ENGENHARIAS, TECNOLOGIA, GESTÃO E TURISMO
ENGINEERING, TECHNOLOGY, MANAGEMENT AND TOURISM
INGENIERÍA, TECNOLOGÍA, ADMINISTRACIÓN Y TURISMO

millenium 83

Millenium, 2(10), 83-90.

en

AMOSTRA DE MONTE CARLO PARA O PROBLEMA DO PROJETO DE VIAGEM TURÍSTICA
MONTE CARLO SAMPLING FOR THE TOURIST TRIP DESIGN PROBLEM
MUESTREO DE MONTE CARLO PARA EL PROBLEMA DEL DISEÑO DEL VIAJE TURÍSTICO

XiaoChen Chou[1]
Luca Maria Gambardella[1]
Roberto Montemanni[1, 2]

[1] IDSIA - Dalle Molle Institute for Artificial Intelligence (USI-SUPSI), Switzerland
[2] Dept. of Engineering Sciences and Methods, University of Modena and Reggio Emilia, Italy

XiaoChen Chou - xiaochen@idsia.ch | Luca Maria Gambardella - luca@idsia.ch | Roberto Montemanni - roberto.montemanni@unimore.it

Corresponding Author
XiaoChen Chou
Dalle Molle Institute for Artificial Intelligence (USI/SUPSI)
Galleria 2, 6928 Manno, Switzerland
xiaochen@idsia.ch

RECEIVED: 27th September, 2019
ACCEPTED: 30th September 2019

Chou, X., Gambardella, L. M. & Montemanni, R. (2019). Monte Carlo sampling for the tourist trip design problem.
Millenium, 2(10), 83-90.
DOI: https://doi.org/10.29352/mill0210.09.00259

RESUMO

Introdução: O Problema de projeto de viagem turística é uma variante de um problema de planeamento de rotas para turistas interessados em vários pontos de interesse. Cada ponto de interesse tem disponibilidades diferentes e um certo índice de satisfação pode ser alcançado quando é visitado.
Objetivos: O objetivo é selecionar um subconjunto de pontos de interesse a visitar dentro de um determinado orçamento de tempo, de modo que a pontuação de satisfação do turista seja maximizada e o tempo total de viagem seja minimizado.
Métodos: No modelo proposto, o cálculo da disponibilidade de um PI é baseado no tempo de espera e / ou na previsão do tempo. No entanto, pesquisas mostram que a maioria dos turistas prefere viajar dentro de uma área lotada e limitada de PIs muito atraentes por razões de segurança e porque sentem um maior controlo.
Resultados: Neste trabalho, demonstramos que o modelo existente do Problema de Orientação Probabilística se encaixa em uma variante probabilística desse problema e que as técnicas de Amostragem de Monte Carlo podem ser usadas dentro de um solucionador de heurísticas para fornecer soluções com eficiência.
Conclusões: Neste trabalho demonstramos que o modelo existente do Problema Probabilístico de Orientação se encaixa no Problema Estocástico de Projeto de Viagem Turística. Propusemos uma maneira de resolver o problema usando técnicas de Amostragem de Monte Carlo num solucionador heurístico e discutimos várias possíveis melhorias no modelo. Uma extensão adicional do modelo será desenvolvida para solucionar problemas mais práticos, no futuro.

Palavras-chave: O Problema do Projeto de Viagem Turística; Problema Probabilístico de Orientação; Amostragem Monte Carlo; Otimização Combinatória.

ABSTRACT

Introduction: The Tourist Trip Design Problem is a variant of a route-planning problem for tourists interested in multiple points of interest. Each point of interest has different availability, and a certain satisfaction score can be achieved when it is visited.
Objectives: The objective is to select a subset of points of interests to visit within a given time budget, in such a way that the satisfaction score of the tourist is maximized and the total travel time is minimized.
Methods: In our proposed model, the calculation of the availability of a POI is based on the waiting time and / or the weather forecast. However, research shows that most tourists prefer to travel within a crowded and limited area of very attractive POIs for safety reasons and because they feel more in control.
Results: In this work we demonstrate that the existing model of the Probabilistic Orienteering Problem fits a probabilistic variant of this problem and that Monte Carlo Sampling techniques can be used inside a heurist solver to efficiently provide solutions.
Conclusions: In this work we demonstrate the existing model of the Probabilistic Orienteering Problem fits the stochastic Tourist Trip Design Problem. We proposed a way to solve the problem by using Monte Carlo Sampling techniques inside a heuristic solver and discussed several possible improvements on the model. Further extension of the model will be developed for solving more practical problems.

