regions (e.g., Tokyo, Kyoto, Osaka, etc.), will have a distinguished data package stored in the database and selectable upon the user's preference. For demonstration purpose, we have implemented the proposed system as a web-based mobile application. Its GUI and various functions are introduced in the next section.

2) Prototype GUI & Integrated Functions

This section describes the implemented web-based prototype of the proposed trip recommendation application for foreign tourists in Japan. Default names of all stations and POIs are shown in English, but the language is intended to be configurable so as to provide a much friendlier user interface.

As shown in Figure 9, when activated, the software asks its user for initial input preferences including trip origin (e.g., current location, specific hotel, train station, etc.), available budget (in Japanese Yen), time window (in hours), level of walking preference (e.g., Definitely, Not Really, or Not At All), whether or not to consider weather forecast, and preferable types of destinations (e.g., those shown in Figure 10). When every parameter is specified, the user can execute the program by clicking the button "Experience Japan" (top button in Figure 9) and receive a recommended touristic trip result, a sample of which is shown in Figure 11.

Since system comprehensively utilizes Google Maps library to visualize the mapping layers above the generated result, the program can also be used as a real-time navigation

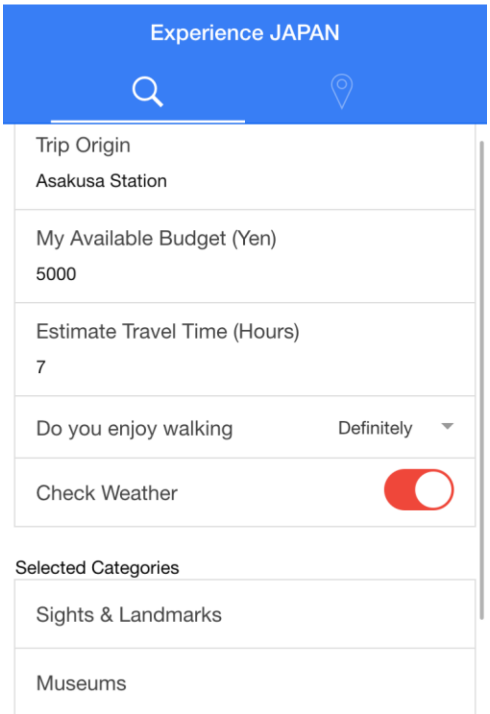

Fig. 9. GUI for preferences input from the user. In this case, the trip starts from user's current location, which is "Asakusa Station" in Tokyo area. A budget of 5000 Yen and 7-hour availability are also specified. "Definitely" is selected as walking preference. The trip would be generated basing on the current weather condition

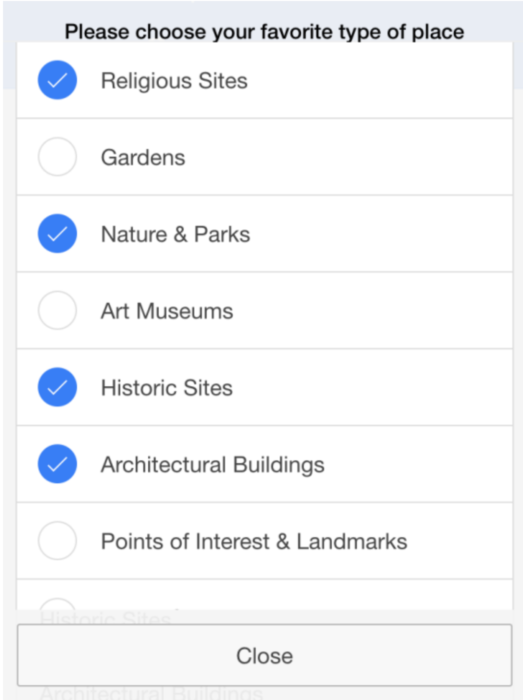

Fig. 10. An example of modal box showing different options for favorite type of destinations.

system during the trip. An array of functional buttons, placed at the bottom of the result screen, enable the user to get the trip's information (Info), look for nearby services or places such as restaurants, tourism office (Search), center the current trip on the navigation map (Center), and tell the system to start navigating the trip from the current location (Start).

