A Genetic Algorithm for Generating Travel Itinerary Recommendation with Restaurant Selection

B. S. Wibowo, M. Handayani
Mechanical and Industrial Engineering Department, Universitas Gadjah Mada, Yogyakarta, Indonesia
(budhi.sholehwibowo@ugm.ac.id)

Abstract – **Experiencing local food while making a trip has a significant impact on the overall tourist experience. No wonder that visiting local restaurants has become an integral part of a tourist itinerary. Nevertheless, manually planning a travel itinerary is a complicated and time-consuming task. This study aims to introduce and solve a planning problem for generating a high-quality itinerary consisting of an efficient route to visit tourist attractions and restaurants at a proper time. We modeled this problem as a rooted orienteering problem with time windows, where the selected restaurants should be scheduled within lunch and dinner time. The objective is to maximize the total collected utility in each visited locations while maintaining the total travel time under a specific constraint. We developed a Genetic Algorithm to solve the problem and presented Yogyakarta city as a case study. The result suggested that the developed algorithm can effectively solve the problem and generate a satisfactory solution with a high total utility value.**

Keywords – **travel itinerary, restaurant selection, orienteering problem, time windows, genetic algorithm**

I. INTRODUCTION

Travel and tourism sector plays a vital role in the world economy. In 2016, the business volume of global tourism surpasses that of oil exports, food products or automobiles [1]. Asia and the Pacific led the tourism sector growth in 2016 with a 9% increase in international arrivals, followed by Africa (8%) and the Americas (3%) [2].

Indonesia is one of the famous tourist destinations in Asia. The Travel and Tourism Competitiveness Report [3] ranked the competitiveness of Indonesia's tourism sector at 42nd out of 141 countries in 2017, climbing eight places from the previous year. In 2017, there were more than 14 millions international tourist arrivals to Indonesia, growing 21% from 2016 [4]. Indonesia has successfully made the most of its globally recognized natural resources as tourist attractions at very affordable prices.

Tourist plays an important role in the travel and tourism sector. Indonesian Ministry of Tourism reports that in 2014, tourists spent more than 18% of their budget on food and beverages [5], the second highest category in spending. This fact indicates that there is significant attention from international tourists to Indonesia cuisine. Seeking experiences with local food has gained growing interest amongst tourists. Recent studies indicated that local food experiences had a significant effect on the overall tourist satisfaction level [6-7]. The study in [8] implied that "heritage" and "ingredients" in food experience are the two main factors which had a significant effect on international tourist satisfaction visiting Indonesia.

Indonesia cuisine is known for its rich and intense flavor. It has more than 5,000 traditional recipes from its indigenous culture. This unique and diverse local food serves as a strong point for promoting Indonesia as a world culinary destination. In 2017, three of traditional Indonesian cuisines were ranked in the World's 50 most delicious food, which by "rendang" was ranked as the top of the list [9]. No wonder that Indonesia cuisine has become a significant attraction and an integral part of a tourist daily itinerary.

On the other hand, planning a travel itinerary is a complicated and time-consuming task. The planner should make the most of the trip while dealing with the extensive available options of destination given a limited time and budget. The planning process includes researching destinations, selecting restaurants, comparing prices, booking accommodations, making routes, estimating time, and organizing schedule in order to maximize overall satisfaction in traveling.

Some attempts have been made to ease the itinerary planning process through automatization [10-12]. One of the approaches is by modeling it as an Orienteering Problem (OP). OP is a routing problem which the objective is to determine a subset of nodes to be visited in a specific order so that the collected score maximized and a given time constraint is not exceeded [13, 14]. OP has been considered as a combination of the Traveling Salesman Problem and Knapsack Problem. The problem has been applied to solve many variants of the itinerary planning problem, e.g., OP with time windows [15], OP with mandatory and optional points [16], OP with hotel selection [17] and OP with transportation mode selection [18].

Despite the wide range of applications, to the authors' knowledge, there is no study that has specifically investigated OP with restaurant selection. As discussed in [6-8], experiencing local food can improve tourist satisfaction. Thus, inserting good restaurants into travel itinerary at the right time should have a better impact on the overall tourist experience. This paper aims to introduce and solve a planning problem for generating travel itinerary recommendation with restaurant selection. By adding selected restaurants in the itinerary, we expected to generate a more viable and enjoyable itinerary for the tourists.

