Online Ridesharing with Meeting Points

Jiachuan Wang¹, Peng Cheng², Libin Zheng³,Lei Chen¹, Wenjie Zhang⁴ ¹Hong Kong University of Science and Technology, Hong Kong, China ²East China Normal University, Shanghai, China ³Sun Yat-sen University, Guangzhou, China ⁴The University of New South Wales, Australia

Outline

- Background and Motivation
- The Meeting-Point-based Online Ridesharing Problem
- Framework Overview
- Methods
- Experimental Evaluation
- Summary

Ridesharing in the World

- Online platforms for ridesharing grows rapidly.
	- Each driver can serve more than one request when their routes have common sub-routes

Route Planning for Ridesharing

- Effective/efficient route planning strategy is highly demanded due to:
	- A large number of dynamically arriving requests
	- A large number of drivers
	- A large number of possible routes allowing share
	- Limited response time

New Mode: Meeting Points

Traditional route planning

- Requests are posted with source locations and destination locations
- Platform organizes drivers to pass these locations and serve riders

New Mode: Meeting Points

Traditional route planning

- Requests are posted with source locations and destination locations
- Platform organizes drivers to pass these locations and serve riders

However, due to the complex topology of the city road network, some locations (e.g., **A** and **B**) are **spatially close** to each other but **hard to access** for drivers.

New Mode: Meeting Points

Route planning with Meeting Points

- Meeting points (MP for short) are introduced as alternative locations for pick-up/drop-off locations of requests.
- E.g., driver and riders now **meet at D**.
- Short walk $(A \rightarrow D)$ of riders, large overall profit!

Problem of Ridesharing with Meeting Point (MP)

- Existing researches [1, 2] for MP are offline
	- **Inefficiency**: cannot serve **large-scale online** applications
- MP is not well-explored in the industry
	- Express Pool (Uber) encourages riders to walk to Express spots (meeting points) for efficient routing
	- **Inflexible**: **wait** until a group of requests has a shareable route and pick up them **together** like at a bus station [3]

Motivation

- Some vertices are more convenient to come and go and thus "**popular**"
	- E.g., vertices close to highways and downtown

• With flexible MPs, it is possible to **serve more** requests at or near those "**popular**" vertices, which makes them even **more frequently** used.

- These vertices serve as the **skeleton** of the road network
	- **Effectiveness**: estimate and select these popular vertices
	- **Efficiency**: fast algorithms especially on popular vertices

Motivation

• The requirement for a **road network skeleton** motivates us to take advantage of k-skip cover V^* [1], which is a selected subset of vertices to be the **skeleton** of a graph G.

• We call a vertex set V^* k-skip cover if for any shortest path of length k on a graph, there is at least one of its vertices $\in V^*$.

• To minimize the size of V^* , we need to find the most "popular" and convenient vertices, which frequently appear in shortest paths, which coincide with our requirement for meeting points.

Outline

- Background and Motivation
- The Meeting-Point-based Online Ridesharing Problem
- Framework Overview
- Methods
- Experimental Evaluation
- Summary

- **Drivers**
	- A set of *n* drivers $W = \{w_1, w_2, ..., w_n\}$
	- Each is defined by $w_i = \langle l_i, a_i \rangle$ with current location l_i and capacity limitation a_i

- Drivers
	- A set of *n* drivers $W = \{w_1, w_2, ..., w_n\}$
	- Each is defined by $w_i = \langle l_i, a_i \rangle$ with current location l_i and capacity limitation a_i
- Requests
	- A set of *m* requests $R = \{r_1, r_2, ..., r_n\}$
	- Traditionally, each is defined by $r_j = \langle s_i, e_i, tr_j, tp_j, td_j, p_j, a_j \rangle$, where:
		- s_i/e_i for source/destination locations;
		- $tr_j/tp_j/td_j$ for time of release/pick-up deadline/drop-off deadline;
		- p_j for rejection penalty;
		- a_j for capacity.

