Piya, Sujan, Triki, Chefi, Al Maimani, Abdulwahab and Mokhtarzadeh, Mahdi (2023) *Optimization model for designing personalized tourism packages*. Computers & Industrial Engineering, 175 . ISSN 0360-8352.

Optimization Model for Designing Personalized Tourism Packages: Case of the Sultanate of Oman

Abstract

The tourism supply chain aims at satisfying the needs of the tourists based on their preferences. However, the preference for each tourist may be different. Some tourists prefer to optimize single criterion while others prefer to optimize conflicting multiple-criteria, in which some criteria need to be maximized and others need to be minimized. Hardly, the tourism service provider can offer the tourists with the itinerary according to their precise preferences. This paper proposes a multi-criteria optimization framework based on which tourists can generate itineraries according to their preferences. Each criterion is presented in terms of its normalized value and integrated into the optimization framework using the weighted-average method. The model helps the tourists to express their preference by setting the most appropriate weight of each criterion in the objective function. The model is tested on a small real case example to generate an itinerary and schedule the activities according to the needs of the tourist. Moreover, a numerical analysis study is conducted on a hypothetical problem to demonstrate the applicability and complexity of the model. The computational analysis shows that the complexity of the model increases with the size of the problem instances. Specifically, an increase in the number of activities and the planned number of days will have the highest impact on the computational timing. Sensitivity analysis shows that an increase of the cost weight will decrease the total number of activities, total cost and travel distance. However, an increase in the number of activity weight and distance will result in an increase in the total number of activities and travel distances but will decrease the total cost.

Keyword: Tourism industry, Tourist preferences, Normalization, Weighted average method, Multi-objective mathematical model.

1. Introduction

The Organization for Economic Co-operation and Development (OECD) defines tourism as the activities or events where people travel to and stay in for some time outside their original environment for official purposes or leisure for no more than one continuous year (WTO, 2001). The tourism industry involves a combination of people, providers, and organizations that integrate their efforts with the aim of serving tourists. It is one of the fastest growing industries and is an increasingly important source of income, employment and wealth in many countries (Neto, 2003). The industry continuously develops in terms of infrastructure and services provided to increase the satisfaction of the tourists. According to the World Travel and Tourism Council (WTTC), the travel and tourism sector generates about 10% of global GDP, which is $7.6 trillion and by 2026 it is projected to reach around 11% of the global GDP (WTTC, 2015). This indicates that the tourism industry is a major contributor to the GDP globally, and it is a potential industry for developed and developing countries. Moreover, the industry contributes significantly to the employment and economic development of the touristic destination and the surrounding area (Rebollo and Baidal, 2003).

The tourism industry involves a set of parties to organize the operations of the tourism supply chain (TSC) and to satisfy the need of tourists (Zhang et al., 2009). Intense competition forces various parties to cluster and cooperate in supply chains to enhance their agility and improve their performance (Sigala, 2008; Piya et al., 2020). The parties in TSC are the service providers that serve the industry by providing their services and products to the end-users. The service providers are clustered into various groups, which are transportation, accommodation, entertainment, heritage and culture, and food. The end users are the tourists that are either domestic or international tourists. Due to the nature of the tourism supply chain, the tourism product cannot be stored for future use since the demand is highly uncertain. The uncertainty of the demand is mainly due to the different tourists' preferences and interests, and it is seasonally affected. Such uncertainty creates challenges for the service providers for designing innovative tourism products. Significant cooperation between multiple service providers in a TSC is detrimental to the success of a product and contributes to the optimization of the whole supply chain (Szpilko, 2017). The preparation of a tourism product often involves many parties providing various tourism related products and services. Due to the frequent interaction between these different parties, the structure of TCS is highly complex. The information flow in the tourism industry is very intensive, and the system works similarly to the customized production system i.e., the information for the required services starts from the end-user and flows through the travel agents. The travel agent then contacts different supply chain tiers to reach the final decision about the information the end-user requests.

The main goal of the tourism industry is to satisfy the needs of the end-users based on their different preferences with respect to transportation, activities and places of interest at the lowest costs. This creates conflicts between different groups of tourists, who may seek different interests at various locations and activities of different nature. One of the important function of the TSC is to interconnect different service providers to offer the best service to the end-users. Tour operators or travel agencies play a significant role in offering different service packages and tour itineraries to the tourists, in coordination with the other service providers. An itinerary represents a planned set of activities and travel route the tourist is going to follow during his/ her journey (Wong and McKercher, 2012). The aim of recommending an itinerary is to provide a sequence of visiting different locations while performing various activities in these locations within a given time frame and cost (Yochum et al., 2020). Most of the travel agencies generate general-purpose itineraries that may not completely take into account the preferences of tourists. This leads to causing wastes (or unsatisfaction) with respect to the time spent in travelling, the distance voyaged, trip costs, and the number of activities, thereby lowering the satisfaction of the tourists.

This research aims at designing personalized tourism packages by allowing the tourists the ability to select the preferred set of activities and defining the major inputs in the generation of the tourism package. The developed model helps in solving the problems related to the tourism planning that the tourists usually face when preparing their trips, and in avoiding any waste in terms of time, distance and cost. The remaining part of the paper is organized as follows. Section 2 summarizes the relevant literature references in the field and highlights the research gaps. Section 3 discusses in details the conceptual framework of this research work. Our mathematical model developed to solve the aforementioned personalized tourism problem is presented in Sections 4. Numerical analysis and the main findings of the study is discussed in Section 5. Finally, the research is concluded in Section 6 along with possible future research directions.

2. Literature Review

The Optimum Personalized Tourism Package (OPTP) has been usually modelled in terms of the well-known Orienteering Problem (Golden et al., 1987). The Orienteering Problem is originated from the homonymous competition, in which the participants visit a set of checkpoints during a given amount of time. Each checkpoint is associated with a known score value, and the participants aim to visit a subset of checkpoints that will guarantee to them accumulating the maximum score. The winner is the participant who succeeds in achieving the highest score through not only selecting the appropriate subset of checkpoints but also by identifying the most efficient sequence of their tour (Yu et al., 2019). The Orienteering Problem has been used to represent a wide set of applications, and researchers have developed over the years many variants in order to face the challenges arising in the different contexts. A recent survey that reviews the most recent variants, enumerates the solutions methods proposed for their solution and describes the most relevant related applications can be found to Gunawan et al. (2016).