978-1-5386-6786-6/18/$31.00 ©2018 IEEE

II. PROBLEM FORMULATION

We modeled the itinerary planning problem with restaurant selection as a rooted orienteering problem with time windows (r-OPTW). In this problem, the starting and ending points of the trip are predetermined (rooted). It means that the planner can select to her like at which location she would like to begin and end the trip. It could be at a specific tourism place, restaurant, bus stop or even an airport. This feature can improve flexibility for the planner to plan the itinerary in practice. Furthermore, each selected tourist attraction must be visited during its opening hours, and the restaurants must be scheduled only within a lunch or dinner time. Note that the total duration of the itinerary plan should not exceed the maximum allowed duration specified by the planner.

Mathematically, this problem can be represented as a complete graph *G(V, E)* where *V* represents the set of vertices and *E* the set of edges. The vertices consist of two types of locations i.e. a set of attractions $N = (v_1, \dots, v_n)$ and a set of restaurants $R = (v_{n+1}, v_{n+2}, \dots v_r)$. Each vertex v_i is associated with positive integer utility value p_i while the utility values at the selected starting point v_{start} and ending point v_{end} is considered to be zero. A service time s_i is spent at each customer vertex v_i, $i \neq 0$. The visit of a vertex v_i can start only within a specified time window $[o_i, c_i]$. We consider a hard time windows, meaning the visitor cannot arrive later than time c_i. In case of the visitor arrive earlier than o_i, she must wait for w_i time until the place is ready to serve. A travel time t_{ij} is associated with each edge $e(v_i, v_j) \in$ E. The total time traveled at location i is denoted as T_i. A constant T_{max} limits the total travel time of a trip. A feasible solution must comprise a trip starting at vertex v_{start} and be ending at v_{end} such that each attraction is visited at most once within its opening hour $[o_i, c_i]$ and two distinct local restaurants are visited within lunch and dinner time, with respect to T_{max} as trip duration limit. The objective is to maximize the total collected utility value p_i on each selected destination between the starting and ending points.

Based on the description above, the problem can be formulated as an integer program. Let us denote variable x_{ij} as integer decision variable, where $x_{ij} = 1$ if a travel from location i to location j is performed, 0 otherwise. We formulated the mathematical problem as follow:

$$\max \sum_{i=2}^{n+r} \sum_{j=2}^{n+r} p_i x_{ij}, \quad i \neq j;\ i \neq n; \tag{1}$$

subject to:

$$\sum_{j=2}^{n+r} x_{1j} = \sum_{k=1}^{n+r} x_{kn} = 1, \quad k \neq n \tag{2}$$

$$\sum_{r} \sum_{k=1}^{n+r} x_{rk} = 2, \quad r = n+1 \tag{3}$$

$$\sum_{i=1}^{n+r} x_{ik} = \sum_{j=2}^{n+r} x_{kj}, \quad k \neq 1; k \neq n; i \neq j; i \neq n \tag{4}$$

$$\sum_{i=1}^{n+r} \sum_{j=2}^{n+r} x_{ij} \leq 1, \quad i \neq j; i \neq n \tag{5}$$

$$\sum_{i=1}^{n+r} \sum_{j=2}^{n+r} (w_i + s_i + t_{ij})\, x_{ij} + w_n + s_n \leq T_{max} \tag{6}$$

$$\sum_{i=1}^{n+r} \sum_{j=2}^{n+r} (T_i + w_i + s_i + t_{ij})\, x_{ij} = T_j \tag{7}$$

$$o_i \leq T_i + w_i + s_i, \quad \forall i \in V \tag{8}$$

$$T_i + w_i + s_i \leq c_i, \quad \forall i \in V \tag{9}$$

$$w_i = o_i - T_i, \quad T_i \leq o_i, \quad \forall i \in V \tag{10}$$

Equation (1) is the objective function which seeks to maximize the total collected utility in a trip. Constraint (2) guarantees that all tours start from location 1 and end at location *n*. Equation (3) ensures that there are only two selected restaurants on the trip. Equation (4) determine the connectivity of a trip. Constraint (5) ensures that every location is only visited once and (6) limit the total duration of a trip. Equation (7) accumulates the total travel time of a trip at location *i*. Constraints (8) and (9) guarantees that the service time at location *i* is performed within the allowed time window. Lastly, (10) explains that the visitor should wait at location *i* if she arrives earlier than the opening hour.