Figure 12 illustrates the info board displaying all the information and available actions relevant to the trip.

Fig. 11. GUI of trip recommendation being integrated with visual navigation map. Blue marker represents user's current location. Red markers show the location of train stations, and yellow one locates POIs during the trip.

Through this interface, its user can check all of the POIs and their traveling order, etc., and as well modify the trip by adding or removing a specific attraction, or re-order them. Other input preferences such as the available budget, time, and favorite categories can be altered on the input screen. All of the changes will be updated on the concurrent trip by considering its user's current context (e.g., new location, remaining budget, etc.).

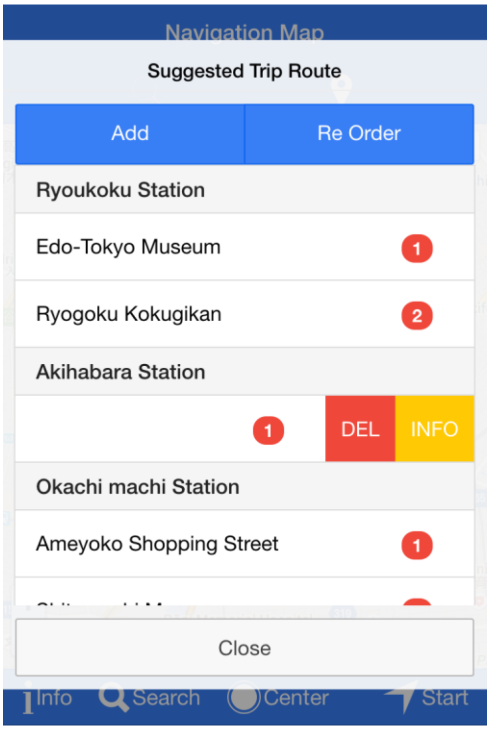

Fig. 12. GUI of modal box displaying the information of the recommended trip. Each of POI is categorized into list heading by the name of the central station. Various tasks can be easily done by the user such as swiping left on an item to delete or search info of a specific attraction, adding a specific place, or re-order the route.

During the trip, the user can activate "Search" function on the main navigation screen to discover the surroundings of the current location. Through a relevant popup GUI shown in Figure 13, its user can find nearby restaurants or coffee shops for lunch or dinner; look for a nearby tourism guidance office for additional help, search for accessible Wi-Fi spots, and seek out for a short resting place (e.g., Natural Park). As shown in Figure 14, each of the newly discovered service location will be indicated on the main map with distinct markers for easier navigation.

3) Connectivity & Energy Consumption

As it is evident from our presented GUIs, the proposed system not only plays the role of trip planning, but also functions as a navigation system throughout the trip. Therefore, the tasks of processing location services (e.g., GPS), database update, and weather forecast impose concerns about the connectivity and energy consumption of the user's device. Battery consumption is a paramount aspect in practical mobile applications [30], especially in real-time or near real-time implementation. There is a strong correlation between the communication protocol and energy consumption of mobile devices, which is actually dependent on the type of communication method (e.g., Wi-Fi, 3G) [30]. Because our system targets foreign tourists in Japan, who most likely visit the country for the first time, there is no guarantee that a sustainable Internet connection is always available, regardless some of the recent efforts being made by

