**Intelligent generation of touristic itinerary**

Radu Galan

# Abstract

The problem we are talking about is a notorious one and it was many names, given mostly by the way in which we choose to solve it. One of the most recurrent is “The orienteering problem” first cited in 1987 in (Bruce Golden 1987). It can go even further back in a bit of a different shape as the “Traveling salesman problem” in 1930 by the mathematician Karl Menger. We seek to analyze any work that is attempting to create an algorithm, system, or application that wants to solve a similar problem, but also takes into consideration a few other parameters related to travelling and the creation of an itinerary. Some such parameters can be a given time-frame for visiting, public transportation route instead of walking, taking into consideration the availability window for a certain point of interest (POI), the user’s profile, multiple travelling days.

# Introduction

Motivations can very easily be found for this particular topic. The cities continue to grow along with the world population and this means both that more and more touristic objectives appear (especially including activities, parks, and other modern places), but also that there are more people that are probable to travel and visit memorable places from. Moreover, the simple fact that so many people are starting to travel has already transformed into a problem itself. Many tourists mean that there will be longer waiting lines, that restaurant and hotels will be more likely to be busy or unavailable all together, and the transportation system will most likely fail in the most crowded cities. This reduces the value of the time spent in these touristic objectives drastically due to longer waiting time, longer travelling time from one point to another, and even impossibility of dining or checking in.

This is an industry worth 1.2 trillion $ and despite this very fact most trips are organized by tourists themselves. This is both very uncomfortable and very bad for the cities themselves. This process takes a lot of time and is also very hard to be done by someone who has no knowledge of the local activities. While there are some local tour guides that attempt to solve this problem by old methods, it is more than clear that they are becoming overwhelmed. Not only that tourists lose their time for organizing, but it is also very likely that they will not choose the best places to visit and it might even give them a false impression of the place. Also, it is usually hard, nearly to impossible, to choose to proper means of public transportation, consider all the visiting time at every attraction, take into account the particular schedule of each attraction, and then be able to visit and experience as much as possible.

Another very simple problem is the individuality of each tourist. Not all of them are interested in the most popular tourist location, and still they are guided towards them by travel guides and agencies. This highly reduces the potential of a trip as being enjoyable.

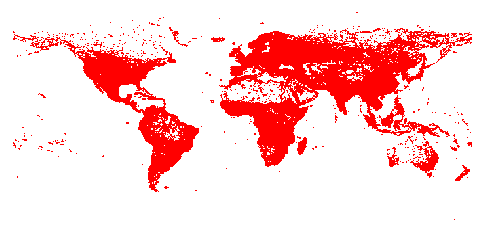
The purpose of most of the work studying this problem is to make travelling easier (at least for when it will be available again) by liberating the user from all the monotonous and repetitive decision he has to take during the planning of a touristic travel. Mainly those tasks include the choosing of a location, of a list of touristic objectives (Points Of Interest), the route and itinerary containing these objectives.

In the next chapters we will explain the big picture of the given problem and then continue by analyzing the problem itself along with some solutions found in the last years and we will end with some conclusions and remarks.

# Big picture

First of all this is a problem that originates in the “Traveling salesman problem” in 1930 by the Austrian mathematician Karl Menger. That is obviously a simpler version given that it doesn’t take into account most of the problem that we are having today. This TSP problem was further studied and extended to multiple variations like the Travelling Salesman Problem with Priority Prizes (TSPPP) (Admilson Alcântara da Silvaa 2018) or Selective Travelling Salesman Problem with Profits (STSPwP) (Koszelew 2014), but most importantly is that it evolved into what we call today “The orienteering problem” (Bruce Golden 1987). The name originates from the sport in which each player seeks to reach multiple points on a map from which he receives various number of points. The winner is that who achieves the most points in a limited amount of time. Although this was usually played in remote outdoor areas like forests, now those principles are heavily applied for the problem that mass tourism poses. This also has multiple variations like Team Orienteering Problem (TOP), Team Orienteering Problem with Time Windows (TOPTW) (Koszelew 2014).

A more complicated version of this problem that corresponds even better to our situation is the Tourist Tour Generation Problem (TTGP) very thoroughly explained by (Koszelew 2014).

The world TSP problem is also worth mentioning ( <http://www.math.uwaterloo.ca/tsp/world/> ). This is an open data library that provides a dataset of 1.9 million cities all over the world which have already been solved with a lower bound of 7’512’218’268 by CPLEX linear-programming solver. It currently keeps track of TSP algorithms that solves the problem and the actual leader is at only 0.0474% error from the optimal tour.

# Problem description

Generally speaking, by having information like:

-a list of possible touristic objectives (POS) in multiple location and multiple details about each of them regarding their schedule, activity type, coordinates, etc;

-a complete profile of the user and his preferences,

we want to construct an intelligent itinerary creator that can deduce this information:

-the ordered list for touristic objectives (sorted using the user’s preferences)

-an itinerary with the touristic objectives for the selected period taking into consideration as many parameters as possible.