III. METHODOLOGY

OP is an NP-hard problem, meaning that an exact solution for the OP is only feasible for a small number of nodes. Therefore, for practical reasons, many studies suggested the use of heuristics approach to solve OP. For instance, [10] proposed the use of guided local search to solve the OP. The study in [15] used an iterated local search to solve OP with time windows, and [17] used a variable neighborhood search to solve OP with hotel selection. However, the studies mentioned above only used individual-based heuristics to solve the problem. The study in [19] had shown that the use of population-based heuristics such as a genetic algorithm can provide better solutions compared to the well-known individual-based heuristics, especially in a large network. Therefore, in this study, we further developed a genetic algorithm to solve the problem.

Genetic Algorithm (GA) is a population-based metaheuristic which uses natural selection and genetics principles to solve an optimization problem. It uses multi-stage processing such as initialization, selection, crossover, and mutation to generate and refine the solution candidate. The detailed implementation of the algorithm is described as follow:

A. *Initialization*

We used random insertion method to initialize feasible solutions. The solution candidate is represented as an array of chromosomes representing a sequence of locations as a travel route. In the beginning, two locations, v_{start} and v_{end}, are pre-selected as a starting and ending point of the trip. Afterward, a random location is inserted in between the two points. While the total duration does not exceed the allowed maximum time, we continuously added a random location to the chromosome adjacent to the previously selected location. Since the solution must have two restaurants for lunch and dinner, therefore, when the duration is approaching to eating time, and there is no restaurant within the chromosome, we randomly selected a restaurant location from the available set and inserted it into the chromosome to maintain solution feasibility. The process of initialization is depicted in Fig. 1.

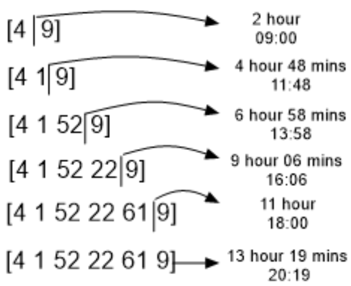

Fig. 1. Illustration of initialization process.

B. *Selection*

Selection is the process of finding out the fittest chromosome to be selected as parents for mating process. We used a roulette wheel method to selects two chromosomes from the initial population. The method ensures that a chromosome is selected through a stochastic process where a chromosome that has a higher fitness value has a higher probability to be selected as a parent.

C. *Crossover*

Crossover is a mating process between the two selected parents to create children for the next generation. In this study, we used a Partially Mapped Crossover method [20] as the crossover operation. It is performed by exchanging a subset of a chromosome from one parent to another with respect to the original position order. Crossover aims to exploit the current search space to get a better fitness value in the next generation so that the optimization process can converge.

D. *Mutation*

Mutation is a genetic operator which aim to maintain diversity in the population and to avoid local optima during the optimization process. We performed the mutation process by inserting or deleting a random location from a selected chromosome. Specifically, we selected a random location from the available set. If the location is not available in the current chromosome, it will be inserted into the chromosome. Otherwise, the location will be deleted from the chromosome.

IV. EXPERIMENT AND RESULTS

A. *Case Description*

For the experiment, we used Yogyakarta city as a case study. Yogyakarta is a cultural city and the second popular tourist destinations in Indonesia. We selected 50 top tourist attractions and 20 top local restaurants in Yogyakarta from TripAdvisor website. We set the time window at a tourist attraction to be similar as its opening and closing hour. However, the time window at a restaurant is set to fit the common eating time, which is between 11 AM-2 PM for lunch and 5 PM – 8 PM for dinner. We assumed that the service time at each tourist attraction to be 2 hours, while the service time at a restaurant to be 1 hour. The maximum duration of the trip is set to be 15 hours. The utility value in each location is assumed to be proportional to its reversed rank as listed in TripAdvisor website. The travel time between locations is retrieved from Google Maps API with a car driving as the transportation mode.

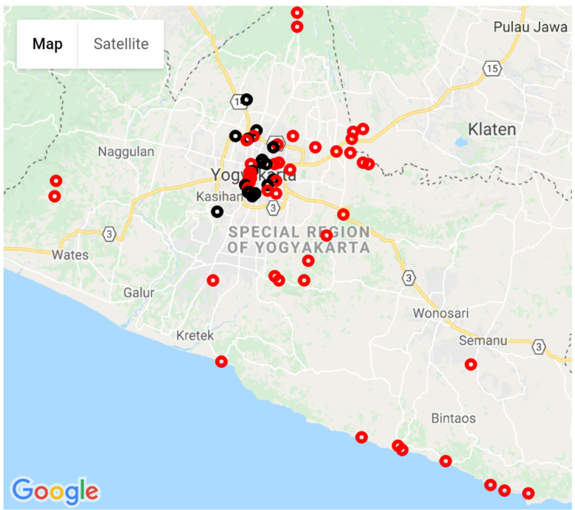

Fig. 2. The locations of 50 top tourist attractions (red signs) and 20 top restaurants (black signs) in Yogyakarta city.