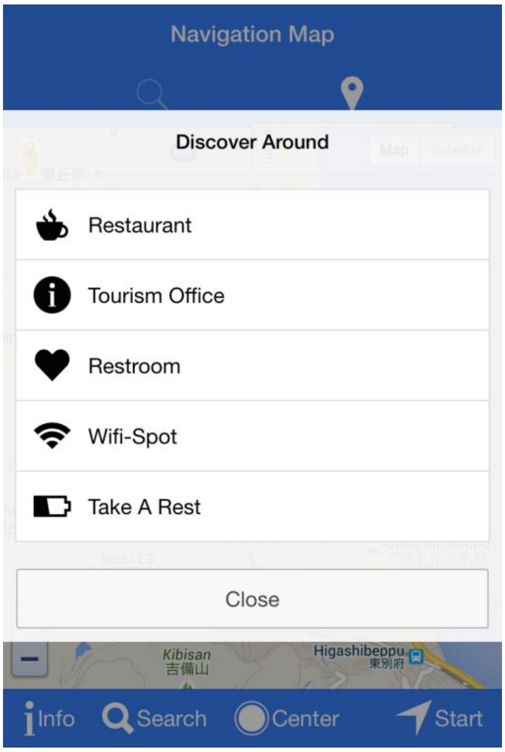

Fig. 13. GUI of "Search" or local discovery function, including the searching for nearby restaurants, tourism office, restroom, Wi-Fi spot, or a surrounding park for taking a rest.

Fig. 13. As a result of the search function, the map shows a green marker representing a tourism guidance office located nearby user's current location (e.g., Blue marker)

the government and domestic service providers. This hence raises our concern about which communication protocol should be employed for the touristic recommendation software, and whether all of the necessary processing should be done on the client or at the server site.

With the recent prevalence of HTML 5 and Web Socket communication technology, we propose the use of Web Socket as the main data transmission protocol (e.g., Table 2). Even though XMPP protocol consumes less energy than Web Socket in the case of Wi-Fi connection, considering its overall high performance and low energy level consumption (e.g., rank 1st in 3G and 2nd in Wi-Fi connection [30]), Web Socket is selected as the single protocol method for a consistent implementation throughout the system.

Furthermore, in order to minimize extensive processing tasks for the mobile devices, the searching and trip generation procedure (module 1), particularly when an Internet connection is unavailable, as well as the regular web-scraping process to update the database, should be done on the central server. This can make the system more efficient, since we could both take advantage of a well-performing server infrastructure and as well optimize the communication process.

TABLE II.
PROPOSED COMMUNICATION METHODS

Tasks	Wi-Fi	3G	Internet Unavailable
Searching & Generating Trip	Server (Web Socket)	Server (Web Socket)	Client
Database Update	Web Socket	Web Socket	Not available
Location Service	GPS	GPS	GPS

IV. DISCUSSION

A. Analytical Result

This subsection examines the result of our various testing attempts in order to demonstrate the practicality and performance of the proposed tourism recommendation system.

We tested our approach on eight different major cities and their downtown areas as recommended by the JNTO which are Tokyo, Osaka, Kyoto, Kobe, Nagoya, Sapporo, Fukuoka, and Yokohama. The results are shown in Tables 2 – 9 and for each of the area, the following three practical scenarios and parameters are presented.

1) General Usage: Travel from a main station to up-to 10 different POIs categorized as "Museum", "Religious Site", or "Historical Site". "Definitely" or "Not Really" is used as walking distance preference.

2) General Usage / Short-term business with long-time window: Travel from a main station or famous hotel (e.g., top 50 regional hotel according to TripAdvisor) to up to 10 different well-known POIs. "Definitely" or "Not Really" is used as walking distance preference.

3) Short-term business with short-time window: Travel from a famous hotel (e.g., top 50 regional hotel according to TripAdvisor) to up-to 4 different popular POIs in the nearby area. "Not at all" is used as walking distance preference.

Even though the suggested trip depends much on the trip origin, the consideration of some main stations in the scenarios partly strengthens the analysis. In each of the test case, percentage of POIs coverage in the top 50, 100, as well as the number of stations for the suggested trip are used as a metric for examination. Performance wise, the system, in general, comes up with the result in less than 10 seconds, but it is much faster for the case of short-term trip with a short-time window. In our view, this scenario is pretty useful in a practical sense, since people attending to short-term business trips, conferences, or transit passengers, etc., would like to tour the country as much as possible, something which has not been recognized and implemented in many commercial tourist packages.