Keywords: The Tourist Trip Design Problem; Probabilistic Orienteering Problem; Monte Carlo Sampling; Combinatorial Optimization

RESUMEN

Introducción: El problema de diseño del viaje turístico es una variante de un problema de planificación de rutas para turistas interesados en múltiples puntos de interés. Cada punto de interés tiene una disponibilidad diferente, y se puede lograr un cierto puntaje de satisfacción cuando se visita.
Objetivos: El objetivo es seleccionar un subconjunto de puntos de interés para visitar dentro de un presupuesto de tiempo determinado, de tal manera que se maximice el puntaje de satisfacción del turista y se minimice el tiempo total de viaje.
Métodos: en nuestro modelo propuesto, el cálculo de la disponibilidad de un PDI se basa en el tiempo de espera y / o el pronóstico del tiempo. Sin embargo, la investigación muestra que la mayoría de los turistas prefieren viajar dentro de un área abarrotada y limitada de puntos de interés muy atractivos por razones de seguridad y porque se sienten más en control.
Resultados: en este trabajo demostramos que el modelo existente del problema de orientación probabilística se ajusta a una variante probabilística de este problema y que las técnicas de muestreo de Monte Carlo se pueden usar dentro de un solucionador heurista para proporcionar soluciones de manera eficiente.
Conclusiones: En este trabajo demostramos que el modelo existente del Problema de Orientación Probabilística se ajusta al Problema estocástico del Diseño del Viaje Turístico. Propusimos una forma de resolver el problema utilizando técnicas de muestreo de Monte Carlo dentro de un solucionador heurístico y discutimos varias posibles mejoras en el modelo. Se desarrollará una extensión adicional del modelo para resolver problemas más prácticos.

Palabras-clave: el problema del diseño del viaje turístico; el problema de orientación probabilística; el muestreo de Monte Carlo; la optimización combinatória

DOI: https://doi.org/10.29352/mill0210.09.00259

INTRODUCTION

Each city has its own tourist attractions, either natural beauty such as lakes and mountains, or cultural places such as museums and historical locations. When a tourist visits a new place, how does she/he usually arrange the trip?

The Tourist Trip Design Problem (TTDP) is a variant of a route-planning problem for tourists interested in visiting multiple points of interest (POI). Suppose that a tourist has her/his own rank of POIs that she/he wants to visit most, that each of the places has different availability, and a certain satisfaction score can be achieved when visited. The objective is to select a subset of POIs to visit within the length of the stay, a given time b, in such a way that the satisfaction score of the tourist is maximized, while the total time spent between attractions and the total travel time is minimized. The planning is normally done in advance, typically in the days immediately before the trip. For POIs with deterministic availability, a simple formulation of the TTDP is proved to be identical with the Orienteering Problem (OP), (Vansteenwegen, Souffriau, Vanden Berghe, & Van Oudheusden, 2009), where a route with maximum score is determined for a subset of locations with fixed depot and destination, limited by the time budget. However, in practice, popular POIs may require long waiting time and open-air POIs may not be visited in case of unseasonable weather. A variety of uncertainties could then affect the availability of the POIs, modelling the availability of POIs with probabilities will therefore lead to more realistic models. Since finding the best route for a deterministic TTDP with large number of POIs is already a rather time-consuming problem, how practical could it be to consider such a stochastic TTDP?

In this paper we aim at demonstrating that the existing model of the Probabilistic Orienteering Problem (POP) Angelelli, Archetti, Filippi, & Vindigni (2017), fits the TTDP well, and we introduce a heuristic method to solve the problem based on Monte Carlo Sampling techniques [3] (Chou, Gambardella, & Montemanni, 2018).

1. LITERATURE REVIEW AND PROBLEM DEFINITION

A detailed definition of the TTDP can be found in Vansteenwegen, et al (2009). The objective function of the POP is defined in Angelelli, et al, 2017). The non-polynomial calculation makes such an objective function hard to compute. Analytical approximations for the expected total travel costs of the POP, leading to exact and heuristic algorithms, have been proposed in Angelelli, et al, 2017) and [8] Campbell & Thomas (2008)

In the [3] Chou, Gambardella, & Montemanni, (2018), we proposed an approximation by using the Monte Carlo Sampling method. As a state-of-the-art approach for several stochastic/probabilistic vehicle routing problems, similar use of the Monte Carlo sampling method can be found in the Probabilistic Traveling Salesman Problem with Deadlines [6] (Weyland, Montemanni, & Gambardella, 2013).and the Orienteering Problem with Stochastic Travel and Service Times [7] (Papapanagiotou, Montemanni, & Gambardella, 2015)..