The Orienteering Problem has often been the optimization engine for defining personalized tourism packages. Table 1 summarizes the contributions of the most related papers that have been published on the topic during recent years. The table highlights how our study follows most of the papers in being application-oriented and confirms that all the studies incorporate time-restrictions within their developed models. However, it is clear that only a few studies explicitly considered the budget cap on the packages to be defined and the flexibility in identifying the starting and ending locations of the tours. More importantly, Table 1 shows that only a few papers have recognized the multi-criteria nature of the OPTP and have proposed a multi-objective optimization model for its solution.

Table 1: Recent papers related to the use of the Orienteering Problem to solve the OPTP Problem

Reference	Application	Flexible Start/End Nodes	Multi-Objective Model	Time Restricted	Budget Restricted	Solution Approach
Sylejmani et al. (2014)	Prishtina			✓		Meta-heuristic
Herzog and Worndl (2014)				✓	✓	Heuristic
Verbeeck et al. (2014)		✓		✓		Meta-heuristic
Malucelli et al. (2015)	Trebon region			✓	✓	Exact
Yu et al. (2015)	Istanbul	✓	✓ (bi-)	✓		Exact & heuristic
Gavalas et al. (2015)	Athens and Berlin			✓		Heuristic
De Falco et al. (2016)	Naples	✓	✓	✓		Meta-heuristic
Brito et al. (2017)				✓		Meta-heuristic
Lim et al. (2018)	Several cities		✓ (bi-)	✓	✓	Heuristic
Mancini and Stecca (2018)	Mediterranean sea			✓		Meta-heuristic
Hapsari et al. (2019)	Besuki			✓		Exact
Expósito et al. (2019)				✓		Meta-heuristic
Yochum et al. (2020)	Several cities		✓	✓		Meta-heuristic
Tenemaza et al. (2020)	Paris, Rome, New York	✓		✓		Meta-heuristic
Persia et al. (2020)	Bolzano			✓		Heuristic
Urrutia-Zambrana et al. (2021)	Three Spanish cities			✓		Meta-heuristic
This Study	Muscat	✓	✓	✓	✓	Exact

More specifically, Yu et al. (2015) solved the OPTP that employs a robot to collect useful information about a set of touristic points of interest (POIs). Their work considered two different objectives to be achieved by the robot: the first, called *budget-minimizing tourist*, seeks to minimize the time spent at the selected POIs and the second, called *reward-maximizing tourist*, attempts to maximize the total information gained from the POIs visits. It is to be noted that even though the authors considered two objectives to be optimized, they used them separately within their optimization framework, without ever defining a multi-objective model, as in our case. The study also developed a solution approach called *anytime planning algorithm* and applied it to generate a day tour among 20 potential POIs spatially distributed in Istanbul.

De Falco et al. (2016) developed an evolutionary algorithm to plan personalized touristic packages but focused only on walking tours within a confined urban zone. The authors defined and optimized five objective functions including maximizing the score of the defined tours and the number of POIs to be visited and minimizing the covered distance as well as any temporal deviation with respect to the planned schedule. They applied their approach to define walking tours in the city of Naples and carried out extensive experiments to assess the effect of the most relevant problem's parameters. However, their study does not take into account the budget restriction and also does not include an optimization model, but rather defined only the objectives used as fitness criteria for their population-based method. Earlier, other works, such as Vansteenwegen et al. (2011), Cotfas et al. (2011) and Gavalas et al. (2015), also dealt with the definition of personalized walking tours but involved less features with respect to De Falco et al. (2016) and, consequently, with respect to our study.

Lim et al. (2018) proposed a tour recommendation framework that incorporates a bi-objective optimization model. The first objective maximizes the tourists' interest preferences with respect to the POIs to be visited, whereas the second one maximizes their popularity score. The paper developed an Orienteering Problem-based solution method and applied the suggested approach to define recommended tours in 10 different big cities sparse around the globe. The last paper that implements a multi-objective approach we are aware of is due to Yochum et al. (2020). However, even this work does not develop a multi-objective model in the commonly known sense, but incorporates a multi-criteria fitness function within a genetic algorithm method. The study considered the following criteria: the total number of POIs, the number of mandatory POIs to be included in the tour, the POIs popularity, their entrance cost and their overall rating. The authors proposed an adaptive genetic algorithm to solve the OPTP problem and used it to identify touristic itineraries for the following six cities: Budapest, Edinburgh, Toronto, Glasgow, Perth, and Osaka.

The above literature review has identified several research gaps that we try to address in this work. Specifically, our study:

- Propose explicitly the first ever multi-objective optimization model for the solution of the OPTP problem with operational constraints;
- Incorporates explicitly the budget restriction on the touristic packages to be identified;
- Allows for starting and ending the tours from/at different locations with respect to the user's accommodation place;
- Applies the developed approach, for the first time, to build OPTP in the city of Muscat, as an emerging touristic destination that still needs to be discovered by the worldwide mass of tourists.

It is, thus, clear that this study fills important gaps with attractive contributions that have never (or scarcely) been addressed in the published references.

3. Conceptual Framework

Before developing our mathematical model for designing the OPTP, a conceptual framework for the problem is developed. The conceptual framework will aid in the definition of the optimization formulation. For this purpose, data pertaining to the tourism activities in the Sultanate of Oman were collected and analyzed. Oman is home to many cultural and eco attractions and, having the benefit of its geographical position, has a significant potential to become a major tourist destination in the region (Muthuraman and Al Haziazi, 2019). The data bank of the Ministry of Tourism and the local travel agents is the major source of the analyzed data. Altogether, ten years of data were collected and these data were mainly related to the flow of tourists, their preferred mode of transportation, accommodation, touristic activities performed, budget and total nights spent by tourists. Moreover, the missing data from the above two sources were collected through a questionnaire that was prepared and distributed to the local inhabitants, international tourists, and expatriates residing in Oman. The collected data helped in identifying the major trends in the Omani tourism industry.