This problem’s complexity is quite high in general, and it is also categorized as NP hard (Garey M 1979).

In her synthetization (Koszelew 2014) present the most important features that algorithms of this type have:

• Selecting and Routing(SaR): General features that generates the most optimal route taking into consideration points given for each objective and a route as short as possible that fits a time frame.

• Possibilities of Returns (POR): The generated route could contain attractions twice, but on the second visit the points given will be 0. This is a helpful feature for algorithms that use graphs.

• Obligatory POIs (OPs): Introduces the notion of “must see” attraction that must be found in the final route.

• OpeningHours (OH): The route will only include an attraction in the route if by the approximated time the tourist visits, it will still be open.

• ScenicRoutes (SR): Attempts creating routes that will pass near objectives like monuments, architectural buildings.

• PublicTransportation (PT): Considers multiple types of transportation aside from personal car ro walking, like: train, plane, bus, etc;

• BudgetLimitations (BL): The objectives on the route will not exceed a certain budget limit.

• DynamicRecalculation (DR): allows the recalculation of the route during the visit in order to cover unforeseen events.

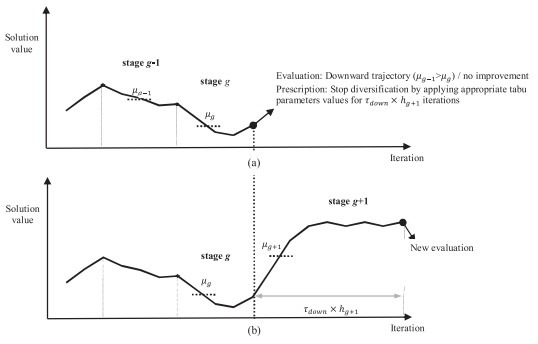
• Max − nT ype (MnT): Constraint that only allows a limited number of a certain category of touristic objectives.

• ObligatoryTypes (OT): Constraint that only allows a route to be valid if it contains a minimum from a certain category of touristic objectives.

# Tabu search for priority prizes - (Admilson Alcântara da Silvaa 2018)

Some of the most important parameters this article takes into account are: the tourist’s profile (preferences for the order of the objectives, importance of certain objectives), other tourists profiles, distances between objectives and the costs.

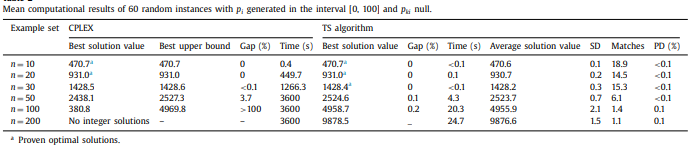
They describe their problem as “combining aspects of these two TSP variants (Travelling salesman problem), PTP (Profitable Tour Problem) and TSPPP (Travelling Salesman

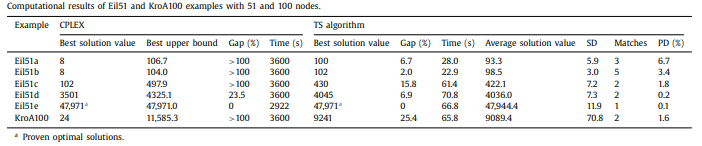
Problem with Priority Prizes)”. They are given a list of touristic objectives, prize values for each of the objectives, cost for each of the travels, and they want to find a route with minimum cost and maximum value.

The algorithm used is “tabu search algorithm“ based on a previous adaptive tabu search approach. Their method attempts to solve the problem step by step choosing the most optimal choice in their current situation based also on the past choices.

More exactly the algorithm begins by selecting all the possible nodes (touristic attractions), it then constructs the route iteratively from a partial route comprised of a randomly selected node. Checks the next node that has the highest profit and repeat until all nodes are routed. With the tabu parameters set a local search will be applied for two consecutive stages and repeat until the maximum number of iterations is reached (or until no improvement has been gained).

After testing the algorithm against CPLEX with branch-and-cut algorithm it obtained satisfying results on multiple databases, but with a runtime of over 1 hour.