According to the definition of the POP, we denote with V= {0,1 ... n,n+1} a set of n points of interests with the depot being node 0 and the destination being n+1. Let t_{ij} be the expected travelling time from POI i to POI j. The given time budget (length of the stay) is T_{max}. The probability of a POI i to be visitable (according to weather forecast and/or waiting time information) is modeled by a Bernoulli variable b_i = {0,1} which takes value 1 with probability π_i. The probability of each POI is considered independent from the others. Each POI has also a satisfaction score p_i representing how valuable is the attraction.

In TTDP, both the depot and the destination are the hotel, which is always available with no satisfaction score to be collected, therefore $\pi_0 = \pi_{n+1} = 1$ and $p_0 = p_{n+1} = 0$. For a POI, the probability of availability π_i is calculated based on weather information and expected waiting time.

A point of interest may be valued differently by different people, therefore the satisfaction score p_i is a set of integers in {1,2,..n} given representing by the personal interests of the person planning the tour. The more you she/he wants to visit a POI, the higher the score is. This personal selection may base on information from websites or the weather conditions. For example, p_i of outdoor POIs decreases when it rains. Besides, if the tourist had already visited a POI before, the corresponding p_i will decrease as well.

A tour τ : $i_0 = 0, i_1, i_2, \ldots, i_q, i_{q+1} = n+1$ is defined as a sequence of q POIs selected to be visited, plus the depot and the destination. The satisfaction score of a tour is $P(\tau)$, and the travel time is $T(\tau)$. We aim at simultaneously maximizing $P(\tau)$ and minimizing $T(\tau)$. Therefore, the objective function is the difference between the expected total satisfaction and the expected total travel time with a coefficient C:

$$u(\tau) = E\big(P(\tau)\big) - CE(T(\tau)) \qquad (1)$$

The coefficient C is a balance between the distance travelled and the satisfaction score. When C is small it means the satisfaction score is more important. In this case, if there is one point of interest that you really want to visit, even though it is far away or it has high probability of being crowded, you still tend to search for a tour that includes this POI. When C takes large values, it is the opposite and the trip will tend to cover attractions grouped together.

Chou, X., Gambardella, L. M. & Montemanni, R. (2019). Monte Carlo sampling for the tourist trip design problem.
Millenium, 2(10), 83-90.
DOI: https://doi.org/10.29352/mill0210.09.00259

Given a tour τ, evaluating the objective function (1) is computationally and time demanding, due to the stochastic nature of the problem. As proved in [3], Chou, Gambardella, & Montemanni, 2018), the Monte Carlo Sampling methods provides a fast and effective practical technique to approximate this kind of objective function.

2. THE MONTE CARLO SAMPLING TECHNIQUE AND A METAHEURISTIC ALGORITHM

In this part we approximate the objective function (1) by using Monte Carlo Sampling. Such an approach is especially useful when not all the POIs are available with probability 1. On the other hand, it has been extensively shown in the literature [3] (Chou, Gambardella, & Montemanni, 2018), [6] (Weyland, Montemanni, Gambardella, 2013), [7] Papapanagiotou, Montemanni, & Gambardella, (2015) that such an approach is able to approximate well the objective function (1) with a very low computational effort.

First, for a TTDP with n POIs, we generate a set of s scenarios with different available POIs by sampling according to the probabilities π_is. Deterministic objective function value can be computed for each scenario, the average value of all scenarios shows the approximation value of a given tour. The approximation value will be more precise when we use more samples, but the computational time will also be longer. Therefore, this procedure requires parameter tuning for s.

In our previous work [3] (Chou, Gambardella, & Montemanni, 2018), we also used the same Monte Carlo Sampling procedure to optimize the subset of attractions that it is worth to visit in the given time, by stopping the evaluation when the deadline T_{max} is incurred. This will save the time for evaluating irrelevant locations and at the same time automatically select the most interesting POIs that it is worth to visit in the given time.