As can be noticed, except for Kyoto city, 60% to 80% of the suggested POIs belong to top 50 and 100 regional rankings, respectively. This demonstrates the balance between the popularity of POIs and distance needed to travel, which turns out to be 3 train stations. Although this paper focuses on JR train system as the main transpiration system, for the case of Kyoto city bicycles and buses turns out to be more convenient, something which should be considered in future work for a much more comprehensive approach.

TABLE III.
TESTING RESULT FOR TOKYO AREA

Criteria	***Result***		
Origin	Tokyo Station	Harajuku Station	Park Hyatt Tokyo Hotel
POIs	Tokyo Central Railway Station, Edo Castle Ruin, Bridgestone Museum of Art, Mitsubishi Ichigokan Museum, Mitsui Memorial Museum, Nihonbashi, Suitengu Shrine, Fukutoku Shrine, Takarada Ebisu Shrine, The National Museum of Modern Art	Shinjuku Gyoen National Garden, Meiji Jingu, Yoyogi Park, Shibuya Pedestrian Scramble, Shibuya Center-gai, Omotesando, Aoyama Cemetery, Shibuya Fureai Botanical Center, Harajuku Takeshita-dori, Meijijingu Gaien	Tokyo Metropolitan Government Office, Shinjuku Golden Gai, Omoide Yokocho, Shinjuku Camera Town
Top 50	20%	**60%**	50%
Top 100	40%	**80%**	50%
Stations	1	3	1
Cluster Radius	1000 meters	2000 meters	800 meters

TABLE IV.
TESTING RESULT FOR YOKOHAMA AREA

Criteria	***Result***		
Origin	Minatomirai Station	Hotel Monterey Yokohama	BEST WESTERN Yokohama
POIs	Cupnoodles Museum, Mitsubishi Minatomirai Industrial Museum, Japan Coast Guard Museum, Iseyama Kotai Shrine, Nipponmaru Memorial Park, Kanagawa Prefectural Museum, Kuan Ti Miao Temple, Yokohama Port Opening Hall, Ma Zhu Miao, Yokohama Museum of Art	Osanbashi Pier, Yokohama Chinatown, Yokohama Stadium, Iseyama Kotai Shrine, Kanagawa Prefectural Museum, Diplomats House, Bashamichi Shopping Destrict, NYK Maritime Museum, Yokohamabashi Shopping District, Yokohama Customs Headqurters	Yokohama Chinatown, Yokohama Stadium, Kanagawa Prefectural Museum, Diplomats House
Top 50	70%	**60%**	50%
Top 100	90%	**90%**	100%
Stations	3	3	2
Cluster Radius	1500 meters	1000 meters	800 meters

TABLE V.
TESTING RESULT FOR FUKUOKA AREA

Criteria	***Result***		
Origin	Hakata Station	Hakata Station	Hotel Nikko Fukuoka
POIs	Kushida Shrine, Tochoji Temple,	Kushida Shrine, Tochoji Temple,	Tochoji Temple, Jotenji Temple,

Criteria		***Result***	
	Jotenji Temple, Hakata Machiya Folk Museum, Hakata Traditional Craft Center, Genko Historical Museum	Jotenji Temple, Fukuoka Asian Art Museum, Shofukuji Temple, Yatai, Hakata Machiya Folk Museum, Rakusuien, Seiryu Park, Itazuke Iseki	Hakata Sennen no Mon, Kakueiji Temple
Top 50	50%	**60%**	50%
Top 100	100%	**80%**	50%
Stations	6	1	1
Cluster Radius	1000 meters	1500 meters	800 meters