From the analyzed data, it was found that the primary visitors to Oman are from the Gulf countries, and the next are from Europe. Even though most visitors are from the Gulf region, it was noticed that European visitors spend more nights in the country. This may be due to the major difference in culture between Oman and the European countries and also the distance separating the two regions. Moreover, other important information collected from the data analysis is summarized in Table 2. In the table, four different choices with the percentage of tourists preferences to each choice are presented. All the percentages in the table are presented in the rounded figure.

Table 2: Result of data analysis

Item	**1st choice (%)**	**2nd choice (%)**	**3rd choice (%)**	**4th choice (%)**
Reason to visit Oman	Nature (38%)	Culture (25%)	Cheap destination (15%)	Weather (12%)
Accommodation	Hotel (42%)	Rented house (28%)	Apartment (18%)	Hostel (6%)
Activity of interest	Sea (42%)	Desert (31%)	Hiking (13%)	City tour: museum, entertainment, etc. (8%)
Duration of stay (Weeks)	<1 (46%)	1 (29%)	1-2 weeks (12%)	>2 (6%)
Mode of transportation	Own vehicle (40%)	Rented Vehicle (33%)	Tour bus (25%)	Others (2%)
Locations visited per trip	2 (41%)	1 (36%)	3 (17%)	>3 (6%)
Planning criterion considered	Maximize # of activities (37%)	Minimize activity cost (26%)	Minimize travel time (21%)	Minimize travel distance (12%)

From the above data analysis, it was made evident that the tours definition problem is of multi-objective nature. Some tourists prefer minimizing activities costs, but at the same time, maximizing the number of activities. Others may consider minimizing travelling time and/or distance. The data analysis suggested that a personalized approach for designing tourism packages is more appropriate than offering predefined packages, which is widely prevalent in the tourism industry. The analysis helped in developing the conceptual framework, as shown in Figure 1. The framework consists of some set of activities. In each activity category, there exist multiple options. For example, inside the heritage exploration category, the activities are forts, castles, traditional villages, etc. The tourists can select their preferred activities and their number to be included in the pool of potential activities. Further, they will input the total number of days they want to spend and their budget. To allow more flexibility, the budget is defined according to three thresholds i.e. high, medium and low. With these inputs, the mathematical model will identify the number of activities the tourists can perform within the given constraints of time and budget, and based on the objectives set by the tourist. Moreover, the model will help in identifying a day-to-day activities schedule and routes plan.

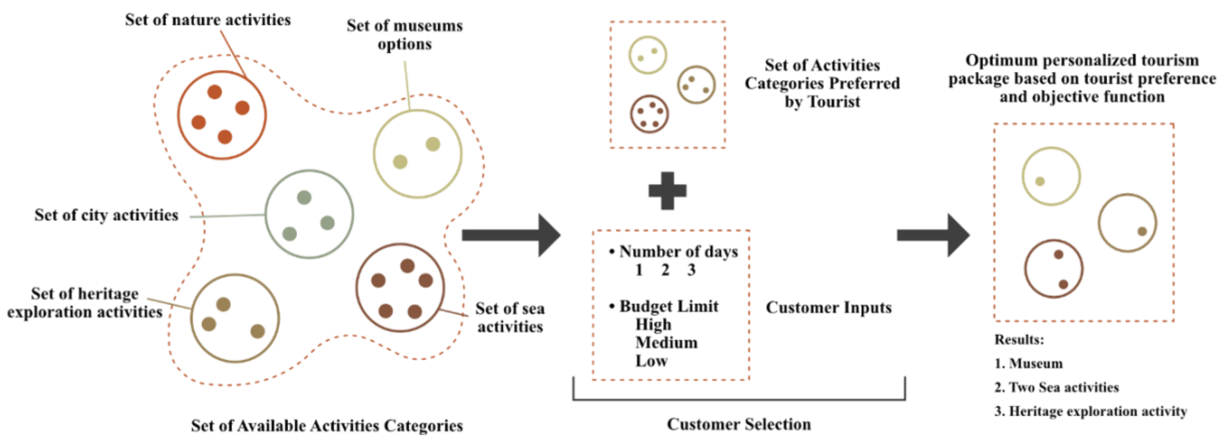


Figure 1: Conceptual framework of our OPTP

4. Mathematical Model

The mathematical model is developed to optimize cost, distance travelled and the number of activities. As the preferences of each tourist are different from the others, these multiple objectives or criteria are combined using a weighted sum. This helps each tourist to express the most suitable weights to each of the defined criteria based on his/her preferences. To develop the mathematical model, the following assumptions are considered:

- The trip can be planned along a multi-day horizon, with T days.
- The departure will be from the tourist accommodation and the final destination may not be necessarily the same accommodation.
- All tourists going to a given activity should leave that activity after spending a certain amount of time.

- The tourist may define a budget limit to control the total cost of the selected activities.
- Only one activity can be done at any given time period and tourists should perform at least one activity per day.

In the model, the set of activities is *M* and each activity may belong to the sea category, sightseeing or hiking, etc. Among the elements of *M* we define *Q* thar represents the set of must-go activities. The inputs of the mathematical model are expressed in terms of subsets of *j* that are preferred by the tourist and the parameters that define the desires of the tourist. The notations used in the development of the mathematical model are illustrated in Table 3.