13

- Traditional route planning:
	- Assign each driver w_i a **route** S_i **, which is a sequence of** s_j/e_j under the time and capacity constraints.
	- Minimizing a unified cost of:

- With **meeting point**, an assigned request has $\langle pi_j, de_j, wp_j, wd_j \rangle$ in addition, where:
	- pi_j/de_j for pick-up and drop-off locations
	- *wp_j, wd_j* for time of riders **walking** before picked up and after dropped off.

- With **meeting point**, an assigned request has $\langle pi_j, de_j, wp_j, wd_j \rangle$ in addition, where:
	- pi_j/de_j for pick-up and drop-off locations
	- wp_j , wd_j for time of riders walking before picked up and after dropped off.
- **Meeting-Point**-based route planning:
	- Assign each driver w_i a route S_i , which is a sequence of $s_i/e_i p_i$ j/de.
	- Minimizing a unified cost of:

cost of routes

 $\overline{r_j\in}\bar{\mathsf{R}}$ The penalty for rejected requests

The walking cost of requests

 $\sqrt{\frac{1}{11}}$ We prove the MORP problem is **NP-hard** by reducing it from the basic route planning problem[1] for $\begin{array}{c} \n\text{else} \\
 \hline\n \text{else} \\
 \end{array}$ shareable mobility services.

furthor chough that no do We further show that **no** deterministic nor randomized algorithm can guarantee a **constant Competitive Ratio.**

• With **meeting point**, an assigned request has [1] Yongxin Tong et al. 2018• /

- **Meeting-Point**-based route planning:
	- Assign each driver w_i a route S_i , which is a sequence of $s_i/e_i p_i$ j/de.
	- Minimizing a unified cost of:

for pick-up and drop-off locations and drop-off locations and drop-off locations of

The penalty for rejected requests

in addition, where:

The walking cost of requests

Outline

- Background and Motivation
- The Meeting-Point-based Online Ridesharing Problem
- Framework Overview
- Methods
- Experimental Evaluation
- Summary

Existing method - Insertion

Insertion is an effective local optimal algorithm for route planning, which has linear ($O(|S_i|)$) time complexity [1]. It involves **shortest path queries** between the current route (S_i) and inserted $\hbox{locations }(s_j,e_j).$

[1] Yongxin Tong et al. 2018

By adapting insertion for MORP, it involves **shortest path queries** between the current route (S_i) and all possible pairs of meeting points near (s_j, e_j) .

 If a vertex has k meeting points on average, the computational cost increases by $k \times k$ times, which is unacceptable.

Selecting optimal meeting points (MPs) **online** is time-consuming

1) Prepare MPs for each vertex **offline** to reduce the search space Design data structure for faster queries.

Outline

- Background and Motivation
- The Meeting-Point-based Online Ridesharing Problem
- Framework Overview
- Methods
- Experimental Evaluation
- Summary

Select Meeting Point Candidates

- Quantize how "convenient" a vertex is for transportation
	- MP candidates should easily get to and conveniently reach other vertices.
	- Given vertex *u*, define its n_r nearest vertices as reference vertices $n_o(u)$.
		- Equivalent Out Cost of *u*: the average distance towards its reference vertices

$$
ECO(u) = \frac{\sum_{v \in n_o(u)} SP_c(u, v)}{n_r}
$$

• Similarly, reverse the graph we can have **Equivalent Inward Cost**

$$
ECI(u) = \frac{\sum_{v \in n_i(u)} SP_c(v, u)}{n_r}
$$

Select Meeting Point Candidates

- Quantize how "convenient" a vertex is for transportation
	- MP candidates should easily get to and conveniently reach other vertices.
	- Now we can rank the candidate MPs $\{v_1, v_2, ..., v_n\}$ of a vertex u by Serving-Cost Score

$$
SCS(u, v_i) = \boxed{\beta \cdot SP_p(u, v_i)} + \boxed{\alpha \big(ECI(v_i) + ECO(v_i) \big)}
$$

Expected cost from walking

Expected cost from driving

Based on these statistics from shortest path queries, an $O(|V|)$ Local-Flexibility-Filter Algorithm is proposed to select MPs for each vertex **offline**.