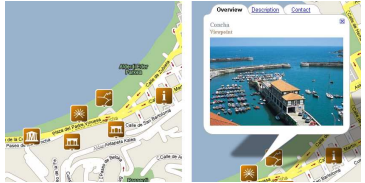


# Operating Research - (Ander Garciaa 2009)

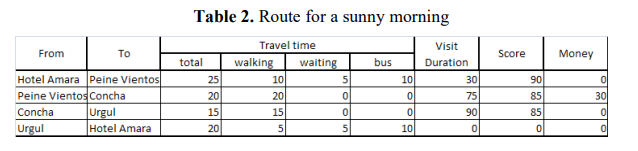
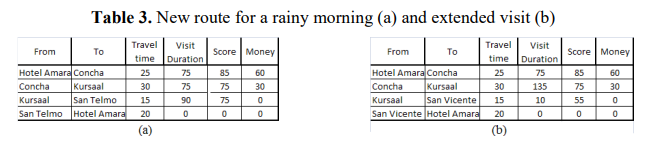
Radu Galan

This method attempts to replace the current “Personalized Electronic Tourist Guide (PET)” with an intelligent routing system and takes into consideration multiple parameters like: public transportation, varying travelling times, rush-hours variation in time, real time computation to react to unexpected events, applies heuristics to create routes efficiently, even for very big dataset of touristic attractions. The algorithm used falls in the category: Multi Path Team Orienteering Problem with Time Windows (MPTOPTW).

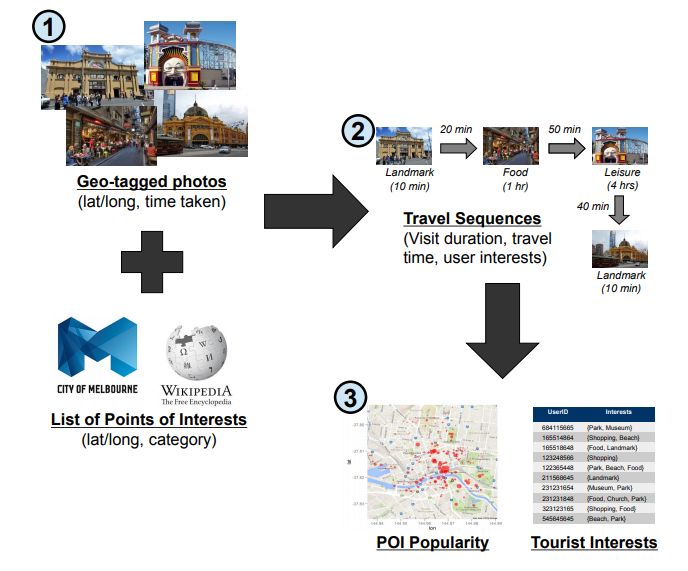
Proposes the usage of Operating Research in order to solve to problem in real time (less than 5 seconds) and adaptive changes (bus miss, longer visit, etc.), but focuses on the routing algorithm.

It applies the algorithm on the city of San Sebastian and uses more precisely information like: a list of attractions (ID, position, customized value for each day, opening and visiting times, cost); information about the public transportation network; preferences of the tourist; the current position of the tourist; and the available resources (time and money).

The actual technique used it Iterated Local Search, metaheuristic method that build sequences generated by an embedded heuristic, which should give better solution then multiple iteration over the same heuristic.

The actual public transportation calculator does not work, but it does take into account the weather, but as it has been tested so far (on human experiences evaluations) it has received decent feedback.

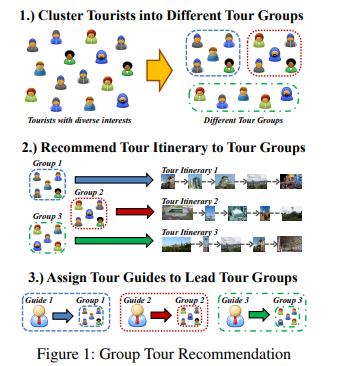
# Group tour recommendation (Lim 2016)

This method attempts to solve the problem of tour operators that only deal with groups of tourists that come in fixed sizes. They propose to: optimize for an appropriate tour group size, creating tours that appeal to multiple people, assign tour guides to each tour group.

The actual method of gathering the data attempts to work in 3 main steps:

-map photos with gps data to a list of points of interests based on distance difference computed with Haversine’s formula.

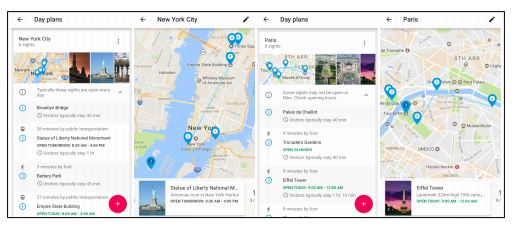
-create the travel history of Points of interest for each tourist from the previous mapped images, also save the visit duration