The second step is to embed the Monte Carlo Evaluator into a 2-opt local search algorithm. We start with a solution generated by random and evaluate it with the Monte Carlo Evaluator. We then cross over to reorder the route and compare the evaluation value of each possible combination in order to find the optimal solution. Since local search methods may get stuck in local minimums, we apply iterated random start 2-opt local search within a given time limit. With this metaheuristic algorithm, the Probabilistic Orienteering Problem can be solved fast and effectively.

3. RESULTS

General Probabilistic Orienteering Problem instances have been proposed in Angelelli, et al (2017) together with a set of exact and heuristic algorithms. Angelelli, et al (2017) We compare our Monte Carlo 2-opt algorithm with the heuristic methods of [2] Angelelli, et al (2017) (we refer the interested reader to this paper for details about the methods). Exact methods are reported to be unable to converge to optimality for all the instances considered, and therefore are not presented in the table. For each method we report the average gap from the best-known solution and the average computational time over the relevant 84 instances from [2] Angelelli, et al (2017).

Table 1 shows that our Monte Carlo 2-opt algorithm is extremely competitive with state-of-the-art heuristics on medium/small POP instances, that are relevant for TTDP applications.

Table 1 – Experimental comparison of heuristic methods on POP instances from [2] Angelelli,et al (2017) with dimension up to 30

Angelelli et al, (2017) [2]							
Smart Chain		Smart Path		Smart-Path two-ways		Monte Carlo 2-opt	
gap (%)	time (s)	gap (%)	time (s)	gap (%)	time (s)	gap (%)	time (s)
2.52	286	3.93	282	2.72	345	0.75	10

When we consider specific applications to the TTDP, there are no specific instances available in the literature. Therefore, we test the evaluator with an example instance generated by the 15 top attractions in Paris, France (Figure 1). The red pinpoint is the location of the hotel from which the tour starts and ends. In the test instance we pick a hotel at one of the attractions.

We collected online information for the travelling time1 between POIs, the average staying time at each POI, and the average queues situation of each POI in order to generate the test instance. The average staying time at each POI varies from 15 minutes to 180 minutes. We set the visitable probability from 0.6 to 0.9 based on the historical statistics of the average queueing time in inverse proportion. The ideal calculation of the probability should also consider an additional small parameter that measures the probabilities of meeting bad weather for outdoor POIs. We did not add this parameter in this test, and this will not influence the evaluation of the performance of the Monte Carlo evaluator.