Table 3: Notation used in our mathematical model

Indices:	
i, j, r	Indices for the activities belonging to set *M* ({0} indicates the set of possible tourist's accommodations and *Q* indicates the set of must-go activities)
t	Indexing the day when an activity is conducted (*t=1, 2, ……,T*)
Input:	
α_j	Expected time in hours to spend enjoying activity *j*
c_j	Cost to conduct activity *j*
d_{ij}	Traversal distance between activity *i* and activity *j*
B	Total budget limit set by the tourist
K	Time limit per day dedicated for activities
w_A	Weight assigned for the activities number
w_C	Weight assigned for the cost
w_D	Weight assigned for the travel distance
Max/Min (C)	Cost incurred if all the activities are (or if only one activity is) performed
Max/Min (D)	Distance travelled if all the activities are (or if only one activity is) performed
Max/Min (A)	Total number of available activities (=1 in the case of the *Min* operator)
Decision Variables:	
X_{ijt}	1, if the tourist travels from activity *i* to activity *j* during day *t*. 0 otherwise
Z_{jt}	1, if the tourist includes activity *j* in the itinerary of day *t*. 0 otherwise

The full space OPTP model can be expressed as follows:

$$\textbf{Maximize score} \;=\; w_A f_1 + w_C f_2 + w_D f_3 \tag{1}$$

where:

$$f_1 = \frac{\sum_{j=1}^{|M|}\sum_{t=1}^{T} Z_{jt} - \min(A)}{\max(A) - 1} \tag{2}$$

$$f_2 = \frac{\max(C) - \sum_{j=1}^{|M|}\sum_{t=1}^{T} c_j Z_{jt}}{\max(C) - \min(C)} \tag{3}$$

$$f_3 = \frac{\max(D) - \sum_{i=1}^{|M|}\sum_{j=1}^{|M|}\sum_{t=1}^{T} d_{ij}X_{ijt}}{\max(D) - \min(D)} \quad (4)$$

Subject to:

$$w_A + w_C + w_D = 1 \quad (5)$$

$$\sum_{i\in\{0\}, j\in M} X_{ijt} = 1 \qquad \forall t = 1, 2, \ldots\ldots, T \quad (6)$$

$$\sum_{i\in M, j\in\{0\}} X_{ijt} = 1 \qquad \forall t = 1, 2, \ldots\ldots, T \quad (7)$$

$$\sum_{j} \alpha_j Z_{jt} \leq K \qquad \forall t = 1, 2, \ldots\ldots, T \quad (8)$$

$$\sum_{t}\sum_{j} c_j Z_{jt} \leq B \quad (9)$$

$$\sum_{i\in M\cup\{0\}, i\neq j} X_{ijt} = \sum_{r\in M\cup\{0\}, r\neq j} X_{jrt} \qquad \forall j \in M; \forall t = 1, 2, \ldots\ldots, T \quad (10)$$

$$\sum_{t} Z_{jt} \leq 1 \qquad \forall j \in M \quad (11)$$

$$\sum_{i} X_{ijt} = Z_{jt} \qquad \forall j \in M \cup \{0\};\ j \neq i;\ \forall t = 1, 2, \ldots, T \quad (12)$$

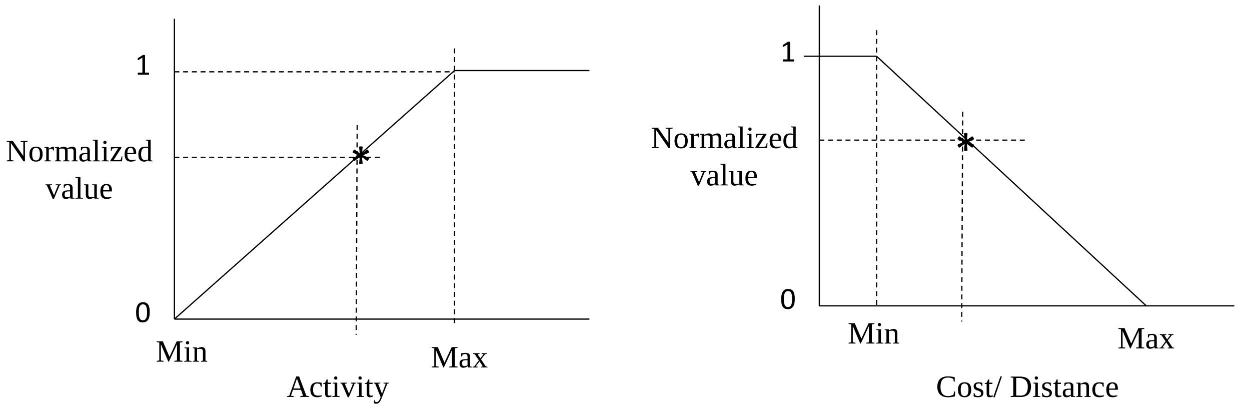
$$\sum_{t} Z_{jt} = 1 \qquad \forall j \in Q \quad (13)$$

$$\sum_{i,j\in S\cup\{0\}} X_{ijt} \leq |S| - 1 \qquad \forall\ S \subset M \cup \{0\}; \forall t = 1, 2, \ldots\ldots, T \quad (14)$$

$$X_{ijt} \in \{0,1\} \qquad \forall\ i, j \in M \cup \{0\};\ \forall t = 1, 2, \ldots\ldots, T \quad (15)$$

$$Z_{jt} \in \{0,1\} \qquad \forall j \in M;\ \forall t = 1, 2, \ldots\ldots, T \quad (16)$$

The objective of the OPTP is to jointly optimize the cost, time and number of activities. Out of these three criteria, cost and distance criteria need to be minimized while the activity number criterion needs to be maximized. Moreover, each of these criteria has a different unit of measure. Therefore, in order to simultaneously optimize the three criteria, the normalized values of the criteria were integrated within the same objective function called *score*, as shown in Equation (1) (Piya et al., 2009). Thus, the objective here is to maximize the normalized weighted sum of the three criteria. Equation (2) normalizes the number of activities, which is to be maximized. On the other hand, Equations (3) and (4) normalize the cost and distance, respectively and both the criteria need to be minimized. The normalization process is carried out as shown in Figure 2. Figure 2 (a) shows that for the criteria "activity", the normalized value increases as the total number of activities to be performed by the tourist approaches towards the total available activities. However, the normalized value decreases if the activity approaches towards its minimum i.e., 1. Conversely, the opposite situation can be seen for the "cost" and "distance". As shown in Figure 2 (b), the normalized value for these criteria increases as the criteria value approaches towards the minimum and vice-versa. The objective function will have the value of 1 if the maximum satisfaction of the weighted sum objective functions is reached and will have a value of 0 "zero" if the satisfaction is at its lowest score.



a. Normalization method for the # of activity b. Normalization method for cost/distance

