iJournals: International Journal of Software & Hardware Research in Engineering
ISSN:2347-4890, ijournals.in/ijshre
Volume 12 Issue 6, June 2024

Proposing an efficient Taxi Share System in Euclidean space based on CTP and skyline

Author: Justin Park [1]; Ryan Hong[2]

Affiliation: Saint Paul Preparatory Seoul[1]; Yongsan International School of Seoul[2]

E-mail: justinparkwork@gmail.com[1]; rhong418@gmail.com[2]

ABSTRACT

Passengers' mobility platform Conatus's "half taxi" has been revived for the first time in 40 years. The city expected that the legislation of shared services would solve transportation problems such as difficulties in riding and increase passenger convenience, but limitations were found in the commercialization stage. The key issue is that the percentage of passengers is too low, the taxi fare is too low, or the above conditions are met. To address this, we propose a more efficient system in this paper. CTP solves the problem of the underworld. This ensures a 70% chance of route intersection sharers within 1 km by allowing users to walk reasonable distances. To ensure a higher probability of matching by adopting the Skyline query, the system recommends several taxi choices through the skyline based on several factors. Group Skyline method is adopted to calculate the optimal passenger group. Finally, we adopted a why-not query to motivate drivers and suggest appropriate guidelines for improving performance. In the result section, Adopting CTP helps the matching during the time complexity isn't increasing compared to the original algorithm. The paper concludes with suggestions in the aspect of real-time query and privacy.

Keywords: Database, Skyline, Group Skyline, Why-not Query, CTP.

1. INTRODUCTION

The taxi ride-on-sharing service, which had been banned for 40 years, has been legalized from 2022 in Korea. Passengers will be able to reduce the burden of taxi fares by utilizing the "half-and-half taxi" of the mobility platform Conatus. According to related industries, the taxi development law revised in July last year will take effect on July 28 after a six-month grace period. As a result, taxi rides, which were banned in 1982, have been revived for the first time in 40 years[1].

The shared ride was a method in which a taxi driver picked up a passenger to share, regardless of passenger's intention. It was banned by law due to frequent stops, some potential criminals, and problems with fare calculation. The Seoul Metropolitan Government tried to push for voluntary shared rides in 2015 as a measure to resolve the shortage of late-night taxi rides, but withdrew the effort due to the opposition of the majority.

The city expects that the legalization of the shared service will solve transportation problems such as difficulties in riding and increase passenger convenience. For the shared service, half-and-half taxis were selected as the regulatory sandbox of the Ministry of Science and ICT in 2019, and were piloted in some areas of Seoul. The regulatory sandbox is a system that suspends regulations when the current regulations put a brake on the launch of products and services of companies using new technologies.

Using the official mobile application, among passengers with a distance of less than 1km between departure points, it automatically connects a vehicle with 70% of the same path. There is also a designated seat separation and an insurance device for the ride.

The calling fee per passenger is 3,000 won. A taxi driver earns 5,000 won from 6,000 won for two people who paid 1,000 won for the platform (Conatus). For example, if a passenger rides at a distance that costs 30,000 won, the passenger can

pay 18,000 won (taxi fee of 15,000 won + call fee of 3,000 won) each. Taxi drivers earn a total of 36,000 won. Excluding the 1,000 won commission, they earn 5,000 won more than if they were carrying a passenger (30,000 won). Riding is only possible when the gender is the same.

However, there's room for improvement. In fact, the usage of this official app shows how the current system is far from the user preference. The core problem is that: too low passenger rate or too low taxi rate that meets the conditions above. To solve this issue, this paper proposes a more efficient system. Here are the contribution points of the research:

1) Adopting CTP: to solve low-passenger problem via CTP – which makes the users walk a reasonable distance so that the probability of the 70% of path intersection sharers within 1km.

2) Adopting Skyline query: to assure the higher probability of matching, the system recommends several taxi choices via skyline based on some factors.

3) Adopting Group Skyline query: to calculate the best group of passengers, the system uses group skyline query.

4) Adopting why-not query: to motivate the drivers and suggests proper guideline for higher performance, the system hires why-not query.

5) Analysis: For each method, efficiency analysis has been described in each section.

In chapter 2, the background knowledge of each method is explained. Description of the proposed system in chapter 3. In chapter 4, the result of evaluation is followed. The paper ends with conclusion and future works in chapter 5.