TABLE VI.
TESTING RESULT FOR OSAKA AREA

Criteria		***Result***	
Origin	Osaka Station	Osaka Station	Swissotel Nankai Osaka Hotel
POIs	Museum of Oriental Ceramics, Osaka City Central Hall, Midosuji Street, Tsuyunoten Shrine, Osaka Prefecture Library, Namban Culture Center, Dojima Yakushido, Taiyuji Temple, Osaka Science Museum, Bank of Japan Osaka Branch Old Building	Floating Garden, Hep Five Ferris Wheel, Museum of Oriental Ceramics, Osaka City Central Hall, Tsuyunoten Shrine, Nakanoshima, Shin Umeda City, Midosuji Street, Shin-Umeda Shokudogai, K itashinchi	Dotonbori, Shinsaibashi, Hozenji Yokocho, Sennichimae Doguyasuji Shopping Street
Top 50	20%	**60%**	100%
Top 100	70%	**100%**	100%
Stations	2	2	0
Cluster Radius	1000 meters	1000 meters	800 meters

TABLE VII.
TESTING RESULT FOR KYOTO AREA

Criteria		***Result***	
Origin	Inari Station	Kyoto Station	Hotel Granvia Kyoto
POIs	Fushimi Inari Shrine, Tofukuji Temple, Unryuin, Sennyuji Temple, Komyoi, Kyoto Municipal Science Center For Youth, Sekihoji Temple, Sanjusangendo Hall, Kyoto National Museum, Toyokuni Shrine Karamon	Sanjusangendo Hall, Tofukuji Temple, Umekoji Steam Locomotive Museum, Nishi Honganji, Sagano, Kyoto National Museum, Sumiya Motenashi Cultural and Art Museum, Toyokuni Shrine Karamon, Kozan Temple, Higashi Honganji	Shinsengumi Mibu Tonjo Kyuseki, Mibudera, Old Maekowa Residence, Chishaku-in
Top 50	30%	**50%**	0%
Top 100	50%	**70%**	50%
Stations	2	2	2
Cluster Radius	1500 meters	1500 meters	800 meters

TABLE VIII.
TESTING RESULT FOR KOBE AREA

Criteria		***Result***	
Origin	Kobe Station	Kobe Station	Kobe Portopia Hotel
POIs	Takenaka Carpentry Tools Museum, Kawasaki Good Times World, Kobe Kitano Museum, Minatogawa Shrine, Kazamidori no Yakata, Kobe Anpanman Children's Museum & Mall, Kobe Kitano Tenman Shrine, Matsuo Inari Shrine, Jodoshianyoji Temple, Puraton Ornament Museum	Kobe Harborland, Kitano-cho, Kobe Kitano Museum, Chinatown, Kobe Port Tower, Kazamidori no Yakata, Former Drewell Mansion, Urokono-ie Uroko Museum of Art, Moegi no Yakata, Moegi no Yakata	Minatogawa Shrine, Shinkobe Ropeway, Urokono-ie Uroko Museum of Art, New Port Fifth Jetty Old Signal Station
Top 50	70%	**60%**	25%
Top 100	80%	**90%**	75%
Stations	2	2	2
Cluster Radius	800 meters	1000 meters	500 meters

TABLE IX.
TESTING RESULT FOR NAGOYA AREA

Criteria		***Result***	
Origin	Nagoya Station	Nagoya Station	Nagoya Marriott Associa Hotel
POIs	Toyota Commemorative Museum of Industry and Technology, Shikemichi, Lucent Avenue, Former Kato Shokai Building, Asama Shrine, Keihoin, Nagoya/Boston Museum of Fine Arts, Hioki Shrine, Osukannon	Nagoya Castle, Yamazaki River, Osu Shopping Street, Hoshoin, Nana-chan Mannequin, Hitsuji Shrine, Sky Promenade, Shikemichi, Great Beckoning Cat Statue, Aichiken Gokoku Shrine	Nunoike Catholic Church, Takamu Shrine, Art Salon Wasabi, Mt. Hachiman Tomb
Top 50	33%	**60%**	0%
Top 100	56%	**90%**	50%
Stations	2	1	2
Cluster Radius	1500 meters	2000 meters	800 meters