Figure 2: Normalization for the maximization and minimization criteria

Constraint (5) ensures that the total weight of the criteria is equal to 1. Constraints (6) and (7) are logical constraints that force the model to start the trip from one of the tourist's accommodation and to end the trip at the same or a different accommodation. The ending accommodation point may be different from the starting point. Equation (8) states that the total hours spent on each daily trip shall not exceed a defined time limit. Equation (9) is the budget limit constraint. Equation (10) is a path continuity constraint, which states that every tourist entering for activity *j* from *i*, should leave *j* for another activity *r* or for his/ her accommodation after spending some time. Equation (11) states that an activity can be conducted at most once during all the trip days. Equation (12) states that a route from the activity's location *i* to location *j* (including accommodation points) should exist only if both activities are conducted at the same day *t*. Equation (13) is related to the predefined set of must-go activities. This constraint will ensure that the tourist will not miss the most important attractions of the city (defined as the must-go set *Q*). Constraint (14) is the well known sub-tour elimination constraint (see for example Triki et al., 2017). Finally, constraints (15) and (16) define the domain of the binary decision variables X_{ijt} and Z_{jt}, respectively.

The developed model (1)—(16) results to be a combinatorial optimization program whose size depends on the number of activities to be considered as well as on the extent of the planning horizon of the problem.

5. Numerical Analysis

A numerical investigation is conducted to validate the mathematical model and to understand the effect of the various parameters on the model. First, a small real case example related to the city of Muscat is solved, followed by solving a bigger size realistic hypothetical problem. Moreover, a sensitivity analysis is also carried out to understand the effect of the weights of the objective functions on the total score. The model is implemented by using the commercial LINGO programming language 18.0.44 (www.lindo.com) and solved by the optimization packages it embeds. The model is executed on a computer having Quad-Core Intel Core i5 processor with a CPU of 2.30GHz and 8 GB of RAM.

5.1 Case example

The case example is conducted by communicating with a group of tourists that recently visited Oman and decided to spend two days to discover its capital city, Muscat. The input data revealed by the tourists consist of the categories of activities of their interest, the budget limit, and the number of days for which they need personalized packages. The selected group of tourists was interested in mainly the heritage activities, museums, and city tour with shopping. As it is their first visit to the country, they did not mention any must-go activities among their destinations. The group mentioned that they are limited to only two days' trip, and are staying at a hotel near the area of their interest. Therefore, they are not willing to change their accommodation during the days covered by the planning horizon. The group requested to generate two alternative itineraries, one with a high budget and another with a medium budget. With the above information, five locations for heritage activities (indexed as 1, 2, .., 5), three for shopping activities (indexed as 6, 7,.., 10) and five museum activities (indexed as 11, 12,…, 15) were considered. Further communication with them highlighted that the group prefers to have as many activities as possible, but considering minimizing the cost to have more priority than minimizing the travel distance. This information helps in identifying the weights (W_A = 0.5; W_C = 0.3; W_D = 0.2) to be considered in the score objective function.

Table 4: Result obtained for the case example

Items	**High Budget**		**Medium Budget**	
	Day 1	***Day 2***	***Day 1***	***Day 2***
Itinerary	0-1-8-3-2-4-7-6-0	0-11-13-5-15-14-0	0-1-9-7-3-2-0	0-6-5-10-14-0
Total number of activities	7	5	5	4
Daily cost ($)	178.6	66.5	65	74.5
Total cost ($)	245.1		139.5	
Travel distance (km)	45	12	16.9	43.4
Score	**0.871**		**0.842**	

The results obtained from the optimization solver are reported in Table 4. The row labelled "Itinerary" in the table shows the activities that were suggested to be performed at each day and their sequence along each day's route. For example, the scheduled activities on day 1 for the high budget scenario, start with going to one of the heritage locations, followed by a shopping centre and so on and concluded with again shopping at location 6, before coing back to the accommodation. There will be no museums visit on day 1. However, day 2 will start with a visit to two museums and concluded by visiting the other two museums with a shopping break in the middle. On day 2, out of the 5 activities, four activities are related to museums visits. This is due to the reason that most of the major museums in Muscat are concentrated in one specific zone of the city and far away from most other attractions. Therefore, scheduling them on the same day will reduce the overall travel distance. From the results of Table 4, it is clear that the number of activities performed with a high budget scenario is higher as compared to its medium budget counterpart, even though the travel distance in the case of medium budget scenario is higher. It is also worth noting that the medium budget itinerary involves visiting one museum only during both days. This is clearly due to the preference of the tourists to assign a higher weight to the cost minimization in the objective function, given that usually the museums entrance has a higher cost. For both the budget scenarios, the objective function score, which represents the comulative weighted level of satisfaction achieved with respect to the tourist preferences, is above 0.8, which is close to the maximum score i.e., 1.

5.2 Hypothetical Test Problem

Next, to understand the complexity level of our model, a further numerical experiment is conducted with a hypothetical instance. The experiment has mainly the aim of assessing the effect of increasing the problem's size and of measuring the complexity of the problem. The experiment allows also measuring the robustness of the obtained solution by evaluating the impact of altering some input parameters of the model such as the number of activity categories, the budget level, and the duration of the trip. It should be noted that, as is made evident from Figure 1, increasing the number of categories will increase the number of total activities to be considered since each category will involve an additional set of activities. For this hypothetical test problem, we consider that for each category, there are 10 activities to be incorporated in the input data set. The set of instances considered in this experiment is reported in the first column of Table 5. The number of activity categories and the number of visit days planned are varied from [1-3] and [2-6] respectively. This means that the total number of activities ranges between 10 and 30. Also, three budget levels, namely, High (H: $500), Medium (M: $350), and Low (L: $200) are considered for the analysis. With the combinations of activity category, planned days and budget levels, altogether, 27 combinations of the problem are generated and analyzed. For the analysis, the cost for each activity, distance from one activity to other and activity time are randomly generated within the intervals [10-30], [8-12] and [1-5] respectively. Moreover, equal weights are considered for three criteria in the objective function. The table shows the model's solution in terms of a selected number of activities, total cost, total distance, score and the CPU timing needed to execute the model by Lingo solver.