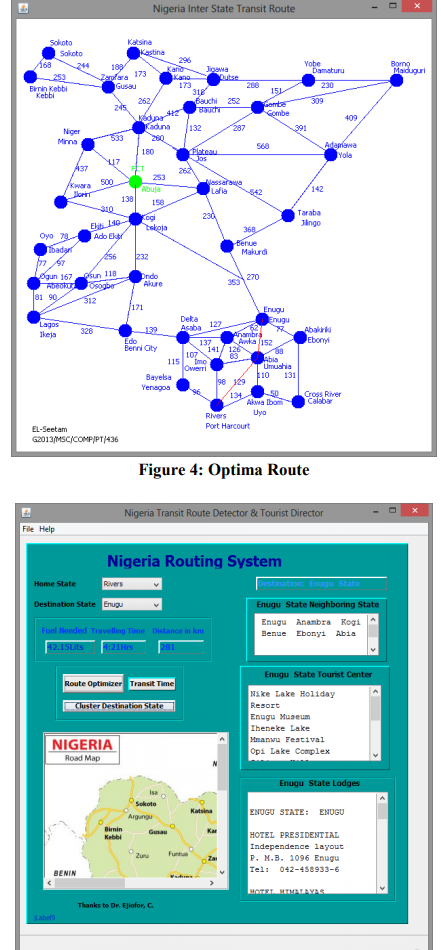
-calculate the popularity of each Point of interest and associate with user’s preferences, this will result in a list of valuable extracted information

The phase of tour recommendation is split in two categories: for individual tourist and for a group of tourists. The first category has two main methods: PERSTOUR and TOURRECINT. PERSTOUR will use visit count and duration to understand user interests for multiple categories which will dynamically recommend how much time someone is likely to stay at a certain place. TOURRECINT optimized orienteering problem that takes into account mandatory Points of interest. The second category has another two main methods: GROUPTOURREC and SST for community detection. GROUPTOURREC attempts to make the best suggestion of a tour guide that fits the people in a group, basically clustering tourists together to a centroid being a guide. SST or special-spatial-temporal links are used in order to detect communities of users that happen to spend time in similar locations.

(Kwan Hui Lim 2016)

# Other mentions

A very interesting study has been conducted by a team of researchers from Google regarding travel itineraries generation. (Zachary Friggstad 2018) They obviously take advantage of the access they have to data and also attempt to improve current generating algorithms for multiple days. They do this by trying to balance out the touristic contents of each day.

A really comprehensive method for group itinerary is studied in (Kadri Sylejmani 2017) that uses a very advanced method of nested tabu search meta heuristic with a neighborhood structure. They create two different version: a fast one and a slow one. Both of them are rigorously tested by users and standard TSP tests alike.

Another interesting paper which is a little different (at least regarding the initial motivations) is trying to find the optimal road between cities in Nigeria using graphs algorithms like Dijsktra’s and A\*. (Emmanuel L. Seetam 2019) The motivation is concentrated towards finding to most optimal route between cities in Nigeria considering how bad the roads are and how much fuel is used between any two cities with direct access.

# Conclusion

We have seen multiple methods solving the same problem that comes under different names (traveling salesman problem, orienteering problem). All of the methods presented choose to solve it different and each of them has a unique point of view towards the problem. Many more research articles will still be written for this problem since the need in the society is growing higher and higher.

# Bibliography

Admilson Alcântara da Silvaa, Reinaldo Morabito, Vitória Pureza. 2018. "Optimization approaches to support the planning and analysis of." *Expert systems with applications* 321-330.

Ander Garciaa, Maria Teresa Linaza, Olatz Arbelaitz, Pieter Vansteenwegen. 2009. "Intelligent Routing System for a Personalised Electronic Tourist Guide." *Information and Communication Technologies in Tourism* 185-197.

Bruce Golden, L. Levy, Rakesh V. Vohra. 1987. "The orienteering problem." *Naval Research Logistics 34(3)* 307-318.

Emmanuel L. Seetam, Ledisi G. Kabari. 2019. "Intelligent Agent Based Itinerary System for Route Optimizer and Tourist Guide." *International Journals of Advanced Research in Computer Science and Software Engineering* 90-97.

Garey M, Johnson D. 1979. *COmputers and intractability.* San Francisco: Freeman.

Kadri Sylejmani, Jurgen Dorn, Nysret Musliu. 2017. "Planning the trip itinerary for tourist groups." *Inf Technol Tourism* 275-314.

Koszelew, Jolanta. 2014. "Optimization Problems and Methods Applicable in Intelligent Tourist Travel Planners." *Trends in Contemporary Computer Science, Podlasie* 127-134.

Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, Shanika Karunasekera. 2016. "Towards Next Generation Touring: Personalized Group Tours." *Proceedings of the 26th International Conference on Automated Planning and Scheduling (ICAPS’16).*

Lim, Kwan Hui. 2016. "Recommending and Planning Trip Itineraries for Individual Travellers and Groups of Tourists." *Association for the Advancement of Artificial.*

Zachary Friggstad, Sreenivas Gollapudi, Kostas Kollias, Tamas Sarlos, Chaitanya Swamy, Andrew Tomkins. 2018. "Orienteering Algorithms for Generating Travel Itineraries." *WSDM’18, February 5-9* 180-188.