TABLE X.
TESTING RESULT FOR SAPPORO AREA

Criteria	*Result*		
Origin	Sapporo Station	Sapporo Station	JR Tower Hotel Nikko Sapporo
POIs	Former Hokkaido Government Office Building, Sapporo Beer Museum, The Hokkaido University Museum, Botanic Garden Hokkaido University, Clock Tower, Sapporo Science Center, Kotoni Shrine, Shin Kotoni Shrine, Teine Shrine, Seikatei	Hokkaido University Sapporo Campus, Sapporo JR Tower Observatory, Former Hokkaido Government Office Building, Sapporo TV Tower, Botanic Garden Hokkaido University, Clock Tower, Ganso Sapporo Ramen Street, Seikatei, Sapporo Ekimaedori, Izumi Statue	Sapporo JR Tower Observatory, Ganso Sapporo Ramen Street, Sapporo Ekimaedori, Izumi Statue
Top 50	50%	**60%**	25%
Top 100	80%	**80%**	50%
Stations	4	0	1
Cluster Radius	1500 meters	1000 meters	500 meters

B. Comparison with Other Tourism Recommendation Systems

This subsection briefly compares our approach with other recommendation systems in order to highlight its unique features.

1) Nature of Recommendation System

While many recommendation systems such as SigTur/E-Destination [31] mainly focuses on the act of POIs filtering, the final result of which is a list of POIs that match with user's preferences, our approach concentrates more on the touristic planning process which embraces logistic consideration and geographical context of the user. Furthermore, our approach introduces more flexibility by considering user's dynamic needs during the trip.

2) Scope of the Recommendation System

Even though Tokyo metropolitan area is chosen as the target vicinity, availability of diverse data sources, as well as examination of Japan's unique traits (e.g., complex train system, free train pass ticket) make the proposed methodologies applicable to other regions as well. Considering that other recommendation systems rarely take the localization aspect into consideration (e.g., [17], [31]), our approach can be considered as a pioneer tourism recommendation system for Japan at a national level.

3) Input Data Space for POIs Filtering

We consider the open accessibility of input data source for POIs filtering quite important. Algorithms such as collaborative filtering require an initial database or previous actual cases, which seems collectible solely through tours agencies. Geo-tagged sightseeing photos collectible by members from various sharing platforms such as Flickr, Panoramio, etc. are also being used to rank tourism attractions [17]. In our case, however, the system makes use of data openly available on the Internet. Because two out of the three main data sources are very active Japanese online services relating to the field of tourism, the input data can be considered more reliable and openly accessible.

V. LIMITS AND FUTURE RESEARCH

The following limitations can provide space for improvement and direction for future research.

1) A systematic approach to update the current database on a regular basis is necessary for providing the most uptodate information (e.g., new events) to the user.

2) The heuristic greedy algorithm finds the local optimal solution and the hope to meet the global optimum has proven its successful in many practical pieces of research and applications [31]. Despite that, there is a need for specific metrics to define the measurement of how good a tour recommendation is according to a specific user's profile.

3) Presently, our proposed system concentrates on short-term touristic plan and without considering hotel accommodation. This ability, however, is available in its architecture and can easily accommodate multi-days trip with hotel recommendation feature. Not only both TripAdvisor.com and Jalan.net contains information on thousands of hotels, the same programming approach used in the system can also extract hotel accommodate information from other well-maintained online sources. In fact, by incorporating hotel accommodation as part of the tour recommendation package, the system can suggest a hotel which complements its user's other input preferences and context.

4) The long-term goal of the research is to launch a system that not only able to draw a good tour plan in accordance with user's preferences, but also to recognize distinguished traveling interests of tourists from different countries while they visit and experience Japan. This differences in cultural preferences often lead to latent conflicts in communications, etc., which is a crucial issue in tourism hospitability. In order to resolve the conflicts, the work of recognizing of patterns, or implicit knowledge, of different users from different cultures needed to be done. However, the work should be not only precise, but also dynamic to cope with social changes. This can hopefully be resolved by applying machine learning and statistical learning approach to collected empirical data from surveys or preliminary mobile application and picture out the repeated pattern of a variety of cultural preferences. For example, Thai tourists tend to visit lots of temples while European travelers enjoy long-distance walk during the tour.