B. *Parameter Optimization*

GA has several parameters, such as population size, crossover rate, and mutation rate, which need to be fine-tuned in order to optimize its performance. In this study, we used a factorial design approach to find the best parameter setting to solve the defined problem. For the experiment, we limited the number of iterations to be 1000. Each combination of parameters is replicated five times. The design of the experiment is shown in Table I. Based on the experiment result, we built a regression model to find the optimal parameter setting for GA. The model suggested that the optimal parameter setting for the problem is 46 for the population size, 0.72 for the crossover rate, and 0.27 for the mutation rate.

TABLE I
DESIGN OF EXPERIMENT FOR PARAMETER OPTIMIZATION

Parameter	Low	Medium	High
Population size	10	30	50
Crossover rate	0,1	0,5	0,9
Mutation rate	0,1	0,5	0,9

C. Results

The output of the proposed algorithm is a travel itinerary consisting of a route to visit selected tourist attractions and restaurants. The example of the generated itinerary is shown in Table II and depicted in Fig. 3. It shows a day-plan to visit the selected locations along with their ranks and utility values. Note that the route was generated with a pre-selected starting and ending points. In this case, we selected “Gembira Loka” Zoo as the starting point, and “Rainbow Garden” as the ending point. Next, the algorithm will automatically assign a number locations in between the starting and ending point to maximize the collected utility value by considering the time windows and the total trip duration.

The result in Table II shows that the proposed method can effectively solve the problem with respect to the available constraints. Each tourist attraction within the route is perfectly planned during its opening hours, and the restaurant is also perfectly scheduled within lunch or dinner time. The algorithm also successfully selects a set of top-ranked locations amongst the numerous available options. This result suggested that the proposed method can solve the problem effectively and generate a satisfactory solution.

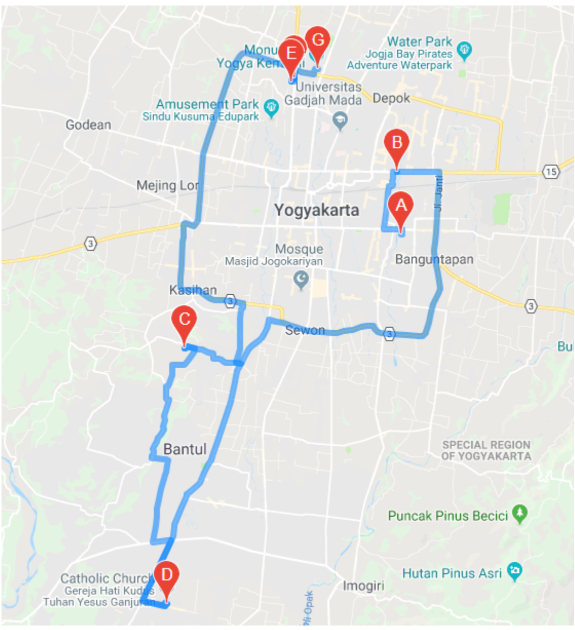

Fig. 3. Visualization of the generated itinerary from the genetic algorithm.

TABLE II
EXAMPLE OF THE GENERATED ITINERARY PLAN FROM THE GENETIC ALGORITHM

ID	Location	Time Schedule		Rank	Utility Value
A	Gembira Loka Zoo (pre-selected)	Arrival	07:00:00	0	
		Departure	09:30:00		
Travel Time			00:12:09		
B	Affandi Museum	Arrival	09:42:09	9	42
		Departure	11:42:09		
Travel Time			00:34:52		
C	The Sawah Restaurant	Arrival	12:17:01	1	20
		Departure	13:17:01		
Travel Time			00:27:07		
D	Ganjuran Church	Arrival	13:44:08	2	49
		Departure	15:44:08		
Travel Time			00:50:32		
E	Jogja City Mall	Arrival	16:34:40	18	33
		Departure	18:34:40		
Travel Time			00:01:44		
F	Sate Ratu Restaurant	Arrival	18:36:24	5	16
		Departure	19:36:24		
Travel Time			00:03:35		
G	Rainbow Garden (pre-selected)	Arrival	19:39:59	0	
		Departure	21:39:59		

V. CONCLUSIONS

This paper contributes to the literature on the formulation of itinerary planning problem with restaurant selection. The problem was modeled as r-OPTW where the restaurants should be planned exactly during the lunch and dinner time. We developed a genetic algorithm to solve the optimization problem and used Yogyakarta city as a case study. The result suggested that the proposed algorithm can effectively solve the problem. It generated a satisfactory travel plan consisting of a set of high-ranked tourist attractions and restaurants with respect to the available constraints.