2. Background

2.1 Normal or Body Text

A group travel planning query (CTP) generates a meeting point, reducing the total distance traveled for multiple travel query points[2]. Nearby query points move to meeting points and collectively to their final destinations. This has often been simulated in group travel of people, with ride-sharing companies following similar query variants. For people, various customer query points that want to travel to similar locations are assigned to meeting points, so they travel together in one vehicle. This reduces the cost and carbon footprint of customers, as multiple cars can take each customer query point to the same destination.

The original cost function is as follows:.

$$C(x) = \sum dist(qj,d) \cdots (1)$$

, where qi represents ith query point and d to the destination.

<Formula 1> above shows the sum of the calculated distances between each query point and distance. The formula reflects the situation in which transportation travels back and forth between all points and destinations, even if the proximity of the query points is different.

The cost of applying the CTP query is as follows:

$$C(x,y) = \sum dist(qi,mj) + \sum dist(mj,d) \cdots (2)$$

, where mj represents the meeting point of the proper query points set.

<formula 2> reflects the CTP query, where the total distance between all query points and each query point is added to the total distance from each query point to the destination. The key difference from the original feature is that the journey from the query point to the destination is shortened to the meeting point, where all query points are moved collectively to the destination.

2.2 Skyline

With the development of Internet accessibility and storage media, modern users can make decisions through vast databases. However, too many candidates can cause confusion for users. Therefore, research has been conducted to help users select candidates that are comfortable and satisfactory. Among them, skyline queries are queries that exclude data with relatively low probability of being the optimal choice in large quantities and diverse dimensions[3].

2.2.1 Group Skyline

The group skyline refers that the skyline spots when it comes to the group. For example, in Fig 2, the skyline points are {p1, p4, p9, p10}. However, suppose that we are trying to choose the 'qualified' groups with the size k = 2.

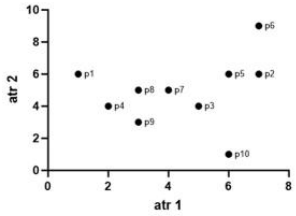


Figure 1. the scatter graph of Table 1.

obj	x	y
p1	1	6
p2	7	6
p3	5	4
p4	2	4
p5	6	6
p6	7	9
p7	4	5
p8	3	5
p9	3	3
p10	6	1

Table 1. example data

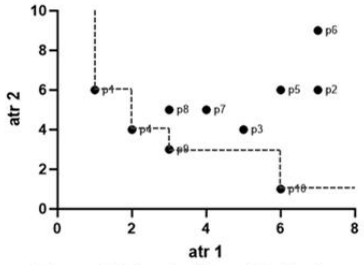


Figure 2. The skyline of Table 1.

First, the relation of the domination should be identified. In the group skyline, the domination relation is determined based on the group element. Suppose the group A = {p1, p8} and the group B = {p1, p4}. We can only compare p8 and p4 since p1 is in common. Because p8 is dominated by p4, we can say, B dominates B. In this way, groups those doesn't dominated by any group becomes the group skyline[4].

2.3 Why-not query

There are also studies to find out why a particular decision was not considered a good decision and was not returned. The question of why it is not the correct answer is called the why-not question. Research on this question is particularly important in that it can be used as a measure of valuation including the potential of specific products and branches in the field of public relations/marketing[3].

For example, let's say that a travel agency combines and commercializes several hotels according to the customer's schedule. If the customer stays for two days (k=2), it is optimal for the travel agency to commercialize the group. In this case, hotel owners who do not belong to the returned groups can determine through the why their hotel was not a target product of the travel agency and whether it is worth investing in order to belong to the product. This is very important in terms of the economic aspect of commerce.

Why-not group skyline queries are not limited to the commercial field. When the number of people who fit into the TO is selected in consideration of all aspects, such as NBA team composition issues, it can also have a great effect in the field of talent selection, such as what points a player who has not been selected should make more efforts for. Applicants will inevitably have numerical scores in various fields after the AI interviewer is mentioned, and a criterion for selecting applicants based on the number of people who can proceed to the face-to-face interview stage is needed. A group skyline query can be one of these criteria, and at this time, a why-not group skyline query from the perspective of failed applicants will be useful.