The results show that the number of activities, the associated total cost and the travelled distance are affected by the combination of all the three parameters considered in the experiment. Generally speaking, with the increase in the value of these parameters, the number of selected activities increases. This is because increasing the number of planning days will increase the time limit allowing more activities to be included within the multi-day trip. Further, increasing the budget limit will allow more activities to be included in the trip plan. However, the result is not true when there are constraints on available activity, budget and limit on planning day. Table 5 shows that when the number of activity is 10, even though the total cost of the selected activities are much less than the budget constraint, the maximum number of selected activities is 10. This is due to the constraints on the number of available activities. Similarly, when the budget is low, then the number of selected activities is in the range of (8-11) irrespective of an increase in the number of available activities and planning day. This is due to the reason that the total cost of selected activities approaches closer to the limit on the budget. Moreover, with 2 days of the planning horizon, the maximum number of selected activities is 12 irrespective of an increase in the number of available activities and budget limit. This is due to the constraint on the time limitation to perform the activities. The maximum number of selected activities reaches 27 (see the last row) when the number of available activities, planning days and the budget limit is at the highest level. In most of the setup combinations, cost and distance increase with the increase in the number of selected activities, with few exceptions (as, for example, experiments 13 and 20). One reason for the exceptional cases is that the selected activities take less cost to be performed. Moreover, it may also be due to the reason that the selected activities are comparatively closer apart so that the total distance to be covered is lower.

The budget limit appears to be a significant constraint on the solution when high-cost activities were selected as part of the model solution. The lower the budget level is, the fewer activities were conducted; hence, less distance is travelled. For all the experiments, the score is above 0.7 with an average score of 0.81 over all the instances, signifying that the model is able to generate the solution with a high customer satisfaction level. Table 5 shows that the computational time increases with the increase in the number of activities and planning days. When the planning horizon is 2 days, the computational time is almost negligible. However, when the number of activities is 30 and planning days is 6 then the computational time reaches more than 12 minutes. This is due to the reason that with the increase in the number of activities and days, the number of possible combinations of activities to be selected and scheduled within the given planning days increases remarkably.

Table 5: Result obtained for the hypothetical test problem

#no	Number of Activities	Planning days	Budget	No of Selected Activities	Cost ($)	Distance (km)	Score	CPU time (min: sec)
1	10	2	L: 200	8	172	89	0.80	00:01
2			M: 350	10	183	85	0.91	00:15
3			H: 500	10	183	85	0.96	01:36
4		4	L: 200	10	169	91	0.82	00:32
5			M: 350	10	169	91	0.92	00:58
6			H: 500	10	169	91	0.94	01:07
7		6	L: 200	10	169	91	0.82	03:11
8			M: 350	10	169	91	0.92	01:09
9			H: 500	10	169	91	0.97	00:19
10	20	2	L: 200	8	172	89	0.80	00:00
11			M: 350	11	199	95	0.77	00:01
12			H: 500	12	206	96	0.79	00:12
13		4	L: 200	10	185	93	0.77	01:06
14			M: 350	19	327	162	0.77	02:00
15			H: 500	20	387	176	0.79	16:00
16		6	L: 200	10	178	86	0.84	06:00
17			M: 350	20	331	178	0.77	07:04
18			H: 500	20	387	176	0.79	10:40
19	30	2	L: 200	10	182	76	0.73	00:01
20			M: 350	11	174	88	0.81	00:11
21			H: 500	12	198	103	0.76	00:12
22		4	L: 200	11	189	84	0.73	00:01
23			M: 350	19	324	159	0.75	02:03
24			H: 500	20	387	176	0.79	10:48
25		6	L: 200	11	194	85	0.73	13:01
26			M: 350	20	324	164	0.75	14:16
27			H: 500	27	417	233	0.76	12:12

5.3 Sensitivity Analysis and managerial insights

As the weights of the criteria will have a significant impact on our mathematical model outcomes, a sensitivity analysis is conducted to know the effect of changing the weight of the objective functions. For the sensitivity analysis, the weight of each criterion is randomly varied into three levels (0, 0.333, 0.667). For example, if the weight of the criterion "number of activities" is zero, then the weight of "cost" will be either 0.333 or 0.667 and that of "distance" will be either 0.667 or 0.333, such that the sum of the three weights is always 1. The effect of changes in the weights is analyzed based on the main effect plots and the analysis is carried out using the Mini-Tab software (www.minitab.com) at a 95% confidence interval. The plots are used to observe the effect on the number of activities considered, total cost, distance travelled and score with respect to the change in the weights assigned to the objective functions.

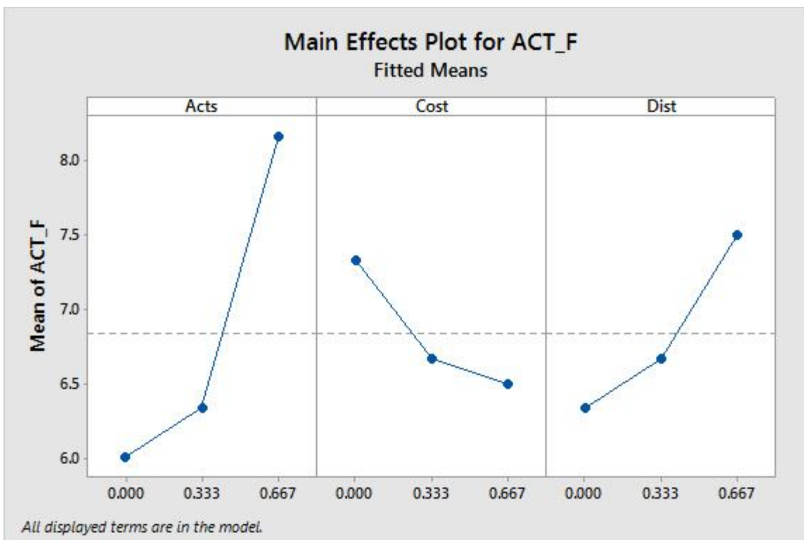


Figure 3: Main effect plot on activities function

Figure 3 shows the main effect plot for the number of activities when the weights of activity, cost and distance in the objective function change. The plot shows that as the weight given for the number of activities increases, this latter increases and the effect is remarkable. Even though the number of activities increases with the increase in the weight for the distance, the effect is moderate. In contrast, increasing the weight of the cost function has the opposite effect on the number of activities and the effect is moderate. Similarly, we carried out several additional experiments with the aim of analyzing the combined effect of the different weights of the total cost, distance travelled and score. For the sake of brevity, not all the plots were reported here, but rather the obtained results of the analysis are summarized in Table 6. From the table, it is evident that the total cost increases with the increase in the weight of activity and distance functions. However, the effects are high and moderate, respectively. It is to be also noted that the total cost decreases with the increase in the weight for the cost function and the effect is significant. The distance travelled decreases with the increase in the weights of all the functions and even here the effect is notable. On the other hand, the score increases as expected with the increase in the weights of all the functions. The effects are high with respect to activities number and distance while it is low for the cost weight.