VI. CONCLUSION

This paper shares the results of an on-going research in recommending tour planning to tourists in Japan under a practical perspective. It has leveraged web scraping technique to collect a huge amount of supporting data, has built a prototype of a system that uses the greedy algorithm for POIs searching, and demonstrated its practicality. Despite its current limitations, potential benefits of such system that is generic to various tourism regions in Japan and customizable to numerous users from different nations is shown to be high. This paper also opens a discussion about the necessity of more localization and customization in building recommendation system, not only in the field of tourism, but in other disciplines as well.

REFERENCES

[1] JLL's Hotels & Hospitality Group, "Tokyo 2020 Olympics: expectations for the hotel industry, November 2014.

[2] "Japan outlines plan to achieve 20m tourists by 2020 Olympics". (n.d.), retrieved May 3, 2015, from http://www.travelweekly.co.uk/ articles/2014/10/03/49577/japan outlines plan to achieve 20m tourists by 2020 olympics.html.e

[3] Japan Tourism Agency, "Efforts to promote a tourism-oriented nation", November, 2009.

[4] "Not just international but'Super Global Universities'", University World News. (n.d.), Retrieved May 3, 2015, from http://www. universityworldnews.com/article.php?story=20141120233337379.

[5] "Promoting two-way student exchange". (n.d.), retrieved May 3, 2015, from http://www.mext.go.jp/english/highered/1303572.htm.

[6] "Transportation - the official guide", Japan National Tourism Organization. (n.d.).

[7] Japan National Tourism Organization, "Japan Rankings", retrieved May 3, 2015, from http://us.jnto.go.jp/survey/japan_rankings.php.

[8] Japan ranked ninth most tourist-friendly nation. (n.d.). Japan Times. Retrieved July 22, 2015.

[9] Han Jiho, Yotsumoto Yukio, "The trends regarding foreign tourists to Beppu, Oita prefecture in Japan", Journal of Ritsumeikan Social Sciences and Humanities, vol. 2, pp. 61-72.

[10] Yuka Mera, Yoshiyuki Kurachi, and Naoko Ozaki, "Recent Increase in Foreign Visitors and Impact on Japan's Economy", Bank of Japan Review, December 2013.

[11] João Romão, Bart Neuts, Peter Nijkamp, Asami Shikida, "Determinants of trip choice, satisfaction and loyalty in an eco-tourism destination: a modelling study on the Shiretoko Peninsula, Japan", Ecological Economics, vol. 107, November 2014, pp. 195-205.

[12] Iponics Japan, "Marketing to Japan - appealing to the Japanese tourist". (n.d.). Retrieved May 3, 2015.

[13] "Attracting more tourists to Japan", The Japan Times, January, 2014. retrieved May 4, 2015.

[14] Damianos Gavalas, Charalampos Konstantopoulos, Konstantinos Mastakas, Grammati Pantziou, "Mobile recommender systems in tourism", Journal of Network and Computer Applications, vol. 39, March 2014, pp. 319-333.

[15] Ricardo Anacleto, Lino Figueiredo, Ana Almeida, Paulo Novais, Mobile application to provide personalized sightseeing tours, Journal of Network and Computer Applications, Volume 41, May 2014, Pages 56-64, ISSN 1084-8045

[16] Hsiu-Sen Chiang, Tien-Chi Huang, "User-adapted travel planning system for personalized schedule recommendation", Information Fusion, vol. 21, January 2015, pp. 3-17.

[17] Kai Jiang, Huagang Yin, Peng Wang, Nenghai Yu, Learning from contextual information of geo-tagged web photos to rank personalized tourism attractions, Neurocomputing, Volume 119, 7 November 2013, Pages 17-25, ISSN 0925-2312

[18] Gülçin Büyüközkan, Buse Ergün, "Intelligent system applications in electronic tourism", Expert Systems with Applications, vol. 38(6), June 2011, pp. 6586-6598.