[1] maps.google.com

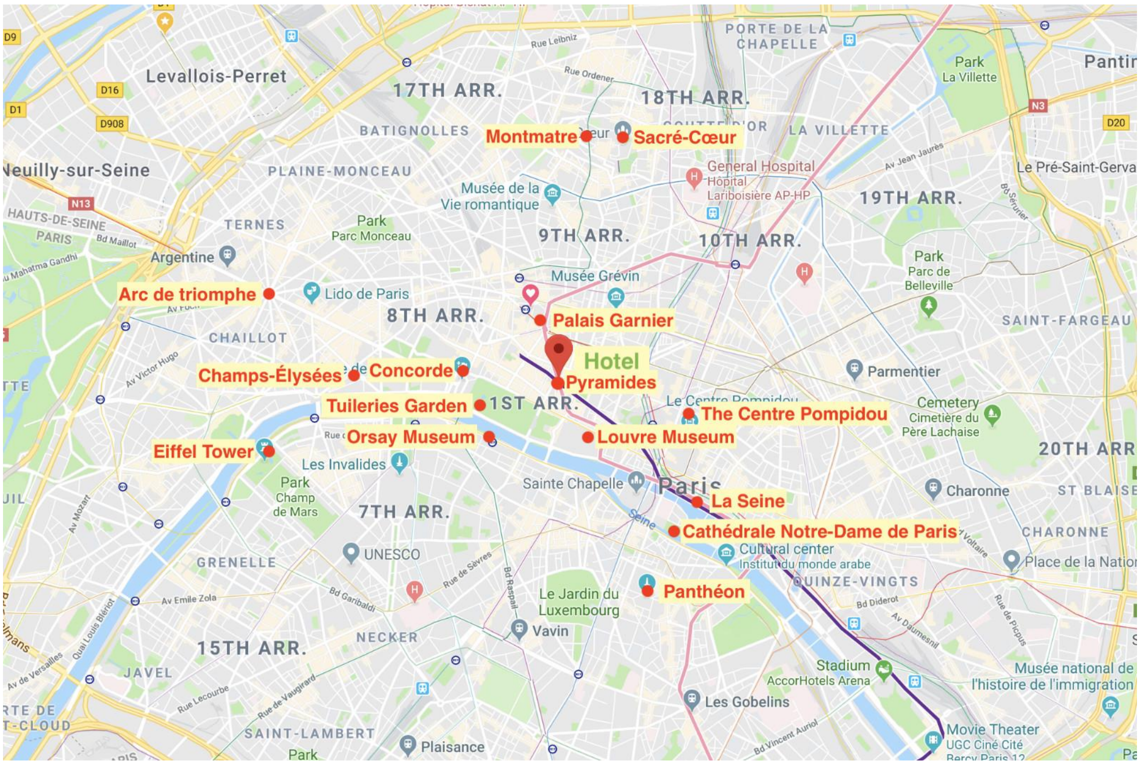

Figure 1 - Locations of 15 Top Attractions in Paris

First, we set equal satisfaction score ($p_i = 1$) for all POIs, Figure 2 shows the evolution over time of the heuristic solution cost provided by the method described in Section 2 with 50 samples. The objective function value is on y-axis and the time is on x-axis. We can observe that the evolution stabilizes and achieves result close to the exact value in 0.01 seconds. The evaluator will then iterate the process and keep searching for a better solution until it reaches the given time limit.

In this process, we are also interested in computational speed with different values for the number of samples s used. Since it only takes 0.01 second for one iteration, we count the number of iterations per second. In Figure 3, the number of samples s is on x-axis and the number of iterations per second with different s is on y-axis. We can see that, for $s < 100$, the number of iterations per second drops dramatically, and for $s > 100$ there is a trend of gradual decrease. Considering the accuracy and consistency of the results, we choose $s = 100$ for the next experiments.

Now we move on to the case when each POI has a different satisfaction score. Among the 15 top attractions, the Louvre Museum is a POI that is always crowded and requires long visiting time. When we set equal satisfaction score for all POIs, it will never appear on the optimal route, for the reason that we care more about saving time and visiting more places (Figure 4.a). However, the Louvre Museum is a very worthwhile place to visit. Therefore, we give Louvre Museum a relatively high score when setting different satisfaction score for each POI, and we obtain the personalized optimal solution that covers less POIs but includes places that the tourist really wants to see (Figure 4.b.).

Chou, X., Gambardella, L. M. & Montemanni, R. (2019). Monte Carlo sampling for the tourist trip design problem.
Millenium, 2(10), 83-90.
DOI: https://doi.org/10.29352/mill0210.09.00259

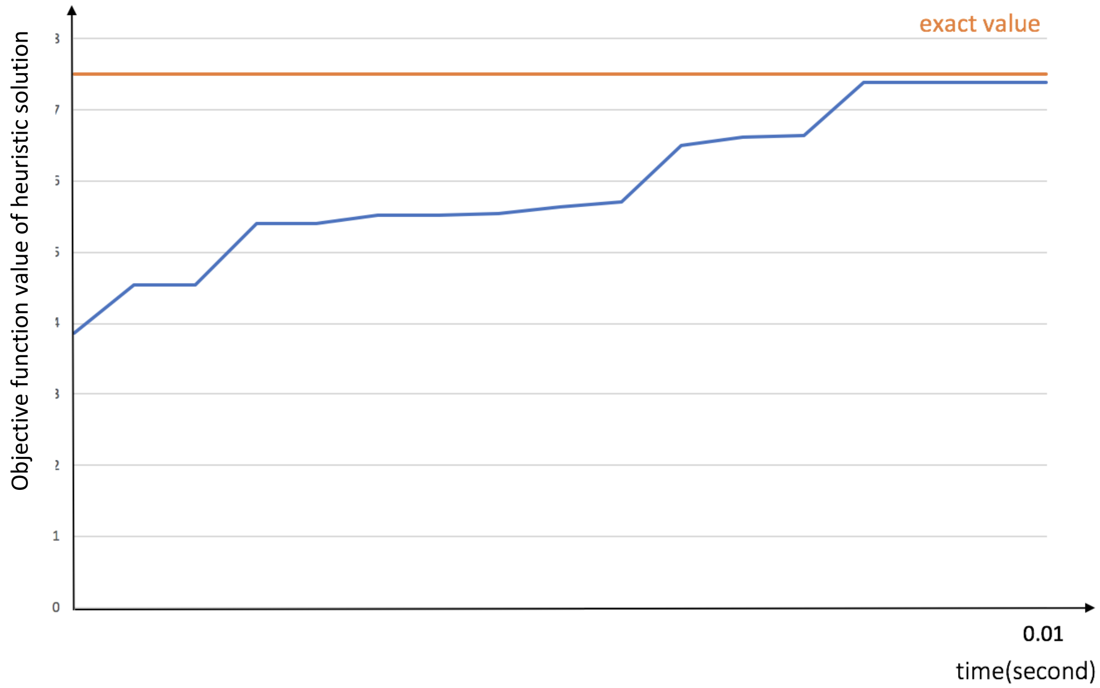

Figure 2 - Evolution of the heuristic solution cost over time for an example instance