3. Materials and Methods

3.1 Adopting CTP for higher matching rate

3.1.1 previous approach

In the previous approach, the match of users and the taxi evokes in those conditions:

1) The users should be in 1KM range
2) The common path should be over 70 percentage
3) The overall path shouldn't be over 10min distances compare to the original path (each)

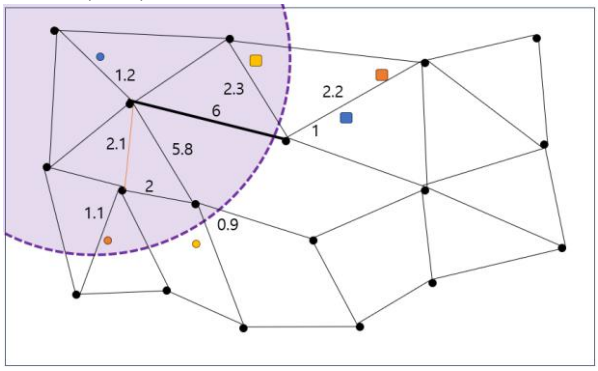


Figure 3. the previous approach without CTP

For example, suppose the situation of Fig 2. The dots represent the users, and the square represents the destination. There are two possible share – {blue and orange} or {yellow and orange}. Suppose that the group with blue dot. The violet circle represents the range of 1KM. Because the

orange dot is the only one within 1km. However, the common path (the bold line) doesn't over 70% of the whole path of the orange one. Also, the total path of taxi driving neither seems to be short enough. Thus, the match fails in those similar situations.

3.1.2 Adopting CTP

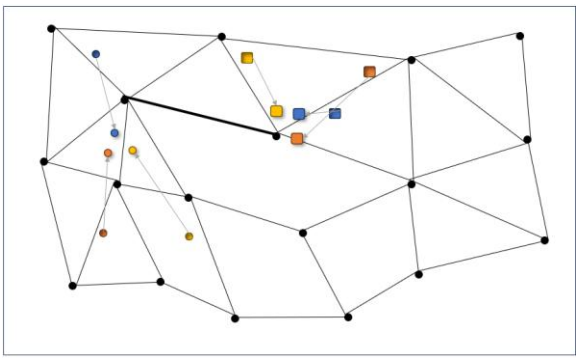


Figure 4. the proposed system with CTP

Via CTP, the users meet at the meeting points, and then ahead to the co-destination. From the co-destination the users should walk some. This not only makes the probability of matching higher but also gives advantage in the aspect of sustainability. In addition to the current application (give the users mileage points due to the environment protection via the car sharing), the carbon footprint would reduce since the total distance that taxi should drive reduces.

proceed to the face-to-face interview stage is needed. A group skyline query can be one of these criteria, and at this time, a why-not group skyline query from the perspective of failed applicants will be useful.

3.2 Adopting the Skyline Query

To assure the higher probability of matching, the system recommends several taxi choices via skyline based on some factors.

3.2.1 Considerable factors

To adopt skyline query for users to choose the proper taxi, there are some factors to consider:

1) The distance between the taxi and the meeting point of the users.
2) Average rates from the previous users.
3) Predict cost based on the number of current passenger

3.2.2 Analysis

The time complexity of this process be $O(n \log n)$, where n is dependent to the number of taxi. The system needs the range limit to avoid over calculation.

3.3 Adopting the Group Skyline

To apply skyline query for taxi drivers, the user data can not be used in single since the match should be made in group.

3.3.1 making group

To adopt group skyline query for taxi to choose the proper users, there are some factors to consider:

1) The distance between the taxi and their meeting point of the users.
2) Average rates from the previous taxi drivers.
3) The walking distance

3.3.2 Analysis

The time complexity of this process also be $O(n \log n)$, where n is dependent to the number of possible groups.

3.3 Why-not query

For the drivers and the application holder, the system provides why-not query. For the driver, it can be motivation. For the application holder, this process gives some idea about low matching rate.