Table 6: Summary of the main effect plots

Effect on:	**Weight increase of:**		
	Activities	**Cost**	**Distance**
Number of Activities	Increase (High)	Decrease (Moderate)	Increase (Moderate)
Total Cost	Increase (High)	Decrease (High)	Increase (Moderate)
Distance traveled	Decrease (High)	Decrease (High)	Decrease (High)
Score (objective value)	Increase (High)	Increase (Moderate)	Increase (High)

The above sensitivity analysis results provide very valuable managerial insights to the tourism service provider that will take advantage from the observed trends with the aim of increasing the customers' satisfaction. For example, the service provider should discourage his/her customers from any weights combination that may produce a high effect on the cost component. This is because, usually customers are sensitive to the cost component and will feel (unfulfilled or frustrated) to observe significant cost increase (even though still within their pre-specified budget threshold) as a result of a minor change in some of the criteria weights. Usually, sales managers are perfectly aware of these kinds of techniques but need to be supported by quantitative insights, as those reported in Table 6, in order to ensure that customers are satisfied with the touristic package they are selecting, even compared with other available and feasible ones.

6 CONCLUSIONS

The tourism supply chain aims at satisfying the needs of the end-users based on their preferences. However, the preferences of each individual tourist may be different from others. Some tourists would like to give more importance to the cost reduction, while others would be satisfied if they can enjoy more activities within the given planning period and preset budget limit. Moreover, there are tourists who may have several preferences and want to concurrently optimize multiple criteria. Consequently, from the optimization point of view, the tourist preferences can be expressed through a multi-objective framework, in such a way that some objectives need to be minimized while others need to be maximized.

This research aims at designing optimum personalized tourism packages based on the preferences of the tourists. The research developed a mathematical model that optimizes multiple objectives by using the weighted average method based on the tourist's preferences when they plan a trip. The method allows tourists to express their preferences of one objective over others by assigning the most appropriate weight to each criterion. Moreover, as the integrated objectives have different units of measurements, the normalization method is implemented to ensure accurate optimization outcomes. At first, the mathematical model was tested on a small case example to check its validity and applicability. The model was able to generate meaningful itineraries for the two-day planning horizon with the activity sequence for each day under high and medium budget constraints. Further, the model was tested on a hypothetical, but realistic, test setting for an extended planning horizon and with a higher number of activities. The analysis shows that an increase in the value of all the considered parameters i.e., the number of activities, days and budget has an impact on the number of selected activities, cost and distance travelled. For all the test instances, the score is above 0.7, indicating that the model can generate touristic plans with a high satisfaction level. Moreover, a sensitivity analysis was conducted to check the effect of varying the weights assigned to each criterion on the model outcomes and to draw useful managerial insights.

Our reported results related to the hypothetical test settings have shown that the execution time to solve the problem increases significantly with the increase in the number of activities and planning horizon. The optimization model dealt with in this research involves a routing problem, which is known to be NP-hard (Triki et al., 2020). It is expected that the developed full space mathematical model cannot solve the problem of large-scale instances (for e.g. when planning horizon is above 2 weeks with 100's of activities to be selected from), related to bigger (countries or cities??), in a reasonable computational time. Therefore, there will be a need to develop efficient heuristic or meta-heuristic methods to solve big size problems efficiently. Further, even though the tourism industry can ensure significant economic development, the research has shown that the industry has negative environmental and social impacts (Northcote and Macbeth, 2006). Moreover, people are increasingly concerned about the impact of their behaviour on the environment even during their entertainment activities. Therefore, the concept of sustainability can be integrated to generate the so-called green touristic itineraries for the OPTP. Finally, the effect of lockdown periods and travelling disruptions due to exceptional natural and health phenomena on the tourism industry of a city like Muscat should be studied and assessed (Hosseini et al., 2021).

References

1. Brito, J., Expíósito-Márquez, A., & Moreno, J. A. (2017). A fuzzy GRASP algorithm for solving a Tourist Trip Design Problem. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)* (pp. 1-6). IEEE.
2. Cotfas, L. A., Diosteanu, A. N. D. R. E. E. A., Dumitrescu, S. D., & Smeureanu, A. L. E. X. A. N. D. R. U. (2011). Semantic search itinerary recommender system. *Int. Journal of Comput*, 5(3), 370-377.
3. De Falco, I., Scafuri, U., & Tarantino, E. (2016). Optimizing personalized touristic itineraries by a multiobjective evolutionary algorithm. *International Journal of Information Technology & Decision Making*, 15(06), 1269-1312.
4. Expósito, A., Mancini, S., Brito, J., & Moreno, J. A. (2019). A fuzzy GRASP for the tourist trip design with clustered POIs. *Expert Systems with Applications*, *127*, 210-227.
5. Gavalas, D., Kasapakis, V., Konstantopoulos, C., Pantziou, G., Vathis, N., & Zaroliagis, C. (2015). The eCOMPASS multimodal tourist tour planner. *Expert systems with Applications*, *42*(21), 7303-7316.