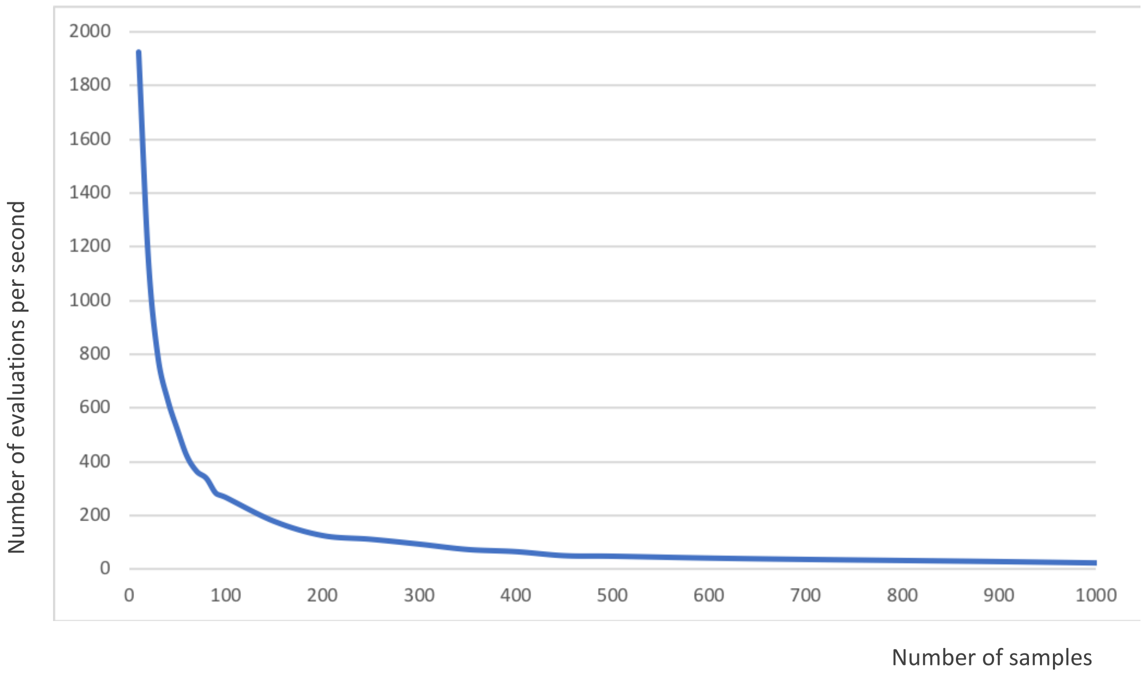

Figure 3 - Computational Speed with different number of samples s for an example instance

Chou, X., Gambardella, L. M. & Montemanni, R. (2019). Monte Carlo sampling for the tourist trip design problem.
Millenium, 2(10), 83-90.
DOI: https://doi.org/10.29352/mill0210.09.00259