- With MPs, assigned routes can be concentrated on the convenient vertices
- We give the hierarchical order over the vertex set V :
	- **Defective vertices** V_{de} They are inconvenient to access.
	- **Core vertices** V_{co} They are used as MPs frequently.
	- **Sub-level vertices** V_{su} The remaining vertices are classified as sub-level vertices.

Defective vertices

- We aim to remove the unwelcomed vertices in traditional ridesharing, which can be alternatively served by meeting points now.
- We propose a method to avoid two potential costs from vertex removal:
	- **The detour cost** A path containing removed vertex *u* no longer exists.
	- **The inaccessibility cost** The potential reject penalty of requests at the removed vertex.

Defective vertices

• A 3-phase $O(NlogN)$ DVS algorithm is proposed with theoretical guarantee

LEMMA 6.2. Removing all vertices selected by the DVS algorithm from G_c with their edges leads to no detour cost.

LEMMA 6.3. $\forall u \in V$ is accessible after removing vertices selected by the DVS algorithm from G_c with MPs.

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.

A graph and its **2-skip cover** ∗ (Any shortest path longer than **2** contains at least one of its vertex)

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.

A graph and its **2-skip cover**

A car picks up 2 requests in the **traditional** mode

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.

A graph and its **2-skip cover**

A car picks up 2 requests in the traditional mode

Both of *k*-skip cover and meeting **points** expect "convenient" vertices A car picks up 2 requests in the **meeting point** mode (meet at v_{23})

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.
	- V_{co} is a *k*-skip cover on the updated graph without V_{de}
	- Proportion factor ϵ of vertices have at least one vertex $u \in V_{co}$ as its MP candidate

Meeting point -> more queries between k -skip cover

*k***-skip cover** \rightarrow faster inner queries to improve efficiency

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.
	- V_{co} is a *k*-skip cover on the updated graph without V_{de}
	- Proportion factor ϵ of vertices have at least one vertex $u \in V_{co}$ as its MP candidate

34

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.
	- V_{co} is a *k*-skip cover on the updated graph without V_{de}
	- Proportion factor ϵ of vertices have at least one vertex $u \in V_{co}$ as its MP candidate

Core vertices

- Select "convenient" vertices as the skeleton of the graph
- k -skip cover has good compatibility with the meeting point: use it as backbone.
	- V_{co} is a *k*-skip cover on the updated graph without V_{de}
	- Proportion factor ϵ of vertices have at least one vertex $u \in V_{co}$ as its MP candidate

Graph construction

- Vertices: previously obtained V_{de} , V_{co} , V_{su}
- Edges: since V_{co} forms a k-skip cover, we can build super edges among $V_{co} \bigcup V_{su}$ following existing theory [1]:
	- E_{cc} : super edges between core vertices
	- E_{cc} : super edges from core to sub-level vertices
	- E_{cc} : super edges from sub-level to core vertices

Graph construction

- Vertices: previously obtained V_{de} , V_{co} , V_{su}
- Edges: since V_{co} forms a k-skip cover, we can build super edges among $V_{co} \bigcup V_{su}$ following existing theory [1]:
	- E_{cc} : super edges between core vertices
	- E_{cc} : super edges from core to sub-level vertices
	- E_{cc} : super edges from sub-level to core vertices

Recall that we select MP Candidates $MC(u)$ **for each vertex** u

- Vertices∈ $MC(u)$ are reachable via short walking \Rightarrow they are close to each other
- One interesting problem is, if inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?

Recall that we select MP Candidates $MC(u)$ **for each vertex** u

- Vertices $\in MC(u)$ are reachable via short walking \Rightarrow they are close to each other
- One interesting problem is, if inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- ⇒ We define a new distance correlation, which **bounds the time saving** of switching from v to any other vertices.
- If deducting the saving still cannot meet the time limitation:
	- \Rightarrow prune the whole set $MC(u)$!

Recall that we select MP Candidates $MC(u)$ **for each vertex** u

- Vertices∈ $MC(u)$ are reachable via short walking \Rightarrow they are close to each other
- One interesting problem is, if inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- ⇒ We define a new distance correlation, which **bounds the time saving** of switching from v to any other vertices.
- If deducting the saving still cannot meet the time limitation:
	- \Rightarrow prune the whole set $MC(u)$!