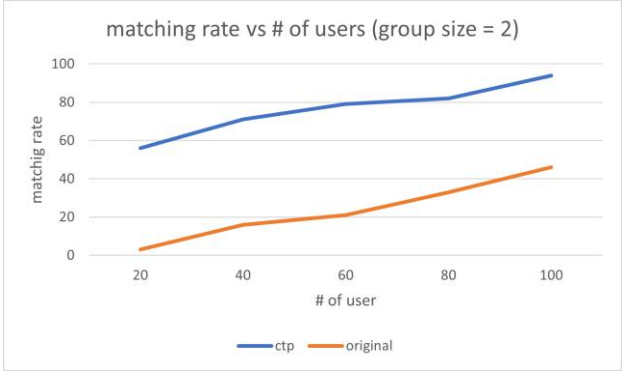
4. Experimental Result

4.1 Environment

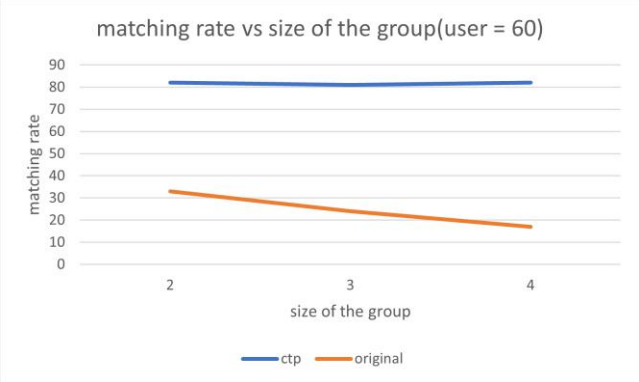
The experiment has been conducted with processer intel i9 in python. The Road network generated synthetic. The distribution of the user data is random.

4.2 Results

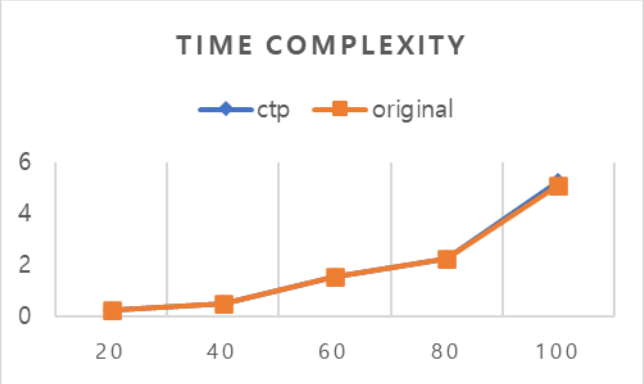
4.2.1 matching rate performance (varying the user)



4.2.2 matching rate performance (varying the size)



4.2.3 time complexity of CTP adopted algorithm



5. Conclusion

In this research, we proposed a system that adopts CTP to solve the low rate of matching in the current taxi sharing service. To apply CTP, the sacrifice of the user isn't unavoidable. Thus, we proposed a proper limit and the compensational system. In addition, we adopted skyline query, group skyline for users and taxi to select their best preference option. In this process, we defined the factors which should be considered.

6. Discussion

There is room for some improvement. First, we should consider the real-time system. Because the taxi moves in real time, and perhaps another customer takes the taxi.

Second suggestion is based on the idea of privacy protection. Because the start point and the destination could be recalculated (in range), we will continue our research with differential privacy system - which add noise to the data.

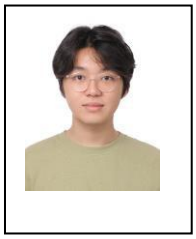
References

[1] Hayan Choi(2023), "Kakao Taxi is allowed to share a ride from tomorrow… Only the same sex is allowed for 5 passengers or less", news article in https://www.hani.co.kr/arti/economy/economy_general/1046939.html

[2] Junku Lee, "한정된 자원을 고려한 집단 여행 계획 질의 처리 기법." 국내석사학위논문 서강대학교 대학원, 2020.

[3] Kun Young Park, & Seok Park (2021). Why-not Group Skyline Query and Pruning Heuristic Method with Group Skyline Groups. 한국정보과학회 학술발표논문집, 개최지.

[4] Y. Yang, W. Lu and C. Tang, "A fast top-group skyline query method based on skyline layer," 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), Sanya, China, 2020, pp. 146-151, doi: 10.1109/ICHCI51889.2020.00040.



Justin Park
He has graduated from the Engineering Summer Academy at UPenn in Computer Science

Ryan Hong
He is attending Yongsan International School of Seoul\