[19] G. Tumas and F. Ricci, "Personalized Mobile City Transport Advisory System", Proc of ENTER'2009, pp. 173-183, 2009.

[20] Beatriz Rodríguez, Julián Molina, Fátima Pérez, Rafael Caballero, Interactive design of personalised tourism routes, Tourism Management, Volume 33, Issue 4, August 2012, Pages 926-940, ISSN 0261-5177.

[21] A. Garcia, P. Vansteenwegen, O., Arbelaitz, W. Souffriau and M.T. Linaza, "Integrating Public Transportation in Personalised Electronic Tourist Guides", Comput & Oper Res, 40(3), pp. 758-774, 2013

[22] NAVITIME. (n.d.). Retrieved July 22, 2015, from http://www.navitime.co.jp

[23] TripAdvisor: Read Reviews, Compare Prices & Book. (n.d.). Retrieved July 22, 2015, from http://tripadvisor.com

[24] 宿・ホテル予約 - 旅行ならじゃらん net. (n.d.). Retrieved July 22, 2015, from http://jalan.net

[25] M. Ye, P. Yin, W.-C. Lee, D.-L. Lee, Exploiting geographical influence for collaborative point-of-interest recommendation, in: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information, SIGIR '11, ACM, New York, NY, USA, 2011, pp. 325–334.

[26] L. Backstrom, E. Sun, C. Marlow, Find me if you can: improving geographical prediction with social and spatial proximity, in: Proceedings of the 19th International Conference on World Wide Web, WWW '10, ACM, New York, NY, USA, 2010, pp. 61–70.

[27] Wallace, Floyd, Associates Inc, "Massachusetts pedestrian transportation plan", 1998.

[28] Kylie Bryant, 1., & Arthur Benjamin. (n.d.). "Genetic Algorithms and the Traveling Salesman Problem".

[29] Sylejmani, K., & Dika, A. "Solving touristic trip planning problem by using taboo search approach". International Journal of Computer Science Issues (IJCSI), vol. 8(5). 2011.

[30] da Silva, V. C. O., Oliveira, D. M., de Araujo, J. C. T., & Maciel, P. R. M. (2014). Energy Consumption in Mobile Devices Considering Communication Protocols. Advances in Information Sciences & Service Sciences, 6(5).

[31] Antonio Moreno, Aida Valls, David Isern, Lucas Marin, Joan Borràs, SigTur/E-Destination: Ontology-based personalized recommendation of Tourism and Leisure Activities, Engineering Applications of Artificial Intelligence, vol. 26(1), January 2013, pp. 633-651.

Thai Quang LE was born in Khanh Hoa Province, Vietnam in 1993. He received his university degree from Ritsumeikan Asia Pacific University (APU) in the field of Business Administration focusing on Strategic Management. He has received numerous scholarships including, APU Tuition Reduction Scholarship from 2011-2015, JASSO (Japan Student Services Organization) Scholarship from 2011-2012, and Oita Prefecture Scholarship in 2012-2013, 2014-2015. His research interests include applied programming in business, modeling for decision-making and machine learning, on which he has carried out a few presentations. The theme of his thesis is retail activity optimization and he has become an IEEE member since 2014.

Davar Pishva is a professor in ICT at the College of Asia Pacific Studies, Ritsumeikan Asia Pacific University (APU) Japan. In teaching, he has been focusing on information security, technology management, VBA for modelers, structured decision making and carries out his lectures in an applied manner. In research, his current interests include biometrics; e-learning, environmentally sound and ICT enhanced technologies. Dr. Pishva received his PhD degree in System Engineering from Mie University, Japan. He is Secretary General of IAAPS (International Association for Asia Pacific Studies), Senior Member of IEEE, and a member of IEICE (Institute of Electronics Information & Communication Engineers), IAAPS and University & College Management Association.