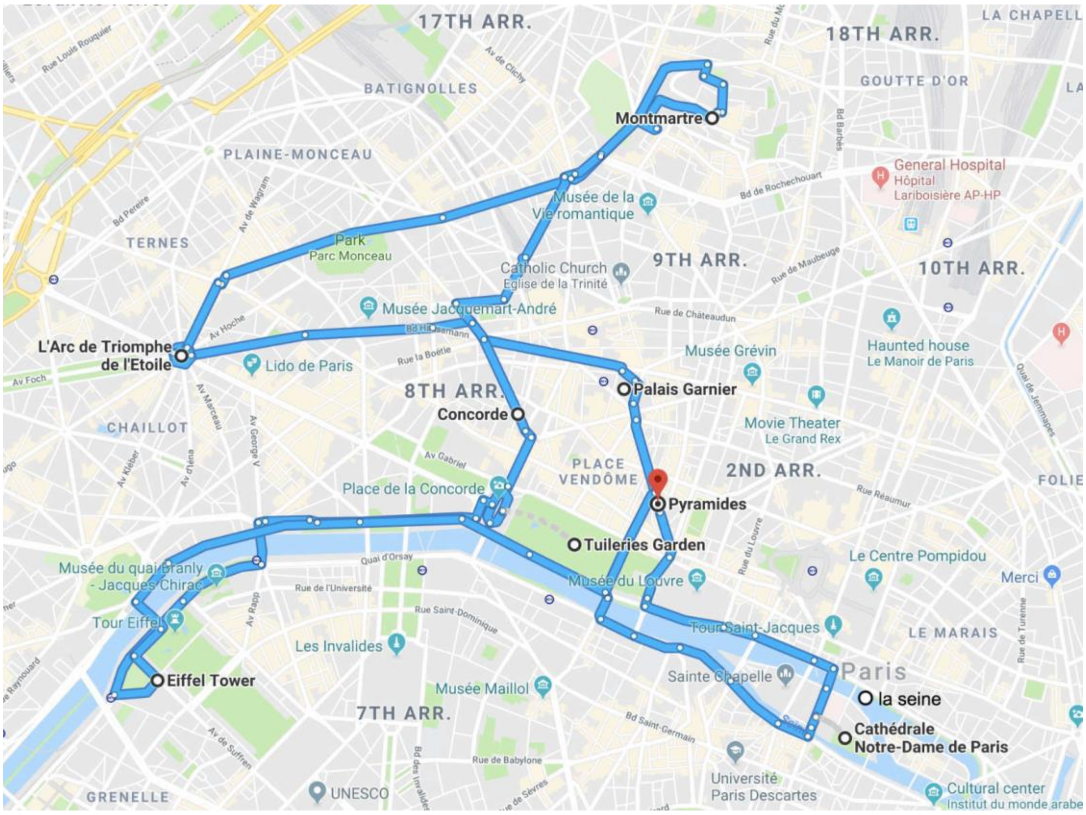

Figure 4.a - Optimal solution for POIs with equal satisfaction score for all POIs

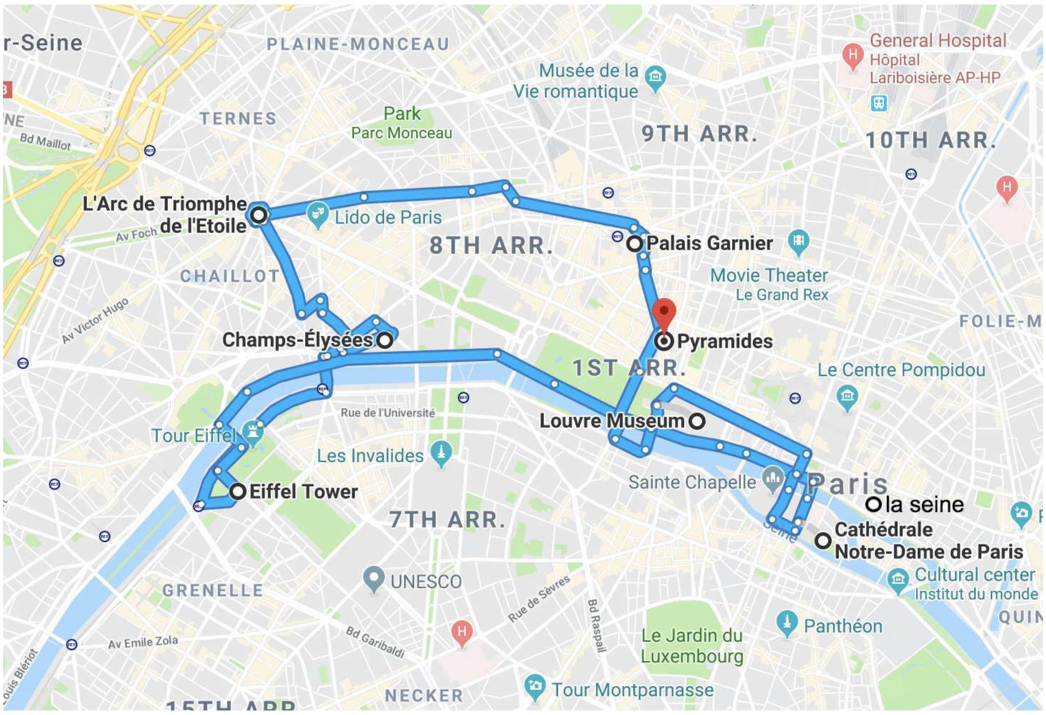

Figure 4.b - Optimal solution for POIs with different satisfaction scores for the POIs

Chou, X., Gambardella, L. M. & Montemanni, R. (2019). Monte Carlo sampling for the tourist trip design problem.
Millenium, 2(10), 83-90.
DOI: https://doi.org/10.29352/mill0210.09.00259

It is worth noting that some of the attractions are adjacent, therefore the optimal solution might not be unique, being different solutions virtually equivalent.