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{32}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{23}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

 $\mathbb{E}(v_{22})$ / Traditionally, we need to derive **all** the time costs from graph to v_{22} and v_{23} though they are **close** to each other

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{23}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{23}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{23}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{23}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.
	- A driver want to serve a request at v_{22} , which has MP candidates $\{v_{22}, v_{23}\}$
	- If v_{22} exceeds the time limitation for 3 minutes, do we still need to test v_{32} ?

- If inserting a candidate $v \in MC(u)$ fails to meet the time limitation, do the rest candidates help?
- **Bounds the time saving** of switching from v to any other vertices.

- Design SMDBoost algorithm.
- For each pair of driver and request, we test one vertex for insertion and prune the rest if the time limitation cannot meet with the bounded saving.

```
Algorithm 2: SMDBoost
  Input: a driver w_i with route S_{w_i}, request r_i, MP candidate set
          MC, set maximum difference SMD, checker set Ch, dead
          vertices DV
  Output: a route S_w^* for the driver w and updated DV
1 if Driver's location l_i \in DV then
      Return S_{w_i} and DV without insertion
 \mathbf{2}3 Generate arriving time \arg[\cdot] for S_{w_i}4 Collect all sub-level vertices which have super-edges to vertices in
    MC(s_i) into set Ne
5 The largest index to insert pick-up: id^* = |S_{w_i}|6 foreach v \in S_{w_i} do
       if v \in Ne then
            Continue
 8
       if \text{arv}[v] + SP_h(v, Ch(s_j)) - SMD(Ch(s_j)) \ge tp_j then
9
            if v=l_i then
10
               Add l_i to DV. Insertion fails and returns Null
11
           Record id^* = idx(v) - 112
           Break
13
14 Insert r_i with adapted insertion algorithm where insertion indexes
    of pick-ups larger than id^* are pruned.
15 return S^*_{\mathbf{w}}, DV
```
Outline

- Background and Motivation
- The Meeting-Point-based Online Ridesharing Problem
- Framework Overview
- Methods
- Experimental Evaluation
- Summary

- Road Network
	- NYC ($|V|=57,030, |E|=122,337$)
- Real-World Dataset
	- Taxi Trips (yellow and green) in NYC (277,410 trip records)
- Synthetic Dataset
	- Generated according to the distribution of NYC (100k to 1m trip records)

- Compared parameters
	- \cdot e_r : the deadline coefficient.
	- a_w : the capacity of workers.
	- \cdot α : the weight for driving cost.
	- \cdot β : the weight for walking cost.
	- p_o : the ratio of penalty cost
	- W|: number of workers
	- : number of requests

- Tested Algorithms
	- Traditional
		- **GreedyDP [1]**: the state-of-art route planning algorithm using insertion. No demandrelated information is used.
		- **Kinetic Tree [2]**: it saves all the possible routes for the assigned request using a structure called Kinetic and inserts requests by traversing and updating the tree.
	- Meeting-Point-Based
		- **BasicMP**: It is an extension from GreedyDP by adapting MPs to solve the MORP problem.
		- **First Serve.** A variant of BasicMP, where each request is directly assigned to the first driver who can serve it.
		- **HSRP**. It uses the HMPO Graph to improve the effectiveness of BasicMP without pruning.

Performance of varying number of workers $|W|$

Performance of varying number of requests $|R|$

Outline

- Background and Motivation
- Framework Overview
- The Cache Replacement Problem
- Theoretical Guarantees
- Experimental Evaluation
- Summary

Summary

- We formulate the online route planning problem with MPs mathematically, namely MORP. We prove that it is NP-hard and has no algorithm with a constant competitive ratio.
- We propose an algorithm to select MP candidates for riders, which is based on a unified cost function considering the travel cost from additional walking.
- We propose a novel hierarchical structure of the road network, namely hierarchical meeting-point oriented (HMPO) graph, to fasten the solution for MORP.
- Based on the HMPO graph, we propose an effective and efficient insertor, namely SMDB, to handle the requests in MORP.

Thank You!

The code and datasets https://github.com/dominatorX/open.