6. Golden, B. L., Levy, L., & Vohra, R. (1987). The orienteering problem. *Naval Research Logistics (NRL)*, *34*(3), 307-318.
7. Gunawan, A., Lau, H. C., & Vansteenwegen, P. (2016). Orienteering problem: A survey of recent variants, solution approaches and applications. *European Journal of Operational Research*, *255*(2), 315-332.
8. Hapsari, I., Iswadi, H., & Cahyadi, Y. O. (2019). Tourism Route Planning Route Based on Tourist Category using Orienteering Problem with Time Windows. *GARUDA (Global Research on Tourism Development and Advancement)*, *1*(1), 42-49.
9. Herzog, D., & Wörndl, W. (2014). A travel recommender system for combining multiple travel regions to a composite trip. In Proceedings of the new trends in content-based recommender systems (CBRECSYS) workshop, 8th association for computing machinery ACM conference on recommender systems, California, USA.
10. Hosseini, S. M., Paydar, M. M., Alizadeh M. & Triki C. (2021). Improving the ecotourism centers services during the Covid-19 pandemic: A real case study. Submitted to *Computers & Industrial Engineering*.
11. Lim, K. H., J. Chan, C. Leckie, & S. Karunasekera (2018). Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. *Knowl. Inf. Syst.,* 54(2), 375–406.
12. Malucelli, F., Giovannini, A., & Nonato, M. (2015). Designing single origin-destination itineraries for several classes of cycle-tourists. *Transportation Research Procedia*, *10*, 413-422.
13. Mancini, S., & Stecca, G. (2018). A large neighborhood search based matheuristic for the tourist cruises itinerary planning. *Computers & Industrial Engineering*, *122*, 140-148.
14. Muthuraman, S., & Al Haziazi, M. (2019). Smart Tourism Destination-New Exploration towards Sustainable Development in Sultanate of Oman. In *5th International Conference on Information Management (ICIM-2019)* (pp. 332-335). IEEE.
15. Neto, F. (2003). A new approach to sustainable tourism development: Moving beyond environmental protection. *Natural resources forum*, 27(3), 212-222).
16. Northcote, J., & Macbeth, J. (2006). Conceptualizing yield: Sustainable tourism management. *Annals of Tourism Research*, *33*(1), 199-220.
17. Persia, F., Pilato, G., Ge, M., Bolzoni, P., D'Auria, D., & Helmer, S. (2020). Improving orienteering-based tourist trip planning with social sensing. *Future Generation Computer Systems*, *110*, 931-945.
18. Piya. S., Takahashi, K., & Morikawa, K. (2009). Introducing quotation and negotiation structure in the order acceptance decision model. *International Journal of Operations and Quantitative Management*, *15*(4), 293-318.
19. Piya, S., Shamsuzzoha, A., Khadem, M., & Al-Hinai, N. (2020). Identification of critical factors and their interrelationships to design agile supply chain: special focus to oil and gas industries. *Global Journal of Flexible Systems Management*, *21*(3), 263-281.
20. Rebollo, J. F. V., & Baidal, J. A. I. (2003). Measuring sustainability in a mass tourist destination: pressures, perceptions and policy responses in Torrevieja, Spain. *Journal of Sustainable Tourism*, *11*(2-3), 181-203.

21. Sigala, M. (2008). A supply chain management approach for investigating the role of tour operators on sustainable tourism: the case of TUI. *Journal of cleaner production*, *16*(15), 1589-1599.
22. Sylejmani, K., Muhaxhiri, A., Dika, A., & Ahmedi, L. (2014). Solving tourist trip planning problem via a simulated annealing algorithm. In *2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1124-1129). IEEE.
23. Szpilko, D. (2017). Tourism Supply Chain–overview of selected literature. *Procedia Engineering*, *182*, 687-693.
24. Tenemaza, M., Luján-Mora, S., De Antonio, A., & Ramírez, J. (2020). Improving Itinerary Recommendations for Tourists Through Metaheuristic Algorithms: An Optimization Proposal. *IEEE Access*, *8*, 79003-79023.
25. Triki, C., Mirmohammadsadeghi S. and Piya. S. (2017). Heuristic Methods for the Periodic Shipper Lane Selection Problem in Transportation Auctions. *Computers and Industrial Engineering*, 106: 182-191.
26. Triki, C., Piya, S., & Fu, L. L. (2020). Integrating production scheduling and transportation procurement through combinatorial auctions. *Networks*, *76*(2), 147-163.
27. Urrutia-Zambrana, A., Tirado, G., & Mateos, A. (2021). Variable neighborhood search to solve the generalized orienteering problem. *International Transactions in Operational* Research, 28 (2021) 142–167.
28. Vansteenwegen, P., Souffriau, W., Berghe, G. V., & Van Oudheusden, D. (2011). The city trip planner: an expert system for tourists. *Expert Systems with Applications*, 38(6), 6540-6546.
29. Verbeeck, C., Sörensen, K., Aghezzaf, E. H., & Vansteenwegen, P. (2014). A fast solution method for the time-dependent orienteering problem. *European Journal of Operational Research*, *236*(2), 419-432.
30. Yu, V. F. Redi, A.A.N. P., Jewpanya, P. & Gunawan, A. (2019). Selective discrete particle swarm optimization for the team orienteering problem with time windows and partial scores. *Computers & Industrial Engineering* 138. 106084.
31. Wong, C. U. I., & McKercher, B. (2012). Day tour itineraries: Searching for the balance between commercial needs and experiential desires. *Tourism Management*, *33*(6), 1360-1372.
32. World Tourism Organization (WTO), Organization for Economic Cooperation and Development (OECD) and Statistical Office of the European Commission (EUROSTAT), 2001. Tourism Satellite Account: Recommended Methodological Framework. ST/ESA/ STAT/SER.F.80. United Nations, New York.
33. WTTC - World Travel & Tourism Council (2015). Travel & Tourism Economic Impact 2015, Oman, Url: http://www.wttc.org/-/media/files/reports/economic%20impact%20research/regional%202015/world2015.pdf
34. Yochum, P., Chang, L., Gu, T., & Zhu, M. (2020). An Adaptive Genetic Algorithm for Personalized Itinerary Planning. *IEEE Access*, *8*, 88147-88157.

35. Yu, J., Aslam, J., Karaman, S., & Rus, D. (2015). Anytime planning of optimal schedules for a mobile sensing robot. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 5279-5286). IEEE.
36. Zhang, X., Song, H., & Huang, G. Q. (2009). Tourism Supply Chain Management: A New Research Agenda. *Tourism Management*, 30(3), 345-358.