4. MODEL WITH EXTRA FEATURES

As proved in Section 2, the POP model is adaptive to the Tourist Trip Design Problem. By taking personal interests, time limits, the POIs' locations and real-time availabilities as input, our solver will be able to return the selection and the routing of the POIs as output.
There is also some additional flexibility in this model: the starting and ending points do not have to coincide. For example, the tour can start at the hotel and finish at a station or airport.
With up-to-date POI information and changeable personal preference, a feasible selection of POIs and a tour can always be suggested in a reasonable time. On the other hand, a drawback of the POP model we propose is that it can only solve the problem for one day, or a fraction of it (e.g. the afternoon of a business trip when all the work is carried out in the morning).
In case the tour covers multiple days, since hotel is a special location with unpredictable staying time, it could be natural to extend the approach we propose by solving several sub-problems for each day and each POI appears only once in one day. Unfortunately, the pre-allocation of POIs to days of such an approach might compromise the optimality of the solution. In order to overcome this issue, a possible extend of this model would be to consider the Probabilistic Team Orienteering problem [4] where the POIs are split to the different days of the stay automatically by the solver.
Besides, in our current model the calculation of the availability of a POI is based on waiting time and/or weather forecast only. However, investigation shows that most tourists prefer to move within a limited crowded area of very attractive POIs for safety reasons and because they feel more in control [5]. When dealing with similar situation with our model, it is high probable that our solver suggests a less attractive POI with less people, but this does not meet tourists' needs. Therefore, putting more specified preference constraints apart from satisfaction scores into the model will be another improvement.

CONCLUSIONS

In this work we demonstrate the existing model of the Probabilistic Orienteering Problem fits the stochastic Tourist Trip Design Problem. We proposed a way to solve the problem by using Monte Carlo Sampling techniques inside a heuristic solver and discussed several possible improvements on the model. Further extension of the model will be developed for solving more practical problems.

ACKNOWLEDGEMENTS

Xiaochen Chou was supported by the Swiss National Science Foundation through grant 200020-182360: Machine Learning and Sampling Techniques for Stochastic Vehicle Routing Problems.

REFERENCES

Angelelli, E.; Archetti, C.; Filippi, C., & Vindigni, M. (2017). "The probabilistic orienteering problem". *Computers and Operations Research*. 81, 269-281.

Campbell, A.M. & Thomas, W. (2008). "Probabilistic traveling salesman problem with deadlines". *Transportation Science*. 42(1), 1-21.

Chou, X., Gambardella, L. M., & Montemanni, R. (2018). "Monte Carlo Sampling for the Probabilistic Orienteering Problem". *New Trends in Emerging Complex Real Life Problems.* AIRO Springer, Serie s 1, 169-177.

Gunawana, A., Laua, H. C., & Vansteenwegenb, P. (2016). "Orienteering Problem: A Survey of Recent Variants, Solution Approaches and Applications". *European Journal of Operational Research*. Volume 255, Issue 2, 315-332.

Kramer, R., Modsching, M., & Ten Hagen, K. (2006). "A city guide agent creating and adapting individual sightseeing tours based on field trial results". *International Journal of Computational Intelligence Research*. 2(2), 191–206.

Papapanagiotou, V., Montemanni, R., & Gambardella, L. M. (2015). "Hybrid sampling-based evaluators for the orienteering problem with stochastic travel and service times". *Journal of Traffic and Logistics Engineering*. 3(2), 108-114.

Vansteenwegen, P., Souffriau, W., Vanden Berghe, G., & Van Oudheusden, D. (2009). "Metaheuristics for tourist trip planning". *In* Lecture Notes in Economics and Mathematical Systems Springer, 15-31.

Weyland, D., Montemanni, R., & Gambardella, L. M. (2013). "Heuristics for the probabilistic traveling salesman problem with deadlines based on quasi-parallel monte carlo sampling". *Computers and Operations Research.* 40(7), 1661-1670.