As the future work, we intended to expand experimentation by comparing the performance of the proposed method to the available benchmarks for the OPTW so that the efficiency of the proposed algorithm can be further verified. We also intended to extend the problem to include multi-period planning and hotel selection as this can be helpful in practice to simplify the complexity of planning an itinerary.

REFERENCES

[1] United Nations World Tourism Organizations, “UNWTO Tourism Highlight 2017 Edition”, 2017.
[2] World Travel and Tourism Council, “Travel and Tourism: Economic Impact 2017 World”, 2017.
[3] World Economic Forum. “The Travel & Tourism Competitiveness Report 2017”, 2017.
[4] Badan Pusat Statistik Indonesia, “Statistik Kunjungan Wisatawan Mancanegara 2017”, 2017.

[5] Deputy Assistant of Research Development on Tourism Policy, “Statistical Reports on Visitor Arrivals to Indonesia," Indonesian Ministry of Tourism, 2014.

[6] Hall, C.M., Sharples, L., Mitchell, R., Macionis, N. and Cambourne, B. “Food Tourism Around The World: Development, Management, and Markets," Routledge, New York, NY, 2003.

[7] Kivela, J.J., and Crotts, J.C., "Tourism and gastronomy: gastronomy's influence on how tourists experience a destination," Journal of Hospitality & Tourism Research, Vol. 30 No. 3, pp. 354-377, 2006.

[8] R.B. Hendijani, "Effect of food experience on tourist satisfaction: the case of Indonesia, " International Journal of Culture, Tourism and Hospitality Research, Vol. 10 Issue: 3, pp.272-282, 2006.

[9] T. Cheung, “World’s 50 best foods: Reader’s choice,” CNN Travel. July 12, 2017. [Accessed 10 May 2018]

[10] W. Souffriau, P. Vansteenwegen, J. Vertommen, G.V. Berghe, D. Van Oudheusdhen, “A personalized Tourist Trip Design Algorithm for Mobile Tourist Guides,” Applied Artificial Intelligence: An International Journal, 22:10, 964-985, 2008.

[11] H. Chang, Y. Chang, and M. Tsai, “ATIPS : Automatic Travel Itinerary Planning System for Domestic Areas,” vol. 2016, 2016.

[12] P. Vansteenwegen, W. Souffriau, G.V. Berghe, and G.V. Oudheusden, “The City Trip Planner: An Expert System For Tourists," Expert Systems with Applications, Vol. 38, pp. 6540-6546, 2011.

[13] T. Tsiligirides. “Heuristic methods applied to orienteering,” The Journal of the Operational Research Society, 35(9):797–809, 1984.

[14] P. Vansteenwegen, W. Souffriau, and D. Van Oudheusden, “The orienteering problem: A survey,” Eur. J. Oper. Res., vol. 209, no. 1, pp. 1–10, 2011.

[15] A. Gunawan, H.C. Lau., and K. Lu, “Iterated Local Search Algorithm for Solving Orienteering Problem with Time Windows," Lecture Notes in Computer Science, Vol. 9026, pp. 61-73, 2015.

[16] S. Kotiloglu, T. Lappas, K. Pelechrinis, and P.P Repoussis, “Personalized Multi-Period Tour Recommendations”, Tourism Management, Vol.62, pp.76- 88, 2017.

[17] A. Divsalar, P. Vansteenwegen, and D. Cattrysse, D., “A Variable Neighborhood Search Method For The Orienteering Problem With Hotel Selection," Int. J. Production Economics, Vol. 145, pp. 150-160, 2013.

[18] A. Garcia, P. Vansteenwegen, O. Arbelaitz, W. Souffriau, and M.T. Linaza, “Integrating Public Transportation In Personalised Electronic Tourist Guides," Computers & Operation Research, Vol. 40, pp. 758-774, 2013.

[19] J. Karbowska-chili, J. Koszelew, K. Ostrowski, and P. Zabielski, “Genetic Algorithm Solving Orienteering Problem in Large Networks,” Adv. Knowledge-Based Intell. Inf. Eng. Syst. M., pp. 28–38, 2012

[20] D.E. Goldberg, R. Lingle, Jr., "Alleles, Loci, and the Traveling Salesman Problem," Proceedings of the 1st International Conference on Genetic Algorithms, pp. 154-159, 1985.